Empirical Model of Dynamic Merger Enforcement – Choosing Ownership Caps in U.S. Radio

Przemysław Jeziorski*

June 4, 2013

Abstract

This paper provides a framework for an empirical analysis of the efficacy of merger enforcement rules in markets with differentiated products. The analysis is conducted using a computable, continuous time, dynamic oligopoly model in which mergers, entry/exit, and product repositioning are endogenous. The model is a generalization of a Rubinstein (1982) framework and casts mergers as a multilateral dynamic bargaining process in which both acquirer and acquiree are forward looking. Solutions to this model can explain and predict merger waves and post-merger product portfolio management; thus, they allow computation of a long-run welfare impact of counterfactual merger enforcement rules. I apply the model to compute long-run response to the change of ownership rules in the U.S. radio industry. I find that deregulation of this industry imposed by the 1996 Telecom Act generated substantial cost efficiencies and resulted in a negligible amount of extra market power, and I demonstrate that one of the reasons for the limited market power is post-merger product repositioning. Moreover, I show that both myopic and naively forward-looking regulators would over-block mergers. For example, I evaluate the policy that uses consumer welfare criteria coupled with looser ownership caps. I find such policies raise total welfare; however, I also find companies can circumvent static welfare antitrust criteria by strategically proposing mergers that are likely to be accepted, and by repositioning after the merger. In such situations, a static consumer surplus criterion would fail to prevent long-run losses to consumer surplus.

*Haas School of Business, UC Berkeley

1 Introduction

A broad goal of merger enforcement is to prevent mergers that significantly reduce competition and lead to a decrease in consumer or total welfare. Most of the economic literature that studies this problem focuses on regulator's decision about a particular merger that came under review (see Williamson (1968), Farrell and Shapiro (1990) for theoretical analysis, and Nevo (2000), Benkard, Bodoh-Creed, and Lazarev (2008) for empirical analysis). However, in reality, proposed mergers are not one-time events and are endogenous as a function of the pre-announced enforcement rules (see U.S. Department of Justice Horizontal Merger Guidelines, European Commission Guidelines on The Assessment of Horizontal Mergers, or U.S. Telecommunication Act of 1996). In particular, the changes in these enforcement rules incentivise companies to alter the set of proposed mergers and affect the dynamics of the industry after the merger (see Lyons (2002) and Nocke and Whinston (2010)). Therefore, a myopic antitrust agency might misestimate the impact of the change in merger policy on consumer and producer surplus. This paper proposes a dynamic model that endogenizes firms' short- and long-run responses to changes in regulation in a way that is robust to the aforementioned issues. The model allows for endogenous mergers, product repositioning, and entry/exit, by which it can explain and predict merger waves and post-merger product portfolio readjustments in the markets with differentiated products. The framework is designed with an empirical analysis in mind, so it has a low estimation burden and is fully computable. Thereby, it can be relatively easily applied to empirically evaluate hypothetical enforcement rules. The paper estimates the model using the data on the 1996-2006 merger wave in the U.S. radio industry and simulates response to retrospective and hypothetical changes in Federal Telecommunication Commission ownership caps and welfare criteria used by the Department of Justice.

To conduct empirical analysis of merger enforcement rules, one needs a tractable dynamic model in which mergers are endogenous. However, because past models of endogenous mergers are rather complex (see Gowrisankaran (1999), Stahl (2011), and Jeziorski (2013)), their estimation and computation proves difficult. These difficulties reflect the fact that mergers are complex decisions involving multilateral dynamic bargaining among heterogeneous agents and generating a multitude of possible industry configurations. This paper proposes a simple but empirically relevant continuous time model of endogenous mergers (for other applications of continuous time to dynamic oligopoly, see Kryukov (2008), Doraszelski and Judd (2012), Arcidiacono, Bayer, Blevins, and Ellickson (2010)). The key advantage of modeling the problem using continuous time is that the probability of two companies taking action at the same time is negligible. The low probability of simultaneous actions is particularly important in a merger model because it allows abstracting from situations in which two companies simultaneously bid to acquire the same firm. Moreover, it allows for simplification of the decision tree to one merger at a time, while generating a positive likelihood of multiple mergers in any time interval, which is important for empirical applications.

I apply the proposed model to analyze retrospective and hypothetical antitrust policy changes in the U.S. radio industry. The industry has undergone an unprecedented regulation change known as the 1996 Telecom Act, which doubled local ownership caps and abolished strict national ownership restrictions. This deregulation enabled entry of large corporations into this previously fragmented industry and spurred a large merger wave that amounted to more than 6,000 acquisitions over the 10-year period. Specifically, the Telecom Act resulted in the following two qualitative developments: (i) the possibility and execution of mergers that were infeasible before 1996, (ii) entry of large conglomerate companies with possibly very different cost structures compared to the incumbents. These dramatic changes in the operating rules of the industry act as a quasi-experimental shock that identifies supply and demand incentives to merge. As a result, I can estimate determinants of timeliness, likelihood, and the size of mergers and product repositioning in this industry, which I use to conduct policy counterfactuals.

According to the anecdotal evidence, the 1996 deregulation was partly motivated by potential efficiency gains of operating larger radio station portfolios. Indeed, I find the consolidation of the ownership generated substantial fixed and marginal cost savings. Namely, operating multiple radio stations jointly is more efficient, particularly for stations providing a similar type of programming. For example, jointly operating two pop-music stations jointly is 14% cheaper than operating them separately. However, conversely, Jeziorski (2012) finds that both mergers and substantial repositioning generated extra market power, which makes the consequences of the Telecom Act less straightforward. Moreover, maybe more importantly, the coexistence of cost synergies and market power complicates assessments of what kind of further deregulation (if any) would be beneficial. I tackle these questions using a series of counterfactual experiments.

First, I evaluate the consequences of the Telecom Act by recomputing the equilibrium in the counterfactual world in which the deregulation did not take place. Using a new equilibrium strategies, I simulate industry paths that involve hypothetical entry of corporations, mergers, station repositioning, and advertising prices. I find that, in the long run, the counterfactual world with pre-1996 local ownership caps has 10.1% lower producer surplus, 0.07% lower listener surplus, and at the same time has only 1.7% higher advertiser surplus. All in all, reversing the deregulation leads to overall decrease in total surplus.

Next, I investigate the consequences of further deregulation, by exploring the possibility of raising a local cap to from three to five stations (depending on the market size) to a uniform cap of seven FM stations. I find this change would lead to additional 4% increase in producer surplus, a 0.01% increase in listener surplus and about a 1% decrease in advertiser surplus. I note this policy raises total surplus; however, the gains are smaller than those brought by the Telecom Act. Thus, a consumer-centered agency might consider additional regulation on top of raising the cap. One candidate is a static merger simulation (see Nevo (2000), and Ivaldi and Verboven (2005)) in which the agency recomputes post-merger advertising prices and rejects the merger if it lowers listener or advertiser surplus. Raising the cap to seven FM stations coupled with the static merger simulation based on listener surplus leads to a long-run 2.24% increase in producer surplus, a 0.08% increase in listener surplus, and a 0.66% decrease in advertisers surplus. Thus, it does a fairly good job in selecting mergers that would benefit listeners. On the other hand, the myopic policy based on advertiser surplus works only in the short run. Indeed, in the first five years, the companies propose mergers that lead to a 0.1% increase in advertiser surplus, knowing such mergers are likely to be accepted. However, in the longer run, the mergers are followed by repositioning that ultimately hurts advertisers. In particular, in 10 to 20 years, the advertiser-centered policy leads to a 0.79% decrease in advertiser surplus and hurts advertisers more than listener-centered policy would. This somewhat counterintuitive result is a direct consequence of the agency using a myopic policy that results in a looser merger criterion than initially intended. This result provides a direct empirical evidence that dynamics is important and should be incorporated into antitrust policy. These numbers are also consistent with the fact that, according to judicial documents, a version of advertiser-centered policy was sporadically used during the 1996-2000 merger wave and did not prevent losses to advertiser surplus, as documented by Jeziorski (2012).

The closest theoretical paper to this work is Nocke and Whinston (2010), who study the optimality of myopic merger enforcement in the dynamic model with homogeneous product Cournot competition. In their model, the mergers are endogenous, and similarly to the model in this paper, merger opportunities appear randomly. Their main result is that under the assumption that the mergers are disjoint a version of myopic merger policy is optimal. In this paper, I relax this assumption and additionally allow for product differentiation, product repositioning and cost efficiencies from mergers. This paper shares some similarities with Armstrong and Vickers (2010), who study a static principal and agent problem. In their case, the principal does not observe full characteristics of unproposed projects and commits to the acceptance policy ex-ante. Nocke and Whinston (2013) extend these results to merger review, allowing for bargaining among firms and multiple agents. This paper also extends arguments of Lyons (2002), who highlights that the regulator cannot choose which mergers to execute, but rather it can approve or reject mergers from the set chosen by strategic players. Following that logic, Lyons gives examples in which the regulator should announce a consumer surplus criterion instead of a total surplus criterion even if the goal is to maximize the total surplus.

The proposed framework is also compatible with a vast literature on mergers in general. For example, similarly to Kamien and Zang (1990), Rodrigues (2001), and Gowrisankaran and Holmes (2004), it assumes both sellers and buyers are fully forward looking. Therefore, it can generate and identify hold outs that arise in a merger process, because the seller has to be paid his dynamic opportunity cost of non-merging which is usually greater than seller's static profits. The paper is also related to the literature studying merger waves (see Harford (2005) and Qiu and Zhou (2007)) by allowing the mergers to be strategic compliments. Another related paper is Mazzeo, Seim, and Varela (2012), who show numerically using a static model that post-merger repositioning can significantly alter the welfare assessment of the merger. In this paper, I empirically demonstrate a similar phenomenon using a dynamic model.

The paper is organized as follows. In the next section, I present the dynamic model in the example of the radio industry. In the third section, I demonstrate the impact of dynamics on the optimality of merger enforcement using a few numerical examples. The forth section contains a description of the data. In the fifth section, I describe the estimation algorithm. The sixth section contains results. The seventh section presents counterfactual. I conclude in the eighth section.

2 Model

Consider a market over an infinite continuous time horizon. The market consists of maximum K active radio owners and N possible broadcast frequencies. Each frequency has an assigned owner and can host one radio station. This technical restriction effectively caps the number of active stations to N. I assume the radio station can be fully characterized by a type coming from a finite type space $\mathcal{F} = \{1, \ldots, F\}$.

The market is modeled as a dynamic game between radio-station owners. The portfolio of the radio-station owner k is characterized by a vector $\omega_k^t = (\omega_{k1}^t, \ldots, \omega_{kF}^t)$, where ω_{kf}^t is the number of radio stations of format f owned by a player k. The state of player k is given by the vector $\mathcal{J}_k^t = (\omega_k^t, z_k^t)$, where z_k^t are the remaining payoff-relevant variables. For convenience, I denote the total number of stations owned by player k as n_k^t. The instantaneous variable profits and fixed cost for firm k are given by $\pi_k(\mathcal{J}^t)$ and $F_k(\mathcal{J}^t)$, respectively.

2.1 Actions

The firms' actions are either mergers or repositioning with entry and exit as special cases of thereof. Opportunities to acquire and reposition arrive randomly according to a collection of Poisson processes. In particular, an opportunity for a company k to acquire other players arrives as a Poisson process with an arrival rate $\lambda_k^A(\mathcal{J}^t)$. Similarly, an opportunity to reposition has an arrival rate of $\lambda_k^R(\mathcal{J}^t)$. To avoid a curse of dimensionality, I assume these Poisson processes are independent (across firms and event types) conditional on $\mathcal{J}^t$. Doraszelski and Judd (2012) demonstrated that those assumptions parallel the independent state transitions frequently used in the discrete-time games.

Conditional on the arrival of an opportunity to merge, a player can choose one company to acquire. [1] During the acquisition, the acquirer makes one take-it-or-leave-it offer P to the chosen acquiree. The acquiree accepts or rejects the offer, taking into account his dynamic opportunity

[1] Because the model is written in continuous time, it can still generate an arbitrary number of mergers within any fixed time period. Moreover, these multiple merger events are going to be correlated in an endogenous way. By contrast, an equivalent assumption in a discrete-time model has more restrictive consequences. It would effectively cap the number of acquisitions by a potential firm to one per period.

cost. The cost of executing a merger[2] with k' is given by $\zeta^{A,t}(\mathcal{J}^t, k')$, which contains legal and procedural expenses associated with executing a merger bid. If the offer is accepted, the regulator blocks the merger of k and k' with a probability $\mathbf{G}(\mathcal{J}^t, k, k')$. Conditional on the arrival of an opportunity to reposition, a player can take a repositioning action $r = (f, f')$, which changes a format of an owned station from f to f'. Such repositioning involves paying a repositioning cost $\zeta^{R,t}(\mathcal{J}^t, r)$. Both mergers and acquisitions are implemented instantaneously, resulting in a new industry state $\mathcal{J}^{t+}$. In this application, I keep z_k^t constant over time; however, changes to z_k^t might arrive as a Poisson process as well (see Arcidiacono, Bayer, Blevins, and Ellickson (2010)).

Because the number of active broadcast frequencies is limited, entry is possible only through acquisition of other active firms. Similarly, exit is modeled as selling off all owned stations. Potential entrants are firms that hold empty portfolios of stations; that is, $\omega_k^t = \overrightarrow{0}$. [3] Note that such modeling of entry and exit endogenizes entry cost and scrap value, which are usually assumed to be primitives (see Ericson and Pakes (1995)). Specifically, in my model, acquisition involves paying an endogenous acquisition price, which acts as an endogenous sunk entry cost for the acquirer and an endogenous scrap value for the acquiree. As a result of this endogenous sunk cost and the fact that large players operate in multiple markets simultaneously, potential entrants frequently delay entering into a particular local market, and wait for favorable market conditions. Consequently, the usual assumption that potential entrants are short lived needs to be modified. Instead, I allow the potential entrants to be long lived, which allows for postponing entry as well as re-entry.

Action-specific payoff shocks are assumed to consist of persistent and idiosyncratic parts. In particular, I set the cost of making an offer to k' to be

$$\zeta_k^{A,t}(\mathcal{J}^t, k') = \mu_k^A(\mathcal{J}^t, k') + \sigma_k^A(\mathcal{J}^t, k')\epsilon_k^{A,t}(k'). \tag{2.1}$$

Similarly, the repositioning payoff shock is given by

$$\zeta_k^{R,t}(\mathcal{J}^t, r) = \mu_k^R(\mathcal{J}^t, r) + \sigma_k^R(\mathcal{J}^t, r)\epsilon_k^{R,t}(r). \tag{2.2}$$

[2] I cannot separately identify the cost of making a merger bid from the cost of executing the merger. This fact is not consequential for estimation and counterfactuals because all bids are accepted in the equilibrium, so these costs coincide on the equilibrium path.

[3] One way to allow the possibility of more traditional entry is to endow the type space $\mathcal{F}$ with an inactive state. This extension is possible but is not implemented because the model without an inactive state captures the first-order dynamics in my data.

Values of ϵ_k^A for all k' are revealed to player k immediately after the arrival of an opportunity to merge. Similarly, repositioning payoff shocks, $\epsilon_k^R(r)$ for all r, are revealed upon the arrival of the opportunity to reposition. The payoff shocks ζ are realized only if the action is taken. In other words, ζ is equal to zero if the player decides not to acquire or reposition upon the arrival of the opportunity to do so. The idiosyncratic parts $\epsilon_k^{A,t}$ and $\epsilon_k^{R,t}$ are private information and are independent across time and players. On the other hard, μ^A and μ^R are time persistent and public information. The reason for having both μ and ϵ is that μ picks up time-persistent action patterns from the data, and ϵ picks up remaining local fluctuations in actions.

A crucial advantage of a continuous time model is that the way to resolve merger conflicts does not have to be specified. Consider a possibility of conflicting merger attempts a_k, $a_{k'}$ (e.g., when two companies bid to acquire the same firm), and let $\text{CON}_{k,k'}$ be the probability that the deal k would be executed. Over a short period of time Δ, the probability of an execution of an attempt a_k is equal to

$$\lambda_k^A(\mathcal{J}^t)\Delta(1-\lambda_{k'}^A(\mathcal{J}^t)\Delta)) + \text{CON}_{k,k'}\lambda_k^A(\mathcal{J}^t)\Delta\lambda_{k'}^A(\mathcal{J}^t)\Delta + O(\Delta^2) = \lambda_k^A(\mathcal{J}^t)\Delta + O(\Delta^2)$$

Doraszelski and Judd (2012) show that only the linear terms of the arrival rates matter for optimality; therefore, in the equilibrium, the conflicting events would not play any role. By contrast, when using discrete time, one usually has to model conflicting mergers explicitly. Because such events are rarely observed in the data, identifying this part of the model is usually hard. In practice, it would force the modeler to assume them away, for example, by putting a structure on a sequence of moves (see Gowrisankaran (1999), Gowrisankaran and Holmes (2004) and Jeziorski (2013)).

