Effects of Mergers in Two-sided Markets: Examination of the U.S. Radio Industry *

Przemysław Jeziorski †

October 30, 2012

Abstract

This paper studies mergers in two-sided markets by estimating a structural supply-and-demand model using data from the 1996-2006 merger wave in U.S. radio. It makes two main contributions. First, it identifies the conflicting incentives of merged firms to exercise market power on both sides of the market (listeners and advertisers). Second, it disaggregates the effects of mergers into changes in product variety and changes in supplied ad quantity. I find that between 1996 and 2006 listener welfare increased by 0.2% (+0.3% from extra variety, -0.1% from changes in ad quantity) and advertiser welfare decreased by 21% per-year (it is composed of 17% drop from variety changes, and extra 5% drop from ad quantity adjustments).

*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.

†Haas School of Business, University of California at Berkeley

1 Introduction

Between 1996 and 2006, the U.S. radio industry experienced an unprecedented merger wave. This merger wave was prompted by the 1996 Telecommunication Act, which raised ownership caps in local markets and abolished cross-market ownership restrictions. At the height of merger activity, more than 20% of stations changed ownership each year and about 14% changed programming format. I use this merger wave to study the consequences of consolidation in two-sided markets, and 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). Specifically, 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. Because similar issues arise in other two-sided markets, such as credit cards, newspapers, or computer hardware, the framework put forward in this paper can be adjusted to analyze these and similar industries.

In two-sided markets, firms face two interrelated demand curves from two distinct types of consumers. In the case of a merger, these demand curves generate conflicting incentives. Namely, exercising market power on one side of the market lowers profits on the other side. In the case of radio, a company provides free programming to listeners and draws revenue from selling advertising, which is priced on a per-listener basis. On the listener side, a merged firm would like to increase post-merger advertising, because it captures some switching listeners. This extra advertising decreases the welfare of listeners but increases the welfare of advertisers. However, from the perspective of the advertising market, increasing ad quantity would lower prices; thus the merged firm might like to supply less advertising. Such a decision would have the opposite impact, raising listener and lowering advertiser welfare. The firm's ultimate decision depends on the relative demand elasticities in the two markets. In this paper, I separately identify these two elasticities, and subsequently compute the impact of the 1996 deregulation on listener and advertiser welfare.

Post 1996, consolidation in radio was characterized by a set of patterns in the data, which support inconsistent predictions about the aforementioned anticompetetive effects on both sides of the market. Namely, during the 1996-2006 period, aggregate advertising quantity increased by about 10%, and the average time spent listening to the radio declined by about 15%.[1] These results suggest that market power could have been exercised on listeners. However, at the same time, we observe a sharp 40% increase in advertising prices per-listener, which suggests exercising market power on advertisers. The potential reasons for this conflicting evidence include exogenous trends in radio listenership (demographic and technological changes), macroeconomic trends in the advertising markets, and economies of scale in supplying advertising. One of the goals of this paper is to develop and estimate a structural model of demand and supply of radio programming and advertising that can control for these confounding factors. After estimating the model, I isolate the impact of market structure on consumer welfare using counterfactual experiments.

I find the merger wave results in an 11% drop in ad quantity and a 6% increase in prices, which translates into an aggregate listener welfare gain of 0.2% and a 21% advertiser welfare loss. By comparison, two reduced form studies by Brown and Williams (2002) and Chipty (2007) find respectively 4% price decrease, and no systematic impact of mergers on ad prices. Additionally, I find market power on the listener side is similar across geographical markets, whereas the amount of market power on the advertiser side depends on market population. In particular, firms have considerable control over advertising prices in smaller markets and less control in larger markets. As a result, the firms disproportionally decrease the supply of advertising in smaller markets. Such behavior extracts a higher fraction of advertiser surplus in these markets (a 32% drop in advertiser surplus in markets with less than 0.5M population, compared to a 17.1% drop in the markets with more than 2M population).

Apart from quantifying the effects of mergers on both sides of the market, the second contribution of this paper is the decomposition of these effects into changes in product variety and extra market power from joint ownership. This exercise is motivated by the fact that in most cases, consumers have a preference for variety, so the extra variety created by mergers might mitigate the negative effects of extra market power. To verify the above claim, I quantify consumer 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 the tastes of yogurt, and Brynjolfsson, Hu, and Smith (2003) who per-

[1] Source: Time Spent Listening by Season (Hours and Minutes per Week), Mon-Sun 6AM-Mid, Total U.S., 2006-1997, Radio Today Report by Arbitron.

form a similar exercise for the variety of books offered in on-line bookstores. These papers assume away the fact that changes in variety are followed by readjustments in equilibrium prices. In this paper, I take their analyses one step further: I incorporate such strategic responses by performing counterfactual experiments in which new equilibrium prices are computed.

Berry and Waldfogel (2001) and Sweeting (2009) 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 0.3% increase in listener welfare. However, because product repositioning softened competition in the advertising market and caused some stations to switch to a "Dark" format[2], advertiser welfare decreased by 17% per year. Additionally, I find that ownership consolidation and product repositioning are followed by advertising quantity readjustments. I estimate this effect leads to a 0.1% decrease in listener welfare and 5% decrease in advertiser welfare. The two effects combined total a 0.2% increase in listener welfare and a 21% decrease in advertiser welfare. Although extra variety mitigates the negative effects of mergers on listeners, it strengthens the negative impact on advertisers.

My work is related to several theoretical papers studying the 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 a standard supply-and-demand framework of single-sided markets might not be sufficient to capture the economics of two-sided markets. Additionally, several empirical studies have examined 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. In particular, I include 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 programming. 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.

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

Few other papers use structural models to analyze two-sided markets. For example, Song (2011) studies magazines , whereas Van Cayseele and Vanormelingen (2009) and Fan (2010) study newspapers. An additional complication for analysis of these industries is a presence of two prices instead of a single price, as in radio. The main difference between this work and Song (2011) as well as Van Cayseele and Vanormelingen (2009) is these papers treat advertisers as local monopolists. In particular, they assume away the potential impact of mergers on the market power in the advertiser side of the market. These papers find an ambiguous or negligible impact of consolidation on both sides of the market. In contrast, I find a large impact on advertisers and a low impact on listeners. This paper's analysis complements Fan (2010), who allows for a richer model of characteristics choice but does not observe advertising quantities. Therefore, she can endogenize the variety in the counterfactuals; however, she is not able to study post-repositioning quantity readjustments. Fan finds that the mergers decreased reader surplus, whereas I find that the listener surplus increased because of lower post-merger advertising quantity. Finally, we both find that mergers decrease advertising surplus.

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 describes the data. Section 4 outlines estimation techniques used to identify the parameters of the model. Section 5 presents the results of the structural estimation. Section 6 describes the results of counterfactual experiments. Section 7 contains the robustness checks of different modeling assumptions. Section 8 concludes.

2 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 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 the quantity one consumer buys determines the market clearing price for the other consumer. In this subsection, I argue that this feedback complicates the process of determining whether the supplied quantities are strategic substitutes or complements (as defined in Bulow, Geanakoplos, and Klemperer (1985)). As a result, this feedback 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, three types of agents are present: 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

$$r_j = r_j(q|s, d, \theta^L), \tag{2.1}$$

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. Because radio programming is free, this equation does not contain price explicitly. However, because listeners have disutility for advertising, its effect is similar to price; that is, $\frac{\partial r_j}{\partial q_j} < 0$.

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

$$p_j = p_j(q, r_j(q)|s, d, \theta^A), \tag{2.2}$$

where θ^A are parameters[3]. 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 no marginal cost exists (I relax these assumptions later). In equilibrium, each radio station chooses its optimal ad quantity, keeping the quantities of the other stations fixed, namely,

$$\max_{q_j} p_j(q, r_j(q)|q_{-j})q_j. \tag{2.3}$$

[3]Note the equation (2.2) implicitly leverages on the existence of the unique inverse demand (market clearing prices) for company j as a function of all quantities q and a market share r_j. A constructive proof of existence for my model is given in the Appendix B. The model is tailored to the radio industry and supports pricing-per-listener in the equilibrium (which we observe in all U.S. radio markets). For the more general result, which allows prices to depend on the full vector r, see White and Weyl (2010).

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 model (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.

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 these strategic interactions ambiguous, 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. Less advertising quantity benefits listeners but hurts advertisers. However, if the feedback loop is stronger than the direct effect then mergers lead 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 (2.1), (2.2) and (2.3), 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.

2.1 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 t in market m is therefore fully characterized by a station/format configuration and ownership structure s^{tm}, a 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.

2.2 Listeners

This subsection describes the details of the demand for listenership introduced in equation (2.1). The model is a variation on the random coefficient discrete choice setup proposed by Berry, Levinsohn, and Pakes (1995).

I assume each listener chooses only one radio station to listen to at a particular moment. Suppose 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 $\iota_{jt} = (0, \ldots, 1, \ldots, 0)$ where 1 is placed in a position that indicates the format of station j. Also, let FM_j be a dummy equal to 1 if station j is an FM station and 0 otherwise.

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

$$u_{ijt} = \theta^L_{1it}\iota_j - \theta^L_{2it}q_{jt} + \theta^L_3\mathrm{FM}_j + \xi_{jt} + \epsilon_{jit} \tag{2.4}$$

where θ_{2it}^L is the individual listener's demand sensitivity to adverting, q_{jt} the amount of advertising, ξ_{jt} the unobserved station quality, ϵ_{jit} an unobserved preference shock (distributed type-1 extreme value), and θ_{1it}^L is a vector of the individual listener's random effects representing preferences for formats.

I assume the random coefficients can be decomposed as

$$\theta_{1it}^L = \theta_1^L + \Pi D_{it} + \nu_{1it}, \quad D_{it} \sim F_{mt}(D_{it}|d), \quad \nu_{1it} \sim N(0, \Sigma_1)$$

and

$$\theta_{2it}^L = \theta_2^L + \nu_{2it}, \quad \nu_{2it} \sim N(0, \Sigma_2),$$

where Σ_1 is a diagonal matrix, and $F_{mt}(D_{it}|d)$ is an empirical distribution of demographic characteristics. Because $F(\dot{})$ is allowed to vary by time and market, it captures trends in demographics that can affect profitability of mergers. The term ν_{it} represents an unobserved taste shock for formats, and Π is the matrix representing the correlation between demographic characteristics and format preferences. I assume draws for ν_{it} are independent across time and individuals.

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

$$\delta_{jt} = \delta(q_{jt}|\iota_j, \xi_j, \theta^L) = \theta_1^L \iota_{jt} - \theta_2^L q_{jt} + \theta_3^L \mathrm{FM}_j + \xi_{jt}$$

an interaction part

$$\mu_{jit} = \mu(\iota_{jt}, q_{jt}, \Pi D_{it}, \nu_{it}) = (\Pi D_{it} + \nu_{1it})\iota_{jt} + \nu_{2i} q_{jt}$$

and error term ϵ_{jit}.

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,

$$r_{jt}(q_t|s_t, \xi, d, \theta^L) = \int\int \frac{\exp[\delta_{jt} + \mu_{jit}]}{\theta^{Lt} + \sum_{j' \in s} \exp[\delta_{j't} + \mu_{j'it}]} dF(\nu_{it}) dF_{mt}(D_{ti}|d),$$

where θ^{Lt} is an exponent of the utility of not listening to radio. This specification allows for a trend in outside option than can be caused by an expansion of internet and satellite radio technologies.

2.3 Advertisers

In this subsection, I present the details of the demand for advertising introduced in equation (2.2). The model captures several important features specific to the radio industry. In particular, the pricing can be approximated by per-listener rates, so the price for a 60sec slot of advertising is a product of cost-per-point (CPP) and station rating (market share in percents). In reality, the approximation is not exact. For example, prices are negotiated ex-ante and depend on forecasts of ratings as opposed to actual realized ratings. If the realized ratings are much smaller than predicted ratings, many radio stations would compensate the advertisers with discounted rates, or by offering extra offsetting advertising spots. Thus, effectively, I am assuming forecasting error is negligible, or the offsets take care of the difference.

