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

Przemysław Jeziorski †

August 26, 2013

Abstract

This study examines mergers in two-sided markets using a structural supply-and-demand model that employs data from the 1996-2006 merger wave in the U.S. radio industry. In particular, it identifies the conflicting incentives for merged firms to exercise market power on both listener and advertiser sides of the market, and disaggregates the effects of mergers into changes in product variety and advertising quantity. Specifically, it finds 0.2% listener welfare increase (+0.3% from increased product variety, and -0.1% from decreased ad quantity) and 21% advertiser welfare decrease (-17% from changes in product variety, and -5% from decreased ad quantity).

*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 broadcasting industry experienced an unprecedented merger wave. This phenomenon 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 per year and about 14% changed programming formats. The present study utilizes this merger wave to examine the consequences of consolidation in two-sided markets, and makes two main contributions. The first is identification of conflicting incentives for stations to exercise market power on both sides of the market (in the case of radio broadcasting, the two sides are advertisers and listeners). Specifically, the study separates the impact of consolidation on listener and advertiser surplus. Second, this impact is decomposed into two parts: changes in product variety and market power, with particular emphasis placed on the evaluation whether extra variety is able to mitigate the negative effects of a decrease in competition. Because similar issues arise in other two-sided markets, including 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 mergers, 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 advertising sales, which are priced on a per-listener basis. Number of advertising minutes acts as a price for radio programming because listeners are averse to advertising; thus, a merged firm has an incentive to increase ad quantity. 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 lowers ad prices; thus a merged firm prefers to supply less advertising. Such a decision would have the opposite impact, raising listener welfare and lowering that of the advertiser. The firm's ultimate decision depends on the relative demand elasticities in the two markets. In this study, I separately identify these two elasticities and subsequently compute the impact of the 1996 deregulation on listener and advertiser welfare.

After 1996, consolidation in the radio market was characterized by a set of patterns in 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 the exercise of 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 advertising supply. One of the goals of this paper is to develop and estimate a structural model of demand and supply for 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 resulted 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% loss in advertiser welfare. By comparison, two reduced form studies by Brown and Williams (2002) and Chipty (2007) find, respectively, a 4% price decrease and no systematic impact of mergers on ad prices. Additionally, I find that the market power on listeners is similar across geographical markets, whereas the market power on advertisers depends on a market's population. In particular, firms have considerable control over advertising prices in smaller markets but 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 populations lower than 0.5M, compared to a 17.1% drop in the markets with populations greater than 2M).

In addition to quantifying the effects of mergers on both sides of the radio programming market, the second contribution of this study is the decomposition of these effects into changes in product variety and that extra market power that arise from joint ownership. This exercise is motivated by the fact that in most cases, consumers have a preference for variety, so the increased variety created by mergers might mitigate the negative effects of additional market power. To verify the above claim, I quantify consumer value for increased variety and compare it to the loss in surplus coming

[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.

from the expanded market power. This approach is similar to that of Kim, Allenby, and Rossi (2002), who compute the compensating variation for the changes of variety in the taste of yogurt, and Brynjolfsson, Hu, and Smith (2003) who perform a similar exercise for the variety of books offered in on-line bookstores. These studies ignore the fact that changes in variety are followed by readjustments in equilibrium prices. I take these analyses one step further: I incorporate these strategic responses by performing counterfactual experiments in which new equilibrium prices are computed.

Berry and Waldfogel (2001) and Sweeting (2009) document that the merger wave occurring after 1996 resulted in an increase in product variety. I investigate their results with a structural utility model and find that the increased variety led 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 that this effect leads to a 0.1% decrease in listener welfare and 5% decrease in advertiser welfare. The two effects have a 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 studies that examine the complexity of pricing strategies in two-sided markets. The closest studies are Armstrong (2006), Rochet and Tirole (2006), Evans (2002) and Dukes (2004). The general conclusion in this literature is that the standard supply-and-demand framework of single-sided markets may 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 on 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 for differentiated products. In particular, I include the richer consumer heterogeneity and substitution patterns (e.g., Berry, Levinsohn, and Pakes (1995), and Nevo (2000)) that

[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.

are necessary to capture complicated consumer preferences for the 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.

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 (2012) study newspapers. An additional complication for an analysis of these industries is the presence of two prices in place a single price, as in radio. The main difference between this work and Song (2011) and that of Van Cayseele and Vanormelingen (2009) is that these authors treat advertisers as local monopolists. Consequently, they assume away the potential impact of mergers on 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. The analysis in this study complements Fan (2012), who allows for a richer model of characteristics choice, but does not observe advertising quantities. Therefore, while she can endogenize the variety in the counterfactuals, she is not able to study post-repositioning quantity readjustments. Where Fan finds that mergers decreased reader surplus, I find that listener surplus increased as the result of lower post-merger advertising quantity. Finally, we both agree that mergers decrease advertising surplus.

This study is organized as follows. Section 2 provides a formal outline of investigation and describes the structural model of the radio broadcasting industry. Section 3 describes the data, Section 4 outlines the estimation techniques used to identify the parameters of the model, and Section 5 presents the results of the structural estimation. Section 6 describes the results of counterfactual experiments, Section 7 contains robustness checks of different modeling assumptions, and Section 8 concludes.

2 Radio as a two-sided market

The radio industry is 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: there are two types of consumers, as previously discussed, 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 any given consumer buys determines the market clearing price for the other consumers. In this subsection, I argue that this feedback complicates the process of determining whether the supplied quantities are strategic substitutes or complements, as these terms have been 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. I will now consider this mechanism in detail using the example of radio broadcasting, although, the same paradigm 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 by selling advertising slots. To begin, 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 listener demand. Because radio programming is free, this equation does not contain price explicitly. However, because listeners have a disutility for advertising, the effect of programming devoted to ads 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 station j's listeners. 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, q, and indirectly, through the listener demand feedback loop

[3]Note the equation (2.2) implicitly leverages 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. In case of radio the pricing is done on the per listener basis, thus such inverse demand exists and is linear in r_j as long as the utility of advertisers in linear in the number of listeners that hear an ad. For the more general result, which allows prices to depend on the full vector r, see White and Weyl (2010).

represented by the second argument, $r_j(q)$.

Suppose, for the time being, that each owner possesses a single station and no marginal cost exists. In equilibrium, each radio station chooses its optimal ad quantity, while keeping the ad quantities of the other stations fixed, namely,

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

In contrast to a differentiated products oligopoly, the firm has just one control, that is 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 which is reminiscent of the standard quantity-setting model (Cournot). On the other hand, 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 these quantities are strategic complements or substitutes. Without knowing the relative strengths of the direct effects and the feedback loop, we cannot conclude whether a merger will increase or decrease ad quantity. Moreover, in the borderline case in which the effects cancel each other, a merger does not affect quantity at all. In this case, even though companies have market power over both kinds of consumers, they cannot 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 to higher prices. Reduced advertising quantity benefits listeners but hurts advertisers. However, if the feedback loop is stronger than the direct effect, then mergers lead to increased advertising and lower ad prices, benefiting advertisers but 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 the separate identification of both sets of demand elasticities. I discover that the relative strength of the direct and feedback effects and perform counterfactuals which 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. The format is assigned by a specialized consulting company and describes the majority of the station's programming. It includes, among others, such designations as Rock, Urban, Country or News/Talk. Beyond the usual categories, station formats include the so-called "dark" format, in which a frequency assigned to a broadcaster is kept off the air. 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 a station/format configuration $s^{mt} \in \mathbb{F}^{\mathbb{J}^m}$. With some 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 the 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 of Berry, Levinsohn, and Pakes (1995).