2.2 Timing

The model has the following sequence of events:

(1) All players observe the state variables $\mathcal{J}^t$.

(2) Players collect the pay off $\pi_k(\mathcal{J}^t) - F_k(\mathcal{J}^t)$ until a merger/repositioning opportunity arises.

(3a) If a merger opportunity arrives for player k, then

(i) Player k observes a vector of costs ζ_k^A of merging with any of the active competitors.

(ii) Player k chooses whether he wants to make a merger bid. If he chooses to make a bid, he puts forward a single take-it-or-leave-it acquisition offer to the chosen acquisition target.

(iii) The acquisition target accepts or rejects the bid. If the bid is accepted, the payoff shock $\zeta_k^A(k')$ is internalized and the merger goes under antitrust review. If the bid is rejected, both companies continue as separate entities and the game moves to (1); that is, the buyer cannot make additional offers at this time.

(iv) The antitrust decision to approve or reject the merger is revealed instantaneously. If the merger is approved, the companies merge instantaneously, and the merger bid is transferred to the seller. The game then moves to stage (1). If the merger is rejected, the game simply moves to stage (1).

(3b) If a repositioning opportunity for player k arises, he observes payoff shocks ζ_k^R for the repositioning of any owned station to a different format. Then he immediately makes a decision to reposition one station or not to reposition at all. Relevant switching costs are paid, the state space is updated, and the flow goes back to point (1).

2.3 Strategies and equilibrium

A strategy consists of four components: a merger strategy, a pricing strategy, a bid accept/reject strategy, and a repositioning strategy. A merger strategy has the following form: $\mathbf{a}_k(\mathcal{J}^t, \zeta^{A,t}) \in \{0, \dots, K\}$. It specifies which merger bid (if any) is proposed, conditional on the arrival of a merger opportunity. The set of feasible acquisitions $\Gamma_k^A(\mathcal{J}^t)$ is the set of active competitors and action 0, which represents no merger bid. Upon deciding to make a merger bid k', the buyer makes a take-it-or-leave-it offer to seller k', given by the pricing strategy $\mathbf{P}_k(\mathcal{J}^t, \zeta^{A,t}, k') \in \mathbb{R}_+$. Temporarily suppose all merger bids are accepted, so that accept/reject function is a constant for all players and can be omitted. I relax the assumption later in the paper. The repositioning strategy $\mathbf{r}_k(\mathcal{J}^t, \zeta^{R,t}) \in (F \times F) \cup \{0\}$ prescribes which station would be repositioned. The feasible repositioning actions $\Gamma_k^R(\mathcal{J})$ allow for staying idle or repositioning any currently owned station to any possible format.

Let $\mathbf{g}_k = (\mathbf{a}_k, \mathbf{P}_k, \mathbf{r}_k)$ be a strategy of player k. For every initial state $\mathcal{J}^0$, a strategy profile

$(\mathbf{g}_k, \mathbf{g}_{-k})$ and regulator's enforcement rule $\mathbf{G}$ prescribe a continuous time jump Markov process on states $\mathcal{J}^t$, actions (a_k^t, P_k^t, r_k^t), decisions of the regulator $G_k^t \in \{0,1\}$, and private shocks $(\zeta^{A,t}, \zeta^{R,t})$. The jumps in the process occur if a move opportunity arrives for any of the players, and a non-empty action is implemented.

Let $\tau_k^{A,(l)}$, $\tau_k^{R,(m)}$ be stopping times that represent respectively arrivals of l-th merger, m-th repositioning opportunity for player k. With some abuse of notation denote new private information shocks revealed at $\tau_k^{A,(l)}$ and $\tau_k^{R,(m)}$ by $\zeta_k^{A,(l)}$ and $\zeta_k^{R,(m)}$ respectively. Similarly, denote the prescribed actions by $a_k^{(l)}$, $P_k^{(l)}$, $G_k^{(l)}$, and $r_k^{(m)}$. Because the moves are implemented immediately, the resulting Markov process on $\mathcal{J}^t$ would have right-continuous paths. However, note the actions are prescribed by the strategies evaluated at the left-side limit of the state space process; for example, $a_k^{(l)} = \mathbf{a}_k(\mathcal{J}^{\tau_k^{A,(l)}-}, \zeta_k^{A,(l)})$. The value function for company k is given by the following equation (for now I ignore the events by which company k is acquired):

$$V_k(\mathcal{J}^0; \mathbf{g}_k, \mathbf{g}_{-k}, \mathbf{G}) = E_{\mathbf{g}} \Bigg\{ \int_0^\infty e^{-\rho t} \left[\pi_k(\mathcal{J}^t) - F_k(\mathcal{J}^t)\right] dt + \\ \sum_{l=1}^{\infty} e^{-\rho \tau_k^{A,(l)}} \left[\zeta_k^{A,(l)} \left(a_k^{(l)}\right) - G_k^{(l)} P_k^{(l)}\right] + \sum_{m=1}^{\infty} e^{-\rho \tau_k^{R,(m)}} \zeta_k^{R,(m)} \left(r_k^{(m)}\right) \Bigg\}. \tag{2.3}$$

The equilibrium of the game is defined as follows.

Definition 2.1 (Markov Perfect Equilibrium). *A strategy profile* $\mathbf{g}^*$ *is a Markov perfect equilibrium (for a given enforcement rule* $\mathbf{G}$*) if*

(i) Strategies maximize a stream of discounted profits at any state,

$$\mathbf{g}_k^*(\mathcal{J}, \zeta_k) \ \in \arg\max_{\mathbf{g}_k} V_k(\mathcal{J}; \mathbf{g}_k, \mathbf{g}_{-k}^*, \mathbf{G}); \quad \forall k, \mathcal{J}, \zeta_k. \tag{2.4}$$

(ii) Acquisition price covers acquiree's long-run discounted profits; that is, for any $k' > 0$*,*

$$\mathbf{P}_k^*(\mathcal{J}, \zeta_k^A, k') \geq V_{k'}(\mathcal{J}; \mathbf{g}^*, \mathbf{G}); \quad \forall k, \mathcal{J}, \zeta_k^A. \tag{2.5}$$

The first equation means each player best responds to the opponents' strategies and a pre-announced enforcement rule. The second condition specifies equilibrium acquisition prices. An acquiree has to be compensated for an option value of rejecting the merger bid and continuing as a separate company until a new merger bid arrives, which endogenizes the bargaining position of a seller in a dynamic way.

In this paper, I look only at equilibria in which the acquisition price is exactly equal to the acquiree's value function, thus it can be ignored in the acquiree's Bellman equation. This restriction is without much loss of generality for two reasons: (i) acquirees do not have private information at the moment of receiving a merger bid, and (ii) acquirees receive only one merger offer at a time (almost surely). Note the reservation value of rejecting the offer is equal to the acquiree's continuation value, which is exactly known to the acquirer. In such a case, the acquirer can propose the acquisition price exactly equal to the reservation value of the acquiree. Moreover, if we assume offers that would be rejected for sure are not made (one way to ensure rejected offers are not made would be to presume any non-zero cost of making an offer), then all merger bids are accepted in the equilibrium. Such formulation endogenizes the bargaining power of the acquiree and acquirer in a way similar to Rubinstein (1982) model. In particular, the possibility of making or receiving a better offer in the future, as well as a possibility of repositioning, influence the bargaining power.

2.4 Existence

To apply an existence result from Doraszelski and Judd (2012), the game needs to be recast as one with continuous actions, which can be done by noting that choosing actions after observing payoff shocks $\zeta_k^{A,t}$ or $\zeta_k^{B,t}$ is mathematically equivalent to choosing conditional choice probabilities (CCP) of actions (see Magesan and Aguirregabiria (2013)).

Let $\mathrm{CCP}_k^A(a|\mathcal{J})$ be an ex-ante probability of company k acquiring a company k' conditional on the arrival of a merger opportunity. Similarly, define $\mathrm{CCP}_k^R(r|\mathcal{J})$ to be an ex-ante probability of repositioning from f to f'. After a small adjustment to continuous time, the results contained in the proof of Theorem 1 from Hotz and Miller (1993) apply for this model. Following the notation in that paper, consider the expectation of $\zeta_k^{A,t}$, when the optimal action conditional on arrival on the right to merge at state $\mathcal{J}^t$ is

$$W_a^A(\mathrm{CCP}_k^A, \mathcal{J}^t) = E[\zeta_k^{A,t} | \mathcal{J}^t, a_k^t = a].$$

A similar expression can be written for the repositioning action:

$$W_r^R(\mathrm{CCP}_k^R, \mathcal{J}^t) = E[\zeta_k^{R,t} | \mathcal{J}^t, r_k^t = r].$$

The above expressions are equal to 0 if no action is chosen. The key fact is that these expectations depend only on the relevant CCPs in a way that does not depend on value functions. Hotz and Miller (1993) established the result for single-agent discrete-time models, and their proof can be repeated with minor adjustments for the continuous-time game studied in this paper. Subsequently, maximizing the value function with discrete choices is equivalent to solving the following Bellman equation with continuous actions:

$$\begin{aligned}\rho V_k(\mathcal{J}) = \max_{\mathrm{CCP}_k^A, \mathrm{CCP}_k^R} \Bigg\{ & \pi_k(\mathcal{J}) - F_k(\mathcal{J}) - \left(\lambda_k^A(\mathcal{J}) + \sum_{k=1}^{K} \lambda_k^R(\mathcal{J}) \right) V_k(\mathcal{J}) - \\ & \lambda_k^A(\mathcal{J}) \left[\sum_{a \in \Gamma_k^A(\mathcal{J})} \mathrm{CCP}_k^A(a) \Big(V_k(\mathcal{J}'(k,a)) - V_a(\mathcal{J}) + W_a^A(\mathrm{CCP}_k^A, \mathcal{J}) \Big) \right] + \\ & \lambda_k^R(\mathcal{J}) \left[\sum_{r \in \Gamma_k^R(\mathcal{J})} \mathrm{CCP}_k^R(r) \Big(V_k(\mathcal{J}'(k,r)) - V_{a'}(\mathcal{J}) + W_r^R(\mathrm{CCP}_k^R, \mathcal{J}) \Big) \right] + \\ & \sum_{k' \neq k} \lambda_{k'}^A(\mathcal{J}) \sum_{a \in \Gamma_{k'}^A(\mathcal{J})} \mathrm{CCP}_{k'}^A(a) V_k(\mathcal{J}'(k',a)) + \\ & \sum_{k' \neq k} \lambda_{k'}^R(\mathcal{J}) \sum_{r \in \Gamma_{k'}^R(\mathcal{J})} \mathrm{CCP}_{k'}^R(r) V_k(\mathcal{J}'(k',r)) \Bigg\},\end{aligned} \tag{2.6}$$

where $\mathcal{J}'(k,k')$ is the future industry state after k, k' merger and $\mathcal{J}'(k,r)$ is the future industry state after company k takes a repositioning action r. Using this formulation, one can directly apply the existence result in Doraszelski and Judd (2012).

2.5 Computational strategy

Equation 2.6 can be used to compute an equilibrium of the game, and the algorithm has relatively low computational requirements. Suppose the idiosyncratic parts of the payoff shocks, as defined in equations (2.1) and (2.2), have the following structure: $\epsilon_k^{A,t}(a) = \tilde{\epsilon}_k^{A,t}(a) - \tilde{\epsilon}_k^{A,t}(0)$ and $\epsilon_k^{R,t}(r) = \tilde{\epsilon}_k^{R,t}(r) - \tilde{\epsilon}_k^{R,t}(0)$, where $\tilde{\epsilon}$s have IID type-1 extreme value distributions (recall that if no action occurs, $\epsilon_k^{A,t}(0) = 0$ and $\epsilon_k^{R,t}(0) = 0$). Then the optimal merger CCPs is given by a closed-form formula,

$$\mathrm{CCP}_k^A(a|\mathcal{J}) = \frac{\exp\left\{\sigma_k^A(\mathcal{J},a)^{-1}\left[V_k(\mathcal{J}'(k,a)) - V_a(\mathcal{J}) + \mu_k^A(\mathcal{J},a)\right]\right\}}{\sum_{a' \in \Gamma_k^A(\mathcal{J})} \exp\left\{\sigma_k^A(\mathcal{J},a')^{-1}\left[V_k(\mathcal{J}'(k,a')) - V_{a'}(\mathcal{J}) + \mu_k^A(\mathcal{J},a')\right]\right\}}, \tag{2.7}$$

where V_a is the value function of the acquiree (equilibrium acquisition price) and μ_k^A is the persistent part of the acquisition payoff shock defined in (2.1). Repositioning CCPs are given by the following formula:

$$\text{CCP}_k^R(r|\mathcal{J}) = \frac{\exp\left\{\sigma_k^R(\mathcal{J},r)^{-1}\left[V_k(\mathcal{J}'(k,r)) + \mu_k^R(\mathcal{J},r)\right]\right\}}{\sum_{r'\in\Gamma_k^R(\mathcal{J})}\exp\left\{\sigma_k^R(\mathcal{J},r')^{-1}\left[V_k(\mathcal{J}'(k,r')) + \mu_k^R(\mathcal{J},r')\right]\right\}}. \tag{2.8}$$

The computational algorithm involves iterating on the value function using a Bellman equation (2.6) and equations (2.7) and (2.8). The procedure can be summarized as follows:

Init: Initialize the value function $V^{(0)}$.

(1) For every state $\mathcal{J}$,

  (i) use $V^{(j)}$ to compute CCPs of all players at $\mathcal{J}$, given by equations (2.7) and (2.8),

  (ii) use the CCPs from (i) to obtain a new value function $V^{(j+1)}(\mathcal{J})$ by iterating a Bellman equation (2.6).

(2) Stop if $\|V^{(j)} - V^{(j+1)}\| <$ tolerance; otherwise, go to (1).

Several features of this algorithm facilitate the computation of large games. Primarily, iteration steps (i) are (ii) are relatively cheap because the integration in the Bellman equation is done player by player, instead of jointly (see discussion in Doraszelski and Judd (2012)). Therefore, its complexity does not grow exponentially but only linearly, as the number of active players increases. Additionally, best response CCPs depend on strategies of other players only through the value functions. In such case one does not need to remember a full set of CCPs at every state. Note that storing all CCPs in a reasonable amount of memory might be infeasible if the action space has large support (large enough support is frequently needed to match the data). Also, because only one player changes state at each instant, the state transitions $\mathcal{J}'(k,a)$ and $\mathcal{J}'(k,r)$ are relatively simple. Therefore, state encoding and decoding routines (which can take up to 60%-70% of the execution time depending on the problem) can be replaced with look-up tables. Lastly, a closed-form of conditional expectations $W_a^A(\text{CCP}_k^A, \mathcal{J}^t)$ and $W_r^R(\text{CCP}_k^R, \mathcal{J}^t)$ for more than two feasible actions is unknown. Instead, these expectations have to be simulated (see section 5 for details).

3 Numerical examples

In this section, I use numerical examples to show how various long-run processes can affect the dynamics of merger bids and the efficacy of a merger enforcement process. The first two examples demonstrate that myopic regulation could be suboptimal. Specifically, I identify cases for which a myopic antitrust agency over-blocks (Example 1) and under-blocks (Example 2) mergers relative to a dynamic optimum. In addition, I show the efficacy of predicting merger waves using a myopic model might be limited (Example 3).

I conduct the numerical analysis on the example of the radio market using an over-simplified single-sided model of competition. The advantage of such a model in the context of numerical experiments is that the results do not depend on the features specific to the radio markets, such as two-sidedness, making the extrapolation of the numerical findings to other industries more straightforward. I lift these restrictions during the empirical implementation.

Suppose the market is composed of radio stations that hold some degree of market power within their formats. I assume there are C types of consumers and the utility of a consumer i of type c of listening to station j of type f is given by

$$u_{ij} = \alpha_{cf} - \beta p_j + \epsilon_{ij},$$

where p_j is the price of listening to the station j and ϵ_{ij} is an idiosyncratic taste shock that is distributed extreme value. The price of listening to the station could be a direct subscription fee or a dollar value of avoiding broadcasted advertising. Additionally, consumers can choose an outside option with zero mean utility.

Conditional on prices p and industry state ω, the market share of station j is given by

$$s_j(p, \omega) = \sum_{c=1}^{C} \mathcal{P}(c) \frac{\exp(\alpha_{cf(j)} - \beta p_j)}{1 + \sum_{j'=1}^{n} \exp(\alpha_{cf(j')} - \beta p_{j'})},$$

where n is the number of active radio stations and $f(j)$ is a format of station j, both prescribed by ω. $\mathcal{P}(c)$ is a proportion of the consumers of type c.