Radio stations have a direct market power over advertisers, such that CPP is a decreasing function of the ad quantities offered by a station and its competitors. The p_{jt} is the cost-per-point at station j, so the price of a 60sec slot is given by $p_{jt}r_{jt}$. The simplest model that captures the above features and is a good approximation of the industry is a linear inverse demand for advertising, such as

$$p_{jt} = \theta_1^{Am}\left(1 - \theta_2^{Am}\sum_{f'\in\mathbb{F}}\omega_{ff'}^{m}q_{f't}\right), \tag{2.5}$$

where f is a format of station j, θ_1^A is a scaling factor for the 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'. Unobserved market-level heterogeneity is captured by the fact that θ_1^{Am} is allowed to be different for each market and θ_2^{Am} is allowed to differ between subsets of markets depending on their size.

The slope parameter θ_2^{Am} measures the amount of changes in per-listener cost in response to changes in the supplied advertising quantity weighted by the competition factors ω. This response in per-listener price is on top of the response in ratings that determine full 60sec price $p_{jt}r_{jt}$. Small values of θ_2^{Am} would indicate that the advertisers are price sensitive, or elastic. In such cases the radio station has difficulty control the market-clearing per-listener price, which translates to low market power in the advertising marker. In the converse case of large θ_2^{Am}, advertisers are inelastic, which generates extra market power for radio stations.

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, doing so here is not feasible because the available data do not contain enough variation in radio-station-level advertising prices. Instead, I make additional assumptions that enable me to compute the weights using publicly available data. The remainder 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; that is, an 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. The fact that most of the advertising is purchased by small local firms this assumption. 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

$$\omega_{ff'} = \frac{1}{\sum_{a \in \mathcal{A}} r_{a|f}^2} \sum_{a \in \mathcal{A}} r_{a|f} \left(r_{a|f} r_{f'|a} \right). \tag{2.6}$$

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 conditional probabilities of advertisers' arrival, which are equal to the conditional probability of listeners' arrival $r_{a|f}$. 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: (i) the change is proportional to the probability of correct targeting in format f, given by $r_{a|f}$, because advertisers are expected utility maximizers; and (ii) the change 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 (2.6).

To illustrate how these weights work in practice, consider the following example. Suppose only two possible formats of programming are available: Talk and Hits; and two types of consumers are present: Teens and Adults. Teens like the Hits format and Adults like the 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 5.

In Table 5, the impact of Hits on the price of Talk is greater than the impact of Talk on the price of Hits, because 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, the own effect of Talk must be weaker than that of Hits. This example illustrates the essence of the mechanism behind equation (2.6). More examples from the data with an extensive discussion are given in section 5.

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

2.4 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 (2.3) that allows for non-zero cost in selling advertising and common radio station ownership. Given the advertising quantity choices of competing owners q_{-kt}, the profit of radio-station owner k is given by

$$\begin{aligned}\bar{\pi}_{kt}(q_{kt}|q_{-kt},\xi,\theta) &= \max_{\{q_{jt};j\in s_{kt}\}} \sum_{j\in s_{kt}} r_{jt}(q_t|\xi,\theta^L)p_{jt}q_{jt} - \mathrm{C}_{jmt}(q_{jt}|\theta) = \\ &= \max_{\{q_{jt};j\in s_{kt}\}} \sum_{j\in s_{kt}} q_{jt}r_{jt}(q|\xi,\theta^L)\theta_1^{Am}\left(1-\theta_2^{Am}\sum_{f'\in\mathbb{F}}\omega_{ff'}^m q_{f't}\right) - \mathrm{C}_{jmt}(q_{jt}|\theta).\end{aligned} \tag{2.7}$$

where $C_j(q_j)$ is the total cost of selling advertising. Advertising marginal cost is likely to be non-zero because many advertising sellers work on commission-type compensation schemes. For the purpose of this paper, I assume sellers are compensated per minute of advertising sold. Because the station's revenue per minute of advertising depends on the quality of its programming, the sellers' commission is also likely a function of the station's quality. According to equation (2.7), revenue is a linear function of θ_1^A, which is also likely to affect the sellers' commission. Another source of marginal cost is the in-house production of advertising. In contrast to newspapers and TV, radio stations manufacture many of their ads. In-house ads might take the form of voice announcements, endorsements by one of the radio hosts, or more elaborate spots which incorporate music and multiple voices. Production of these spots is a part of the advertising service and its costs should be proportional to the length of the spot and its quality.

The simplest marginal cost specification that captures the above features and allows for station-level unobserved heterogeneity can be summarized by the following equation:

$$\mathrm{C}_{jmt}(\theta^A, \theta^C) = \theta_1^{Am}[\theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Cm}\xi_{jt}, +\theta_3^{Cm}\mathrm{SYN}_{jt} + \eta_{jt}]q_{jt}$$

This specification additionally allows station-level unobserved heterogeneity captured by η_{jt}. Terms θ^{Cmt} are time dummies capturing aggregate shocks to marginal cost. Unobserved market level heterogeneity is captured by the fact that θ_1^{Am} is allowed to be different for each market and θ^{Cm} is allowed to differ between subsets of markets depending on their size. The parameter θ_3^{Cm} measures the extent of marginal cost synergies between stations of the same format owned by the same owner. Such cost synergies occur because some of the ads with similar target groups are sold jointly and by the same sales agent, or generate economies of scale in manufacturing. The term θ_3^{Cm} is multiplied by a dummy variable SYN_{jt} that is equal to 1 if the current owner owns more than one station in the same format.

The formulation 2.7 does not allow cross-station bundling of advertising slots, which could be a potential problem if the radio stations offered a large amount of exclusive packages of ads across different stations that cannot be unbundled. There are several reasons why Such bundling should not be an issue for several reasons. Primarily, these kind of translations are likely taking place with large advertisers, but not with smaller clients. Thus, because most of the clients are small, bundling could affect only a fraction of translations. Additionally, I find little discussion of bundling in the regulatory reports that investigated a couple of the larger mergers in the industry, which suggests such practices are not of first-order importance. Finally, bundling is more likely to take place within than across formats, since the sets of advertisers across formats usually hardly overlap. In such case, bundling would be picked up by the cost synergy SYN_{jt} dummy and netted out in the counterfactual. However, the welfare interpretation of SYN_{jt} might change. Thus, this paper asserts that cost synergies have a larger effect on the equilibrium outcomes and in practice assumes away bundling.

I assume the markets are in a Nash equilibrium. The first-order conditions for profit optimization become

$$r_{jt}p_{jt} + \sum_{j' \in s_{kt}} q_{j't} \left[\frac{\partial r_{j't}}{\partial q_{jt}} p_{j't} - r_{j't}\theta_2^{Am}\omega_{jj'}^m \right] - \theta^{Cmt} - \theta_1^{Cm} - \theta_2^{Cm}\xi_{jt} - \theta_3^{Cm}\text{SYN}_{jt} - \eta_{jt} = 0. \quad (2.8)$$

This formulation assumes away strategic timing of advertising between stations. This assumption is necessary because the exact time of airing an ad is unobserved. For a comprehensive study of that aspect of the radio market, see Sweeting (2010).

Finally, I assume station unobserved quality is exogenous but serially correlated, similar to Sweeting (2011). It evolves according to an AR(1) process such that

$$\xi_{jt} = \rho\xi_{jt-1} + \zeta_{jt}, \quad (2.9)$$

where ζ_{jt} is an exogenous innovation to station quality.

3 Data description

I have constructed a panel of data on radio stations and radio-station ownership, merging data from two sources: *BIA Kelsey* and *SQAD Media Market Guide*.

BIA Kelsey provided 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 (2011),only 50 stations 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.

An observation in my data is a radio station operating in a specific half-year and in a specific market. BIA 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.” Taking into account that most of the advertising is local, I assume no interdependence between markets. To further assure no overlap between markets, I use only the 88 market sub-selection developed in Sweeting (2011). Table 11 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 [4]. 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.

3.1 Advertising prices and quantities

The main dependent variables in this paper are advertising prices and quantities. This subsection aims to provide detailed information about these variables. Because the direct data on station advertising quantities are unavailable, the advertising quantities are imputed from station-level revenues, station ratings, and per-listener ad prices. Below I describe the strengths and weaknesses of these three data sets, that directly translate into the strengths and weaknesses of the imputed quantity data.

I obtain station-level revenues from BIA Kelsey, a large consulting company dealing with the radio markets. The BIA data set on revenues is considered the most complete and reliable resource of radio station financial performance. The data is compiled from mail and telephone surveys of radio stations, which account for local, regional, and national advertising sales. Barter and production revenues are not included. BIA data is a balanced 1996-2006 panel including all active and inactive radio stations in the United States as of 2006. Because the survey data is self-reported, it might suffer from self-reporting and self-selection bias. To solve some of these problems, BIA corrects the numbers by a direct consultation with radio partners and by applying proprietary statistical tools. For example, the missing responses are imputed using historical data and estimated industry trends. Numerous academic papers use this data, including Sweeting (2011), Mooney (2010), Waldfogel and Wulf (2006). However, any systematic measurement error that is correlated with the changes in the market structure could potentially affects the results presented in this paper. In particular, the revenues of larger stations and stations that change ownership (these two sets largely overlap) must be measured with a reasonable accuracy. For this

[4] Source: Arbitron Format Trends Report

reason, I have obtained verbal reassurance from BIA that the missing data points are predominantly small stations with low revenue shares. Moreover, BIA has a long-term relationship with large owners, many of which are publicly traded; thus, the most relevant self-reported values are likely to be reliable.

The second source of the data are station-level market shares (ratings) coming from Arbitron, which is the biggest consulting company reporting on radio markets. The company has been in the radio business for 60 years and its ratings support $19 billion of transactions every year. The ratings are measured using diary based surveys of the representative population. The survey panel changes periodically, and the results are averaged quarterly. These manipulations are aimed to reduce the individual-level reporting bias and diversify the population to net out unobserved heterogeneity. Because the data is self-reported and the panel of respondents varies, one might worry about the comparability of responses between time periods. However, the qualitative conclusions of this paper would not be affected if the relevant measurement error is uncorrelated with changes in the ownership. In such a case, the reporting bias could affect nominal values of consumer welfare. The percentage changes in welfare should be less susceptible.

The third source of the data are average market prices per listener coming from SQAD. The prices come directly from advertising agencies that report actual transactions. The upside of this collection procedure is that it is not based on surveys and is not adjusted by proprietary formulas. The downside is that one cannot directly observe station-level prices. The data is organized by demographics, which means that I observe price per listener separately for different gender and age groups. Using these numbers and the ratings of stations across different demographic group, I can impute station-level prices. The assumption behind these imputations is that listeners' demographics are a main driver of station-level per-listener prices. I chose this method because it is known in the industry that precise targeting on demographics is a strength of radio as an advertising medium. Such targeting involves sorting high- and low-value consumers, which drives the differences in per-listener advertising prices across stations.

Combining the above three numbers for each station – revenue, rating and per-listener ad price – I can impute the advertising quantity. For the results in this paper to be meaningful, these quantities must respond to the changes in the market structure. I investigate this issue in the next section, using descriptive analysis.

3.2 Descriptive analysis

Table 1 contains basic aggregate statistics. The mean advertising quantity is 37.5 minutes per day. The standard deviation is 40 and the median is 28.5. It shows the sample is fairly dispersed and many, usually niche, stations broadcast small amounts of advertising. Because the average number of ads might seem fairly small, I obtained additional data to validate it. The data comes from the Federal Communication Commission Research Studies on Media Ownership. This FCC study directly categorized the programming content of a subset of stations. It consists of $1,000$ stations in the year 2007. Each station was recorded for two hours, and each five seconds was labeled as ad/non-ad content. This database might be a noisy measure of the station-level advertising quantity because the intensity of advertising varies significantly during the day and across days. However, because the study is fully randomized, the aggregate statistics are unbiased. For example, I computed that for my sub-selection of markets, according to the FCC database, stations play about 23 minutes of advertising each day, which is slightly lower than the average in my sample but is reasonably close. I can account for some the difference by the fact that my 37.5 average does not include stations that did not meet Arbitron measurement standards (rating below 0.5%), and these stations are likely to play a negligible amount of advertising. If I include those stations, my average gets closer to the FCC number and drops to 28 minutes per day.