I assume each listener chooses only one radio station to listen to at any 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 variable 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

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

where θ^L_{2it} is that listener's demand sensitivity to advertising, q_{jt} is the amount of advertising, ξ_{jt} is an unobserved station quality, ϵ_{jit} is an unobserved preference shock (distributed type-1 extreme value), and θ^L_{1it} is a vector of format specific random effects.

I assume the above random coefficients can be decomposed as

$$\theta^L_{1it} = \theta^L_1 + \Pi D_{it} + \nu_{1it}, \quad D_{it} \sim F_{mt}(D_{it}|d), \quad \nu_{1it} \sim N(0, \Sigma_1)$$

and

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

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 with time and across markets, it captures trends in demographics that can affect the 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 individual listeners.

The random coefficients 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 decompose the utility into: a part that does not vary with consumer characteristics

$$\delta_{jt} = \delta(q_{jt}|\iota_j, \xi_j, \theta^L) = \theta^L_1 \iota_{jt} - \theta^L_2 q_{jt} + \theta^L_3 \text{FM}_j + \xi_{jt}$$

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 as well as the fact that ϵ_{ji} is distributed as an extreme value, one can derive the expected station rating conditional on an advertising quantity vector q, market structure s, an unobserved station characteristics vector 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 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 advertising model proposed in this section captures several important features specific to the radio industry. In particular, the pricing can be approximated with per-listener rates; that is, the price for a 60sec slot of advertising can be approximated by a product of cost-per-point (CPP) and station rating, which is the market share specified as a percentage. 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 the predicted ratings, many radio stations compensate the advertisers with discounted rates, or offer additional advertising spots to offset the difference. Thus, effectively, I assume that either the forecasting error is negligible, or the offsets make up for 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 substitution of advertising across different formats can be caused, among others, by listeners' multihoming (see Gentzkow, Petek, Shapiro, and Sinkinson (2013)) or advertising congestion (see Anderson, Foros, Kind, and Peitz (2012)). Although I am unable to distinguish between these possibilities, the bottom line of this study is unaffected as long as the reduced form of the inverse demand for advertising is specified correctly.

Let p_{jt} be the cost-per-point at station j, then the price of a 60sec slot is given by $P_{jt} = p_{jt} r_{jt}$. 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 the 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. Small values of θ_2^{Am} would indicate that the advertisers are price sensitive, or elastic. In such cases the radio station has difficulty controlling the market-clearing per-listener price, and this results in low market power in the advertising marker. Conversely, for a large θ_2^{Am}, advertisers are inelastic, which generates additional market power for radio stations.

The weights ω are key factors determining advertiser's substitution patterns across formats and, thus, radio owners' market power. These weights reflect the fact that formats can be further or closer substitutes for advertisers because the demographic composition of listeners varies from format to format. 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 adequate 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 determining the weights and provides an example supporting this intuition.

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, otherwise he gets zero utility. 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 to format f is of type a. Advertisers take these numbers, along with station ratings r_j, as given and make their advertising choices. The fact that most of the advertising is purchased by local small business confirms this assumption. The advertising decisions of such firms are unlikely to influence listener sorting 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

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

These weights correspond to the marginal change in market clearing prices in response to supplied quantities, thus they are proportional to the change in the marginal advertisers' valuation. 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 proportional to two things: (i) the probability of correct targeting in format f, given by $r_{a|f}$; and (ii) the share of advertising purchased by advertiser a in format f', given by $r_{f'|a}$. The former relationship is a consequence of the assumption that advertisers are expected utility maximizers. The latter is motivated by an observation that advertisers who purchase more ads in format f' should be more affected by changes in $q_{f'}$. In order to compute the average impact of changes in q on the overall price of ad in format f, I assume that sales representatives approach the advertiser of type a with a probability equal to corresponding listener share given by $r_{a|f}$. Assembling these pieces together and normalizing the weights to sum to 1 yields 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 1.

In Table 1, the impact of Hits on the price of Talk is greater than the impact of Talk on the price of Hits, because the ad quantity supplied in the Hits format affects advertisers who target Adults to a much greater extent than ad quantity in Talk affects advertisers who target Teens. These two sets of advertisers drive the Talk and Hits format pricing, respectively. Moreover, because the weights sum to 1, the own effect of Talk must be weaker than that of Hits. This example provides insight into the nature of the mechanism addressed by 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: an adaptation of equation (2.3), which allows for non-zero cost of 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

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

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 study, I assume that the compensation for sellers is tallied by the minute of advertising sold. Because a station's revenue per minute of advertising depends on the quality of its programming, the seller's commission is also likely to be 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 seller's commission. Another source of marginal cost is the in-house production of advertising. In contrast to newspapers and TV, radio stations produce many of their ads. In-house ads might take the form of voice announcements, endorsements by 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 while allowing for unobserved heterogeneity at the station level 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 also allows for station-level unobserved heterogeneity captured by η_{jt}. Terms θ^{Cmt} are time dummies that capture aggregate shocks to marginal cost. Unobserved market level heterogeneity can be captured because θ_1^{Am} is allowed to vary for each market, and θ^{Cm} is allowed to vary 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 ads with similar target groups can be sold jointly by the same sales agent and such ads ofter create economies of scale during production. The term θ_3^{Cm} is multiplied by the dummy variable SYN_{jt} that is equal to 1 if the owner possesses more than one station with the same format.

Formulation 2.7 does not take into account cross-station bundling of advertising slots, which could be a problem if the owners of multiple stations were to offer large quantities of exclusive ad packages that do not permit unbundling. There are several reasons why bundling should not be an issue. Primarily, these kinds of transactions are likely to take place with large advertisers, but not with smaller clients. Consequently because most of the clients are small, bundling is likely to affect only a fraction of transactions. Additionally, I find little discussion of bundling in the regulatory reports that investigated two of the larger mergers in the industry, which suggests such practices are not of first-order importance. Finally, bundling should be more important within than across formats, because the advertisers are likely to buy multiple ads from stations of the same format. In such cases, bundling would be picked up by the cost synergy SYN_{jt} dummy and netted out in the counterfactual. However, because the welfare interpretation of SYN_{jt} might change, this study assumes away bundling and focuses on cost synergies instead.

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

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

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

Finally, I assume that a station's unobserved quality is exogenous, but serially correlated, similar to the approach employed by Sweeting (2012). The quality 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 by merging data from two sources: *BIA Kelsey* and *SQAD Media Market Guide*.

BIA Kelsey provides the 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 (2012), only 50 stations exited during this period mostly for violating FCC regulations. Because this figure is small relative to the 11,000 stations in the sample, the omission is unlikely to influence the results significantly.

An observation in my data is a radio station operating in a specific half-year and in a specific market. I follow BIA and SQAD who use Arbitron market definitions, which in most cases and a county or metropolitan area. According to the 2009 survey by Radio Advertising Bureau national advertising contributed only 15% of overall advertising revenue. Moreover surveys conducted by CRA International (2007) for the Canadian market, which resembles US market, have shown that "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 exclude overlap between markets, I use only the market sub-selection developed in Sweeting (2012) in which I drop markets that have no SQAD data. Table 2 presents a list of the 88 markets used in this study, along with their populations.

To achieve a sharper identification of the random effects covariance matrix, for each of the formats I utilize the listenership shares of the demographic groups that have been aggregated from the 100 largest 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 observe what the share of rock stations is among black, educated males. Instead, I have shares for blacks, educated people, and males.

[4] Source: Arbitron Format Trends Report

3.1 Advertising prices and quantities

The main dependent variables in this study are prices and quantities of advertising. This subsection provides 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 translate directly into the strengths and weaknesses of the imputed quantity data.