For simplicity, in the remainder of this section, I assume $C = F = 2$. I also assume each consumer type has a favorite format; that is, $\alpha_{cc} = \gamma_1$ and $\alpha_{fc} = \gamma_2$ if $c \neq f$, where $\gamma_1 > \gamma_2$. The difference between γ_1 and γ_2 measures a consumers' willingness to switch to a format different from their favorite, and determines the degree of within-format market power.

In the reminder of this section, I divide active firms into corporations and local firms. A corporation can hold multiple radio stations, whereas a local firm can hold only one radio station. Both types of firms can reposition radio stations, and in particular, a corporation can reposition individual stations within its portfolio. Importantly, corporations can acquire local firms, whereas local firms cannot acquire corporations. Both types of firms are fully forward looking when repositioning and bargaining over the acquisition price. In the first example, I show that a myopic regulator can behave sub optimally and over-block mergers.

Example 1: Post-merger repositioning (entry). Mergers usually result in an increase of market power and markups, which benefits the merged entity but can also spill over to competitors (see Salant, Switzer, and Reynolds (1983) for an analysis of the Cournot model). These spillovers encourage post-merger entry and repositioning, which could mitigate the negative effects of the merger (see section 9 of Horizontal Merger Guidelines). A myopic regulator, who does not account for entry, would over-estimate the negative effect of the merger, which could result in over-blocking.

A simple setting in which I can show over-blocking is with a single corporation and four local firms. Consider a state $\mathcal{J}^0$ in which a corporation owns one station in format 1, and additionally one local firm is format 1, and three local firms are in format 2. Formally, $\mathcal{J}^0 = [(\mathbf{1}, \mathbf{0}), (1, 0), (0, 1), (0, 1), (0, 1)]$. Now consider the corporation's acquisition of company 2. The industry moves into a state $\mathcal{J}^1 = [(\mathbf{2}, \mathbf{0}), (0, 1), (0, 1), (0, 1)]$. If a static welfare of $\mathcal{J}^0$ is greater than $\mathcal{J}^1$, because of the monopolization of format 1, then a myopic regulator rejects the merger. However, in the long run, higher rents in format 1 could invite repositioning from format 2 leading to $\mathcal{J}^2 = [(\mathbf{2}, \mathbf{0}), (1, 0), (0, 1), (0, 1)]$. Note that format 1 is no longer monopolized, so enough economies of scale would make $\mathcal{J}^2$ preferable over $\mathcal{J}^0$. Thus the merger should be approved provided the repositioning is timely, likely, and sufficient (see Horizontal Merger Guidelines, section 9).

I analyze the aforementioned situation numerically in the context of both total-surplus criterion and consumer-surplus criterion. Table 1 presents the results of computational experiments for an industry with fixed cost synergies and a regulator using a total-surplus criterion. The static deadweight loss varies between 0.6% (of the total surplus) and 4.2%, for high and low levels of fixed cost synergies, respectively. Thus the myopic total-surplus maximizer would reject this merger. However, depending on the repositioning cost and the size of cost synergies, a forward-looking regulator might want to approve the merger. For the highest level of fixed cost synergies, the merger should be approved even for large values of repositioning cost, bringing between a 0.7% and 2.2% increase in total welfare. Note that if the cost synergies are smaller, the regulator should approve the merger only if the entry is likely and timely, that is, when the repositioning cost is low. Thus, to make a correct decision, the regulator should have estimates of both cost synergies and repositioning cost.

In Table 2, I present a similar analysis for the industry with marginal cost efficiencies and for the regulator who enforces a consumer-surplus criterion. I consider the situation in which the corporation has lower marginal cost than the local firm. The top of the table contains a case in which marginal cost efficiencies from the merger are large enough to counteract a static increase in market power. Consequently, both myopic and forward-looking regulator take the same decision and approve the merger. However, if the cost synergies are smaller, the static impact of the merger on consumer surplus is negative and a myopic regular rejects the merger. At the same time, the dynamic regulator approves the merger if the repositioning cost is low enough.

Example 2: Predatory mergers. Consider a situation in which a merger results in substantial marginal cost synergies. Such cost synergy would benefit the conglomerate but might hurt the smaller competitors and prompt exit. Following this logic, I call a merger predatory if it is aimed at driving competitors out of the market by using substantial marginal cost advantages. This example illustrates how myopic regulators can act sub optimally by approving predatory mergers they should reject.

A modification of Example 1 provides a simple setting for analyzing this kind of dynamics. The industry starts in state $\mathcal{J}^0 = [(\mathbf{1}, \mathbf{0}), (1, 0), (1, 0), (0, 1), (0, 1)]$. The corporation proposes an acquisition of the second firm, which leads to $\mathcal{J}^1 = [(\mathbf{2}, \mathbf{0}), (1, 0), (0, 1), (0, 1)]$. A myopic regulator approves the merger because it calculates that the drop in consumer surplus would not be significant. However, the local firm repositions to format 2, leading to a suboptimal configuration $\mathcal{J}^2 = [(\mathbf{2}, \mathbf{0}), (0, 1), (0, 1), (0, 1)]$.

Table 2 presents the results of multiple computational experiments exploring this dynamics. A corporation with two products has a marginal cost of 0.75, whereas a corporation with one product and any other company has a marginal cost of 1. The tables contain percentage changes in consumer surplus caused by the merger and the probability that at least one firm exits the first format during a unit interval following the merger. I demonstrate that the myopic regulator consistently underestimates the impact of mergers on consumer surplus and in many cases would under-block the mergers. In the most notable case, the myopic regulator would actually predict a 1.3% increase in consumer surplus, but in reality, the merger would result in more than a 3.6% drop in consumer surplus.

Example 3: Merger waves and hold-out. This example demonstrates the difference between a myopic and dynamic model concerning merger waves. Consider an industry with one format, no repositioning, one active corporation owning one station, and three local firms available for acquisition. The utility of the product is equal to 2, and the price coefficient is equal to -1. Jointly operating two and three firms results in 50% and 66% efficiency gains in marginal cost, respectively. The regulator always forbids a merger to monopoly, thus only four-to-three and three-to-two mergers are feasible. I consider two levels of player sophistication: a myopic case in which both a buyer and a seller maximize static profits, and a forward-looking case in which they maximize a discounted stream of profits.

Table 4 presents equilibrium probabilities of mergers for different primitives and levels of player sophistication. Intuitively, the likelihood of the merger is a decreasing function of a merger execution cost and an increasing function of marginal cost (because of the marginal cost efficiencies). First, I note that a probability of a merger in a myopic model is usually higher than in a dynamic model. To explain the higher probability, looking at the three-to-two merger first is useful. In this case, no further mergers are possible, so the marginal market power of merging in a myopic world is equal to the marginal market power of merging in a forward-looking world. However, the same is not true for acquisition prices and the difference between the myopic and forward-looking acquisition prices is depicted in Figure 1. In particular, I find the acquisition price is substantially higher in the forward-looking case resulting in a *hold out*. The reason is that a merger increases the markups of competitors that are not acquired, so a forward-looking acquiree has an incentive to reject a static merger bid and internalize these spillovers. Note these incentives disappear once the economies of scale increase and competing with a conglomerate becomes harder.

The case of a four-to-three merger is more complicated. Figure 2 depicts the difference in acquisition price between myopic and forward-looking cases. For small marginal cost synergies, the results for a four-to-three case are similar to those of a three-to-two case. However, for large values of the marginal cost the acquisition price in the myopic model is greater than in the forward-looking model. To understand that, consider the firm's incentives to reject a four-to-three merger bid. In the case of a high marginal cost, the potential acquiree recognizes that rejecting the bid would likely lead to competing against a highly efficient conglomerate. As a result, the acquirer might credibly threaten the acquiree to make a bid to a competitor and subsequently come back with a lower offer leveraging on its size. Thus, in the extreme case, the acquiree agrees to a haircut, that is, selling out below current static profits. Such a haircut explains why a four-to-three merger is more probable than a three-to-two merger (see the last rows of Table 4).

4 Data

The data used to estimate a dynamic model consists of (i) a complete set of radio station acquisition transactions between 1996 and 2006 with monthly time stamps, and (ii) formats of every radio station in the United States with half-year time stamps. Additionally, the paper uses a pre-estimated static mapping, $\pi_k(\mathcal{J}_t)$, between market structure and station revenues. The mapping is estimated for a subset of 88 non-overlapping markets, using a panel data set on listenership shares, advertising quantities, advertising prices and revenues. Appendix B contains the details on this data set and static estimation. The reminder of this section concerns the data used to estimate the dynamic model.

During the estimation, I introduce several data simplifications that reflect the main features of the radio industry. Primarily, I divide the set of players into three groups: national owners, local owners, and fringe. National owners include such companies as Clear Channel, Cumulus, or Univision that own a complicated network of stations nationwide. I allow these companies to own multiple stations in local markets as well as acquire new stations. I also allow national owners to reposition stations within their portfolios. The second group of companies are local owners. They are not allowed to own multiple stations; however, they are allowed to reposition. Both national and local owners are forward looking about repositioning and bargaining about the acquisition prices. The remainder of companies compose the fringe. Companies in the fringe are myopic and cannot reposition or be acquired but still play in the static game.

In each market, I label three companies with the largest national revenue share in 2006 (last year of the data) as national owners. I label as local owners the next 22 stations with the highest market share, including those that were ever owned by a national owner are labeled as local owners. One exception to the above rule is that I put all AM stations in the fringe in the bigger markets. In such markets, FM stations generate a dominant part of total revenues. In rural markets, AM stations become important, so I allow both AM and FM stations to be outside of the fringe. In practice, the fringe stations are small and usually have less than a 0.5% market share.

Dividing owners into the aforementioned groups has some important consequences. The upside is that it captures the important features of the radio market and reduces the complexity of the estimation. In particular, it enables estimating acquisition price, that is, a value function the acquiree, without tracking the possibility that the acquiree can make merger offers himself. This procedure enables me to use a simple two-step estimator to recover the parameters of the dynamic model. I note that dividing players into groups is not a limitation of the model per se and can be relaxed if the application requires it. I can also relax it when computing counterfactuals and if using a nested-fixed-point estimator, but given my data, these extensions would require further assumptions and are not implemented.[4]

Modeling national owners, local owners, and the fringe is a minimum necessary compromise chosen to capture first-order dynamics of the radio industry. For example, putting all local owners in the fringe would be a strong assumption. First, as shown in examples 1 and 2 in section 3, the regulator must track product repositioning of smaller players in response to the merger. Also, as shown in example 3 in section 3, the acquirees should be forward looking. However, at the same time, the smallest stations rarely change formats and are almost never acquired by larger owners in the data. Thus, in practice, modeling forward-looking decisions of every small owner has little benefit, while increasing the complexity of the estimation. Nevertheless, dropping the smallest owners altogether is unrealistic because they collectively affect markups in the pricing game. Thus, fully tracking only large and medium, and only partially tracking small local owners is a realistic compromise.

One artifact of not allowing smaller players to make merger bids is prohibiting spin offs. I do observe spin offs in the data, but they are mostly a consequence of lumpy cross-market mergers

[4]I did not use a nested-fixed-point algorithm for two reasons: (i) it requires strong assumptions on equilibrium selection, and (ii) it would require further and unrealistic simplifications to the model.

that can violate the local ownership caps. Consequently, the owners spin off certain stations to stay within the regulatory rules. According to the anecdotal evidence, the candidates for spin offs are determined in advance and are unlikely to be fully integrated in the new owner's portfolio in the first place. Thus, counting them in the new owner's portfolio would most likely overestimate the market power of the merged entity. I use this convenient fact and ignore acquisition of stations that were spun off later.

The data contains information about more than 100 possible formats. I aggregate these formats into three meta formats: (i) "Adult Music," containing such formats as Adult Contemporary, Jazz, Rock, and Country, (ii) "Hits Music," containing such formats as Contemporary Hit Radio, Urban, and Alternative, and (iii) "Non-music," containing such formats as Talk, News, Ethnic and Religious formats. Such choice is dictated by taking into account substitution patterns described by Jeziorski (2012). The "Adult Music" caters to a more mature population of listeners, while the "Hits Music" attracts mostly younger crowd. The details on the substitution patterns can be found in section 6.1. The aggregation is a way to trade off a static realism for a dynamic realism. Namely, I sacrifice some accuracy in capturing within-period behavior by dropping second order format designations. However, such aggregation allows me to describe cross-period behavior in more detail. At the same time, I note that the inaccuracy in describing within-period behavior can translate into inaccurate cross-period predictions. Keeping this caveat in mind, I proceed to estimate the model and come back to this issue when discussing the results.

5 Estimation

The estimator used in this paper belongs to the class of two-step methods pioneered by Hotz and Miller (1993). These methods enable estimation of large dynamic systems without re-solving for an equilibrium at each parameter value. Hotz and Miller (1993) developed the estimator for discrete-time single-agent problems, and many paper have extended their method to discrete-time dynamic games (see Pakes, Ostrovsky, and Berry (2007), Bajari, Benkard, and Levin (2007), Aguirregabiria and Mira (2007) and Arcidiacono and Miller (2011)). This paper adapts the method of Aguirregabiria and Mira (2007) to the case of continuous time games, and for the data that comes in discrete intervals. The method developed in this paper is similar to the CCP inversion method developed by Arcidiacono, Bayer, Blevins, and Ellickson (2010), but does not rely on the existence of the terminal state with a normalized terminal value, and relaxes the functional form of payoff shock distribution. [5]

The section is divided into three parts. The first describes the pre-estimation of a one-shot profit function; the second describes the estimation of acquisition and repositioning strategies; and the third describes simulated pseudo-likelihood estimation of structural parameters.

5.1 Estimation of one-shot profits

The single-period profit function is identical to one used in Jeziorski (2012) with the exception that I use three meta-formats instead of eight. Below I describe the parametrization in order keep to the paper self-contained; however, the discussion is rather brief to avoid duplications.

Firms receive a continuous stream of advertising variable profits from the station portfolio they own. The infinitesimal variable profit flow is summarized by a function $\pi_k(\mathcal{J}^t)$. These profits are a result of a static competition, and account for marginal cost with a possibility of post-merger synergies. Variable profits of the firm in the radio market have the following general form:

$$\pi_k(\mathcal{J}^t) = \sum_{\substack{\text{j owned by k} \\ \text{in market m}}} \bigg(p_j(\bar{q}_j^t|\mathcal{J}^t) r_j(\bar{q}^t|\mathcal{J}^t) - \text{MC}_j(\mathcal{J}^t) \bigg) \bar{q}_j^t. \tag{5.1}$$

$p_j(\cdot)$ is price per listener (advertising inverse demand) of one ad slot, $r_j(\cdot)$ is a listenership market share (demand for programming), and $\bar{q}_j^t$ is the equilibrium number of advertising slots at station j. MC_j is marginal cost of selling advertising at station j. Dependence of the marginal cost on the state $\mathcal{J}^t$ signifies a possibility of marginal cost synergies from joint ownership.

I compute the station's market share using a logit model with random coefficients, following Berry, Levinsohn, and Pakes (1995). Let $\iota_j = (0, \ldots, 1, \ldots, 0)$, where 1 is placed in a position that indicates the format of station j. Denote the amount of broadcasted advertising minutes in station j as q_j. For a given consumer i, the utility from listening to a station j is given by

$$u_{ij} = \theta_{1i}^L \iota_j - \theta_{2i}^L q_j + \theta_3^L \text{FM}_j + \xi_j + \epsilon_{ji}, \tag{5.2}$$

[5] In the case of mergers, a terminal state occurs when the player is acquired. The terminal value, which is an acquisition price, cannot be normalized because it is endogenous; that is, it depends on the option value of rejecting the merger bid.

where θ_{1i}^L is a set of format fixed effects, θ_{2i} is a disutility of advertising, and θ_3^L is an AM/FM fixed effect. I assume the random coefficients can be decomposed as

$$\theta_{1i}^L = \theta_1^L + \Pi D_i + \nu_{1i}, \quad D_i \sim F_m(D_i|d), \quad \nu_{1i} \sim N(0, \Sigma_1)$$

and

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

where Σ_1 is a diagonal matrix, $F_m(D_i|d)$ is an empirical distribution of demographic characteristics, ν_i is an unobserved taste shock, and Π is the matrix representing the correlation between demographic characteristics and format preferences. I assume draws for ν_i are uncorrelated across time and markets. The term ξ_j represents the unobserved quality of station j. The assumptions on ξ_j are the same as in Berry, Levinsohn, and Pakes (1995).[6] The model allows for an outside option of not listening to radio u_{i0}, which is normalized to zero in the years 1996 and 1997. For subsequent years, u_{i0} contains time dummies to control for the influx on new broadcasting technologies such as satellite radio and internet.