In addition to ad quantities, I report statistics on revenue, station market share in the listenership market, and station power. All these variables are fairly dispersed and generally skewed to left.

Table 2 documents changes in the concentration of radio-station ownership. The average number of stations owned in our dataset grew from 1.64 in 1996 to 2.41 in 2006. It was computed on the market level and averaged across markets. This ownership consolidation resulted in a growth of the market share of the three biggest owners (C3) from 52% in 1996 to 62% in 2006, peaking at 64% in 2000. The middle part of the table contains the average percentages of stations that switched owners and switched formats. Between 1996 and 2000, more than 10% of stations switched owners yearly. After 2000, the number dropped to below 10%. Greater concentration activity in the 1996-2000 period was also associated with more format switching. The percentage of stations that switched formats peaked in 1998 at 14%.

Next, in Table 3, I report statistics on interactions between acquisitions and ad quantity. I computed an average change in station-level ad quantity after the merger and I compare it to the overall trend. I find that for an average station, the quantity of advertising goes up by 0.3 minutes per day, from one half-year to the next. However, for recently acquired stations, ad quantity goes down by 0.7 minutes. Additionally, the effect is stronger if one looks one year forward. That is, an average acquired station cuts its advertising by 2.6 minutes (year to year), compared to a market trend increase of 0.4. These numbers suggest strategic interactions occur between mergers and the imputed measure of advertising quantity. Because the quantities decrease, the initial assessment is that market power is exercised on advertisers. However, without properly taking into account endogeneity of advertising quantity this assessment is correlational.

In Figure 1, I depict a national trend in total advertising spending in News, Magazines and TV. After the peak in 2000, the spending declines because of the recession. The recovery is slow and the spending does not reach levels from the year 2000. By contrast, in Figures 2 and 3, I depict average revenue and prices per point in radio. We can observe a steady increase in both of these quantities, which does not follow a national advertising trend. One hypothesis that would explain this discrepancy is an increase in the importance of radio as an advertising medium. However, judging from my conversations with industry insiders, as well as declining listenership trends, this hypothesis is unlikely. In this paper, I identify another explanation, an increased market power, which is also consistent with the numbers in Table 3.

Figure 4 contains trends in the national levels of radio advertising quantity. The quantity is fairly volatile. This observation is consistent with the 2002 FCC study, called Radio Industry Review, regarding the profitability of the radio industry. The study reports median profit margins for a large subset of radio owners, which is directly related to the ability to sell advertising. These margin numbers fluctuate in a way similar to my ad quantities. For example, median EBIT margin, as reported by FCC, can move by as much as 30% within a year. Because advertising prices have small variance over time, variation in ad quantities is a natural candidate to explain high volatility in the EBIT margin found by FCC.

Considerable degree of volatility can be observed in station- and market-level data as well. I report three representative markets by size: Los Angeles, Albuquerque, and Bismark. I also investigated other markets, but do not report them because they show similar patterns. Figure 5 contains a graph of ad quantities for the most popular station in each of these markets.[5] Market-level average ad quantity for several of the largest stations (20, 10, and 5 depending on the market size) is depicted on Figure 6. I find the overall variance is smaller for the market average than for individual stations, however the difference is small. Moreover, individual and average graphs show a large degree of similarity. Thus I expect the size of idiosyncratic station-level shocks is small relative to market trends. Small contribution of station-level unobservables is reassuring, because the rationalization and decomposition of the market trends is one the main goals of this paper.

One may worry the larger radio stations might be able to charge an unobserved price premium over smaller stations, beyond the one incorporated in demographic composition and higher ratings, in which case, my advertising data would use smaller than actual CCPs to compute ad quantities of large stations. As a consequence, I would overestimate the ad quantities for these stations. Although, I cannot formally rule out this possibility, I note the ad quantities of large stations in my data are lower than the market average.

Part of the analysis in this paper is describing the impact of format-switching patterns. A number of studies have investigated the impact of mergers on product variety (e.g. Berry and Waldfogel (2001) or Sweeting (2009)). The general conclusion is that mergers increase variety, which is consistent with avoiding cannibalization or creating spacial preemption. In this paper, I take these results as given and do not try to replicate them. Instead, I describe the raw data by providing a format-switching matrix in Table 4. The matrix contains the probabilities of switching across formats for the whole population of stations as well as for those stations that switched owners. I find the probability of switching a format is higher for stations that change owners, which suggests that format switching and mergers might be compliments. In which case, format switching should amplify the effects of the mergers on consumer surplus. Another observation is that format-switching patterns are different for acquired and average stations, which highlights strategic portfolio considerations. For this reason, welfare calculations should include format switching. Interestingly, for some formats, acquisitions trigger switching to Dark, which might mean that some mergers are aimed at reducing the number of active stations.

[5] I computed an average station 12+ rating and picked the station with the largest value.

4 Estimation

I conduct the estimation of the model n two steps. In the first step, I estimate the demand model that includes parameters of the consumer utility θ^L (see equation (2.4)) and the unobserved station quality lag parameter ρ (see equation (2.9)). In the second step, I recover parameters of the inverse demand for advertising θ^A, $w_{jj'}$ (see equation (2.5)) and cost parameters θ^C (see equation (2.7))[6]. I adjust the asymptotic variance-covariance matrix of the estimates for the second step for the estimation error in the first step by treating the problem as a large GMM system.

4.1 First stage

This stage provides the estimates of the demand for radio programming θ^L, which are obtained 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:

$$E[\xi_{jt} - \rho\xi_{jt-1}|Z_1, \theta^L] = 0, \tag{4.1}$$

This moment condition follows Berry, Levinsohn, and Pakes (1995) with an extension of explicitly modeling auto-correlation of ξ. I use instruments for advertising quantities because these quantities are 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 assumptions: (i) ξ_t follows an AR(1) process and (ii) 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 average market share of format f among listeners possessing certain demographic characteristics

[6] I divided the estimation was divided into two steps because the second step can be estimated by two-stage least squares. This greatly decreases the computational burden of the whole procedure because it avoids minimization over (θ^C, θ^A) which has 128 dimensions. The minimization procedure in the first step has 78 parameters, whereas the joint estimation would require minimization over 206 parameters. At the same time, a first-stage readjustment to second-stage standard errors is minor, which suggests a minimal loss in efficiency.

c. The average is taken across markets and time. The population moment conditions are

$$\int_m \int_t \int_{D_{it}} \int_{\nu_i} \frac{\exp[\delta_{jt} + \mu_{jit}]}{\theta^{Lt} + \sum_{j' \in s_{mt}} \exp[\delta_{j't} + \mu_{ij't}]} dF(\nu_{it}) dF_{mt}(D_{it}|c) dt dm = R_{fc} \tag{4.2}$$

where $F_{mt}(D_{it}|c)$ is a distribution of people who possess characteristic c at time t in market m. To lower the computational burden, I simulate the moment conditions (4.1) and (4.2) jointly by using the same draws for demographics for both equations and summing across a sub-population possessing characteristic c in the equation (4.2). This gives $\mathcal{I}_{mtc}$ number of draws for each characteristic.

I formulate the problem using Mathematical Programming with Equilibrium Constraints in a similar way as Dube, Fox, and Su (2012):

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

Subject to:

$$\begin{aligned}
&\frac{1}{\mathcal{I}} \sum_i \frac{\exp[\delta_{jt}(\theta) + \mu_{jit}(\theta)]}{\theta^{Lt} + \sum_{j' \in s_{mt}} \exp[\delta_{j't}(\theta) + \mu_{ij't}(\theta)]} = r_{jt} \quad \forall t, j \\
&\frac{1}{T} \sum_t \frac{1}{M} \sum_m \frac{1}{\mathcal{I}_{mtc}} \sum_{i \in c} \frac{\exp[\delta_{jt}(\theta) + \mu_{jit}(\theta)]}{\theta^{Lt} + \sum_{j' \in s_{mt}} \exp[\delta_{j't}(\theta) + \mu_{ij't}(\theta)]} - R_{fc} = g_1 \quad \forall c \\
&\frac{1}{\text{size of } \xi} Z_1 (\xi - \rho L \xi) = g_2,
\end{aligned} \tag{4.3}$$

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.

4.2 Second stage

The second stage of the estimation obtains the competition matrix Ω, the parameters of demand for advertising θ^A and marginal cost θ^C. To compute the matrices Ω^m for each market, I use the specification laid out in section 2.3. The elements of the matrix Ω are specified as

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

following equation (2.6). The term $r_{f|a}$ represents advertisers' beliefs about listeners' preferences for formats and are constant across markets. To recognize that advertisers know the demographic composition of each market, I allow for market-specific conditional probabilities of listeners' arrival for each format $r^m_{f|a}$. However, I assume the advertisers compute those values by using Radio Today reports and the Current Population Survey. I treat Ω^m as exogenous and fixed in all of the following steps.[7]

After computing matrices Ω, I estimate θ^A and θ^C. 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 FOCs for owner's profit maximization (see equation (2.7)) to set up a system of linear equations:

$$\begin{aligned} &r_{jt} + \sum_{j' \in s_{kt}} q_{j't} \frac{\partial r_{j't}(q_t)}{\partial q_{jt}} = \\ &\quad \theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Am} \left[r_{jt} v_j + \sum_{j' \in s_{kt}} \left(r_{j't}(q_t) \omega^m_{jj'} + v_{j'} \frac{\partial r_{j't}(q_t)}{\partial q_{jt}} \right) \right] + \theta_2^{Cm} \xi_{jt} + \theta_3^{Cm} \mathrm{SYN}_{jt} + \eta_{jt} \end{aligned} \tag{4.4}$$

where $v_j = \sum_{j' \in s_{kt}} \omega^m_{jj'} q_{j't}$.

Because the equation does not depend on θ_1^A, I can use it to estimate θ_2^A and θ^C. Two sources of heterogeneity in marginal cost and slope coefficients exist across markets. Effective marginal cost parameters for each station in market m are given by $\theta_1^{Am}\theta^{Cm}$, and θ_1^{Am} is allowed to be different across markets. Moreover, to control for potential heterogeneity that is not captured by a level of revenues, I allow for three different set of values of all parameters in θ^{Cm}: for small (up to 500 people), medium (between 500 and 1500), and large (more than 1500) markets. To avoid having a full set of dummies and to facilitate identification, I set time dummies for years 1996 and 1997 to zero.

Similar specification is true for the slope of the inverse demand for ads and its effective slope is given by $\theta_1^{Am}\theta_2^{Am}$. To control for the fact that stations might have different market power in the advertising market depending on its size, I allow for four different values for the slope of inverse demand, depending on the population of the market (up to 500 people, between 500 and 1500, between 1500 and 4500, and more than 4500).[8]

I calculate ratings and derivatives of ratings in equation (4.4) using the estimates of θ^L and ξ

[7] 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.

[8] One could potentially allow for CPPs to depend on station quality. This allowance would introduce a non-linear

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_2^{Am} and θ^C, I can back out θ_1^{Am} by equating the observed average revenue in each market with its predicted counterpart.

To control for the fact that ratings depend on quantity, which is likely to be correlated with η, I estimate the model with two-stage least squares using the following instruments: number of stations in the same format and ad quantities of competitors. Additionally, the instruments were lagged one period to control for potential serial correlation in η.