I obtain station-level revenues from BIA Kelsey, a large consulting company serving the radio broadcasting markets. The BIA data set on revenues is considered the most complete and reliable resource for 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 that includes all active and inactive radio stations in the United States as of 2006. Because the survey data is self-reported, it may suffer from self-reporting and self-selection bias. To compensate for these problems, BIA corrects the numbers through a direct consultation with radio partners and through the application proprietary statistical tools. For example, the missing responses are imputed using historical data and estimated industry trends. Numerous academic papers utilize this data, including Sweeting (2012), Mooney (2010), Waldfogel and Wulf (2006). However, any systematic measurement error that is correlated with the changes in the market structure 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 reasonable accuracy. For this reason, I have obtained verbal assurance 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) from Arbitron, which is the largest consulting company reporting on radio markets. Arbitron 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 a representative population. The survey panel changes periodically, and the results are averaged quarterly. These manipulations are intended to reduce the individual-level reporting bias and to diversify the population sample in order to net out unobserved heterogeneity. Because the data is self-reported and the panel of respondents varies, one might be concerned about the comparability of responses between time periods. However, the qualitative conclusions of this paper would not be affected unless the relevant measurement error is correlated with changes in the ownership. In such a case, the reporting bias could affect nominal values of consumer welfare. Nevertheless, the percentage changes in welfare should be less susceptible.

The third data source is SQAD, which provides average market prices per listener. These figures are provided by advertising agencies, which report directly on actual transactions. The advantage of this collection procedure is that it is not based on surveys and is not adjusted by proprietary formulas. The downside is that station-level prices cannot be observed directly. The data is organized by demographics, which means that price per listener is given separately for different gender and age groups. Using these numbers and the ratings of stations across different demographic groups, one 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 have chosen 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 3 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. The sample is fairly dispersed and many, usually niche, stations broadcast small amounts of advertising. Because the average number of ads appears to be fairly small, I obtained additional data to validate this observation. The data comes from a Federal Communication Commission Research Study on Media Ownership, which directly categorized the programming content of a subset consisting of $1,000$ stations in the year 2007. Each station was recorded for two hours, and each five second increment was identified as containing either ad or non-ad content. This database might constitute 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, using this FCC database, I computed that, for my sub-selection of markets, stations play about 23 minutes of advertising per day. This figure is slightly lower than the average in my sample, but it is reasonably close. I can account for some of the difference by the fact that my 37.5 minute average does not include stations that did not meet Arbitron measurement standards (rating below 0.5%), and these stations are likely to air a negligible amount of advertising. If I include those stations, my average gets closer to the FCC number and drops from 37.5 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 4 documents changes in the concentration of radio-station ownership. The average number of stations per owner 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 increased format switching. The percentage of stations that switched formats peaked in 1998 at 14%.

Next, in Table 5, I report statistics on interactions between acquisitions and ad quantity. I computed an average change in station-level ad quantity after the merger and compared 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 in the first year after the acquisition, compared to a market trend increase of 0.4 minutes. 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 as a result of the recession. The recovery is slow, and the spending does not reach the levels of 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 study, I identify another explanation, an increased market power, which is also consistent with the numbers in Table 5.

Figure 4 contains trends for national levels of radio advertising quantity. The quantity is fairly volatile, and this observation is consistent with the 2002 FCC radio profitability study, "Radio Industry Review." The study reports median profit margins for a large subset of radio owners. These margin numbers are directly related to the ability to sell advertising and their fluctuations resemble those of my ad quantities. For example, the median EBIT margin, as reported by the FCC, can move by as much as 30% within a year. Because advertising prices show little variance over time, a fluctuation in ad quantities is a natural candidate to explain high volatility in the EBIT margin found by the FCC.

A considerable degree of volatility can be observed in station- and market-level data as well [5]. I find that the overall variance is less 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 to be small relative to market trends. The small contribution of station-level unobservables is reassuring, because the rationalization and decomposition of the market trends is one the main goals of this study.

[5] Online appendix, available on `http:\\jeziorski.me`, contains figures depicting ad quantity for three representative markets by size: Los Angeles, Albuquerque, and Bismark; other markets show similar patterns. In particular I report a graph of ad quantities for the most popular station in each of these markets and market-level average ad quantity for a number of the largest stations (20, 10, and 5 depending on the market size).

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

A component of the analysis in this study is a description of the impact of format-switching patterns. A number of studies have investigated the consequences of mergers for product variety (e.g. Berry and Waldfogel (2001) or Sweeting (2009)). The general conclusion is that mergers increase variety, which is consistent with the avoidance cannibalization and the creation spatial 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 6. The matrix contains the probabilities of switching across formats for the entire population of stations as well as for stations where ownership changed. I find the probability of switching a format is higher for stations that change owners, which suggests that format switching and mergers might be complementary. If true, 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 calls attention to strategic portfolio considerations. For this reason, welfare calculations should include format switching. Interestingly, for some formats, acquisitions trigger switching to the Dark format, which might mean that some mergers are intended to reduce the number of active stations.

4 Estimation

I conduct the estimation of the model in two steps. In the first step, I estimate the demand model, which 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 second stage estimates by treating the problem as a large GMM system in order to accommodate the first stage estimation error.

4.1 First stage

This stage provides the estimates of the demand for radio programming θ^L, which are obtained through the generalized method of simulated moments. I employ two sets of moment conditions, the first of which is based on the fact that innovation to a station's unobserved quality ξ_j has a mean of zero conditional on the instruments employed:

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

This moment condition extends Berry, Levinsohn, and Pakes (1995) by explicitly first differencing auto-correlation of ξ. This procedure produces efficient estimates, but more importantly it allows me to use instruments based on the timing assumptions, while allowing for serial correlation in ξ. Instrumentation for advertising quantities is necessary because these quantities are likely to be correlated with unobserved station quality. My instruments include lagged means and second central moments of competitors' advertising quantity, lagged market HHIs, lagged numbers and cumulative shares 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.

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

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

[6] I divided the estimation 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.

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 in both equations and summing across a sub-population possessing characteristic c in the equation (4.2). This gives an $\mathcal{I}_{mtc}$ number of draws for each characteristic.

I formulate the problem using Mathematical Programming with Equilibrium Constraints through a procedure similar to that of Dube, Fox, and Su (2012):

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

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 the 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 these preferences are assumed to be constant across markets and equal to the values by reported by Arbitron's Radio Today reports. Specifically, the advertisers' beliefs about the listener aggregate preferences for formats conditional on demographics are assumed to be constant nation wide. However, to recognize that advertisers know the demographic composition of each market, I allow for market-specific composition of listeners in each format $r_{a|f}^m$. For example, in markets with large Hispanic population I allow the share of Hispanic listeners to be greater for every format. I obtain $r^m_{a|f}$ by employing the Bayes' rule $r^m_{a|f} \propto r_{f|a} r^m_a$, where r^m_a is the proportion of listeners of type a obtained from the Current Population Survey that was aggregated over the period of 1996-2006. I use 64 listener types, which are a combination of the following binary variables: gender, Hispanic, Black, income over \$40,000 a year, Bachelor degree or higher, and 44 or more years of age. 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 owners' profit maximization (see equation (2.7)) to set up a system of linear equations:

$$\begin{aligned} r_{jt} + \sum_{j' \in s_{kt}} q_{j't} \frac{\partial r_{j't}(q_t)}{\partial q_{jt}} &= \\ \theta^{Cmt} + \theta^{Cm}_1 + \theta^{Am}_2 \left[r_{jt} v_j + \sum_{j' \in s_{kt}} \left(r_{j't}(q_t) \omega^m_{jj'} + v_{j'} \frac{\partial r_{j't}(q_t)}{\partial q_{jt}} \right) \right] &+ \theta^{Cm}_2 \xi_{jt} + \theta^{Cm}_3 \mathrm{SYN}_{jt} + \eta_{jt} \end{aligned} \tag{4.4}$$

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

Because the equation does not depend on θ^A_1, it may be used to estimate θ^A_2 and θ^C. Two sources of heterogeneity in marginal cost and slope coefficients exist across markets because the effective marginal cost parameters for each station in market m are given by $\theta^{Am}_1 \theta^{Cm}$. I allow θ^{Am}_1 be different for each market, and I allow for three different sets of values for 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.