The market share of the station j is given by

$$r_j(q|\mathcal{J}^t) = \text{Prob}\big(\{(\nu_i, D_i, \epsilon_{ij}) : u_{ij} \geq u_{ij'}, \text{for } j' = 1, \ldots, J\}\big|q, \mathcal{J}^t\big). \tag{5.3}$$

The radio-station owners are likely to have market power over advertisers. Moreover, because of heavy ad targeting, the stations with different formats are not perfect substitutes, which might be a result of multihoming of advertisers, as well as advertising congestion. The simplest reduced-form model that captures these features, and is an approximation of the industry, is a linear inverse demand for advertising, such as

$$p_j = \theta_1^A \left(1 - \theta_2^A \sum_{f' \in \mathbb{F}} w_{ff'}^m q_{f'}\right), \tag{5.4}$$

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 $w_{ff'} \in \Omega$ are weights indicating competition closeness between formats f and f'.

[6] The assumptions on ξ are a simplification compared to the specification used by Jeziorski (2013) and Sweeting (2013), who both assume ξ_j follows an AR(1) process. This decision was made to keep the dynamic model computable.

To capture potential marginal cost synergies, a marginal cost of station j is allowed to depend on the portfolio of stations ω_k^t of its owner. In particular, I set

$$\mathrm{C}_{jmt}(\theta^A, \theta^C) = \theta_1^{Am}[\theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Cm}\xi_{jt} + \theta_3^{Cm}\mathrm{SYN}_{jt} + \epsilon_{jt}^C]q_{jt}, \tag{5.5}$$

which allows for station-level unobserved heterogeneity captured by ϵ_{jt}^C. The term θ^{Cmt} represents time dummies capturing aggregate shocks to marginal cost. Unobserved market-level heterogeneity is captured by the fact that θ_1^{Am} is allowed to be different for each market, and θ^{Cm} is allowed to differ between subsets of markets depending on their size. The parameter θ_3^{Cm} measures the extent of marginal cost synergies between stations of the same format owned by the same owner, and is interacted with a dummy variable SYN_{jt} that is equal to 1 if the current owner owns more than one station in the format. Cost synergies are likely to occur because of scale economies in producing and selling advertising for multiple stations with similar target groups.

Given the advertising quantity choices of competing owners, each radio-station owner k chooses q_j^t for all owned stations to maximize its variable profits. The market is assumed to be in a quantity-setting Nash equilibrium.

The profit function varies across markets because of market-specific parameters and because the demographic composition of listeners is heterogeneous. Note that the static model is non-stationary because it contains time dummies in the demand in the demand and supply equation, and because the distribution of listeners' demographics varies from year to year. Consequently, I estimate it as a non-stationary model to obtain more robust measures of listener and advertiser price elasticity. However, after the estimation, I detrend the static profits to fit it into the dynamic model presented in section 2. Specifically, I remove the trends in supply and demand by using an average value of the time dummies and draw listeners from a joint 1996-2006 market-specific empirical distribution instead of year-by-year distributions. In the case of radio, most of the possible non-stationarity would result is underestimation of cost synergies, because of overestimation of the gains from mergers. For example, a possible long-run downward trend in listenership should make mergers less profitable in the absolute sense. In such a case, my model would likely overestimate the long-run returns from the merger and underestimate the cost synergies. Another example is possible changes in technology that make running bigger companies more efficient, and would also bias my cost estimates downwards. I return to these points when discussing the results and counterfactuals.

Additionally, I simplify the state space to make the dynamics manageable. I use average station quality and tower power computed for each market-format combination, and separately for the fringe to control for size differences. Such procedure further standardizes the profit function in a way to be consistent with the model laid out in section 2; in particular, this standardization makes stations homogenous with an exception of the meta-format and fringe designation. Note this assumption is stronger than the one in Jeziorski (2013), who allows for persistent but exogenous heterogeneity. If the radio stations are indeed heterogeneous in more dimensions and such heterogeneity creates a large amount of extra market power, this procedure could pollute the estimates of cost synergies as well as counterfactuals. To alleviate some of the concerns, I compare cost synergy estimates to those in Jeziorski (2013); however, all results in the paper should be interpreted keeping the standardization in mind.

5.2 Estimation of acquisition and repositioning strategies

The data I use to estimate the dynamic model is summarized by two sets. The first set describes merger decisions

$$X^A = \{a^{mhi} \subset K \times K : 1 \leq i \leq 6, h \in H, m \in M\},$$

a^{mhi} is an observed set of mergers, m is a local market, h is a half-year period, and i is the month in which the mergers took place. Several instances of multiple mergers exist in the same half-year, as well as multiple mergers in the same month. I can observe the sequence of mergers across months; however, I do not observe the sequence of mergers within the month. Therefore, for the periods that have multiple mergers in the same month, the state space at the time of taking an action is only partially observed.

The second set describes repositioning decisions:

$$X^R = \{b^{mh} \subset J \times F \times F : h \in H, m \in M\},$$

b^{mhd} the observed set of repositioning events during half-year h. The formats are observed once every half a year, as opposed to mergers that are observed on a monthly basis. Multiple merger and repositioning actions during the same data period create complications. For example, if the station was acquired and repositioned in the same half-year, I do not see which player took the repositioning action. Furthermore, I do not know how many active players were present when the repositioning action was taken. For these reasons, the state, set of players, and players' actions are only partially observed which has to be taken into account during the estimation.

One can integrate out the aforementioned unobservables in several ways. One option is to use simulations and perform either maximum likelihood or generalized method of moments. Unfortunately, both methods are impractical for my application. The former would require too many simulations to obtain a reasonably precise likelihood; the latter would lead to substantial loss in efficiency. Another option is to obtain partially analytical likelihood using Chapman-Kolmogorov equations describing state transitions, as suggested by Arcidiacono, Bayer, Blevins, and Ellickson (2010). However, this method also cannot be directly applied, because the full intensity matrix for my largest markets can contain up to 4 million by 4 million entries. The matrix would be quite sparse; however, one still would not be able to store it in the memory, or recompute it "on the fly." Instead, I develop a method based on partial Chapman-Kolmogorov equations. The partial equations use the fact that only a small subset of feasible latent industry states are relevant for the estimation. The method proceeds in steps.

First, I construct an augmented set of latent states during the half year h and denote it by Ω^h. This set contains the feasible latent states that do not contradict the observed data and a coffin state. Denote a set of feasible states at the end of month i by $\Omega^{hi} \subset \Omega^h$. States in Ω^{hi} incorporate all mergers that happened prior to and including month i, that is, $\{a^{hd} : d \leq i\}$, as well as any possible subset of repositioning events b^h that occurred during half-year h. The special cases are: Ω^{h0}, which contains only the fully observed starting state at the beginning of a half-year h, and Ω^{h6}, which contains only the fully observed state at the end of half-year h. A full set of feasible states Ω^h is a union of: (i) all sets Ω^{hi}, (ii) all feasible transitory states between from Ω^{hi} to Ω^{hi+1}, and (iii) a coffin state $\bar{\omega}$. The coffin state encompasses all infeasible states. In practice, constructing the set of feasible states Ω^h might be computationally expensive. In this paper, I employ a backward induction recursion that constructs and examines all feasible paths of the industry between Ω^{h0} and Ω^{h6}. For example, a computation of feasible paths for seven mergers and three repositioning events within one half-year period can take up to a week and require up to 40GB of memory to store the temporary data (Matlab code on 2GHz AMD Opteron CPU). The exercise in this paper is feasible because the process was parallelized. Despite a long preparation time, this computation has to be done only once for each data set. The final augmented state space is a thousands times smaller than the full state, which dramatically reduces the size of the intensity matrix.

Upon arrival of the merger and repositioning actions at time t, the equilibrium strategies induce transitions according to instantaneous conditional choice probabilities of acquisition $\text{CCP}^A(\mathcal{J}^t)$ or repositioning $\text{CCP}^R(\mathcal{J}^t)$. Together with action arrival rates λ^A and λ^R, these CCPs generate an intensity matrix Q^h on the augmented state space Ω^h. The overall goal is to use the Markov process on Ω^h to compute the conditional likelihood of the data, that is, $L(\Omega^{h6}, \ldots, \Omega^{h1}|\Omega^{h0})$. The exact states from Ω^{hi} (expect for the beginning and the end of the half-year) as well as transitory states between Ω^{hi}s are unobserved to the econometrician and have to be integrated out.

Denote the time that passed since the beginning of h by $s \in [0, 6]$. Let $\iota^h(s)$ be a stochastic process of the latent state of the system conditional on $\{\Omega^{hi} : i < s\}$. Conditioning prevents the $\iota^h(s)$ from contradicting the data by killing the infeasible paths. Note that $\iota^h(0)$ is a degenerate distribution at Ω^{h0}. First, I compute the distribution after the first month, $\iota^h(1)$, by numerically solving a Chapman-Kolmogorov system of differential equations

$$\frac{d\iota^h(s)}{ds} = \iota^h(s)Q^h, \tag{5.6}$$

subject to the initial condition of $\iota^h(0)$ being degenerate at Ω^{h0}. Knowing $\iota^h(1)$, I can obtain $L(\Omega^{h1}|\Omega^{h0})$ by taking the mass of states that belong to Ω^{h1}. The next step is obtaining $L(\Omega^{h2}|\Omega^{h1}, \Omega^{h0})$.[7] For this purpose, I compute $\iota^h(2)$ by solving equation (5.6) with $\iota^h(1)$ conditioned on Ω^{h1} used as an initial condition. The likelihood is the mass of the set Ω^{h2} obtained according to $\iota^h(2)$. By repeating the procedure, we can obtain any of $L(\Omega^{hi}|\Omega^{hi-1}, \ldots, \Omega^{h1}, \Omega^{h0})$, and as a result, I get the joint likelihood $L(\Omega^{h6}, \ldots, \Omega^{h1}|\Omega^{h0})$ by using Bayes rule.

The above procedure can be repeated to obtain a conditional likelihood for every market m and half-year h as a function of CCPs. In my model, equilibrium CCPs are given by equations 2.7 and 2.8, and depend on the state through unknown value functions, which have to be estimated semi-parametrically. In particular, the acquisition CCPs are given by

$$\widehat{\text{CCP}}^A(k'|k, \mathcal{J}, \theta^A) = \frac{\exp\left\{\Upsilon^A(k, k', \mathcal{J})\right\}}{\sum_{k''} \exp\left\{\Upsilon^A(k, k'', \mathcal{J})\right\}},$$

[7]Note that $L(\Omega^{h2}|\Omega^{h1}, \Omega^{h0}) \neq L(\Omega^{h2}|\Omega^{h1})$ even though the latent state ω^{hi} is Markovian, because Ω^{h1} is a set, and the value of Ω^{h0} is informative about the distribution of the latent states in Ω^{h1}.

where $\Upsilon^A(k, k'\mathcal{J})$ are unknown functions of the state $\mathcal{J}$.

Denote the fraction of the total number of active non-fringe stations in format f and owned by player k as $\eta_{f,k}$. Formally,

$$\eta^t_{f,k} = \frac{\omega^t_{fk}}{J}.$$

Additionally, I denote a set of national owners as $\mathbf{K^N}$ and a set of local owners as $\mathbf{K^L}$. These sets must meet an adding-up constraint given by $K = \#\left(\mathbf{K^N} \cup \mathbf{K^L}\right)$,where $\#$ denotes the number of elements in the set. The above notation is useful for expressing statistics from the state that determine acquisitions and repositioning. For example, a fraction of stations that are locally owned and have format f is given by $\sum\limits_{k\in\mathbf{K^L}} \eta_{k,f}$.

After introducing the above notation, I define the approximations of Υ^A and Υ^R by polynomials of η. The coefficients of these polynomials satisfy a certain set of restrictions imposed by the availability of the data, namely: (i) symmetric equilibrium and (ii) no mergers across the national owners. With more data, I could potentially relax the first restriction by estimating Υ^A and Υ^R separately for each player. Similarly, if I observed many mergers of national owners, I could potentially estimate Υ^A separately for those types of actions. In practice, despite the fact that the merger data is rich, relaxing either of these restrictions is infeasible. Imposing the above restrictions, I approximate the above indexes with polynomials

$$\Upsilon^A(k, k', \mathcal{J}) \approx \mathcal{P}(\theta^A_{f(k')}, \eta),$$

where $\mathcal{P}$ is a polynomial of the statistics η, and $\theta^A_{f(k')}$ are coefficients that are specific to the format $f(k')$ of the only radio station owned by a local owner k'. Similarly, the repositioning policy index function could be written as

$$\Upsilon^R(k, f, f', \mathcal{J}) \approx \mathcal{P}(\theta^R_{f,f'}, \eta).$$

The unknown functions Υ^A and Υ^R are estimated using a sieve minimum distance estimator (see Ai and Chen (2003)). In a finite sample, I simply choose the polynomial coefficients that maximize a pseudo-likelihood of the data.

I do not observe the events in which the players decide not to take an action, so I can only identify the product of a move arrival rate λ and CCPs in the first stage. However, as long the move arrival rate is not too large, I can estimate the first stage by choosing a reference value of λ

set to 1. Relevant CCPs for a desired value of an arrival rate could be obtained by dividing the estimates by λ.

5.3 Estimation of structural parameters

This subsection contains a description of the simulated pseudo-likelihood estimator that is used to recover the structural parameters of the game. The estimator is based on simulated instantaneous value functions, that is, conditional on the arrival of the right to move. The simulations are based on the arrival rates of executed actions recovered in the previous stage. The simulated value functions are subsequently used to form new CCPs, which are used as inputs to generate the new pseudo-likelihood using the sequential procedure based on Chapman-Kolmogorov equations described in the previous section. Some unknown primitives of the structural model are further parametrized, and I describe these parametrizations in the reminder of this section.

Fixed cost. The fixed cost of player k of owning a station j in format f is parametrized as follows:

$$F_{kj}^{m}(\mathcal{J}^t|\theta^F) = \bar{F}_f^m \times F^S(\omega_{kf}^t, z_k|\theta^F) \times F^E(n_k^t, z_k|\theta^E). \tag{5.7}$$

The cost is composed of three terms: (i) term $\bar{F}_f^m$ is a fixed cost of owning a single station of format f in market m, without owning additional stations in this or any other market; (ii) the function F^S represents a fixed cost discount caused by synergies of operating multiple stations in the same format and the same local market; and (iii) the function F^E represents a fixed cost discount caused by within- and cross-market economies of scale. Note that for local owners, F^E and F^S are equal to 1.

The market-level fixed cost of owning one station $\bar{F}_f^m$ is assumed to be proportional to average variable profits (before fixed cost) in the market, calculated separately for each format. I compute this average by simulating an industry path for each observed data point and averaging over time. The simulation is done using the first-stage estimates.

I postulate that for national owners:

$$F^S(\omega_{kf}^t, z_k = N|\theta^F) = \frac{\left(\omega_{kf}^t\right)^{\theta^F}}{\omega_{kf}^t}.$$

Parameter θ^F captures the synergy and is expected to lie between 0 and 1. I allow for economies of scale by setting F^E in the following way[8]

$$F^E(n_k^t, z_k = N|\theta^E) = \theta_N^F \frac{(n_k^t)^{\theta^E}}{n_k^t},$$

where θ_N^F is a discount for being a national owner and θ^E is a parameter that captures local economies of scale.

Acquisition cost. The acquisition cost has a persistent part $\mu_k^A(\mathcal{J}^t, k'|\theta^A)$ and an idiosyncratic part with volatility $\sigma_k^A(\mathcal{J}^t, k'|\theta^A)$. The persistent part is parametrized in the following way:

$$\mu_k^{A,m}(\mathcal{J}^t, k'|\theta^A) = \theta^{A,m} + \theta_\pi^A \pi_k(\mathcal{J}^t).$$

The acquisition cost might depend on the company's size, because integrating into a bigger company can be more costly, which is captured by the dependence of μ_k^A depends on the variable profits as a proxy for size. I postulate a similar relationship for the idiosyncratic volatility:

$$\sigma_k^{A,m}(\mathcal{J}^t, k'|\theta^A) = \theta_\sigma^{A,m} + \theta_{\sigma,\pi}^A \pi_k(\mathcal{J}^t).$$

Because acquisition cost is likely heterogeneous across markets, I allow the intercepts $\theta^{A,m}$ and $\theta_\sigma^{A,m}$ to vary across four market categories. The first category consist of markets in which a single station has average variable profits greater than \$150,000, the second category has variable profits in the range \$150,000-\$60,000, the third in the \$60,000-\$20,000, range and the fourth less than \$20,000. Additionally, $\theta^{A,m}$ might vary across formats, because layoff costs as well as other integration costs (human and physical resources reallocation) might vary with the type of programming. I try this specification and find the differences are economically small (less than 5%) and statistically (1%-size test) insignificant.