4.3 Identification

First, I discuss the identification of listener demand slope, separately from advertiser demand slope. The listener demand slope is identified from variation in advertising quantity and ratings. During this procedure, I do not impose optimality of quantities; instead I trace out the size of rating response to changes in ad quantity caused by exogenous factors. The structural model enables me to control for confounding factors that include exogenous time trends in radio listenership, changes in demographics, changes in market structure, and changes in unobserved station quality.

Having the estimates of listener elasticity, I am able to compute the listener feedback effect to changes in the quantity of advertising, which I incorporate into supply-side optimality conditions. Using these optimality conditions, I predict the response in advertising quantity to exogenous variables (e.g., demographics, macro time trends, or shocks to competitors' quality) Finally, I find the advertiser demand slope and marginal cost that minimizes the prediction error.

The slope parameter θ_2^A is identified separately from marginal cost using the observed response of advertising quantity to mergers. In particular, assuming the CCP slope is flat, the estimated slope of listeners' demand would predict large increases in ad quantity after the merger. However, in the data, I frequently observe a decrease in the quantity supplied. This fact can be rationalized by a negative value of CPP slope, θ_2^A, but cannot be rationalized by marginal cost.

relationship between ad price per minute and ratings. Because I am unsure whether stations with big market shares can charge more per listener (they do not necessarily attract wealthier and more educated people), I decided not to include dependence of CCPs on quality in the specification. Moreover, the identification of this effect separately from θ_2^{Am} and θ_2^{Cm} using equation (4.4) would crucially depend on the values of $\omega_{jj'}^m$. In the specification without non-linear effects, the results are robust to the weights (see section 7).

The time trends in advertising demand and marginal cost are identified using a triple-panel structure of the data. I have observations on multiple time periods and markets, and I can track stations over time. To obtain the fixed effect estimates in equation 4.4, I compare the size of the residual quantity responses across time periods, after netting out the observed time-varying factors. One limitation of this procedure is that the trends in advertising demand and marginal cost are not separately identified given the current data. However, for my counterfactuals, this fact is not a limiting factor, because I am interested in netting out both of these effects simultaneously.

Cross-price elasticities are identified under the assumptions given in Appendix B. Because these assumptions are fairly strong, the values of individual substitution coefficients $\omega_{ff'}$ should be treated as approximations. For this reason, in section 7, I demonstrate the policy experiments are invariant to these approximations.

5 Results

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

5.1 Listeners' demand

Table 6 contains estimates of demand parameters for radio programming. The estimate of the mean effect of advertising on listeners' utility is negative and statistically significant, which is consistent with the belief that radio listeners have a disutility for advertising. Regarding the mean effects of programming formats, the Contemporary Hit Radio format gives the most utility, whereas the News/Talk format gives the least.

The second column of Table 6 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 tastes for the News/Talk and Country formats are the most persistent.

Table 8 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, whereas older people go for News/Talk. The negative coefficients on the interaction of Hispanic format with education and income suggests less-educated Hispanic people with lower incomes 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 finding is consistent with the fact that Urban radio stations play mostly rap, hip-hop, and soul music performed by black artists.

5.1.1 Discussion of the instruments

To check if the instruments used in the demand estimation are helpful in fixing the endogeneity bias, I perform an estimation of the model without random effects by 2SLS and compare it with an OLS estimation. The results[9] are reported in Table 9. I find endogeneity of quantity produces a biased OLS estimator. The estimated disutility from advertising is almost two times smaller when using OLS, which flattens the demand curve. This bias goes in the right direction if quantity q was positively correlated with quality ξ. The extensive discussion of the mechanism of such bias is contained in Berry (1994).

Additionally, I tested for weak instruments by doing a joint F-test on parameters from the first-stage regression in the 2SLS procedure. I obtain an F-statistic equal to 3932, and overwhelmingly reject the hypothesis that no linear dependence of quantity on instruments exists. Moreover, I obtain a first-stage R^2 of 0.32, which suggests that instruments can explain about 30% of variation in quantity.

5.2 Advertisers' demand

Tables 10 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. I also compute a total impact coefficient that is the sum of all the columns of the table for each format. Not surprisingly, general interest formats

[9]The fact that demographic random effects are not mean zero causes the large differences between format dummies in models with and without random effects.

such as AC and News/Talk have the biggest impact on the price of advertising, whereas 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, which suggests the ad quantity in the same format is a strong determinant of per-listener price. In accordance with an intuition, the formats with the most demographically homogeneous listener pools, Urban/Alternative and Spanish, have the highest values of the own effects. On the other hand, general interest formats such as CHR and Rock are characterized 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 News/Talk is close to AC and Urban is close to CHR. This observation can be explained by, for example, the age of the listeners. In the former case, the formats appeal to an older population, whereas in the latter case, to a younger one.

Estimates of intercepts θ_1^{Am} can be found in Table 11. The between-market variation of these intercepts picks out large variation of revenue between markets in the data. Because the ad prices are measured on the listener basis, one can expect a positive relationship between the market population and the intercept. This relationship, however, is not exactly linear and reflects unobserved factors that affect the value of the ad listener in different markets. The non-linearity can be caused by two factors: (i) unobserved characteristics of listeners, such as income or propensity to purchase, as well as competition from other advertising markets, (ii) the amount of market power in the listenership market. Because of the second factor, I postpone further discussion until I present the slope coefficients.

The estimates of the slope of the inverse demand for advertising θ_2^{Am} can be found in Table 12. The variation in the data allows me to separately identify θ_2^{Am} for four groups of markets, clustered by total population. Slopes are relative to the intercept θ_1^{Am} and are normalized by the standard deviation of supplied ad quantity on the data. A convenient interpretation of the slopes is the percentage of θ_1^{Am} decrease in the price of advertising, caused by increasing ad quantity by a standard deviation.

I find that in larger markets, the advertisers are more price sensitive than in smaller markets, which means that the radio stations have more market power over advertisers in smaller markets. This difference is caused by unobserved factors that affect the advertising-demand slope, such as the value of the marginal ad listener and competition from other sources of advertising.

Having discussed the slope coefficients, I come back to rationalizing the intercept parameters. In an extreme example, Dallas (TX) and Austin (TX) have the intercepts of respectively \$342 and \$337, which is more than four times higher per-listener in Austin. This difference is partly driven by the fact that the raw 2006 SQAD per-radio-listener price is 120% higher in Austin than in Dallas (the SQAD prices are based on actual transactions). Another factor is that, the inverse demand for advertising is much flatter in Dallas than in Austin. Thus, the equilibrium price in Dallas is likely to be close to the intercept, whereas in Austin, it is likely to be much smaller than the intercept. These factors cumulate and generate much higher intercept estimates in Austin than in Dallas.

5.2.1 Discussion of the instruments

As I do in the demand model, I check if the instruments are helpful in fixing the endogeneity bias, by comparing the results with an OLS model. In this case, the bias is hard to sign because q appears on both the right and left sides of the equation (4.4). The results are summarized in Tables 12, 13, and 15. I find endogeneity of quantity biases both the advertiser demand and marginal cost estimates.

Because I allow for multiple values of θ_2^A, the first-stage regression has three parameters on the left side that correspond to markets of different sizes. I perform a test for weak instruments by regressing an endogenous regressor for each market group on an appropriate instruments (I use the term instruments to denote all exogenous variables) vector for that group. The F-stats are respectively 714, 924, 827, and 202 for small, medium-small, medium-large, and large markets. The R^2 statistics are 0.22, 0.19, 0.26, and 0.11, which suggests the instruments can explain a reasonable amount of variation in the endogenous regressors. [10]

[10] A lower value of R^2 for large markets illustrates that the variation in the instrument might not be enough to allow for four groups simultaneously in marginal cost and slope of advertising demand. The choice to allow the fourth group in advertising demand is motivated by the fact that it adds only one parameter. An additional group in marginal cost would add 12 more parameters.

5.3 Supply

Estimates of a marginal cost level θ_1^{Cm} and quality coefficient θ_2^{Cm} can be found in Table 13. I can identify unobserved differences in marginal cost between three groups of markets, depending on the population. The numbers reported in the table are relative to the intercept θ_1^{Am}. They can be bigger than one because the ad price is additionally multiplied by station market share measured in percentages. One convenient interpretation of the level is the lower bound on the rating of the station with average quantity that makes it profitable. I find that in small markets, an average quality station has to have at least a 3% market share to break even, whereas in big markets, this number drops to 1.1%. Because there are many more stations in the bigger markets, this difference confirms my prior intuition. Note the above numbers apply to a station of an average quality. Because, smaller stations have lower unobserved quality, they also enjoy smaller marginal cost. Thus the model predicts positive variable profits even for stations with ratings that are smaller than the above-average break-even thresholds. [11] Moreover, during the supply side estimation, the first order conditions are enforced. In means that the station that looks unprofitable, but supplies positive amount of advertising, is automatically assigned a negative shock to the marginal cost. This rationalizes the station's action. In the counterfactual, I employ a constrained optimization algorithm that in practice assigns zero quantity to stations that would earn less than zero gross profits otherwise.

The coefficients on station quality are positive, which suggests stations pass-through some of the additional revenue from an additional advertising minute to the ad agents. Higher wages for agents can be directly generated by higher ratings of stations with higher quality as well as different listener composition. The relationship is more pronounced in smaller markets.

[11] I note positive variable (or gross) profits do not necessarily mean that the radio station is profitable. Indeed, Jeziorski (2011) finds stations bear a considerable amount of fixed costs, which would sometimes translate to operating losses. The possibility of losses is consistent with the low and sometimes negative median industry EBIT margins that are reported in the FCC Research Studies on Radio Industry. This FCC study shows that the median EBIT margin was negative for the part of 1998, 2001, and 2002, and was below 5% for the part of 1999.

6 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. The dollar values in counterfactuals were extrapolated from a subset of 88 markets into the whole country, by assuming the analyzed markets are a random selection of small, medium and large markets from the whole Arbitron market pool.

Before performing counterfactual calculations, I consider descriptive relationships between concentration and prices. First, I regressed market CCPs 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 with a standard error of 0.3, which is an 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. These counterfactuals also quantify changes in both listener and advertiser surplus, due to market power as well as programming variety.

6.1 Impact of mergers on consumer surplus

To isolate the impact of the Telecom Act on a surplus division between advertisers and listeners, I perform counterfactuals in which I recompute new equilibrium ad quantities and prices for each half year between 1996 and 2006 under the old 1996 ownership structure and the 1996 formats. This methodology is motivated by the fact that in 1996, many markets were at their ownership caps. During this computation, I fix the values of stations' quality ξ and shocks to marginal cost η at the estimated values. I account for marginal cost synergies and adjust the estimates to account for a changing ownership structure using the appropriate pre- or post-merger values for dummies SYN_{jt}. I net out time effects by setting year dummies on the demand and supply side to their 1996 levels, and I use 1996 draws in common population survey to detrend demographics. Note I do not observe listeners paying to remove advertising, so I am unable to quantify listener welfare in dollars. Instead, I am going to use person-day-minutes (pdm), which quantifies welfare difference in disutility units of being exposed to one minute of advertising every day. Although I use advertising as units for listener welfare, the numbers reflect full welfare changes or compensating variation (including changes in choice sets caused by quantity readjustment by competitors as well as by product repositioning), not just the simple change in advertising exposure. Changes in advertising exposure are also reported under "Average ad load."

The total impact of consolidation on advertiser and listener welfare is presented in the last row of Table 16. It is a difference in listener and advertiser welfare between the equilibria computed using 1996 ownership and formats and current ownership and formats. I find mergers decreased total ad quantity by roughly 15,000 minutes, resulting in lowering average ad exposure by 7.3 pdm, which is about 17% of the total ad load[12]. The changes translated to about a 1 pdm increase in consumer welfare. Because we do not observe dollar prices in the listenership market, I cannot compute the dollar value of this compensating variation. However, I can compute a rough estimate using the prices for the satellite radio. If I assume people buy satellite radio just to avoid listening to advertising, we get a rough estimate of 1.5 cents per minute, or $730 million dollars for each pdm per year. This number is of course a very ad hoc and loose upper bound on the overall welfare gain. For advertisers, a decrease in quantity supplied leads to about a 6.5% increase in per-listener prices,[13] or a $223 million (or 21%) decrease in advertiser surplus. I therefore conclude the Telecom Act led to a reallocation of surplus from advertisers to listeners.