A similar specification is used for the slope of the inverse demand for ads, and the effective slope of this inverse demand is given by $\theta^{Am}_1 \theta^{Am}_2$. To control for the fact that the market power of stations might differ across markets depending on their 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

[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.

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 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 utilizing 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 the listener demand slope, separately from advertiser demand slope. The listener demand slope is identified from the variation in advertising quantity and ratings. During this procedure, I do not impose optimality of quantities. Instead, I utilize the size of response in ratings to changes in ad quantity caused by exogenous factors. The structural model enables me to control for confounding factors, which 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, $\frac{\partial r_j}{\partial q_j}$, which I incorporate into supply-side optimality conditions given by equation 4.4. Using these optimality conditions I find advertiser demand slope and marginal cost that rationalizes the observed response in advertising quantity to exogenous variables, including demographics, macro time trends, or shocks to competitors' quality.

The slope parameter θ_2^A is identified separately from marginal cost using the observed response

[8]One could potentially allow for CPPs to depend on station quality, which would introduce a non-linear relationship between ad price per minute and ratings. Because I am unsure whether stations with large 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).

of advertising quantity to mergers. In particular, assuming the CPP slope is flat, the estimated slope of listener demand would predict large increases in ad quantity after the merger. However, in the data, I frequently observe a decrease in the quantity supplied after the merger. This fact can be rationalized by postulating a negative value of CPP slope, θ_2^A, but cannot be rationalized by marginal cost.

The time trends in advertising demand and marginal cost are identified using a triple-panel structure of the data, which contains observations on multiple time periods and markets. 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 on Ω^m presented earlier in this section. 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 show that 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 7 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 7 contains variances of random effects for station formats. The higher a format's variance, the more persistent are the listeners' tastes for that format. For example, in response to an increased amount of advertising, listeners tend to switch to a station of the same format if the variance of the random effect for that format is high. 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, while older people prefer News/Talk. The negative coefficients on the interaction of a Hispanic format with education and income suggest that less-educated Hispanic people with lower incomes are more inclined 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.

Table 9 presents the estimates of the time shocks to the utility of an outside option for the listeners. The trend in the outside option captures the influx if the internet and satellite radio, which results in the decline in listenership. Such downward trend lowers the incentives of the radio stations to exercise market power in the listenership market, which can be an alternative explanation for the increase in advertising prices. For this reason it is necessary to account for these trends when computing the merger counterfactuals.

5.1.1 Discussion of the instruments

To determine whether the instruments used in the demand estimation are helpful in fixing the endogeneity bias, I perform an estimation of the model without random effects using 2SLS for comparison with an OLS. The results[9] are reported in the third and fourth columns of Table 7. I find the endogeneity of quantity produces a biased OLS estimator. The estimated disutility from advertising declines nearly by half when using OLS, which flattens the demand curve. This bias goes in the right direction if ad quantity q is positively correlated with programming quality ξ. The extensive discussion of the mechanism of such bias is contained in Berry (1994).

[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.

Additionally, I tested for weak instruments by performing a joint F-test on parameters from the first-stage regression in the 2SLS procedure. I obtain an F-statistic equal to 3932, which overwhelmingly rejects the hypothesis that there is no linear dependence of quantity on instruments. 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

Table 10 presents the weights for selected markets representing large, medium, and small listener populations. These weights were computed using the 1999 edition of the Radio Today publication and the Common Population Survey aggregated from 1996 to 2006. I also compute a total impact coefficient that is the sum of all the columns of this table for each format. Not surprisingly, general interest formats such as AC and News/Talk have the greatest impact on the price of advertising, whereas the Spanish format has the least. The values on the diagonals of the matrices represent the formats' own effect of the quantity of advertising supplied on per-listener price. These diagonal values are usually larger than the off-diagonal values, which suggests that ad quantity in the same format is a strong determinant of per-listener price for stations within the same format. In accordance with intuition, the formats with the most demographically homogeneous listener pools, Urban/Alternative and Spanish, have the highest values for own effects. On the other hand, general interest formats such as CHR and Rock are characterized by the smallest values of the own effect, which indicates that their target population has greater dispersion across other formats. For cross effects, one notices that News/Talk is close to AC and that Urban is close to CHR. This observation can be explained, for example, by the age of the listeners. In the former case, the formats appeal to an older population, and in the latter case, to a younger one.

Estimates of intercepts θ_1^{Am} can be found in Table 2. The between-market variation of these intercepts picks out large variation of revenue across the markets in the data. Because the ad prices are measured on per 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; and (ii) the amount of market power in the advertising market. Because of the second factor, I postpone further discussion until I have presented the slope coefficients.

The estimates of the slope of the inverse demand for advertising θ_2^{Am} can be found in Table 11. 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 using the standard deviation of the 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, corresponding to an increase in 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 intercepts of \$342 and \$337, respectively, which is more than four times higher per-listener in Austin because Dallas has much larger population. 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 interact and generate much higher intercept estimates in Austin than in Dallas.

5.3 Supply

Estimates of a marginal cost level θ_1^{Cm} and quality coefficient θ_2^{Cm} can be found in Table 12. 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 greater than one because the ad price is additionally multiplied by station market share measured in percentages. One convenient interpretation of the marginal cost level is that is presents the lower bound on the rating of the average station required for that station to be profitable. I find that in small markets, an average station needs at least a 3% market share to break even, whereas in larger markets, this number drops to 1.1%. Because there are many more stations in the bigger markets, this difference confirms my prior intuition. Note that the above numbers apply to single-owned stations of average quality. Because smaller stations have lower unobserved quality, they also enjoy reduced marginal costs. Thus the model predicts positive variable profits even for stations with ratings that are less than the above-average break-even thresholds. [10] Moreover, during the supply side estimation, the first order conditions are enforced. This means that the station that appears to be unprofitable, but supplies a positive amount of advertising, is automatically assigned a negative shock value to the marginal cost, an assumption which 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 revenue increase from an additional advertising minutes to the ad agents. Higher wages for agents can be generated by both higher station ratings and higher quality. The relationship is more pronounced in smaller markets.

Table 13 presents time effects in the marginal cost. The effects are fairly large which suggests fluctuations in the advertising labor markets. However, as noted earlier, these coefficients can also be interpreted as shocks to aggregate willingness to pay for advertising. In particular, large coefficients indicate low willingness to pay. This interpretation is compatible with the estimated 2001-2003 peak in the time coefficients which coincides with the recession.

Table 14 presents the estimates of cost synergies from joint ownership of multiple stations in the same format. I find statistically significant synergies in the largest and smallest markets. In particular, in the largest markets, the marginal cost of selling 60 seconds of advertising in second station in the same format is about 20% smaller.

[10] I note that positive variable (or gross) profits are not necessarily indications of a radio station profitability. Indeed, Jeziorski (2013) finds that 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 reported by the FCC Research Studies on Radio Industry. For example, the FCC's Radio Industry Review shows that the median EBIT margin was negative for parts of 1998, 2001, and 2002, and was below 5% for the part of 1999.