Repositioning cost. Similarly to the acquisition cost, the repositioning cost has a persistent part $\mu_k^R(\mathcal{J}^t, k'|\theta^R)$ and an idiosyncratic part with volatility $\sigma_k^R(\mathcal{J}^t, k'|\theta^R)$. It is reasonable to expect that national owners face different repositioning costs than local owners. For example, voice-tracking technology can allow the national owners to temporarily bring announcers from other markets to streamline format switching. However, local owners might have better access to

[8] I also try other specifications, such as fixed effects for discounts when n_k^t is greater than 3 or 4, and arrive at similar results.

local labor markets and a more flexible workforce. These differences might additionally vary by format. One example is the Hits Music format, which requires a large tower and costly marketing to gain enough listenership. For this reason, switching to the Hits format is likely to require more capital investments and access to specialized production factors. Thus I expect national owners to have lower switching cost into this format. To accommodate that expectation, I postulate the following parametrizations:

$$\mu_k^{R,m}(\mathcal{J}^t, f, f'|\theta^R) = \theta^{R,m}\left[\mathbf{1}(z_k = L)\theta^R_{L,f',f} + \mathbf{1}(z_k = N)\theta^R_{N,f',f}\right] + \theta^R_\pi \pi_k(\mathcal{J}^t)$$

and

$$\sigma_k^{R,m}(\mathcal{J}^t, f, f'|\theta^A) = \theta^R_{\sigma,m} + \theta^R_{\sigma,\pi}\pi_k(\mathcal{J}^t).$$

The intercepts of repositioning costs are allowed to be from/to format specific and company-type specific. I find that allowing for this flexible specification is critical to fitting the model to the data. Specifically, I estimate separate $\theta^R_{\sigma,m}$ for national and local owners, and allow for mean differences and heteroscedasticity by size, which is captured by parameters θ^R_π and $\theta^R_{\sigma,\pi}$, respectively. To control for differences in switching costs across markets, I allow for market-category multiplicative fixed effects in the persistent part $\theta^{R,m}$, and the idiosyncratic part $\theta^R_{\sigma,m}$.

The above specification is used to simulate the value function

$$\begin{aligned} V_k(\mathcal{J}^t|\theta) = &\int_{s=t}^{\infty} e^{-\rho s}\pi_k(\mathcal{J}^s)ds - \int_{s=t}^{\infty} e^{-\rho s}F_k(\mathcal{J}^s|\theta)ds + \sum_{l=1}^{\infty} e^{-\rho\tau_k^{A,(l)}} P(a_k^{(l)}, \mathcal{J}^{\tau_k^{A,(l)}}|\theta) + \\ &\sum_{l=1}^{\infty} e^{-\rho\tau_k^{A,(l)}} W^A_{a_k^{(l)}}(\mathrm{CCP}^A_k, \mathcal{J}^{\tau_k^{A,(l)}}|\theta) + \sum_{m=1}^{\infty} e^{-\rho\tau_k^{R,(m)}} W^R_{r_k^{(m)}}(\mathrm{CCP}^R_k, \mathcal{J}^{\tau_k^{R,(m)}}|\theta). \end{aligned} \tag{5.8}$$

The acquisition prices $P_k^{(l)}$ (value functions of the local firms) are simulated using a nested routine. The routine is triggered upon the arrival of a merger action at time $\tau_k^{A,(l)}$, and it simulates the continuation value of the local owner conditional on rejecting the merger bid. This value includes future mergers between rivals, as well as potential repositioning of the firm and its rivals. One can show by backward induction (on the number of active rivals) that the option value of the local firm must be equal to the value of rejecting all subsequent merger bids. The nested-simulation routine simulates this value using the following formula:

$$V_k(\mathcal{J}^t|\theta) = \int_{s=t}^{\infty} e^{-\rho s}\pi_k(\mathcal{J}^s)ds - \int_{s=t}^{\infty} e^{-\rho s}F_k(\mathcal{J}^s|\theta)ds + \sum_{m=1}^{\infty} e^{-\rho\tau_k^{R,(m)}} W^R_{r_k^{(m)}}(\mathrm{CCP}^R_k, \mathcal{J}^{\tau_k^{R,(m)}}|\theta).$$

The closed-form solution for the conditional expected value of shocks W is unknown for the number of alternatives that is larger than 1 (not including empty action), which is a consequence of the fact that idiosyncratic shocks are not type-1 extreme value. Instead, I simulate the idiosyncratic part of W on the grid of CCPs and fit the 4^{th}-degree complete Chebyshev polynomial. I fit a separate polynomial for each number of feasible alternatives. Such interpolation provides a good approximation along the equilibrium path, with a maximum error of about 1% and lower.

To obtain a second-stage pseudo-likelihood by repeating a sequential Chapman-Kolmogorov procedure described in section 5.2 using updated CCPs given by equations (2.7) and (2.8). Note the value function has to be simulated for every potentially feasible state in Ω^h as well as any state that can be reached by one action. For example, in the case of Los Angeles, 46 feasible latent states exist overall, which generate $1,208$ potentially accessible states. Overall, 88 markets contain $106,304$ accessible states, and the value function needs to be separately simulated for each. Each simulation is composed of $1,000$ draws, so the procedure involves obtaining $106,304,000$ industry paths (the industry path is assumed to evolve for 40 years and is kept constant thereafter). For this reason the simulation procedure must be efficient. Two features facilitate this efficiency: (i) Using my functional-form specification, one can simulate the industry path once and compute the value function for different candidate values of structural parameters θ by using a set of sufficient statistics (details in an Appendix C). In such case the computation of the pseudo-likelihood takes about the same amount of time as the first-stage likelihood. (ii) Continuous time enables simulating the arrival rate of the next change in the industry structure (executed move) instead of drawing move probabilities every period. This feature is important if the moves are rare (such as mergers and product repositioning). Note that I am not drawing opportunities to move, but rather execute moves directly. In an extreme case, when the draw of the waiting time for the first executed move exceeds 40 years, the draw of the value function collapses to perpetual static profits.

Several components of the model need to be identified, namely, (i) repositioning cost θ^R, (ii) merger cost θ^A, (iii) cost efficiencies from mergers (θ^F, θ^E), (iv) level of the fixed cost $\bar{F}_j^m$, and (v) arrival rate λ. Repositioning cost is identified as the residual from endogenizing format repositioning in a way similar to Sweeting (2013) and Jeziorski (2013) – I omit the details for brevity. Merger cost is a residual from entry decisions and is identified in a similar way as entry cost. Because entrants are long lived, they can decide not to enter now, to wait, and to enter in the future, which identifies the volatility of the merger cost. Cost efficiencies are identified as residuals from endogenizing mergers (see Jeziorski (2013) for details). Lastly, the level of the fixed cost (or the fixed cost of owning one station) is bounded from above by the fact that the entry is profitable in all markets, and from below by the level of cost discount required to justify the mergers.

Identifying rate λ separately from other structural parameters is hard without observing move opportunities that were not executed. The identification might still be possible with an exclusion restriction that shifts the continuation value but does not affect current payoffs. One candidate for such exclusion is ownership caps. If the owner is below the cap, the cap does not affect current profits from a merger; however, it shifts future profits through the ability to execute mergers. I tried this approach and found it infeasible, because the variation in the ownership caps in my data is not sufficient. Note the difficulty of identifying λ is not specific to continuous time and is present, but not prominently exposed, in discrete-time games. Specifically, an arrival rate in a continuous-time game is an analog of a period length in a discrete-time game. Identifying such period length is similar to identifying a discount factor, which is know to be hard (see Rust (1994)). For this reason, the length of the period is usually not estimated but fixed, for example, to one action per year. I make a similar simplification and set the arrival rate to once per month; however, I estimated the model with an arrival rate of once per year and obtained qualitatively and quantitatively similar results.

6 Results

In this section, I report the estimates of the structural parameters of the model. I start by describing the estimates of the static pricing game. Then I discuss first- and second-stage estimates of the dynamic model.

6.1 Static model

This subsection contains a brief description of the static profit function estimates. The model of a profit function is a simplified version of Jeziorski (2012) and to minimize the duplication I provide only a brief description of the parameters.

Table 5 presents the estimates of listenership demand. The first and second column contain the mean and the standard deviation of the random coefficients. The numbers match the intuition. Namely, I find that advertising has a negative effect on listenership, and this effect is fairly homogenous among listeners. I also find that listeners prefer FM to AM stations, and greater power of the tower translates to higher listenership. Format dummies are negative; however, by construction, they capture preferences of a specific demographic group, that is, male, uneducated, low income, white, non-Hispanic teenagers. To obtain preferences of other demographic groups, one has to add appropriate demographic interactions, which are presented in Table 6. These numbers match the intuition as well. The first-meta format, which delivers adult-oriented music, such as, classic rock, country, or jazz, is the most popular among middle-aged listeners, with a slight majority of women. The second meta-format, which delivers contemporary pop and alternative music, appeals to younger people. It contains popular pop and urban formats, as well as hip-hop, which explains the large and highly significant African-American fixed effect. The last meta-format, which contains talk radio and ethnic or religious stations, is popular among older listeners, who are mostly higher-educated females with relatively large incomes. A positive Hispanic dummy is related to Hispanic and religious stations in this meta-format.

Table 7 describes the national trends in radio listenership. The values represent the residual trend in radio listenership that is beyond the changes in the demographic composition. In general, I find the trend in the utility of the outside option is non-monotonic. I use these numbers to detrend the profit function in the dynamic estimation.

Table 8 presents the coefficients of the inverse demand for advertising. The inverse demand curve is downward sloping, which is evidence of the direct market power of radio stations over advertisers. The slope is bigger in smaller markets.

Tables 9 and 10 contain the marginal cost estimates. The marginal cost is larger in smaller markets, and I find evidence of marginal cost synergies in small and large markets, but not in medium markets.

6.2 Dynamic model: First stage

The merger and repositioning strategies are estimated as one joint maximum likelihood run described in section 5.2. I grouped the results into multiple tables to facilitate the exposition. Tables 11, 12, 13, and 14 contain the estimates of the acquisition strategy. Tables 15, 16, 17, and 18 present the estimates of the repositioning strategy of national owners. Finally, Tables 19, 20, and 21 contain the coefficients of the repositioning strategy of the local owners. As mentioned in section 5.2, I parametrized the industry state using fractions of the total number of active non-fringe stations in each format f owned by each national owner k, which is denoted as $\eta_{f,k}$.

6.2.1 Acquisition strategy

Table 11 contains the estimates of format-acquisition dummies and interactions between format acquisition and demographic composition of the local market. Format-acquisition dummies, contained in the first row of the table, are large and negative, which reflects the fact that mergers are relatively rare events. The values of these dummies are similar across formats, which suggests that observable characteristics can explain most of the variation in format acquisition. I find the interactions between the demographic composition and acquisition propensity represent the preferences for formats described in Table 6, which serves as a sanity check for the first-stage specification. For example, the Adult Music format has a positive (however statistically insignificant) interaction with age, Hits Music has a positive interaction with Black, and Non-music has a positive interaction with Hispanic.

Tables 12 and 13 present the coefficients describing the relationship between the industry state and a propensity to acquire a particular format f. Specifically, Table 12 contains coefficients on the state variables corresponding to the format f of a potential acquiree. The term $\eta_{f,k}$ represents the coefficient on the number of owned stations in the format of the acquiree. The positive number suggests firms acquire in the formats they already own, which could be the result of demand- or supply-side complementarities. By contrast, Sweeting (2013) and Jeziorski (2013) find that owners avoid acquiring stations similar to their current portfolios. The explanation for this difference is a broader definition of the format in this paper. For example, the owner that owns a Rock station might not acquire another Rock station. However, he might acquire another station in the Adult music format, such as Country or Adult Contemporary. The coefficient on the square of $\eta_{f,k}$ is negative, which could mean the synergies from mergers have decreasing returns.

The third column of Table 12 contains a coefficient capturing national competitors in the same format as a potential acquiree. I find that a larger number of national competitors in a particular format is correlated with a higher propensity to acquire; however, the result is not significant. The fourth column reports a coefficient on ownership concentration (similar to the Herfindahl index) of stations in format f. In general, the less concentrated the ownership, the bigger the propensity to acquire. The last two numbers describe the impact of the portfolio of the local owners on competitors. I find that, opposite to the number of national competitors, the number of local competitors is negatively correlated with an acquisition.

Table 13 contains more state covariates explaining the acquisition decision. I find that a larger number of owned stations is positively correlated with higher likelihood of acquisition. Also, the more stations competing national owners own, the less probable the acquisition. The last two numbers in the table contain higher-order terms that represent concentration of ownership across competitors and distribution of nationally owned stations across formats.

Table 14 contains dummies reflecting being close to the ownership cap. Everything else equal, if the acquisition results in being closer to the cap, it is less probable, which means owners close to the cap become choosier in order keep an option value.

6.2.2 Format-switching strategy

Table 15 contains from-to format-switching dummies, as well as interactions between demographics and the target format. As with acquisition strategy, I find format switching largely reflects listeners' tastes. Tables 16 and 17 describe the impact of the industry state on format switching. In general, the number of stations owned in the format positively correlates with switching to that format, and the opposite is true for the number of stations owned by competitors.

Table 18 contains an impact of closeness to the cap on format switching. Note that the closer the owners are to the ownership cap, the more probable the format switching. Namely, the format switching acts as a substitute for acquisition. For example, if the owner is at the cap and wants to respond to competitors' actions, the only choice is to reposition. If he is not at the cap, he can also acquire.

Tables 19, 20, and 21 present the covariates of format switching by local owners. The numbers are similar to those of national owners, so I omit the discussion.

6.3 Dynamic model: Second stage

In this subsection, I present the estimates of the structural parameters, including fixed cost, as well as persistent and variable components of acquisition cost and repositioning cost.

According to equation (5.7), the fixed cost of operating a portfolio of stations is composed of three parts: a market-level fixed-cost multiplier $\bar{F}_f^m$, a multiplier representing cost synergies of owning multiple stations in the same format $F^S(\mathcal{J}_{kf}^t|\theta^F)$, and a multiplier representing economies of scale of owning multiple stations of any format $F^E(n_k^t|\theta^E)$.

Table 22 presents the fixed cost estimates of owning a single station, averaged across formats. The level of the fixed cost varies across markets and is roughly proportional to the population. Table 24 contains the estimates of the within-market economies of scale. I find operating two stations together is 14% cheaper than operating them separately, regardless of their formats. The last column contains the estimate of the cost advantage of being a national owner, which captures cross-market cost synergies. I find national owners have a 4% lower fixed cost than local owners, but the result is not statistically significant. I also document further cost synergies of operating stations of the same format. According to Table 23, operate stations of the same format it is an additional 14% cheaper, and the discount is applied on top of the economies of scale from Table 24.

Table 25 presents the estimates of the merger cost. I find the mean acquisition cost varies by market type and company size. Moreover, I find the cost has relatively high volatility, which is homoscedastic. Note that firms obtain a new draw from the idiosyncratic component every month, and mergers are tail events; thus, the combination of a large mean and a high volatility usually leads to a low cost for realized mergers.

Repositioning cost estimates are contained in Tables 26,27, 28, and 29. I allow the repositioning costs to depend on the market category, source-target formats, and the ownership structure of the firm. I operationalize the estimation by using multiplicative market category, source-target, and national owner fixed effects. I find the repositioning cost varies considerably across the market categories. In particular I find higher switching cost in more profitable markets. Similar to the merger actions, switching is a tail event, and the estimates reveal high average switching costs with fairly high volatility over time. Additionally, the costs are statistically different depending on the source and target format, suggesting that switching is cheaper between some formats than others. These differences in switching cost are driven by the patters in the raw data. For example, Hits stations are quite profitable, but I do not observe much switching into this format, which can be rationalized by high switching costs. One can similarly explain the other switching-cost estimates.

7 Counterfactuals

Using the estimates of the structural parameters of the model, I perform several counterfactuals that study alternative merger enforcement policies.