Looking separately at small, medium, and large markets can reveal more details about the surplus reallocation. As mentioned in the previous section, radio stations have considerable control over prices in small markets and less control in the large markets. Motivated by this fact, I present counterfactuals for markets with populations less than 0.5m, between 0.5m and 2m, and more than 2m. In smaller markets (see first row of Table 16), stations contract advertising to exercise market

[12]Ad load is defined as the expected number of advertising minutes heard by an average listener (it is different from total amount of minutes broadcasted).

[13]I measure prices by computing a mean price index, which is a price per average listener (average of station CPPs weighted by market shares and market population) – computed in 1996 units to control for different levels between markets.

power on advertisers. They supply more than 13,000 fewer minutes of advertising, which translates into a 7.3 pdm decrease in ad exposure, increasing consumer surplus by 1.3 pdm. However, prices rise by 7% and cause a $56m (or 33%) loss in advertiser surplus.

In medium markets (see first second of Table 16), mergers lowered the number of broadcasted advertising minutes by 3,500 minutes (or 8.8%) and ad load by 8.7 pdm (or 22%). This decrease resulted in lowering advertising surplus by $68 million (or 23.5%) which is bigger in nominal terms than for small markets (because of bigger demand intercepts in medium markets), yet much smaller in percentage terms. The consumer surplus goes down despite a smaller advertising load, which is a combination of an impact of quantity readjustment and format repositioning.

In large markets (see third row of Table 16), firms supply only 950 (or 6.5%) fewer minutes of advertising, which translates to lowering the average ad load by 6.4 pdm (or 14%), increasing listener surplus by almost 2 pdm. At the same time, the smaller supply of advertising decreased advertising welfare by $100m, which is nominally more than in the medium and small markets. However, this decrease constitutes only 17% of total surplus of advertisers, which is less than in the other two market groups. Large intercepts for bigger markets (which drives the nominal number) and smaller slope coefficient (which drives the percentage number) cause this duality. Overall, I conclude that on average, mergers transferred surplus from advertisers to listeners. The loss in the listener surplus was higher nominally in larger markets; however, it was higher percentage-wise in smaller markets.

In the model without marginal cost synergies the total effect of the merger wave on advertisers was 10% stronger and the effect on listeners was 40% stronger. This difference is caused by the fact that ownership consolidation provided additional marginal cost savings. To get the correct counterfactual, these savings need to be nullified when computing the pre-merger equilibrium. Otherwise, although the ad quantities post-merger would still be correctly computed, ad quantities pre-merger would on average be biased upwards. Therefore, the firms would seem to have contracted the quantities much more and have led to overstating the change in consumer welfare.

6.2 Effects of product variety and market power

Berry and Waldfogel (2001) suggest the negative effects of ownership consolidation on listeners might be mitigated by format switching. They find post-merger repositioning results in spatial competition, leading to more variety, which they assume is beneficial for the listeners.[14] To quantify this effect, I compute the ad quantity imposing 1996 ownership and formats, which serves as a proxy for a pre-regulation regime. Then I use this quantity to compute the surpluses by imposing 1996 ownership and formats and current ownership and formats. The difference between these surpluses measures an impact of just post-regulation changes in market structure on welfare without taking into account quantity adjustments. The results of this experiment are presented in the first row of Table 17. I confirm the result of Berry and Waldfogel (2001). In particular, listeners have a 0.3% larger surplus (about 1.6pdm) 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 the number of ad minutes fixed, because I allow the listeners to reoptimize. Given more variety is available on average, they are able to choose a station with less advertising that suits their programming taste. I note these results are opposite to the findings of Fan (2010). In contrast to the impact on listeners, I find a negative effect of repositioning on advertiser surplus. Advertisers lose more than $180 million which constitutes 17% of their total surplus. I note the literature has previously documented the positive effects of variety on the listener (or equivalent) side of the market. For example, George (2007) uses reduced-form arguments to show that mergers in the newspaper industry led to more variety. Also, she finds that mergers combined with extra variety did not lower circulation. She concludes that these two facts suggest that variety mitigates the negative effects of mergers on reader surplus. I find and quantify similar effects in radio. However, more importantly, I quantify the effect on the advertiser side, which is largely negative. I note the negative impact of repositioning on advertisers is consistent with the findings of Fan (2010).

In the real world, repositioning changes firms' incentives to set ad quantity, because it softens competition in the advertising market. To quantify that effect, I compare the consumer welfare values without quantity adjustments computed in the previous step with consumer welfare computed under the full equilibrium response (current formats and ownership and equilibrium quantities). The difference between these surpluses is a proxy for the impact of ad readjustments. The results are contained in the middle row of Table 17. Both listeners and advertisers are worse off due to quantity adjustments. Listeners lose 0.8 pdm and advertisers lose an additional $44 million in

[14] Similar results obtained using direct analysis of station play lists can be found in Sweeting (2009).

surplus. Both listeners and advertisers lose welfare because the number of advertising minutes goes down, whereas average ad exposure goes up. I find that the companies that own multiple stations engage in complicated ad readjustments. They lower the number of advertising minutes in the stations with low market share by a large amount and increase the number of minutes in popular stations by a smaller amount. These changes in advertising supply allow them to get additional revenue from higher prices, which depend on the total number of minutes supplied. In addition, they are able to increase revenue with greater average ad exposure. [15]

The quantity readjustment does not decrease the listener surplus enough to offset the positive impact of extra variety. However, in the advertiser case, the two effects work in the same direction, so the repositioning strengthens the effect of ad readjustment. About 80% of adverse effects on advertisers was caused by product repositioning, which softened within format competition. The size of these effects suggests that repositioning is much more effective in extracting social surplus than quantity readjustment. It reconciles with high repositioning cost estimates found by Sweeting (2011) and Jeziorski (2011). Also, it is consistent with a conjecture made using the raw data in section 3 that format switching amplifies the market power effects. Another finding, that reflects patterns in the raw data described at the end of section 3, is that some mergers trigger switching to the Dark format. It happens in more niche formats and leads to price increases.

7 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 stylized decision model for advertisers that uses publicly available data to predict substitution patterns between formats. These patterns directly determine the market power of stations over advertisers and can potentially alter the results of counterfactual experiments.

To investigate the robustness of the results, I re-estimated the model under two alternative assumptions. The first scenario represents the extreme situation in which formats compete only

[15] This effect could be weaker if the price per ad slot were a non-linear, in particular convex, function of the ratings.

between themselves. In particular, suppose advertiser types get utility from only one particular format. In this case, equation (2.6) has $\omega_{ff} = 1$ and $\omega_{ff'} = 0$ if $f \neq f'$. The second scenario represents another extreme in which formats are perfect substitutes; that is, there exists only one type of advertiser who values all formats in the same way. Formally, this means $\omega_{ff'} = 1/8$, because eight formats are possible. The estimated model is in a sense between these extreme alternatives because it assumes formats are imperfect substitutes.

Estimates of the slopes of the inverse demand for advertising are presented in Table 18. 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. The structure of the oligopoly within format grants stations more market power because it turns off between-format advertiser switching. The smaller value of the slope coefficient compensates for that extra market power to match the level of ad adjustment in the data. On the other hand, in the perfect-substitutes model, the estimated slope tends to be higher and it compensates for higher between-format advertiser switching. Despite the fact that small but statistically significant differences exit between the different models, the main qualitative assertion, that stations have more power in smaller markets, still holds. Similarly,the difference in the estimates of marginal cost between different specifications of advertising demand is small. The results are presented in Table 19.

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 20. The baseline again lies between the new counterfactuals. No large differences are present in the results. The only qualitative difference is a lower listener surplus in an oligopoly within-format model, caused by a smaller decrease in the average ad load. However, the percentage changes in consumer and advertiser are largely of the same magnitude.

8 Conclusion

In this paper, I analyze mergers in two-sided markets using the example of the 1996-2006 consolidation wave in the 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 that 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 decrease advertising, listeners benefit but advertisers are hurt. Second, I disaggregate the impact of a merger on consumers into changes in the variety of available products and changes in the supplied quantity of ads.

Radio is an important medium in the United States, reaching each week about 94% of Americans twelve years old or older. Moreover, the average consumer listens to about 20 hours of radio per week, and between 6am and 6pm, more people use radio than TV or print media.[16] In 1996, the Telecommunication Act deregulated the industry by raising local ownership caps. This deregulation caused a massive merger wave, which reshaped the ownership structure by moving from family-based ownership into more corporate structures. I estimate this consolidation raised listener surplus by 0.2% but lowered advertiser surplus by $223 million. I find the mergers created extra variety that increased listener welfare by 0.3%. On the other hand, they softened competition and decreased advertiser welfare by $180 million per year. Subsequent ad quantity adjustments led to a 0.1% decrease in listener welfare (with the variety effect a total increase of 0.2%) and an additional $43.8 million decrease in advertiser welfare (combined with the variety effect for a total of $223 million).

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*, 36(4), 771–788.

Anderson, S. P., and S. Coate (2005): "Market Provision of Broadcasting: AWelfare Analysis," *Review of Economic Studies*, 72(4), 947–972.

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.

[16]Source: A.Richter (2006)

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.

Brown, K., and G. Williams (2002): "Consolidation and Advertising Prices in Local Radio Markets," Discussion paper, Federal Communications Commission.

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.

Chipty, T. (2007): "Station Ownership and Programming in Radio," Discussion paper, Federal Communications Commission.

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

DUBE, J.-P., J. T. FOX, AND C.-L. SU (2012): "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, 80(5), 2231–2267.

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*.

FAN, Y. (2010): "Ownership Consolidation and Product Quality: A Study of the U.S. Daily Newspaper Market," Discussion paper, University of Michigan.

GAL-OR, E., AND A. DUKES (2006): "On the Profitability of Media Mergers," *Journal of Business*, 79(2), 489–526.

GANDHI, A., L. FROEB, S. TSCHANTZ, AND G. J. WERDEN (2008): "Post-merger Product Repositioning," *The Journal of Industrial Economics*, 56(1), 49–67.

GEORGE, L. (2007): "What's fit to print: The effect of ownership concentration on product variety in daily newspaper markets," *Information Economics and Policy*, 19(3-4), 285–303.

JEZIORSKI, P. (2011): "Estimation of cost synergies from mergers: Application to U.S. radio," Discussion paper, UC Berkeley.

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).

Mooney, C. T. (2010): "A Two-Sided Market Analysis of Radio Ownership Caps," Discussion paper, University of Oklahoma.

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.

Song, M. (2011): "Estimating Platform Market Power in Two-Sided Markets with an Application to Magazine Advertising," Working paper.

Sweeting, A. (2009): "The Effects of Horizontal Mergers on Product Positioning: Evidence from the Music Radio Industry," *RAND Journal of Economics*, 40(4).

——— (2010): "The strategic timing of radio commercials: An empirical analysis using multiple equilibria," *RAND Journal of Economics*, 42(2).

——— (2011): "Dynamic Product Positioning in Differentiated Product Industries: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," Discussion paper, Duke University.

Van Cayseele, P. J. G., and S. Vanormelingen (2009): "Prices and Network Effects in Two-Sided Markets: The Belgian Newspaper Industry," Working paper.

Waldfogel, J., and J. Wulf (2006): "Measuring the Effect of Multimarket Contact on Competition: Evidence from Mergers Following Radio Broadcast Ownership Deregulation," *The B.E. Journal of Economic Analysis & Policy*, 0(1), 17.

White, A., and E. G. Weyl (2010): "Imperfect Platform Competition: A General Framework," Working Papers 10-17, NET Institute.