5.3.1 Discussion of the instruments

As in the demand model, I determine whether the instruments are helpful in correcting the endogeneity bias, through a comparison between 2SLS with OLS estimates. In this case, the bias is difficult to sign because q appears on both sides of the equation (4.4). The results are summarized in Table 11. I find that endogeneity of quantity biases the slope of downwards. I also find a negligible negligible bias in time effects and marginal cost, which was not reported and can be found in the online appendix.

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 an appropriate instruments vector for that group; in its use here, the term "instruments" denotes all exogenous regressors vector for that group. The F-stats are 714, 924, 827, and 202, respectively, 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 that the instruments can explain a reasonable amount of variation in the endogenous regressors. [11]

6 Counterfactual experiments

In this section, I investigate the impact of consolidation on listener and advertiser welfare. First, I investigate the changes in listener and advertiser surplus. In particular, I calculate how much market power was exercised on both of those groups. Second, I decompose this market power into adjustments of variety and advertising quantity. The dollar values in the counterfactuals were extrapolated from a subset of 88 markets to the entire US market on the assumption that the analyzed markets are a random selection, weighted by the population, of small, medium and large markets drawn from the entire Arbitron market pool.

I compute the standard errors of the counterfactual results using a parametric bootstrap performed in three stages: (i) I draw model parameters θ from the joint asymptotic distribution of

[11] A lower value of R^2 for large markets illustrates that the variation in the instruments might not be adequate 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.

first and second stage estimates, (ii) I obtain the error terms of the model, ξ and η, for the drawn parameters, and (iii) I recompute the counterfactual for the new parameters and errors terms. Standard errors obtained using this procedure should be treated as conservative approximations, because they are likely to be overstated, which is caused by an accumulation of numerical errors when computing thousands of local market equilibria. The bootstrap is performed using 90 independent draws of θ.[12]

Before performing counterfactual calculations, I consider descriptive relationships between concentration and prices. First, I regress 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 market power on advertisers. Second, I regressed the total advertising supplied on the market's HHI and market fixed effects. Here I obtain a coefficient of 1.65 with a standard error of 0.3, which is evidence of market power on listeners. Because market power appears to be present in both market segments, I cannot definitely conclude which segment had more surplus extracted by radio-station owners. In the following subsection, I present the structural counterfactuals that answer this question.

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 calculate 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 and account for marginal cost synergies by adjusting the marginal cost estimates with appropriate pre- or post-merger values for dummies SYN_{jt}. Finally, I net out demand and supply time effects by setting year dummies to their 1996 levels and detrend demographics by employing 1996 draws from Common Population Survey. Because I do not observe listeners paying to remove advertising, I am unable to quantify listener welfare in dollars. Instead, I am going to use person-day-minutes (pdm) of advertising as the unit of utility. Although I measure

[12]Note that estimates of percentage changes have lower standard errors than absolute changes because the percentage changes do not depend on the estimates the intercepts of the inverse demand for advertising, θ_1^{Am}.

changes in welfare in the units of advertising, the impact of mergers cannot be measured by simple changes in ad exposure, reported as "Average ad load," [13] because listeners reoptimize in response to both product repositioning and adjustments in advertising supply. For example, listeners frequently switch to stations with less favorable programming if such stations decreased the amount advertising. In such situation, changes in the average ad load would overstate changes in listener welfare. The welfare measures reported in this study correct for this potential bias.

The total impact of consolidation on advertiser and listener welfare is presented in the last row of Table 15. This impact, presented as such, is a statement of the difference in listener and advertiser welfare between the equilibria computed computed using both 1996 and current ownership and formats. I find that mergers decreased total ad quantity by roughly 15,000 minutes, resulting in 17% or 7.3 pdm reduction in average ad exposure. These changes translated to an increase of 1 pdm in consumer welfare. As mentioned before do not observe dollar prices in the listenership market, thus I cannot compute the dollar value of this compensating variation. However, I can compute a rough estimate using the prices for the satellite radio. On the assumption that people buy satellite radio only to avoid listening to advertising, this estimate estimate comes to 1.5 cents per minute, or \$730 million dollars for each pdm per year. This number is, of course, an ad hoc, loose upper bound on the overall welfare gain. For advertisers, a decrease in quantity supplied leads to an approximately 6.5% increase in per-listener prices,[14] or a \$223 million (21%) decrease in advertiser surplus. Therefore I conclude that the Telecom Act led to a reallocation of surplus from advertisers to listeners.

Separate observations of small, medium, and large markets reveal more details about the reallocation of the surplus. As mentioned in the previous section, radio stations have considerable control over prices in small markets and less control in large markets. With this fact in mind, I present counterfactuals for markets with populations less than 0.5M, between 0.5M and 2M, and in excess of 2M.

In smaller markets (see first row of Table 15), stations decrease advertising volume to exer-

[13] Average ad load is defined as the expected number of advertising minutes heard by an average listener, which is different from total number of minutes in the airwaves.

[14] 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.

cise market power on advertisers. They supply over 10,000 fewer minutes of advertising, which translates into a 8.1 pdm decrease in ad exposure, increasing the consumer surplus by 1.3 pdm. However, prices rise by 7% and cause a $56M (33%) loss in advertiser surplus.

In medium markets (see second row of Table 15), mergers reduced the number of broadcast advertising minutes by 3,500 minutes (8.8%) and ad the load by 8.7 pdm (22%). This decrease resulted in lowering advertising surplus by $68 million (23.5%), which is larger in nominal terms than for small markets (because of larger demand intercepts in medium markets), yet much smaller in percentage terms. The consumer surplus goes down despite a smaller advertising load, which results from the combination of an impact of quantity readjustment and format repositioning.

In large markets (see third row of Table 15), firms supply only 950 (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 reduced advertising supply decreased advertising welfare by $100M, which is nominally greater than the welfare decrease in medium and small markets. However, this decrease constitutes only 17% of the total surplus of advertisers, which is less than the decease for the other two markets. This dichotomy is caused by larger intercepts for larger markets, which drive the nominal numbers, and smaller slope coefficients, which drive the percentage numbers. Overall, I conclude that on average, mergers transferred surplus from advertisers to listeners. The loss in the advertiser surplus was higher nominally in larger markets; however, percentage-wise, it was higher 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 in the computation of 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 appear to have decreased the ad quantities to a greater extent, which leads to overestimating the change in listener and advertiser 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 greater variety, which they assume is beneficial for the listeners.[15] To quantify this effect, I compute the ad quantity by using 1996 ownership and formats, which serve as proxies for a pre-regulation regime. Next, I use this quantity to compute the surpluses by using 1996 and current ownership and formats. The difference between these surpluses measures a welfare impact of post-regulation changes in product variety without taking into account quantity adjustments. The results of this experiment are presented in the first row of Table 16. 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. This increase was caused by an interaction of greater variety and lower advertising exposure. Note that despite keeping the number of ad minutes fixed the advertising load decreased, because listeners reoptimize. In particular, greater variety allows the listeners to choose stations with desired programming but with less advertising. Other studies have also documented similar effects, but do not agree on their direction. For example, George (2007) uses reduced-form arguments to show that mergers in the newspaper industry led to greater variety. She also finds that mergers combined with extra variety did not lower circulation and concludes that variety mitigates the negative effects of mergers on reader surplus. However, I note that, in another study, Fan (2012) finds negative impact of newspaper mergers on reader welfare.

I addition to quantifying the impact of changes in variety on listener welfare, I compute the impact of these changes on advertiser welfare. In contrast to the positive impact on listeners, I find that repositioning effects advertiser surplus negatively. In particular, advertisers lose more than $180 million (17%) in surplus. I note that these findings are consistent with the findings of Fan (2012).