7.1 Impact of the 1996 Telecom Act

The first set of experiments is aimed at investigating the impact of the 1996 Telecom Act on producer, listener, and advertiser surplus. Table 30 presents counterfactuals evaluating the impact of looser post-1996 local ownership caps. In particular, I recomputed the equilibrium merger and repositioning strategies with old ownership caps, and computed the relevant surpluses for 5, 10, and 20 years into the future. I use static measures of producer and consumer surplus for the following reasons: (i) the results are easily comparable with static analysis, and (ii) the static measures do not contain payoff shocks ζ, whose variance was hard to identify from the move arrival rate λ. I find radio-station owners benefit from the deregulation because, under the old caps, the producer surplus is 10% lower. Roughly 6% of this decrease comes from losing some market power and 4% comes from losing cost efficiencies. Moreover, the deregulation leads to less advertising supplied and higher per-listener prices. As a consequence, reverting to the old caps lowers listener surplus by 0.07% and increases advertiser surplus by 1.7%.

The above exercise addresses the changes in the local cap without imposing a national cap (post-1996 situation). A second exercise is aimed at partially addressing changes in the national cap by nullifying the cross-market cost benefits. In general, lack of national synergies leads to fewer mergers, which can lower the impact of the deregulation. On the other hand, each merger without the deregulation is less efficient, thereby generating smaller gains to producer surplus. The net effect is presented in the bottom three rows of Table 30. Because of fewer mergers, the negative effect of deregulation on advertiser surplus is smaller. However, at the same time, no cross-market synergies exist, so the producer surplus is smaller.

Comparing the findings from dynamic analysis with the findings obtained using a static analysis, which assumes that mergers are exogenous, is useful. In particular, Jeziorski (2012) finds that over the course of 1996-2006, ownership consolidation decreased advertisers' welfare by 21%, whereas I find the decrease to be 1.7%. The following two factors contribute to this difference: (i) a static analysis cannot account for long-run product repositioning, which could correct the negative effect of the mergers, and (ii) a static analysis assumes no mergers or product repositioning would occur without the deregulation. I examine the first effect by computing the net entry of radio stations into the format of the merger. Net entry is defined as the difference between the entry and exit arrival rates instantaneously after the merger, which accounts for churn. Recall from section 3 that the positive impact of the merger on net entry means the merger is likely to be self-correcting. To investigate this possibility, I compare pre- and post-merger net entry on Figure 3. Indeed, I find mergers lead to more entry in all 88 markets. Moreover, note the forward-looking regulator who uses pre-merger entry rates to estimate post-merger entry rates is likely to underestimate entry, which could lead to over-blocking. Instead, the regulator should use the correct entry rates predicted by the model.

7.2 Alternative merger policies

Merger enforcement based on ownership cap is rarely applied in markets other than radio. Instead, the regulator applies policies based on concentration indexes or direct measures of welfare. In the next experiment, I increase the ownership caps to seven FM stations and impose welfare criteria based on static merger simulations.

First, I evaluate the impact of increasing ownership caps to seven FM stations (subsequently CAP7) and present the results in the first three rows of Table 31. I find that in the long run, the relaxation of the caps leads to about a 4.2% increase in producer surplus. Approximately a third of this gain comes from fixed cost efficiencies, and a remaining two-thirds comes from market power. In the long run, the market power is exercised predominantly on advertisers that lose about 1% of their surplus. At the same time, in the long run, the listeners gain 0.01%. Shorter-run analysis (first 5 to 10 years) further demonstrates the tension between exercising market power on listeners versus advertisers. Namely, in the first five years after moving to CAP7, the companies exercise market power, on listeners, and in 10 to 20 years, on advertisers. The reason for this reversal is that the short-run welfare figures are driven by ownership consolidation, whereas the long-run welfare figures are driven by consolidation and post-merger product repositioning. These findings are in line with the previous literature on retrospective post-merger repositioning in the radio industry. In particular, post-merger repositioning can raise variety in this industry (see Berry and Waldfogel (1999), and Sweeting (2009)), and this extra variety can benefit listeners but hurt advertisers by thinning competition (see Jeziorski (2012)). However, prior to this paper, whether these results extend to hypothetical and out-of-sample policies such as CAP7 was unknown.

Next, I impose additional antitrust criteria and recompute the equilibrium of the dynamic game. Rows 4 to 6 of Table 31 present the results of experiments in which mergers that decrease static listener surplus are forbidden. On one hand, the policy is successful in selecting mergers that benefit listeners, raising their surplus by 0.03% in the long run, which is three times the gain from pure CAP7. On the other hand, the listener welfare criterion renders many mergers infeasible, leading to a smaller increase in producer surplus. However, because the executed mergers are in general more efficient compared to CAP7, much of the cost synergy is still realized.

Lastly, I evaluate the policy based on advertiser surplus. Contrary to the listener surplus policy, advertiser surplus policy is unsuccessful in preventing mergers that harm advertisers in the long run. Note that the welfare criterion does well in the short run, leading to about a 0.13% gain in advertiser surplus. However, post-merger repositioning reverts this trend in the long run and leads to a 0.77% loss in advertiser surplus. Moreover, contradictory to the static intuition, the static advertiser surplus criterion delivers a worse outcome for advertisers than the static listener surplus criterion, which demonstrates that myopic merger policy can be dynamically suboptimal and can have somewhat counterintuitive long-run consequences.

8 Conclusions

This paper proposes a model of industry response to different merger enforcement regimes. The regulator proposes and subsequently follows a merger enforcement policy, and companies respond to that policy via mergers and product repositioning. The merger transfer prices are endogenous and are an outcome of a dynamic bargaining process.

I estimate the model using the data on 1996-2006 consolidation wave in U.S. radio. In addition to marginal cost efficiencies identified by Jeziorski (2012), (a second station in the same format has a 10% smaller marginal cost), I find substantial fixed cost synergies from joint ownership. Namely, operating similar stations together within the local markets is cheaper than operating those stations individually. Additionally, significant economies of scale of operating multiple stations exist within markets, and statistically insignificant synergies exist across markets. Operating two stations in the same market is up to 14% cheaper than operating them separately. Moreover, owning two stations of the same format creates a further 14% fixed cost discount.

The cost synergies translate into significant incentives to merger and social benefits from mergers. I use the developed model to quantify the economic impact of these synergies. I start by computing a merger retrospective evaluating an impact of the 1996 Telecom Act. This retrospective compares the industry trajectory without the Act, which includes counterfactual mergers and product repositioning, with the factual trajectory with the Act. I find the deregulation enhanced total surplus by raising producer surplus and generating a negligible impact on listener and advertiser surplus. Small impact of the Act on advertisers contrasts with the large drop in advertiser surplus suggested by the static model and highlights the need to incorporate dynamics into the merger analysis.

Furthermore, I evaluate the counterfactual policy of using looser caps supplemented by welfare criteria. In general, I find that increasing an ownership cap to seven stations increases total surplus. For example, I demonstrate that the mergers in the radio industry are largely self-correcting because they invite a significant amount of repositioning that mitigates the market power. However, I also demonstrate that static welfare criteria are no enough to prevent losses to consumer surplus. Specifically, I show that the criterion that rejects mergers lowering static advertiser surplus does not prevent long-run losses to the advertiser surplus. Such losses to the advertisers surplus are a consequence of the fact that companies can circumvent the welfare rule by proposing the merger that exercises market power on listeners instead of advertisers and altering the product characteristics after the merger. Also, more generally, I show that one has be to cautious when using a pre-merger industry trajectory to estimate the post-merger trajectory. In particular, in the case of the radio industry, using a pre-merger trajectory would underestimate repositioning rates and could result in over-blocking mergers.

References

Aguirregabiria, V., and P. Mira (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75(1), 1–53.

Ai, C., and X. Chen (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71(6), 1795–1843.

Arcidiacono, P., P. J. Bayer, J. R. Blevins, and P. Ellickson (2010): "Estimation of Dynamic Discrete Choice Models in Continuous Time," Discussion paper.

Arcidiacono, P., and R. A. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79(6), 1823–1867.

Armstrong, M., and J. Vickers (2010): "A Model of Delegated Project Choice," *Econometrica*, 78(1), 213–244.

Bajari, P., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

Benkard, C. L., A. Bodoh-Creed, and J. Lazarev (2008): "Simulating the Dynamic Effects of Horizontal Mergers: U.S. Airlines," Discussion paper, Stanford University.

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

Berry, S. T., and J. Waldfogel (1999): "Mergers, Station Entry, and Programming Variety in Radio Broadcasting," Working Paper 7080, National Bureau of Economic Research.

Doraszelski, U., and K. L. Judd (2012): "Avoiding the curse of dimensionality in dynamic stochastic games," *Quantitative Economics*, 3(1), 53–93.

Ericson, R., and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62(1), pp. 53–82.

Farrell, J., and C. Shapiro (1990): "Horizontal Mergers: An Equilibrium Analysis," *American Economic Review*, 80(1), 107–26.

Gowrisankaran, G. (1999): "A Dynamic Model of Endogenous Horizontal Mergers," *The RAND Journal of Economics*, 30(1), 56–83.

Gowrisankaran, G., and T. J. Holmes (2004): "Mergers and the Evolution of Industry Concentration: Results from the Dominant-Firm Model," *The RAND Journal of Economics*, 35(3), 561–582.

Harford, J. (2005): "What drives merger waves?," *Journal of Financial Economics*, 77(3), 529 – 560.

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60(3), pp. 497–529.

Ivaldi, M., and F. Verboven (2005): "Quantifying the effects from horizontal mergers in European competition policy," *International Journal of Industrial Organization*, 23(910), 669 – 691, ¡ce:title¿Merger Control in International Markets¡/ce:title¿.

Jeziorski, P. (2012): "Effects of mergers in two-sided markets: Examination of the U.S. radio industry," Discussion paper, UC Berkeley.

——— (2013): "Estimation of cost synergies from mergers: Application to U.S. radio," Discussion paper, UC Berkeley.

Kamien, M. I., and I. Zang (1990): "The Limits of Monopolization Through Acquisition," *The Quarterly Journal of Economics*, 105(2), 465–499.

Kryukov, Y. (2008): "Dynamic R&D and the Effectiveness of Policy Intervention in the Pharmaceutical Industry," Discussion paper.

Lyons, B. R. (2002): "Could politicians be more right than economists?: A theory of merger standards," Discussion paper.

Magesan, A., and V. Aguirregabiria (2013): "Euler Equations for the Estimation of Dynamic Discrete Choice Structural Models," Discussion paper.

Mazzeo, M., K. Seim, and M. Varela (2012): "The Welfare consequences of mergers with product repositioning," Discussion paper.

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

Nocke, V., and M. D. Whinston (2010): "Dynamic Merger Review," *The Journal of Political Economy*, 118(6).

——— (2013): "Merger Policy with Merger Choice," *American Economic Review*.

Pakes, A., M. Ostrovsky, and S. Berry (2007): "Simple estimators for the parameters of discrete dynamic games (with entry/exit examples)," *The RAND Journal of Economics*, 38(2), 373–399.

Qiu, L. D., and W. Zhou (2007): "Merger Waves: A Model of Endogenous Mergers," *The RAND Journal of Economics*, 38(1), 214–226.

Rodrigues, V. (2001): "Endogenous mergers and market structure," *International Journal of Industrial Organization*, 19(8), 1245–1261.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1), pp. 97–109.

Rust, J. (1994): "Structural estimation of Markov decision processes," *Handbook of econometrics*, 4(4).

Salant, S. W., S. Switzer, and R. J. Reynolds (1983): "Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium," *The Quarterly Journal of Economics*, 98(2), 185–199.

Stahl, J. (2011): "A dynamic analysis of consolidation in the broadcast television industry," Working paper, Federal Reserve Board.

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

——— (2013): "Dynamic Product Positioning in Differentiated Product Industries: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," *Econometrica*, Forthcoming.

Williamson, O. E. (1968): "Economies as an Antitrust Defense: The Welfare Tradeoffs," *The American Economic Review*, 58(1), pp. 18–36.

Appendices

A Tables and Figures

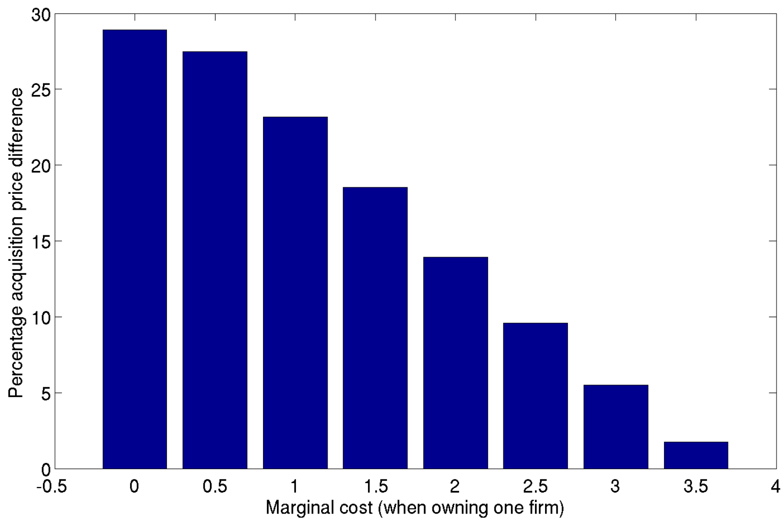

Figure 1: Percentage difference in acquisition price of the three-to-two merger, between a forward-looking and a myopic model. The x-axis represents different levels of marginal cost that determine the importance of marginal cost synergies. Note that a forward-looking acquisition price is greater in the forward-looking model for all considered values of marginal cost.

Fixed-cost synergy			Myopic impact	Long-run impact Switching cost (ρ^R) 18.500	17.500	12.000	7.000	4.000
One station	0.170	Consumer surplus	-13.6%	-9.0%	-7.3%	-4.6%	-5.1%	-4.6%
Two stations	0.170	Total surplus	-0.6%	**0.7%**	**1.2%**	**2.6%**	**2.4%**	**2.2%**
Fringe	0.170	Prob. of entry without the merger	-	0.001	0.002	0.045	0.241	0.231
		Prob. of entry after the merger	-	0.008	0.013	0.209	0.513	0.296
One station	0.170	Consumer surplus	-13.6%	-9.0%	-7.3%	-4.6%	-5.1%	-4.6%
Two stations	0.204	Total surplus	-1.5%	-0.3%	**0.3%**	**1.7%**	**1.5%**	**1.4%**
Fringe	0.170	Prob. of entry without the merger	-	0.001	0.002	0.045	0.241	0.231
		Prob. of entry after the merger	-	0.008	0.013	0.209	0.512	0.296
One station	0.170	Consumer surplus	-13.6%	-9.0%	-7.3%	-4.6%	-5.0%	-4.5%
Two stations	0.255	Total surplus	-2.9%	-1.6%	-1.0%	**0.3%**	**0.3%**	**0.1%**
Fringe	0.170	Prob. of entry without the merger	-	0.001	0.002	0.045	0.241	0.231
		Prob. of entry after the merger	-	0.008	0.013	0.209	0.495	0.296
One station	0.170	Consumer surplus	-13.6%	-9.0%	-7.3%	-4.6%	-4.8%	-4.3%
Two stations	0.306	Total surplus	-4.2%	-3.0%	-2.4%	-1.0%	-1.1%	-1.3%
Fringe	0.170	Prob. of entry without the merger	-	0.001	0.002	0.045	0.241	0.231
		Prob. of entry after the merger	-	0.008	0.013	0.209	0.326	0.308

Table 1: The table illustrates the difference between the myopic and long-run impact of a merger described in Example 1a in section 3 on total and consumer surplus. Additionally, I report pre-merger and post-merger probabilities of repositioning that increases the number of competitors in the relevant format. The relevant events, in which a myopic regulator that enforces a total-surplus criterion would block the merger but a forward-looking regulator would allow it, are underlined.

Marginal cost			Myopic impact	Long-run impact Switching cost (ρ^R) 3.000	2.500	2.000	1.500	1.000
Large owner	1.000	Consumer surplus	-14.1%	-5.1%	-4.8%	-4.5%	-4.3%	-4.3%
Fringe	1.000	Total surplus	-4.0%	-1.0%	-0.9%	-1.0%	-1.3%	-1.3%
		Prob. of entry without the merger	-	0.649	0.636	0.573	0.547	0.617
		Prob. of entry after the merger	-	0.861	0.767	0.666	0.666	0.719
Large owner	1.000	Consumer surplus	-2.9%	-0.5%	-0.0%	**0.4%**	**0.8%**	**1.2%**
Fringe	2.750	Total surplus	8.6%	8.4%	8.3%	8.1%	7.9%	7.2%
		Prob. of entry without the merger	-	0.104	0.161	0.217	0.281	0.417
		Prob. of entry after the merger	-	0.100	0.158	0.213	0.273	0.416
Large owner	1.000	Consumer surplus	-1.2%	**0.3%**	**0.8%**	**1.3%**	**1.7%**	**2.2%**
Fringe	3.000	Total surplus	10.4%	10.0%	9.9%	9.8%	9.5%	8.6%
		Prob. of entry without the merger	-	0.075	0.125	0.182	0.251	0.393
		Prob. of entry after the merger	-	0.063	0.110	0.165	0.232	0.383
Large owner	1.000	Consumer surplus	2.5%	2.5%	3.0%	3.5%	4.0%	4.5%
Fringe	3.500	Total surplus	14.0%	13.6%	13.4%	13.2%	12.8%	11.5%
		Prob. of entry without the merger	-	0.038	0.075	0.125	0.199	0.348
		Prob. of entry after the merger	-	0.025	0.054	0.097	0.166	0.327

Table 2: The table illustrates the difference between the myopic and long-run impact of a merger described in Example 1b in section 3 on total and consumer surplus. Additionally, I report pre-merger and post-merger probabilities of repositioning that increases the number of competitors in the relevant format. The relevant events, in which a myopic regulator that enforces a consumer-surplus criterion would block the merger but a forward-looking regulator would allow it, are underlined.