Appendices

A Data appendix

The data in this paper come from four main sources: two consulting companies, BIA Inc and SQAD, a Common Population Survey, and Radio Today publications by Arbitron. BIA provides two comprehensive data sets on the vast majority of U.S. radio broadcasting. The first data set covers years 1996-2001 and the second, 2002-2006. I combined the data to form a large panel for 1996-2006. For a small number of stations (less than 0.1%) from the 2002 data set I could not find a match in the 1996 data set, most likely a results of stations switching the majority of their characteristics and potential errors in the data. I dropped these stations from the sample; however, because all of them have negligible market share, their absence is unlikely to significantly affect the results. I also drop stations that are inactive and have zero market share or do not meet minimum Arbitron reporting standards (which is true for stations with less than 0.5% market share) for the whole sample period. The ownership of the stations was traced over time by tracking acquisition events reported in the BIA data set. The BIA consultant assured me the data set of acquisitions is complete.

SQAD provides a data set on average prices per rating point (CPP) for each market and half of a year, grouped by demographics and time of the day. Unfortunately, it does not provide data on station-level per-listener pricing. However, because the pricing is done on the per-listener basis, one can still compute a station-level price of an advertising slot by multiplying the CPP by the station rating. According to the anecdotal evidence, many advertisers follow this procedure to figure out the prices they are likely to pay. This procedure does not account for the fact that stations might have different listenership pools and therefore CPP for different stations might vary. I alleviate that concern by computing a proxy for a station-level CPP. I take a weighted average of prices by demographics and time of the day, where weights are relevant ratings of the station. By doing so I assume stations that have most of their listenership at a particular time of the day also set a price that is closer to the average price at that time of the day. Although this estimate of station level prices is not a perfect, it produces a considerable amount of variation within market. I subsequently use these price proxies to compute station-level advertising quantities by dividing estimates of station revenues (provided by BIA) by a product of prices and ratings. Note that ad quantity computed in such a way might carry some measurement error, because it is a function of two estimates. However, if this measurement error is not endogenous within markets – for example, if it only introduces error to an overall level of advertising in each market – it would not affect the results.

To compute the probability of listening to a particular format by different demographic groups I use Radio Today publications. These papers provide a demographic composition for each format. The numbers were inverted using Bayes' rule and demographic distributions in all markets obtained from the Census Bureau. I averaged the probability distributions for gender and age groups across years 1999, 2000, 2001, 2003, and 2004. The Education data is available for 2003 and 2004. Ethnicity data is available only for 2004. Given almost no variation in the national values for these numbers across years, I match these averages to data moments for 1996-2006. Moreover, I supply the data with a share of an outside option for different markets from Arbitron Listener Trends publications.

B Advertising demand: Micro foundations

In this section, I present a model that rationalizes inverse demand for advertising (2.5)

Assume 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 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 Arbitron does not offer such detailed data, it seems unlikely that the advertisers would be able to compute the exact expected value of each slot. Instead, I assume 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; that is,

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

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 whether 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 they are price and rating takers.[17] 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; that is, $q_a = AS_a \sum_f r_{f|a} q_f$. If I make the simplifying assumption that $AS_a \approx LS_a$, 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 format f is given by

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

Finally, I obtain equation (2.5)

$$p_j = \theta^A_{1f} r_j \left(1 - \theta^A_2 \sum_{f' \in \mathbf{F}} \omega^m_{ff'} q_{f'}\right)$$

by setting $\omega_{jj'} = \left(\sum_a (r_{a|f})^2\right)^{-1} \sum_a (r_{a|f})^2 r_{f'|a}$, $\theta^A_2 = \frac{1}{\gamma_2}$ and assuming $\theta_1 = \gamma_{1f} \sum_a (r_{a|f})^2$ for all f. The last assumption basically means 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 (2.5) 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 (4.3), 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, which allows using of a fairly large dataset (about 42,000 observations) and many draws (500 draws from Normal and CPS per date/market) when computing

[17] This assumption 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.

the constraints. When parallelizing the code, I was careful to maintain independence of the draws within and between threads. To achieve the independence, 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.

D Tables

	Mean	Standard deviation	Median
Advertising quantity (minutes per-day)	37.5	39.9	28.5
Station revenue (thousands $s)	3,848	6,303	1,500
Station rating	0.04	0.03	0.03
Station power (kW)	4.3	4.2	2.5

Table 1: Basic descriptive statistics. The statistics were computed using active stations (no DARK, positive ratings) across all markets and available half-years (1996-2006).

	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	16.65	15.60	14.96	14.22	13.40	13.11	13.25	13.01	12.75	12.56	12.60
C3	0.52	0.56	0.60	0.62	0.64	0.64	0.64	0.64	0.63	0.63	0.62
Number of stations owned	1.64	1.77	1.87	2.02	2.16	2.23	2.25	2.31	2.38	2.42	2.41
Fraction of stations that changed the owner	0.11	0.22	0.19	0.19	0.22	0.06	0.07	0.06	0.06	0.06	NaN
Fraction of stations that changed the format	0.04	0.11	0.14	0.11	0.11	0.11	0.08	0.09	0.08	0.09	NaN

Table 2: Descriptive statistics about ownership.

	Mean	Median
Half-year to half-year change	0.3	-0.04
Half-year to half-year change conditional of merger	-0.7	-0.6
Year to year change	0.4	0.3
Year to year change conditional of merger	-2.6	-1.1

Table 3: Statistics about dynamics of advertising quantity (in minutes per day). The statistics were computed using active stations (no DARK, positive ratings) across all markets and available half-years (1996-2006).

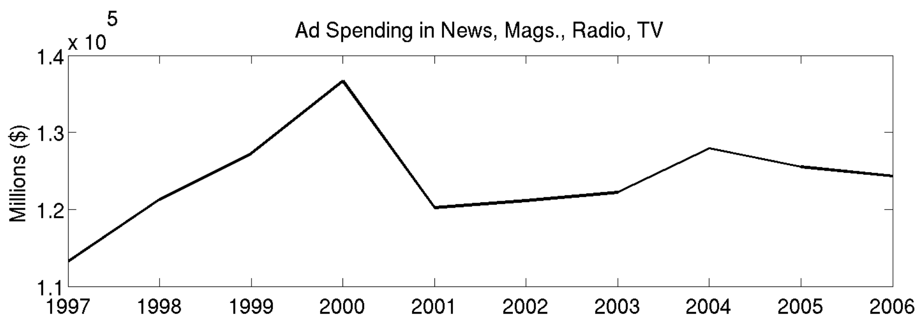

Figure 1: U.S. annual advertising spending in News, Magazines, Radio, and TV. The amount is in 1996 dollars deflated by CPI. Source: Coen Structured Advertising Expenditure Dataset (CS Ad Dataset).

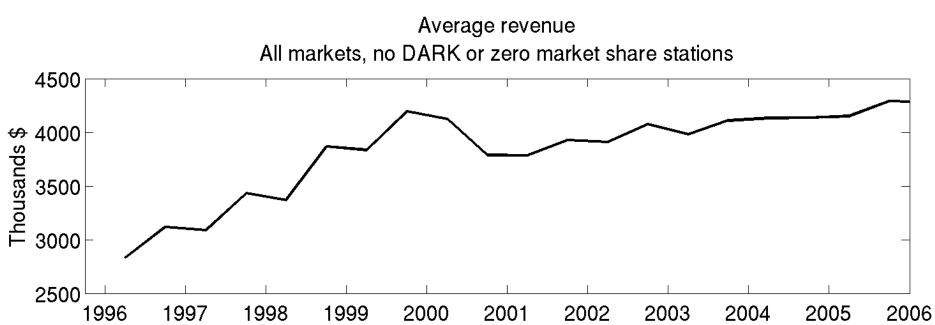

Figure 2: Average station revenue for all active stations (Dark and zero-market-share stations excluded). The amount is in 1996 dollars deflated by CPI.

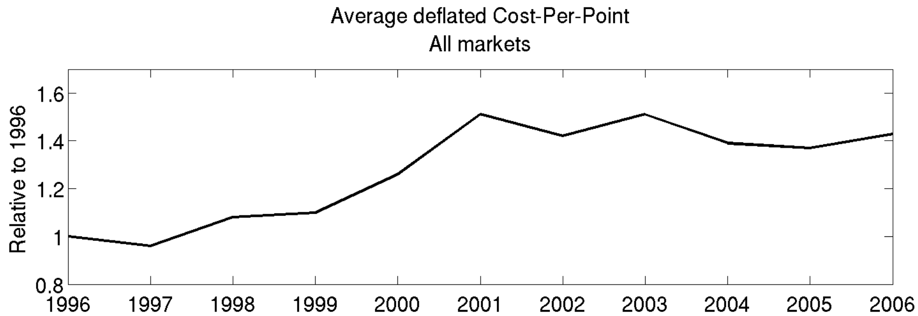

Figure 3: Average cost per point (CCP) relative to 1996 value. The relative value was computed for every market and averaged across markets. The values were deflated by CPI.

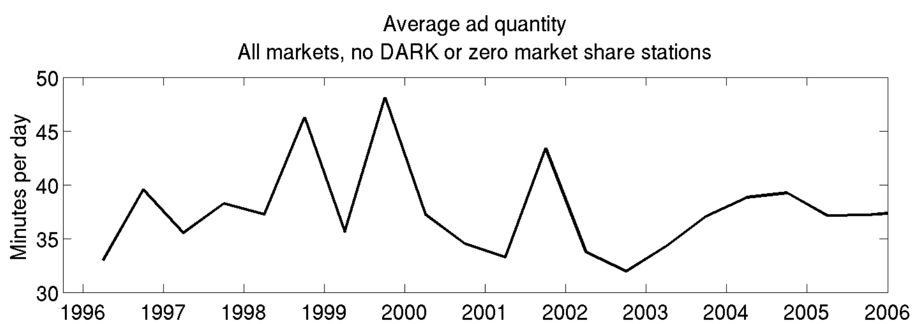

Figure 4: Average quantity for all active stations (Dark and zero-market-share stations excluded). The amount is in 1996 dollars deflated by CPI.

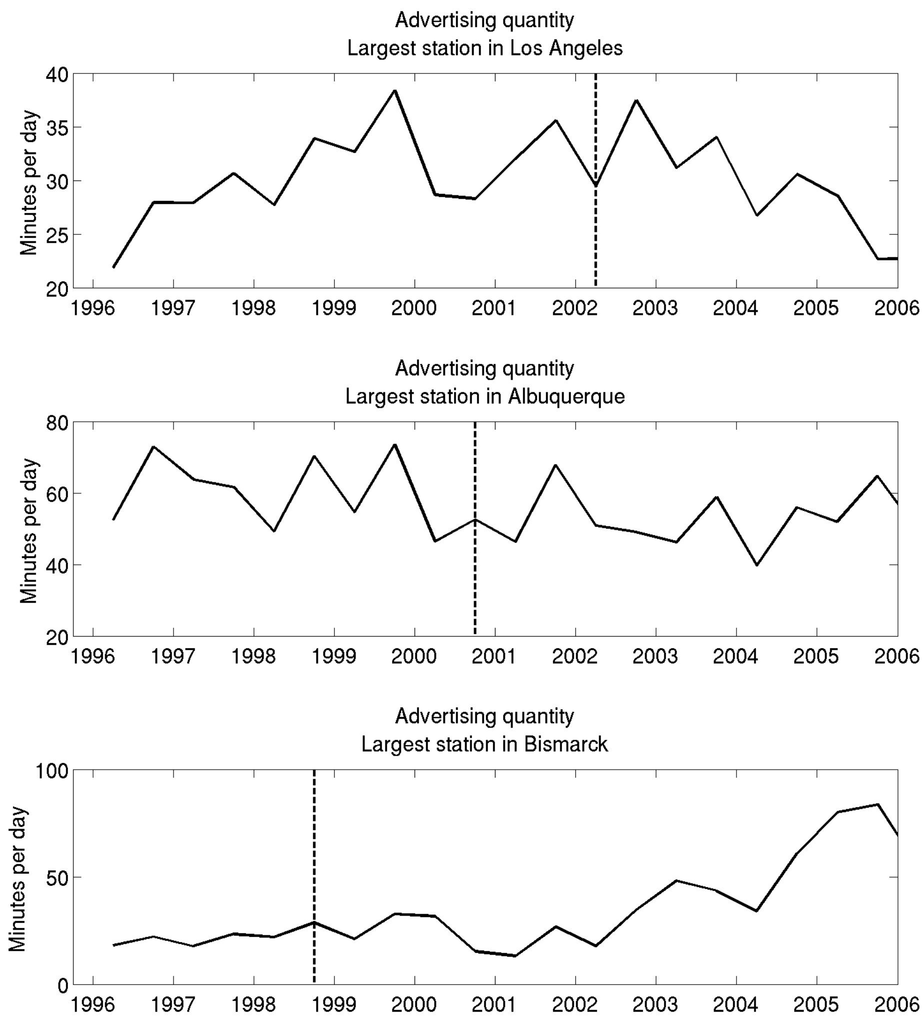

Figure 5: Number of advertising minutes per day for the largest station in the market (largest average 1996-2006 rating, among always active stations). The figures present three representative markets: the largest, Los Angeles; mid size, Albuquerque; and the smallest, Bismarck. The vertical line represents acquisition.