In the real world, repositioning changes firms' incentives to set ad quantity, because it softens competition in the advertising market. To quantify this effect, I compare the consumer welfare values without quantity adjustments computed in the previous step with consumer welfare computed under the full equilibrium response, including current formats and ownership as well as equilibrium quantities. The difference between these surpluses measures the impact of ad readjustments. The results are contained in the middle row of Table 16. In this case, both listeners and advertisers are

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

worse off due to quantity adjustments; listeners lose 0.8 pdm and advertisers lose an additional $44 million in surplus. Both listeners and advertisers lose welfare because the number of advertising minutes goes down, whereas average ad exposure goes up. I find that companies who 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 earn additional revenue from higher prices, which depend on the total number of minutes supplied. Furthermore, they are able to increase revenue with greater average ad exposure. [16]

The quantity readjustment does not decrease the listener surplus enough to offset the positive impact of extra variety. However, for the advertisers, the two effects work in the same direction, so the repositioning strengthens the effect of ad readjustment. About 80% of adverse effects on advertisers were caused by product repositioning, which softened the competition within formats. The magnitude of these effects suggests that repositioning is much more effective in extracting social surplus than readjusting ad quantity. This observation agrees with the large repositioning cost estimates found of Sweeting (2012) and Jeziorski (2013). Also, it is consistent with the conjecture made made with the raw data in section 3 of this study that format switching amplifies 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. This switching happens more often in niche formats and leads to price increases.

The above results have two limitations. Primarily, because I employ a static model, the results should be interpreted as short run approximations. In particular, it is possible that in the long run more repositioning is present, which could either mitigate or strengthen the market power. Lastly, I do not observe the extend of repositioning in the counterfactual world without the deregulation, thus, I assume that in such case no repositioning takes place. Quantifying the extent to which these limitations affect the results requires a dynamic model in which product repositioning is endogenous and is left for further research.

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

7 Robustness analysis

This section examines the robustness of my advertising model to different assumptions about competition among station formats. This investigation is motivated by the fact that the data concerning advertiser deals is incomplete, and I deal with this problem by proposing a stylized specification of the inverse demand for advertising 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 among themselves. In particular, suppose a specific type of advertiser only obtains utility from one specific 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 equally. Formally, this means $\omega_{ff'} = 1/8$, because eight formats are possible. In a sense, the estimated model is between these extreme alternatives because it assumes formats are imperfect substitutes. The estimates of the slopes of the inverse demand for the two hypothetical scenarios presented in Table 17 confirm this intuition[17]. In particular, when I assume an oligopoly within a format, the estimated slope parameter θ_2^L is smaller than the parameter in the baseline model. The weights ω_{ff} for an oligopoly within format grant stations more market power because they preclude the opportunity for the advertisers to switch between formats. The smaller value of the slope coefficient compensates for that extra market power in order 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 compensates for higher between-format advertiser switching. Despite the fact that small but statistically significant differences exist between the different models, the main qualitative assertion, that stations have more power in smaller markets, still holds. The differences in the remaining second stage parameters, such as marginal cost, are economically and statistically negligible and are reported in an online appendix.

To draw final conclusions concerning the strength of the assumption about weights, I recalculate the main counterfactual using the alternative models. The results are presented in Table 18. The

[17]Standard errors were obtained using 90 bootstrap draws. In case of "Perfect substitutes" the non-linear solver failed to compute an equilibrium in 0.25% of cases. These cases were excluded when computing standard errors.

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, however the difference is not statistically significant.

8 Conclusion

In this study, I analyze mergers in two-sided markets using the example of the 1996-2006 consolidation wave in the U.S. radio broadcasting industry. The goal of this study is to describe and quantify how mergers in a two-sided market differ from a differentiated product oligopoly setting. The study makes two main contributions. First, it acknowledges the fact that two-sided markets consist of two types of consumers, who may be affected by a merger in different ways. For example, if additional market power causes the radio station to decrease advertising, listeners benefit but advertisers are hurt. The second contribution is the disaggregation of the impact of a merger on consumers into changes in the variety of available products and changes in the supplied quantity of ads.

Radio broadcasting is an important medium in the United States, reaching every week reaching 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.[18] In 1996, the Telecommunication Act deregulated the industry by raising local ownership caps and by abolishing national caps. This deregulation caused a massive merger wave, which reshaped the ownership from a predominantly family-based into a corporate model. I estimate this consolidation raised listener surplus by 0.2%, but lowered advertiser surplus by $223 million. I find the mergers created extra variety, increasing listener welfare by 0.3%. On the other hand, these merger 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).

[18]Source: A.Richter (2006)

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.

Anderson, S. P., O. Foros, H. J. Kind, and M. Peitz (2012): "Media market concentration, advertising levels, and ad prices," *International Journal of Industrial Organization*, 30(3), 321–325.

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

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

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

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

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

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

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

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

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. (2012): “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.

Gentzkow, M., N. Petek, J. Shapiro, and M. Sinkinson (2013): “Do Newspapers Serve the State?: Incumbent Party Influence on the US Press, 1869-1928,” Nber working paper.

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. (2013): "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).

——— (2012): "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 Arbitron's Radio Today publications. BIA provides two comprehensive data sets that cover the vast majority of U.S. radio broadcasting. The first data set covers 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%) drawn from the 2002-2006 data set I was not able to find a match within the 1996-2001 data set, most likely because these stations changed their market and or as a result of 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 affect the results significantly. I also drop inactive stations that are in the "dark" format or have zero market share throughout the whole sample period. The ownership of the stations was traced over time through the tracking of acquisition events reported in the BIA data set.

SQAD provides a twice yearly data set on average prices per rating point (CPP) for each market, grouped by demographics and time of the day. Unfortunately, it collects no data on station-level per-listener pricing. However, because the pricing is done on a per-listener basis, one can still compute a station-level price for an advertising slot by multiplying the CPP by the station rating. According to the anecdotal evidence, many advertisers follow this procedure to estimate the prices they are likely to pay. This procedure does not account for the fact that stations may have different listenership pools and, therefore, that the CPPs for different stations might vary. I alleviate this concern by computing a proxy for a station-level CPP by taking a weighted average of prices by demographics and time of the day, where weights are relevant ratings of the station. In doing so I assume that the stations with most of their listenership concentrated at a particular time of the day and on the particular demographics also set a price that is closer to the average price at that time of the day and for that demographics. Although this estimate of station level prices is not a perfect, it generates 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 this manner might carry some measurement error, because it is a combination 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 that a member of a particular demographic group will listen to a specific format, I use Radio Today publications, indicate a demographic composition for each format. The numbers were inverted using Bayes' rule using market-level demographic distributions obtained from the Census Bureau. I averaged the probability distributions for gender and age groups across the years 1999, 2000, 2001, 2003, and 2004. Education data is available for 2003 and 2004, whereas ethnicity data is available only for 2004. Moreover, I supply the data with a share of an outside option for different markets from Arbitron Listener Trends publications.

B Numerical considerations

To solve the optimization problem (4.3), I employ a version of the Gauss-Newton method implemented in the commercial solver KNITRO. The use of this state-of-the-art solver avoids certain convergence problems that are common in 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 optimized parallel C code, which allows the use of a fairly large dataset (about 42,000 observations) and many draws (500 draws from Normal and CPS per date/market) when computing the constraints. When parallelizing the code, I was careful to maintain the independence of the draws within and between threads. To achieve this independence, I implemented a version of a pseudo-random number generator described in L'Ecuyer and Andres (1997). This generator enables the creation of 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 involves the inversion of a Hessian estimator within the KNITRO solver; unfortunately this inversion cannot be parallelized because it is done inside the KNITRO solver, apart from the user's control.