Marginal cost			Myopic impact	Long-run impact			
				Switching cost (ρ^R)			
				3.000	2.500	2.000	1.500
One station	1.100	Consumer surplus	3.5%	**-0.5%**	**-0.1%**	0.5%	1.3%
Two stations	0.853	Total surplus	7.5%	20.4%	19.3%	17.8%	15.8%
Fringe	1.000	Prob. of exit without the merger	-	0.001	0.002	0.045	0.241
		Prob. of exit after the merger	-	0.906	0.921	0.937	0.953
One station	1.100	Consumer surplus	2.4%	**-1.6%**	**-1.2%**	**-0.5%**	0.3%
Two stations	0.880	Total surplus	6.2%	17.9%	16.8%	15.3%	13.4%
Fringe	1.000	Prob. of exit without the merger	-	0.001	0.002	0.045	0.241
		Prob. of exit after the merger	-	0.877	0.898	0.920	0.940
One station	1.100	Consumer surplus	1.3%	**-2.6%**	**-2.2%**	**-1.5%**	**-0.7%**
Two stations	0.907	Total surplus	5.0%	15.4%	14.3%	12.9%	11.0%
Fringe	1.000	Prob. of exit without the merger	-	0.001	0.002	0.045	0.241
		Prob. of exit after the merger	-	0.841	0.870	0.899	0.925
One station	1.100	Consumer surplus	0.2%	**-3.6%**	**-3.1%**	**-2.4%**	**-1.6%**
Two stations	0.935	Total surplus	3.7%	12.8%	11.8%	10.4%	8.7%
Fringe	1.000	Prob. of exit without the merger	-	0.001	0.002	0.045	0.241
		Prob. of exit after the merger	-	0.798	0.836	0.873	0.906

Table 3: The table illustrates the difference between the myopic and long-run impact of a merger described in Example 2 in section 3 on total and consumer surplus. Additionally, I report pre-merger and post-merger probabilities of repositioning that decreases the number of competitors in the relevant format. The relevant events, in which a myopic regulator that enforces a total-surplus criterion would allow the merger but a forward-looking regulator would block it, are underlined.

Merger execution cost	Merger type	Player sophistication	Marginal cost 0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5
Low	4⇒3 merger	myopic	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		forward looking	0.044	0.386	0.904	0.987	0.998	1.000	1.000	1.000
	3⇒2 merger	myopic	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		forward looking	0.280	0.627	0.836	0.926	0.965	0.982	0.990	0.994
Medium	4⇒3 merger	myopic	0.000	0.364	1.000	1.000	1.000	1.000	1.000	1.000
		forward looking	0.000	0.018	0.243	0.844	0.974	0.995	0.999	1.000
	3⇒2 merger	myopic	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		forward looking	0.045	0.147	0.457	0.734	0.870	0.933	0.963	0.978
High	4⇒3 merger	myopic	0.000	0.000	0.000	0.891	1.000	1.000	1.000	1.000
		forward looking	0.000	0.000	0.000	0.090	0.708	0.939	0.986	0.996
	3⇒2 merger	myopic	0.000	0.184	0.997	1.000	1.000	1.000	1.000	1.000
		forward looking	0.000	0.009	0.048	0.217	0.546	0.758	0.865	0.920

Table 4: Probability of diffent types of mergers for myopic and forward-looking buyers and sellers for different levels of merger execution cost and marginal cost synergies.

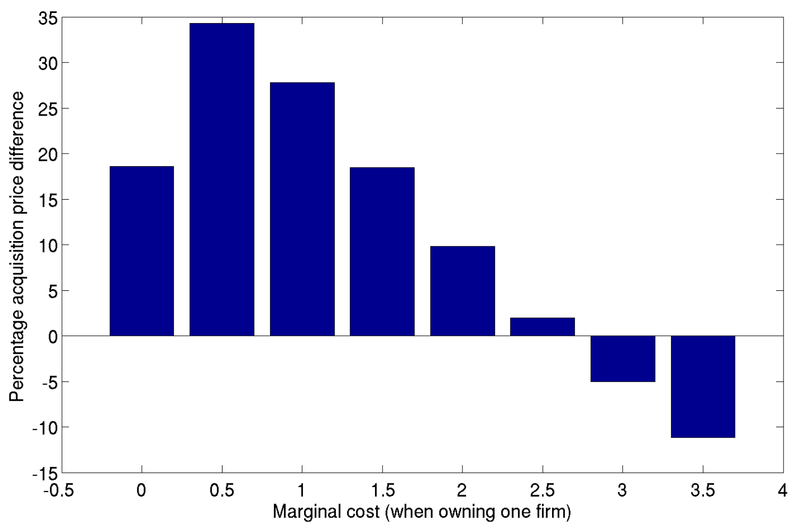

Figure 2: Percentage difference in acquisition price of the three-to-two merger, between a forward-looking and a myopic model. The x-axis represents different levels of marginal cost that determine the importance of marginal cost synergies. On the left side of the graph, the myopic price is smaller than the forward-looking price (hold out). On the right side of the graph the the myopic price is greater (speed-up).

	Mean Effects	Random Effects
Advertising	-1.226^{*} (0.727)	0.083^{*} (0.043)
AM/FM	0.689^{***} (0.135)	-
Power (kW)	0.113^{***} (0.042)	-
AC Rock Country Jazz	-3.348^{***} (0.111)	0.083^{**} (0.043)
CHR Urban Alternative	-1.745^{***} (0.166)	-0.052 (0.150)
News/Talk Religious Ethnic Others	-3.260^{***} (0.108)	0.499^{***} (0.034)

Standard errors in parentheses

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

Table 5: Estimates of the random-coefficients logit model of radio listeners' demand. The first column consists of the mean values of parameters in the utility function. The second row consists of the standard deviations of a random effect ν.

	Demographics Characteritics					
	Age	Sex	Education	Income	Black	Spanish
AC Rock Country Jazz	0.001^{***} (0.000)	-0.217^{***} (0.004)	0.271^{***} (0.001)	-0.116^{***} (0.001)	-0.496^{***} (0.004)	-1.278^{***} (0.003)
CHR Urban Alternative	-1.066^{***} (0.004)	0.540^{***} (0.006)	1.529^{***} (0.005)	-0.796^{***} (0.003)	3.367^{***} (0.012)	-0.612^{***} (0.005)
News/Talk Religious Ethnic Others	0.069^{***} (0.001)	-0.411^{***} (0.005)	0.674^{***} (0.002)	-0.086^{***} (0.001)	0.937^{***} (0.005)	0.725^{***} (0.009)

Standard errors in parentheses

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

Table 6: The table presents estimates of covariances in the random-coefficients logit model of radio listeners' demand. Each cell represents a covariance between specific demographic characteristics and listening to the particular radio =station format.

1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
0.907***	0.766***	1.194***	0.903***	1.081***	1.324***	1.005***	0.946***	1.474***	
(0.022)	(0.029)	(0.059)	(0.051)	(0.070)	(0.093)	(0.076)	(0.075)	(0.122)	

Standard errors in parentheses

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

Table 7: Estimates of utility (exponentiated) of not listening to radio. Value for 1996 is normalized to 1.

	Population <.5	Population .5M-1.5M	Population 1.5M-3.5M	Population >3.5M
OLS	-0.10^{***}	-0.04^{***}	-0.05^{***}	-0.03^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
2SLS	-0.07^{***}	-0.03^{***}	-0.03^{***}	-0.02^{***}
	(0.00)	(0.00)	(0.00)	(0.00)

Standard errors (corrected for the first stage) in parentheses

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

Table 8: The slope of advertising price per rating point (CPP). Intercept is set to 1. Units are standard deviations of quantity supplied on a station level.

	Mean level			Quality intercept		
	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M
OLS	2.32^{***}	2.16^{***}	1.22^{***}	0.22^{***}	0.16^{***}	0.08^{***}
	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)
2SLS	2.99^{***}	2.42^{***}	1.67^{***}	0.26^{***}	0.18^{***}	0.12^{***}
	(0.04)	(0.04)	(0.05)	(0.00)	(0.00)	(0.00)

Standard errors (corrected for the first stage) in parentheses

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

Table 9: The marginal cost per minute of advertising sold. The intercept of advertising price per rating point is set to 1. Note these numbers might be higher than 1 because the final price of advertising is CPP times the station rating in per cent. Units for quality are standard deviations of quality in the sample.

	Cost synergies		
	Pop. <.5	Pop. .5M-1.5M	Pop. >1.5M
OLS	-0.28^{***}	-0.02	-0.10^{***}
	(0.02)	(0.01)	(0.01)
2SLS	-0.21^{***}	0.01	-0.05^{***}
	(0.02)	(0.01)	(0.01)

Standard errors (corrected for the first stage) in parentheses

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

Table 10: Marginal cost synergies from owning multiple stations of the same format.

	Adult Music	Hits Music	Non-Music
Dummy	-3.633 (1.046)	-3.609 (0.886)	-3.743 (1.747)
Age	1.963 (4.247)	0.659 (3.615)	-1.125 (3.069)
Education	0.906 (1.992)	-0.108 (1.697)	-0.539 (1.289)
Income	-0.696 (0.753)	-0.814 (0.662)	-0.623 (0.567)
Black	-0.205 (0.800)	1.475 (0.638)	0.507 (0.711)
Hispanic	-0.496 (0.786)	-0.817 (0.703)	0.608 (1.849)

Table 11: Acquisition CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages; details in the appendix.

$\eta_{f,k}$	$\eta^2_{f,k}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}\right)^2$	$\sum_{k' \in \mathbf{K^L}} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^L}} \eta_{k',f}\right)^2$
2.447 (1.530)	-0.083 (17.924)	0.755 (9.757)	-0.513 (3.842)	-1.321 (3.524)	-0.515 (5.078)	0.076 (3.854)

Table 12: Acquisition CCP: Coefficients on the covariates related to the target acquisition format.

$\sum_{f' \neq f} \eta_{f',k}$	$\sum_{f' \neq f} \eta^2_{f',k}$
$-$	$-$
5.116	-1.224
(9.876)	(8.252)
$\left(\sum_{f' \neq f} \eta_{f',k}\right)^2$	$\sum_{k' \in \mathbf{K^N} \setminus k, f' \neq f} \eta_{k',f'}$
-0.353	-0.131
(5.247)	(4.912)
$\left(\sum_{k' \in \mathbf{K^N} \setminus k, f' \neq f} \eta_{k',f'}\right)^2$	$\sum_{k' \in \mathbf{K^N} \setminus k} \left(\sum_{f' \neq f} \eta_{k',f'}\right)^2$
-1.999	4.618
(3.845)	(0.940)
$\sum_{f' \neq f} \left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f'}\right)^2$	
-4.306	
(2.009)	

Table 13: Acquisition CCP: Coefficients on the covariates related to other-than-acquiree's formats.

One from the cap	Two from the cap
-0.497	0.268
(0.139)	(0.121)

Table 14: Acquisition CCP: Dummies for the closeness to the ownership cap.

	To: Adult Music	To: Hits Music	To: Non-Music
From: Adult Music	-	-6.130 (1.017)	-4.945 (0.756)
From: Hits Music	-4.118 (0.922)	-	-5.374 (0.752)
From: Non-Music	-3.922 (0.972)	-6.634 (0.782)	-
Age	-0.367 (3.769)	-2.420 (2.955)	-2.072 (3.127)
Education	-1.207 (1.140)	0.145 (1.966)	-1.172 (1.260)
Income	0.136 (0.705)	0.574 (0.572)	0.445 (0.659)
Black	-0.958 (0.489)	2.755 (0.801)	0.725 (0.545)
Hispanic	-0.845 (0.697)	1.017 (0.537)	1.803 (1.195)

Table 15: National owner repositioning CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages; details in the appendix.

$\eta_{f,k}$	$\eta^2_{f,k}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f} \right)^2$	$\sum_{k' \in \mathbf{K^L}} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^L}} \eta_{k',f} \right)^2$
6.653 (1.117)	-3.300 (3.828)	-1.315 (0.987)	0.152 (9.708)	0.832 (4.574)	0.760 (2.338)	-1.276 (1.577)

Table 16: National owner repositioning CCP: Coefficients on the covariates related to the current format.

$\eta_{f,k}$	$\eta^2_{f,k}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f} \right)^2$	$\sum_{k' \in \mathbf{K^L}} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^L}} \eta_{k',f} \right)^2$
-1.502 (4.381)	-0.281 (3.417)	-1.793 (6.854)	-0.994 (6.511)	-1.448 (4.874)	-0.807 (0.850)	-1.630 (1.596)

Table 17: National owner repositioning CCP: Coefficients on the covariates related to the target format.

At the cap	One from the cap
0.593 (0.094)	0.436 (0.457)

Table 18: National owner repositioning CCP: Dummies for the closeness to the ownership cap.

	To: Adult Music	To: Hits Music	To: Non-Music
From: Adult Music	-	-6.681 (0.446)	-4.683 (0.545)
From: Hits Music	-3.653 (0.373)	-	-4.769 (0.542)
From: Non-Music	-4.373 (0.380)	-7.554 (1.390)	-
Age	-2.754 (1.403)	1.374 (2.437)	-3.738 (1.362)
Education	1.415 (1.212)	1.175 (0.886)	1.258 (0.646)
Income	-0.333 (0.240)	0.199 (0.445)	-0.134 (0.232)
Black	-0.709 (0.513)	2.519 (0.345)	0.880 (0.406)
Hispanic	-0.865 (0.236)	0.583 (0.424)	1.685 (0.237)

Table 19: Local owner repositioning CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages; details in the appendix.

$\sum_{k' \in \mathbf{K^N}} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N}} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N}} \eta_{k',f}\right)^2$	$\sum_{k' \in \mathbf{K^L} \setminus k} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}\right)^2$
0.514 (3.480)	-1.768 (3.269)	0.245 (1.866)	2.473 (0.555)	-3.786 (1.048)

Table 20: Local owner repositioning CCP: Coefficients on the covariates related to the current format.

$\sum_{k' \in \mathbf{K^N}} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N}} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N}} \eta_{k',f}\right)^2$	$\sum_{k' \in \mathbf{K^L} \setminus k} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}\right)^2$
-0.647 (0.604)	-2.380 (4.875)	-3.307 (1.997)	4.275 (1.412)	-7.163 (0.868)

Table 21: Local owner repositioning CCP: Coefficients on the covariates related to the target format.