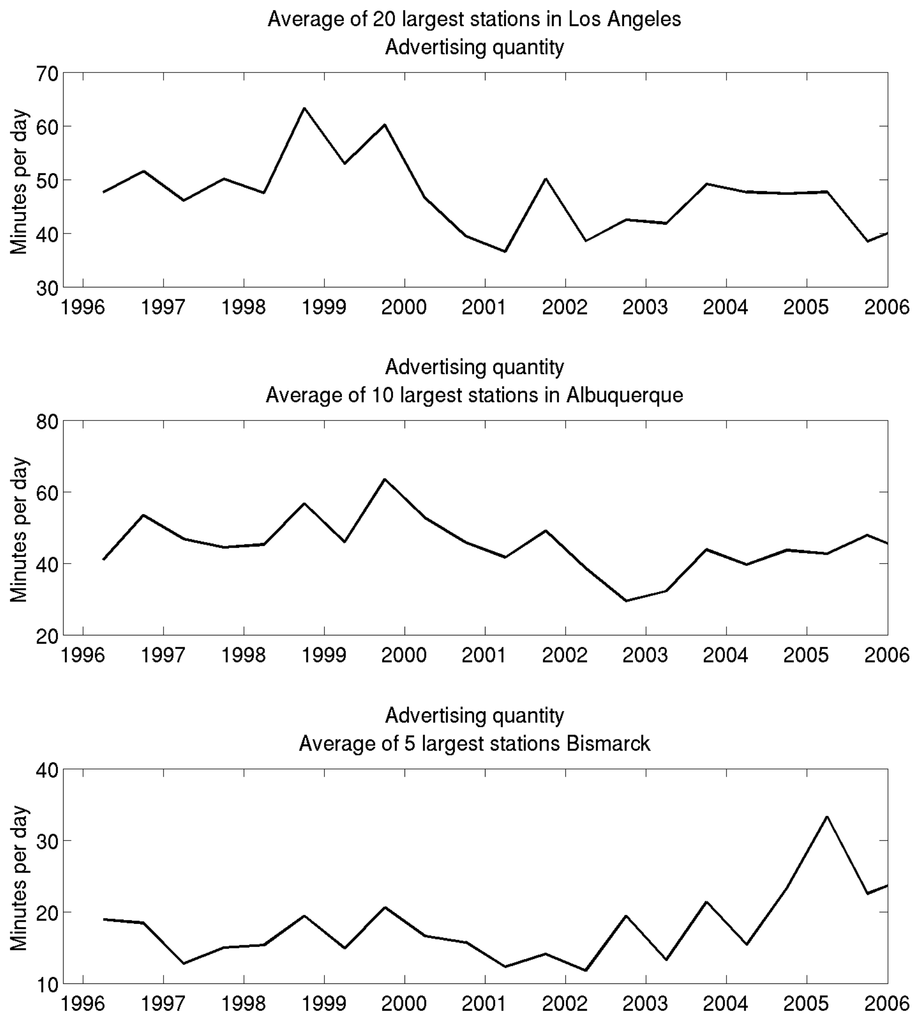

Figure 6: Market average of advertising minutes per day. The figures present 3 representative markets: the largest, Los Angeles; mid size, Albuquerque; and the smallest, Bismarck.

	AC	Rock	CHR	Urban Alt.	News Talk	Country	Spanish	Other	Dark
AC	**0.920** **0.897**	0.009 0.015(*)	0.009 0.029(*)	0.005 0.007(*)	0.005 0.015(*)	0.005 0.000	0.002 0.007(*)	0.014 0.007	0.031 0.022
Rock	0.012 0.042(*)	**0.947** **0.883**	0.002 0.000	0.010 0.008	0.002 0.008(*)	0.003 0.017(*)	0.002 0.017(*)	0.011 0.017(*)	0.011 0.008
CHR	0.015 0.093(*)	0.006 0.047(*)	**0.934** **0.837**	0.016 0.000	0.001 0.000	0.002 0.000	0.005 0.023(*)	0.009 0.000	0.011 0.000
Urban Alt.	0.007 0.024(*)	0.012 0.012(*)	0.004 0.000	**0.933** **0.905**	0.003 0.000	0.003 0.012(*)	0.004 0.000	0.013 0.000	0.022 0.048(*)
News Talk	0.003 0.009(*)	0.001 0.000	0.001 0.009(*)	0.000 0.000	**0.926** **0.913**	0.002 0.000	0.002 0.000	0.005 0.000	0.060 0.070(*)
Country	0.007 0.000	0.005 0.012(*)	0.003 0.000	0.003 0.012(*)	0.006 0.024(*)	**0.915** **0.906**	0.004 0.012(*)	0.006 0.012(*)	0.050 0.024
Spanish	0.001 0.000	0.000 0.000(*)	0.002 0.021(*)	0.002 0.021(*)	0.003 0.021(*)	0.001 0.000	**0.869** **0.667**	0.004 0.021(*)	0.117 0.250(*)
Other	0.015 0.043(*)	0.006 0.007(*)	0.002 0.007(*)	0.007 0.007(*)	0.009 0.014(*)	0.004 0.014(*)	0.002 0.014(*)	**0.860** **0.836**	0.096 0.057
Dark	0.021 0.060(*)	0.007 0.010(*)	0.004 0.010(*)	0.008 0.015(*)	0.039 0.045(*)	0.024 0.035(*)	0.037 0.035	0.075 0.060	**0.786** **0.731**

Upper number: Unconditional transition probability

Lower number: Transition probability conditional on merger

Table 4: Format-switching matrix. Events with higher probability after the merger are marked with a star.

$r_{f|a}$

	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 5: Simple example of advertising weights.

	Mean Effects	Random Effects
Advertising	-1.386^{***} (0.226)	0.235 (0.177)
FM	0.742^{***} (0.043)	-
Power (kW)	0.127^{***} (0.004)	-
AC SmoothJazz New AC	-4.082^{***} (0.058)	0.028 (0.103)
Rock	-3.380^{***} (0.076)	0.188^{***} (0.050)
CHR	-0.969^{***} (0.070)	0.028 (0.085)
Alternative Urban	-4.979^{***} (0.067)	0.314^{***} (0.048)
News/Talk	-11.088^{***} (0.096)	0.559^{***} (0.051)
Country	-4.950^{***} (0.069)	0.571^{***} (0.012)
Spanish	-4.463^{***} (0.090)	0.182 (0.156)
Other	-5.938^{***} (0.053)	0.080 (0.099)
ρ	0.748^{***} (0.008)	-

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Estimates of random effects logit model of radio listeners' demand. First columns consists of mean values of parameters in the utility function. Second row consists of standard deviations of a random effect ν.

1998	1999	2000	2001	2002	2003	2004	2005	2006
1.157***	1.476***	1.170***	1.616***	1.063***	1.513***	1.278***	1.171***	1.633***
(0.031)	(0.053)	(0.047)	(0.070)	(0.047)	(0.068)	(0.058)	(0.054)	(0.076)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Estimates of utility (exponentiated) of not listening to radio. Values for 1996 and 1997 are normalized to 1.

	Demographics Characteristics					
	Age	Sex	Education	Income	Black	Spanish
AC SmoothJazz New AC	0.250*** (0.001)	−0.773*** (0.008)	0.840*** (0.002)	0.160*** (0.002)	0.357*** (0.006)	−1.440*** (0.005)
Rock	−0.288*** (0.001)	−0.471*** (0.006)	0.989*** (0.002)	−0.371*** (0.002)	−11.694 (80.504)	−1.791*** (0.006)
CHR	−1.550*** (0.002)	−1.650*** (0.011)	1.357*** (0.002)	0.072*** (0.001)	1.486*** (0.006)	0.106*** (0.004)
Alternative Urban	−0.208*** (0.002)	0.732*** (0.010)	0.930*** (0.002)	0.434*** (0.002)	4.106*** (0.018)	−0.635*** (0.005)
News/Talk	1.120*** (0.007)	1.381*** (0.016)	1.651*** (0.007)	0.185*** (0.002)	0.965*** (0.006)	−2.962*** (0.021)
Country	0.225*** (0.001)	−0.274*** (0.005)	0.552*** (0.001)	−0.096*** (0.001)	−0.858*** (0.006)	−1.511*** (0.004)
Spanish	−0.450*** (0.004)	1.450*** (0.015)	−2.170*** (0.020)	−0.528*** (0.010)	−5.912*** (0.905)	5.629*** (0.045)
Other	0.566*** (0.001)	−0.325*** (0.008)	1.310*** (0.002)	−0.173*** (0.002)	1.177*** (0.006)	−1.219*** (0.008)

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: Table presents estimated covariances in the random effects logit model of radio listeners' demand. Each cell represents covariance between specific demographic characteristic and listening to the particular radio station format.

	OLS	2SLS
Advertising	-0.720^{***} (0.030)	-1.391^{***} (0.053)
AM/FM	0.611^{***} (0.013)	0.625^{***} (0.013)
Power (kW)	0.079^{***} (0.001)	0.082^{***} (0.001)
AC SmoothJazz New AC	-2.368^{***} (0.018)	-2.309^{***} (0.019)
Rock	-2.330^{***} (0.020)	-2.256^{***} (0.021)
CHR	-2.167^{***} (0.023)	-2.127^{***} (0.024)
Alternative Urban	-2.237^{***} (0.020)	-2.186^{***} (0.020)
News/Talk	-2.184^{***} (0.015)	-2.105^{***} (0.016)
Country	-2.321^{***} (0.019)	-2.270^{***} (0.020)
Spanish	-2.880^{***} (0.020)	-2.812^{***} (0.020)
Other	-2.634^{***} (0.016)	-2.589^{***} (0.016)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 9: Estimates of a logit model of radio listeners' demand without random effects.

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 10: Product closeness matrices for chosen markets.