C Tables

$r_{f|a}$

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

$r_{a|f}$

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

Ω

	Talk	Hits
Talk	0.56	0.44
Hits	0.28	0.72

Table 1: Simple example of advertising weights.

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 2: 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.

	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 3: 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 4: 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 5: 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).

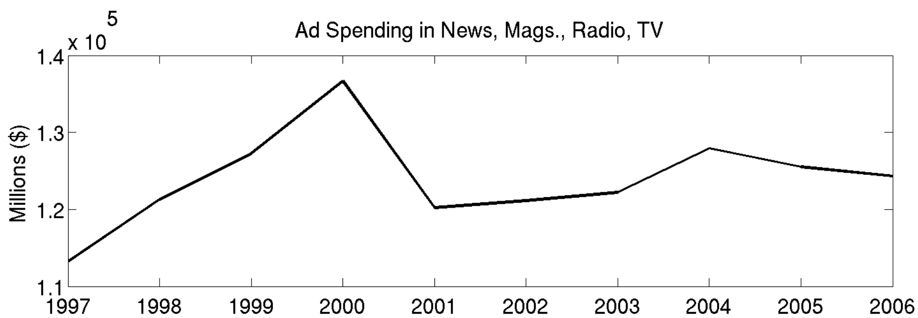

Figure 1: U.S. annual advertising spending in News, Magazines, Radio, and TV. The amount is in 1996 dollars deflated by CPI. Source: Coen Structured Advertising Expenditure Dataset (CS Ad Dataset).

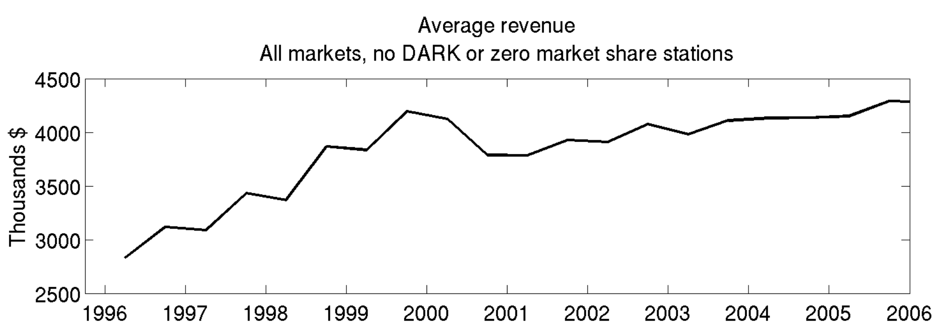

Figure 2: Average station revenue for all active stations (Dark and zero-market-share stations excluded). The amount is in 1996 dollars deflated by CPI.

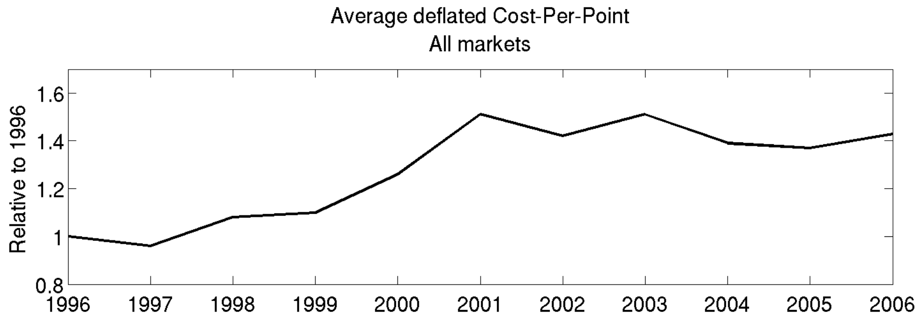

Figure 3: Average cost per point (CCP) relative to 1996 value. The relative value was computed for every market and averaged across markets. The values were deflated by CPI.

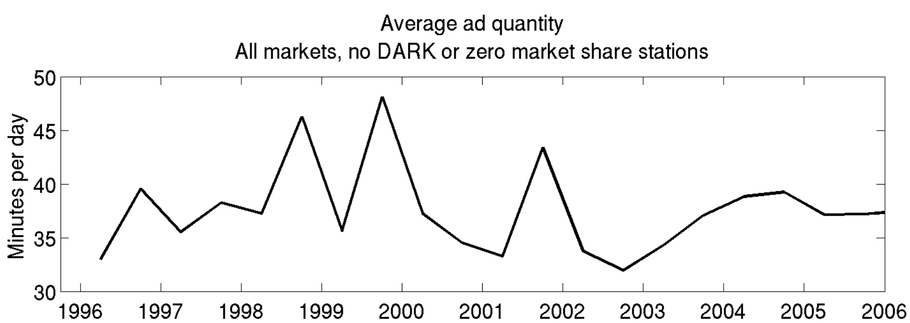

Figure 4: Average quantity for all active stations (Dark and zero-market-share stations excluded). The amount is in 1996 dollars deflated by CPI.

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

	GMM		OLS	2SLS
	Mean Effects	Random Effects		
Advertising	-1.386^{***} (0.226)	0.235 (0.177)	-0.720^{***} (0.030)	-1.391^{***} (0.053)
FM	0.742^{***} (0.043)	-	0.611^{***} (0.013)	0.625^{***} (0.013)
Power (kW)	0.127^{***} (0.004)	-	0.079^{***} (0.001)	0.082^{***} (0.001)
AC SmoothJazz New AC	-4.082^{***} (0.058)	0.028 (0.103)	-2.368^{***} (0.018)	-2.309^{***} (0.019)
Rock	-3.380^{***} (0.076)	0.188^{***} (0.050)	-2.330^{***} (0.020)	-2.256^{***} (0.021)
CHR	-0.969^{***} (0.070)	0.028 (0.085)	-2.167^{***} (0.023)	-2.127^{***} (0.024)
Alternative Urban	-4.979^{***} (0.067)	0.314^{***} (0.048)	-2.237^{***} (0.020)	-2.186^{***} (0.020)
News/Talk	-11.088^{***} (0.096)	0.559^{***} (0.051)	-2.184^{***} (0.015)	-2.105^{***} (0.016)
Country	-4.950^{***} (0.069)	0.571^{***} (0.012)	-2.321^{***} (0.019)	-2.270^{***} (0.020)
Spanish	-4.463^{***} (0.090)	0.182 (0.156)	-2.880^{***} (0.020)	-2.812^{***} (0.020)
Other	-5.938^{***} (0.053)	0.080 (0.099)	-2.634^{***} (0.016)	-2.589^{***} (0.016)
ρ	0.748^{***} (0.008)	-		

Standard errors in parentheses

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

Table 7: 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.

	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.

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

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.

	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 11: 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
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 12: 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.

	1998	1999	2000	2001	2002	2003	2004	2005	2006
<.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 13: Time effects in the marginal cost. 1996 and 1997 values are normalized to zero.

Cost synergies		
Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M
-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.