Name	Pop. 2007	Intercept	Name	Pop. 2007	Intercept
Los Angeles, CA	13155.1	1125.69 (66.73)	Omaha-Council Bluffs, NE-IA	740.3	48.26 (10.31)
Chicago, IL	9341.4	573.13 (38.96)	Knoxville, TN	737.4	49.33 (6.60)
Dallas-Ft. Worth, TX	5846.9	342.11 (12.35)	El Paso, TX	728.2	63.81 (14.30)
Houston-Galveston, TX	5278.5	315.59 (8.21)	Harrisburg-Lebanon-Carlisle, PA	649.4	43.52 (15.05)
Atlanta, GA	4709.7	256.30 (21.19)	Little Rock, AR	618.7	44.43 (7.63)
Boston, MA	4531.8	278.83 (7.41)	Springfield, MA	618.1	34.16 (1.07)
Miami-Ft. Lauderdale-Hollywood, FL	4174.2	268.86 (11.93)	Charleston, SC	597.7	52.52 (3.69)
Seattle-Tacoma, WA	3775.5	228.33 (9.27)	Columbia, SC	576.6	42.08 (4.85)
Phoenix, AZ	3638.1	165.44 (10.50)	Des Moines, IA	576.5	29.74 (12.21)
Minneapolis-St. Paul, MN	3155	230.20 (4.51)	Spokane, WA	569.1	26.30 (6.43)
St. Louis, MO	2688.5	211.09 (2.80)	Wichita, KS	563.9	35.60 (7.65)
Tampa-St. Petersburg-Clearwater, FL	2649.1	192.18 (4.78)	Madison, WI	539.5	75.33 (5.68)
Denver-Boulder, CO	2603.5	283.61 (17.33)	Ft. Wayne, IN	520	31.79 (3.59)
Portland, OR	2352.2	284.25 (30.24)	Boise, ID	509.9	43.84 (1.12)
Cleveland, OH	2133.8	167.19 (2.20)	Lexington-Fayette, KY	509	39.58 (1.06)
Charlotte-Gastonia-Rock Hill, NC-SC	2126.7	121.59 (5.19)	Augusta, GA	498.4	27.65 (3.62)
Sacramento, CA	2099.6	246.04 (24.65)	Chattanooga, TN	494.5	43.11 (0.99)
Salt Lake City-Ogden-Provo, UT	1924.1	150.74 (7.81)	Roanoke-Lynchburg, VA	470.7	40.09 (3.37)
San Antonio, TX	1900.4	158.01 (5.58)	Jackson, MS	468.6	39.13 (2.62)
Kansas City, MO-KS	1870.8	140.34 (1.66)	Reno, NV	452.7	70.07 (0.66)
Las Vegas, NV	1752.4	118.53 (7.07)	Fayetteville, NC	438.9	28.60 (0.88)
Milwaukee-Racine, WI	1712.5	128.64 (3.79)	Shreveport, LA	399.6	25.16 (1.96)
Orlando, FL	1686.1	231.78 (12.84)	Quad Cities, IA-IL	358.8	26.70 (1.88)
Columbus, OH	1685	130.80 (5.48)	Macon, GA	337.1	24.99 (0.44)
Indianapolis, IN	1601.6	104.97 (2.28)	Eugene-Springfield, OR	336.4	23.81 (0.43)
Norfolk-Virginia Beach-Newport News, VA	1582.8	158.54 (0.80)	Portland, ME	276.1	41.42 (4.11)
Austin, TX	1466.3	337.14 (318.09)	South Bend, IN	267	28.71 (1.58)
Nashville, TN	1341.7	158.72 (163.83)	Lubbock, TX	255.3	33.59 (0.37)
Greensboro-Winston Salem-High Point, NC	1328.9	72.84 (10.86)	Binghamton, NY	247.9	21.51 (0.27)
New Orleans, LA	1293.7	82.99 (11.34)	Odessa-Midland, TX	247.8	18.37 (0.31)
Memphis, TN	1278	83.32 (31.29)	Yakima, WA	231.4	18.53 (0.23)
Jacksonville, FL	1270.5	80.84 (14.98)	Duluth-Superior, MN-WI	200.3	24.76 (0.22)
Oklahoma City, OK	1268.3	64.98 (10.06)	Medford-Ashland, OR	196.2	19.47 (0.19)
Buffalo-Niagara Falls, NY	1150	104.51 (9.26)	St. Cloud, MN	191.2	16.05 (0.88)
Louisville, KY	1099.6	91.66 (13.86)	Fargo-Moorhead, ND-MN	183.6	24.36 (0.31)
Richmond, VA	1066.4	65.93 (13.73)	Abilene, TX	159.1	15.62 (0.21)
Birmingham, AL	1030	72.34 (11.61)	Eau Claire, WI	156.5	20.40 (0.36)
Tucson, AZ	938.3	55.66 (12.37)	Monroe, LA	149.2	18.90 (1.40)
Honolulu, HI	909.4	62.81 (8.33)	Parkersburg-Marietta, WV-OH	149.2	14.74 (0.19)
Albany-Schenectady-Troy, NY	902	101.85 (8.79)	Grand Junction, CO	130	11.47 (0.88)
Tulsa, OK	870.2	62.31 (10.25)	Sioux City, IA	123.7	11.70 (0.15)
Ft. Myers-Naples-Marco Island, FL	864.1	113.01 (149.48)	Williamsport, PA	118.3	11.29 (0.15)
Grand Rapids, MI	856.4	56.45 (13.14)	San Angelo, TX	103.8	10.18 (0.06)
Albuquerque, NM	784.9	58.67 (23.95)	Bismarck, ND	99.2	12.80 (0.15)
Omaha-Council Bluffs, NE-IA	740.3	48.26 (10.31)			

Standard errors (corrected for the first stage) in parentheses

Table 11: Intercepts of an advertiser inverse demand function for each market. Units are 1996 US dollars for a 30 second ad slot listened by a 1% of the market population.

	Population <.5	Population .5M-1.5M	Population 1.5M-3.5M	Population >3.5M
OLS	-0.26^{***} (0.01)	-0.14^{***} (0.01)	-0.15^{***} (0.00)	-0.09^{***} (0.00)
2SLS	-0.19^{***} (0.01)	-0.14^{***} (0.05)	-0.13^{***} (0.00)	-0.08^{***} (0.00)

Standard errors (corrected for the first stage) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 12: Slope of advertising price per rating point (CPP). Intercept is set to 1. Units are standard deviations of quantity supplied on a station level.

	Mean level			Quality intercept		
	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M
OLS	2.60^{***} (0.09)	2.08^{***} (0.15)	1.05^{***} (0.09)	0.18^{***} (0.01)	0.11^{***} (0.01)	0.04^{***} (0.00)
2SLS	3.06^{***} (0.10)	2.08^{***} (0.50)	1.22^{***} (0.08)	0.20^{***} (0.01)	0.11^{***} (0.02)	0.05^{***} (0.00)

Standard errors (corrected for the first stage) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 13: Marginal cost per minute of advertising sold. Intercept of advertising price per rating point is set to 1. Note that these numbers might be higher than one because the final price of advertising is CPP times the station rating in per cent. Units for quality are standard deviations of quality in the sample.

	Cost synergies		
	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M
OLS	−0.51*** (0.05)	−0.13*** (0.04)	−0.24*** (0.03)
2SLS	−0.43*** (0.05)	−0.13 (0.08)	−0.21*** (0.04)

Standard errors (corrected for the first stage) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 14: Marginal cost synergies from owning multiple stations of the same format.

		1998	1999	2000	2001	2002	2003	2004	2005	2006
OLS	<.5M	−0.12 (0.08)	−0.70*** (0.08)	−0.80*** (0.08)	−0.64*** (0.09)	−0.70*** (0.08)	−0.52*** (0.09)	−0.61*** (0.08)	−0.47*** (0.08)	−1.09*** (0.09)
	.5M-1.5M	−0.20*** (0.07)	−0.39*** (0.07)	−0.43*** (0.07)	−0.26*** (0.09)	−0.22*** (0.07)	−0.25*** (0.08)	−0.35*** (0.07)	−0.33*** (0.07)	−0.75*** (0.09)
	>1.5M	−0.21*** (0.07)	−0.51*** (0.07)	−0.45*** (0.07)	0.06 (0.07)	−0.13** (0.07)	−0.02 (0.07)	−0.23*** (0.07)	−0.16** (0.07)	−0.18** (0.07)
2SLS	<.5	−0.14 (0.08)	−0.68*** (0.09)	−0.70*** (0.09)	−0.68*** (0.09)	−0.61*** (0.09)	−0.57*** (0.09)	−0.56*** (0.09)	−0.41*** (0.09)	−1.12*** (0.09)
	.5M-1.5M	−0.20*** (0.07)	−0.39*** (0.07)	−0.43*** (0.07)	−0.26* (0.15)	−0.22*** (0.07)	−0.25** (0.12)	−0.35*** (0.07)	−0.33*** (0.07)	−0.75*** (0.12)
	>1.5M	−0.20*** (0.07)	−0.48*** (0.07)	−0.41*** (0.07)	0.03 (0.07)	−0.12* (0.07)	−0.04 (0.07)	−0.21*** (0.07)	−0.15** (0.07)	−0.21*** (0.07)

Standard errors (corrected for the first stage) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 15: Time effects in the marginal cost. 1996 and 1997 values are normalized to zero.

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Small markets <0.5m pop.	1.3pdm +0.3%	-8.1pdm -25.4%	-56.2m -32.6%	-10,707min -16.0%	+7.16%
Medium markets 0.5-2m pop.	-0.3pdm -0.1%	-8.7pdm -21.8%	-68.0m -23.5%	-3,582min -8.8%	+5.75%
Large markets >2m pop.	1.7pdm +0.3%	-6.4pdm -13.7%	-99.1m -17.1%	-954min -2.9%	+7.39%
All markets	1.0pdm +0.2%	-7.3pdm -16.8%	-223.3m -21.4%	-15,243min -10.9%	+6.53%

Table 16: Total impact of the regulation on consumer welfare by market size.

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Impact of ownership change and format switching No ad adjustment	1.8pdm +0.3%	-8.9pdm -20.4%	-179.5m -17.2%	-161min -0.1%	+3.00%
Impact of ad adjustment	-0.8pdm -0.1%	1.6pdm +4.6%	-43.8m -5.1%	-15,081min -10.7%	+3.34%
Total impact of ownership change format switching and ad adjustment	1.0pdm +0.2%	-7.3pdm -16.8%	-223.3m -21.4%	-15,243min -10.9%	+6.53%

Table 17: Decomposition of the consumer welfare changes into product repositioning and quantity readjustments.

	Population <.5	Population .5M-1.5M	Population 1.5M-3.5M	Population >3.5M
Baseline model	-0.19^{***} (0.01)	-0.14^{***} (0.05)	-0.13^{***} (0.00)	-0.08^{***} (0.00)
Oligopoly within format	-0.15^{***} (0.01)	-0.08^{***} (0.03)	-0.10^{***} (0.00)	-0.06^{***} (0.00)
Perfect substitutes	-0.24^{***} (0.01)	-0.15^{**} (0.06)	-0.14^{***} (0.00)	-0.08^{***} (0.00)

Standard errors (corrected for the first stage) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 18: Robustness of the estimates of slope of advertising cost per rating point (CPP).

	Mean level			Quality intercept		
	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M
Baseline model	3.06^{***} (0.10)	2.08^{***} (0.50)	1.22^{***} (0.08)	0.20^{***} (0.01)	0.11^{***} (0.02)	0.05^{***} (0.00)
Oligopoly within format	2.97^{***} (0.10)	2.50^{***} (0.36)	1.31^{***} (0.08)	0.19^{***} (0.01)	0.12^{***} (0.02)	0.05^{***} (0.00)
Perfect substitutes	3.06^{***} (0.10)	2.26^{***} (0.55)	1.31^{***} (0.08)	0.20^{***} (0.01)	0.12^{***} (0.03)	0.05^{***} (0.00)

Standard errors (corrected for the first stage) in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 19: Robustness of marginal cost per minute of advertising sold.

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Baseline model	0.6pdm +0.1%	-7.3pdm -17.1%	-219.7m -21.7%	-15,150min -11.0%	+6.59%
Oligopoly within format	-0.9pdm -0.2%	-5.8pdm -15.1%	-150.7m -21.3%	-17,533min -12.3%	+3.16%
Perfect substitutes	2.2pdm +0.4%	-8.7pdm -19.5%	-251.9m -26.1%	-19,170min -13.7%	+17.68%

Table 20: Robustness of counterfactuals.