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Small markets <0.5m pop.	1.3pdm (2.2pdm)	-8.1pdm (0.8pdm)	-56.2m (8.9m)	-10,707min (1,395min)	+7.16% (1.24%)
	+0.3% (0.3%)	-25.4% (1.6%)	-32.6% (2.8%)	-16.0% (1.7%)	
Medium markets 0.5-2m pop.	-0.3pdm (3.7pdm)	-8.7pdm (0.6pdm)	-68.0m (42.8m)	-3,582min (891min)	+5.75% (3.51%)
	-0.1% (0.4%)	-21.8% (1.9%)	-23.5% (6.0%)	-8.8% (2.6%)	
Large markets >2m pop.	1.7pdm (2.8pdm)	-6.4pdm (0.8pdm)	-99.1m (20.6m)	-954min (422min)	+7.39% (2.13%)
	+0.3% (0.2%)	-13.7% (1.2%)	-17.1% (2.7%)	-2.9% (1.2%)	
All markets	1.0pdm (1.2pdm)	-7.3pdm (0.5pdm)	-223.3m (55.0m)	-15,243min (1,411min)	+6.53% (2.21%)
	+0.2% (0.2%)	-16.8% (0.8%)	-21.4% (2.7%)	-10.9% (0.9%)	

Table 15: Estimated impact of consolidation 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 (1.1pdm) +0.3% (0.2%)	-8.9pdm (0.6pdm) -20.4% (0.8%)	-179.5m (50.3m) -17.2% (1.3%)	-161min (22min) -0.1% (0.0%)	+3.00% (0.85%)
Impact of ad adjustment	-0.8pdm (0.2pdm) -0.1% (0.0%)	1.6pdm (0.2pdm) +4.6% (0.6%)	-43.8m (33.9m) -5.1% (2.8%)	-15,081min (1,401min) -10.7% (0.9%)	+3.34% (6.05%)
Total impact of ownership change format switching and ad adjustment	1.0pdm (1.2pdm) +0.2% (0.2%)	-7.3pdm (0.5pdm) -16.8% (0.8%)	-223.3m (55.0m) -21.4% (2.7%)	-15,243min (1,411min) -10.9% (0.9%)	+6.53% (2.21%)

Table 16: Decomposition of 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 17: Robustness of the estimates of slope of advertising cost per rating point (CPP).

	Consumer surplus	Average ad load	Advertiser surplus	Advertising minutes	Mean price index
Baseline model	1.0pdm (1.2pdm) +0.2% (0.2%)	-7.3pdm (0.5pdm) -16.8% (0.8%)	-223.3m (55.0m) -21.4% (2.7%)	-13,719min (1,411min) -11.1% (0.9%)	+6.53% (2.21%)
Oligopoly within format	-0.9pdm (2.2pdm) -0.2% (0.3%)	-5.8pdm (0.5pdm) -15.1% (1.2%)	-150.7m (88.9m) -21.3% (6.4%)	-17,533min (3,002min) -12.3% (2.1%)	+3.16% (3.10%)
Perfect substitutes	2.2pdm (1.1pdm) +0.4% (0.2%)	-8.7pdm (1.1pdm) -19.5% (1.4%)	-251.9m (102.0m) -26.1% (7.3%)	-19,170min (2,320min) -13.7% (1.2%)	+17.68% (2.53%)

Table 18: Robustness of the counterfactuals.

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

Przemysław Jeziorski *

August 23, 2013

*Haas School of Business, University of California at Berkeley

1 Additional tables and figures

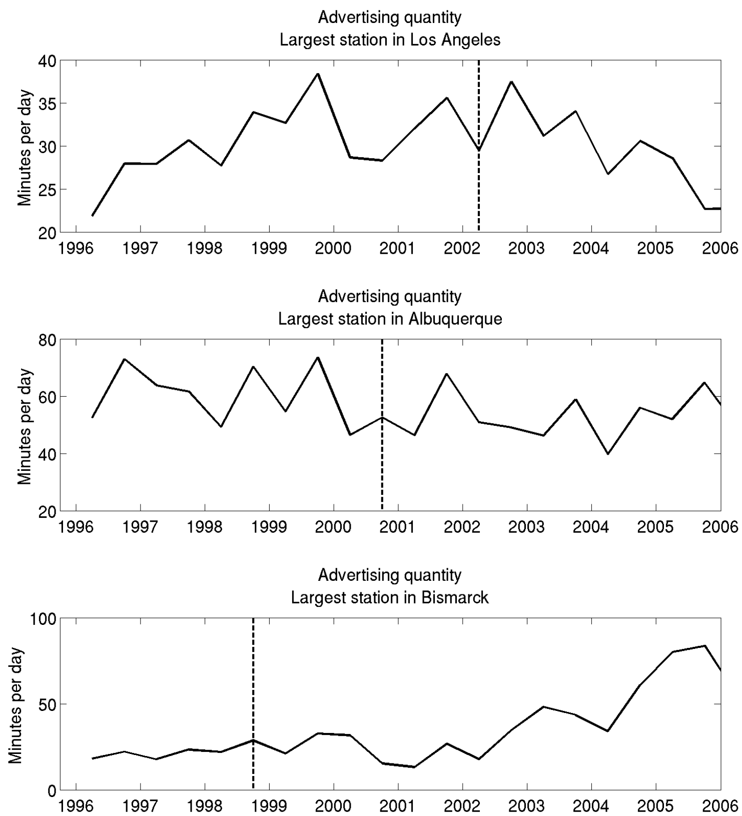

Figure 1: Number of advertising minutes per day for the largest station in the market (largest average 1996-2006 rating, among always active stations). The figures present three representative markets: the largest, Los Angeles; mid size, Albuquerque; and the smallest, Bismarck. The vertical line represents acquisition.

2

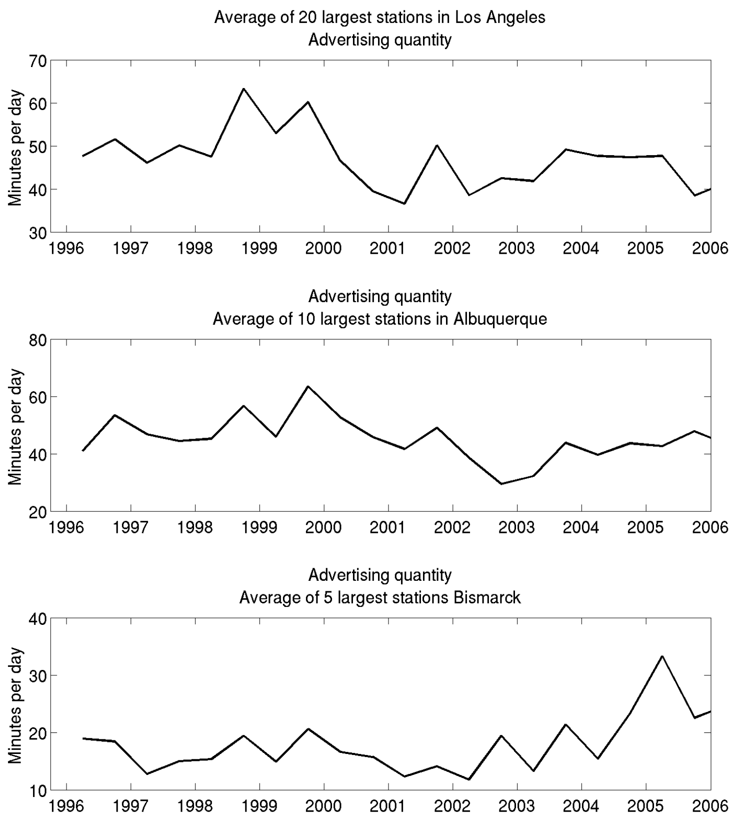

Figure 2: Market average of advertising minutes per day. The figures present 3 representative markets: the largest, Los Angeles; mid size, Albuquerque; and the smallest, Bismarck.

3

	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 1: Edogeneity bias in estimating 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.

		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 2: Edogeneity bias in estimating time effects in the marginal cost. 1996 and 1997 values are normalized to zero.

	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 3: Edogeneity bias in estimating marginal cost synergies from owning multiple stations of the same format.

	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 4: Robustness of marginal cost per minute of advertising sold.