Name	Pop. 2007	Intercept	Name	Pop. 2007	Intercept
Los Angeles, CA	13155.1	0.2324 (0.02413)	Omaha-Council Bluffs, NE-IA	740.3	0.0255 (0.00265)
Chicago, IL	9341.4	0.1138 (0.01182)	Knoxville, TN	737.4	0.0153 (0.00158)
Dallas-Ft. Worth, TX	5846.9	0.0924 (0.00959)	El Paso, TX	728.2	0.0797 (0.00827)
Houston-Galveston, TX	5278.5	0.0695 (0.00721)	Harrisburg-Lebanon-Carlisle, PA	649.4	0.0343 (0.00356)
Atlanta, GA	4709.7	0.0512 (0.00531)	Little Rock, AR	618.7	0.0074 (0.00077)
Boston, MA	4531.8	0.0941 (0.00977)	Springfield, MA	618.1	0.0098 (0.00102)
Miami-Ft. Lauderdale-Hollywood, FL	4174.2	0.1223 (0.01270)	Charleston, SC	597.7	0.0071 (0.00074)
Seattle-Tacoma, WA	3775.5	0.1137 (0.01181)	Columbia, SC	576.6	0.0105 (0.00109)
Phoenix, AZ	3638.1	0.0577 (0.00599)	Des Moines, IA	576.5	0.0046 (0.00048)
Minneapolis-St. Paul, MN	3155	0.0692 (0.00718)	Spokane, WA	569.1	0.0123 (0.00128)
St. Louis, MO	2688.5	0.0214 (0.00222)	Wichita, KS	563.9	0.0144 (0.00150)
Tampa-St. Petersburg-Clearwater, FL	2649.1	0.0768 (0.00797)	Madison, WI	539.5	0.0237 (0.00247)
Denver-Boulder, CO	2603.5	0.0686 (0.00712)	Ft. Wayne, IN	520	0.0077 (0.00080)
Portland, OR	2352.2	0.1153 (0.01197)	Boise, ID	509.9	0.0240 (0.00249)
Cleveland, OH	2133.8	0.0504 (0.00523)	Lexington-Fayette, KY	509	0.0050 (0.00052)
Charlotte-Gastonia-Rock Hill, NC-SC	2126.7	0.0279 (0.00289)	Augusta, GA	498.4	0.0024 (0.00025)
Sacramento, CA	2099.6	0.0415 (0.00431)	Chattanooga, TN	494.5	0.0077 (0.00080)
Salt Lake City-Ogden-Provo, UT	1924.1	0.0269 (0.00279)	Roanoke-Lynchburg, VA	470.7	0.0038 (0.00039)
San Antonio, TX	1900.4	0.0540 (0.00560)	Jackson, MS	468.6	0.0011 (0.00011)
Kansas City, MO-KS	1870.8	0.0432 (0.00448)	Reno, NV	452.7	0.0155 (0.00161)
Las Vegas, NV	1752.4	0.0710 (0.00737)	Fayetteville, NC	438.9	0.0060 (0.00063)
Milwaukee-Racine, WI	1712.5	0.0217 (0.00225)	Shreveport, LA	399.6	0.0018 (0.00019)
Orlando, FL	1686.1	0.0537 (0.00558)	Quad Cities, IA-IL	358.8	0.0115 (0.00119)
Columbus, OH	1685	0.0119 (0.00123)	Macon, GA	337.1	0.0022 (0.00023)
Indianapolis, IN	1601.6	0.0184 (0.00191)	Eugene-Springfield, OR	336.4	0.0137 (0.00142)
Norfolk-Virginia Beach-Newport News, VA	1582.8	0.0173 (0.00179)	Portland, ME	276.1	0.0112 (0.00116)
Austin, TX	1466.3	0.0812 (0.00842)	South Bend, IN	267	0.0226 (0.00234)
Nashville, TN	1341.7	0.0488 (0.00506)	Lubbock, TX	255.3	0.0271 (0.00281)
Greensboro-Winston Salem-High Point, NC	1328.9	0.0185 (0.00193)	Binghamton, NY	247.9	0.0041 (0.00043)
New Orleans, LA	1293.7	0.0195 (0.00202)	Odessa-Midland, TX	247.8	0.0040 (0.00042)
Memphis, TN	1278	0.0045 (0.00047)	Yakima, WA	231.4	0.0099 (0.00103)
Jacksonville, FL	1270.5	0.0112 (0.00116)	Duluth-Superior, MN-WI	200.3	0.0123 (0.00127)
Oklahoma City, OK	1268.3	0.0119 (0.00123)	Medford-Ashland, OR	196.2	0.0076 (0.00079)
Buffalo-Niagara Falls, NY	1150	0.0401 (0.00417)	St. Cloud, MN	191.2	0.0100 (0.00104)
Louisville, KY	1099.6	0.0311 (0.00322)	Fargo-Moorhead, ND-MN	183.6	0.0150 (0.00155)
Richmond, VA	1066.4	0.0082 (0.00085)	Abilene, TX	159.1	0.0059 (0.00061)
Birmingham, AL	1030	0.0104 (0.00108)	Eau Claire, WI	156.5	0.0061 (0.00063)
Tucson, AZ	938.3	0.0317 (0.00329)	Monroe, LA	149.2	0.0054 (0.00056)
Honolulu, HI	909.4	0.0311 (0.00323)	Parkersburg-Marietta, WV-OH	149.2	0.0049 (0.00051)
Albany-Schenectady-Troy, NY	902	0.0323 (0.00335)	Grand Junction, CO	130	0.0091 (0.00094)
Tulsa, OK	870.2	0.0137 (0.00142)	Sioux City, IA	123.7	0.0119 (0.00123)
Ft. Myers-Naples-Marco Island, FL	864.1	0.0712 (0.00739)	Williamsport, PA	118.3	0.0036 (0.00037)
Grand Rapids, MI	856.4	0.0124 (0.00129)	San Angelo, TX	103.8	0.0057 (0.00059)
Albuquerque, NM	784.9	0.0614 (0.00638)	Bismarck, ND	99.2	0.0024 (0.00025)
Omaha-Council Bluffs, NE-IA	740.3	0.0255 (0.00265)			()

Standard errors (corrected for the first stage) in parentheses
*** p¡0.01, ** p¡0.05, * p¡0.1

Table 22: Fixed cost of owning one station in each market.

Number of stations owned in the format in the local market	1	2	3	4	5
Fixed cost discount	1.000 (–)	0.862*** (0.034)	0.790*** (0.064)	0.743*** (0.058)	0.708*** (0.049)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, one-tail test

Table 23: Fixed cost: Table contains estimates of discounts to fixed cost resulting from local cost synergies from owning multiple stations in the same format.

Number of stations owned local market	1	2	3	4	5	National
Fixed cost discount	1.000 (–)	0.863** (0.063)	0.791** (0.120)	0.744*** (0.109)	0.709*** (0.092)	0.963 (0.178)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, one-tail test

Table 24: Fixed cost: Table contains estimates of within- and cross-market economies of scale. The number reflects the per-station discount.

	Mean		Standard deviation	
	Intercept	Variable profits	Intercept	Variable profits
Category 1	7.203*** (1.374)	2.653*** (0.653)	2.039*** (0.290)	0.086 (0.091)
Category 2	3.917*** (0.720)		1.030*** (0.151)	
Category 3	3.724*** (0.793)		0.950*** (0.177)	
Category 4	2.061*** (0.424)		0.510*** (0.095)	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, two-tail test

Table 25: Estimates of the acquisition cost. The table contains an intercept of the mean and standard deviation of the acqusition distribution. It includes market-size fixed effects (relative to the smallest category) and a coefficient on the static variable profits of an acquistion target.

Owner	Source format	Mean: Target format: Adult Music	Mean: Target format: Hits Music	Mean: Target format: Non-Music	Mean: Variable profit	Std. deviation: Intercept	Std. deviation: Variable profit
National	Adult Music	-	18.799*** (1.863)	18.088*** (1.732)	2.084 (1.842)	3.654*** (0.330)	0.069 (0.329)
	Hits Music	14.723*** (1.419)	-	17.757*** (1.732)			
	Non-Music	16.125*** (1.568)	21.194*** (2.099)	-			
Local	Adult Music	-	18.828*** (1.819)	13.665*** (1.306)			
	Hits Music	10.791*** (1.121)	-	10.443*** (1.104)			
	Non-Music	17.160*** (1.606)	22.076*** (2.115)	-			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, two-tail test

Table 26: Market category 1: Estimates of the format-switching cost. The table contains to-from format fixed effects for local and national owners.

Owner	Source format	Mean: Target format: Adult Music	Mean: Target format: Hits Music	Mean: Target format: Non-Music	Mean: Variable profit	Std. deviation: Intercept	Std. deviation: Variable profit
National	Adult Music	-	8.873*** (0.999)	8.538*** (0.922)	2.084 (1.842)	1.713*** (0.180)	0.069 (0.329)
	Hits Music	6.949*** (0.773)	-	8.381*** (0.926)			
	Non-Music	7.611*** (0.858)	10.004*** (1.145)	-			
Local	Adult Music	-	8.887*** (0.963)	6.450*** (0.701)			
	Hits Music	5.093*** (0.616)	-	4.929*** (0.593)			
	Non-Music	8.100*** (0.876)	10.420*** (1.129)	-			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, two-tail test

Table 27: Market category 2: Estimates of the format-switching cost. The table contains to-from format fixed effects for local and national owners.

Owner	Source format	Mean				Std. deviation	
		Target format			Variable profit	Intercept	Variable profit
		Adult Music	Hits Music	Non-Music			
National	Adult Music	-	6.531*** (0.892)	6.284*** (0.838)	2.084 (1.842)	1.226*** (0.163)	0.069 (0.329)
	Hits Music	5.114*** (0.711)	-	6.169*** (0.831)			
	Non-Music	5.602*** (0.775)	7.363*** (1.036)	-			
Local	Adult Music	-	6.541*** (0.881)	4.747*** (0.641)			
	Hits Music	3.749*** (0.537)	-	3.628*** (0.525)			
	Non-Music	5.961*** (0.797)	7.669*** (1.033)	-			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, two-tail test

Table 28: Market category 3: Estimates of the format-switching cost. The table contains to-from format fixed effects for local and national owners.

Owner	Source format	Mean				Std. deviation	
		Target format			Variable profit	Intercept	Variable profit
		Adult Music	Hits Music	Non-Music			
National	Adult Music	-	2.354*** (0.232)	2.265*** (0.220)	2.084 (1.842)	0.435*** (0.042)	0.069 (0.329)
	Hits Music	1.844*** (0.195)	-	2.224*** (0.227)			
	Non-Music	2.019*** (0.208)	2.654*** (0.266)	-			
Local	Adult Music	-	2.358*** (0.231)	1.711*** (0.169)			
	Hits Music	1.351*** (0.153)	-	1.308*** (0.151)			
	Non-Music	2.149*** (0.211)	2.764*** (0.270)	-			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1, two-tail test

Table 29: Market category 4: Estimates of the format-switching cost. The table contains to-from format fixed effects for local and national owners.

Counterfactual regime		Total producer surplus	Variable profits	Fixed cost	Listener surplus	Advertiser surplus
Pre-1996 local caps	5 years	-4.44 (2.30%)	-2.62 (1.36%)	1.82 (0.94%)	0.00 (0.00%)	-0.89 (0.21%)
Pre-1996 local caps	10 years	-9.63 (4.87%)	-5.67 (2.87%)	3.97 (2.00%)	-0.04 (0.01%)	-0.63 (0.15%)
Pre-1996 local caps	20 years	-20.95 (10.13%)	-12.38 (5.99%)	8.57 (4.15%)	-0.21 (0.07%)	7.28 (1.73%)
Pre-1996 local caps No cross-market ownership	5 years	-4.32 (2.24%)	-2.62 (1.36%)	1.70 (0.88%)	-0.02 (0.01%)	-0.75 (0.18%)
Pre-1996 local caps No cross-market ownership	10 years	-9.37 (4.74%)	-5.62 (2.84%)	3.75 (1.89%)	-0.08 (0.03%)	-0.44 (0.10%)
Pre-1996 local caps No cross-market ownership	20 years	-20.41 (9.88%)	-12.24 (5.92%)	8.18 (3.96%)	-0.27 (0.09%)	7.37 (1.75%)

Table 30: Impact of different enforcement regimes on producer, listener, and advertiser surplus. The table reports differences between simulated future states using the equilibrium merger and repositioning strategies at the counterfactual regime and the observed regime. Hence, a positive number means the counterfactual regime yields a higher value.

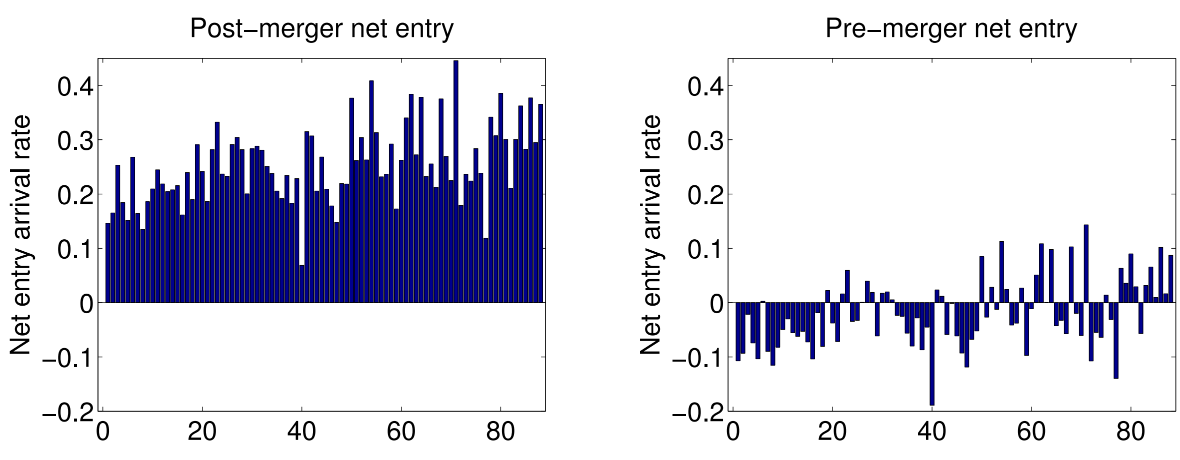

Figure 3: Post- and pre-merger net entry rates by market. Markets are sorted by population, from the smallest (Bismark) to the largest (Los Angeles).

Counterfactual regime		Total producer surplus	Variable profits	Fixed cost	Listener surplus	Advertiser surplus
Cap of 7	5 years	2.21 (1.14%)	1.66 (0.86%)	-0.55 (0.28%)	-0.03 (0.01%)	0.91 (0.22%)
Cap of 7	10 years	4.44 (2.24%)	3.18 (1.61%)	-1.26 (0.64%)	-0.00 (0.00%)	-1.20 (0.28%)
Cap of 7	20 years	8.59 (4.16%)	5.90 (2.86%)	-2.69 (1.30%)	0.03 (0.01%)	-3.71 (0.88%)
Cap of 7 Myopic Listener Surplus Criterion	5 years	0.59 (0.30%)	0.55 (0.28%)	-0.04 (0.02%)	0.05 (0.02%)	0.42 (0.10%)
Cap of 7 Myopic Listener Surplus Criterion	10 years	1.70 (0.86%)	1.20 (0.61%)	-0.50 (0.25%)	0.07 (0.02%)	-0.51 (0.12%)
Cap of 7 Myopic Listener Surplus Criterion	20 years	4.65 (2.25%)	2.80 (1.36%)	-1.85 (0.89%)	0.08 (0.03%)	-2.58 (0.61%)
Cap of 7 Myopic Advertiser Surplus Criterion	5 years	1.72 (0.89%)	1.16 (0.60%)	-0.56 (0.29%)	-0.07 (0.02%)	0.53 (0.13%)
Cap of 7 Myopic Advertiser Surplus Criterion	10 years	3.58 (1.81%)	2.25 (1.14%)	-1.33 (0.67%)	-0.10 (0.03%)	0.44 (0.10%)
Cap of 7 Myopic Advertiser Surplus Criterion	20 years	6.76 (3.27%)	4.12 (1.99%)	-2.64 (1.28%)	-0.05 (0.02%)	-3.25 (0.77%)

Table 31: Impact of increasing FM local ownership cap to 7 stations on producer, listener, and advertiser surplus. The table reports differences between simulated future states using the equilibrium merger and repositioning strategies at the counterfactual regime and the observed regime. Hence a positive number means the counterfactual regime yields a higher value.

B Static payoffs

This appendix section contains a discussion of the data and estimation procedure that is used to obtain static profit function $\pi(\cdot)$.

B.1 Static data

The data come from four main sources: two consulting companies, BIA Inc and SQAD, a Common Population Survey, and Radio Today publications by Arbitron. BIA provides two comprehensive data sets on the vast majority of U.S. radio broadcasting. The first data set covers years 1996-2001 and the second, 2002-2006. I combined the data to form a large panel for 1996-2006. SQAD provides a data set on average prices per rating point (CPP) for each market and half of a year, grouped by demographics and time of the day. Unfortunately, it does not provide data on station-level per-listener pricing. However, because the pricing is done on the per-listener basis, one can still compute a station-level price of an advertising slot by multiplying the CPP by the station rating. According to the anecdotal evidence, many advertisers follow this procedure to figure out the prices they are likely to pay. This procedure does not account for the fact that stations might have different listenership pools and therefore CPP for different stations might vary. I alleviate that concern by computing a proxy for a station-level CPP. I take a weighted average of prices by demographics and time of the day, where weights are relevant ratings of the station. By doing so, I assume stations that have most of their listenership at a particular time of the day also set a price that is closer to the average price at that time of the day. Although this estimate of station-level prices is not perfect, it produces a considerable amount of variation within market. I subsequently use these price proxies to compute station-level advertising quantities by dividing estimates of station revenues (provided by BIA) by a product of prices and ratings. Note that ad quantity computed in such a way might carry some measurement error, because it is a function of two estimates. However, if this measurement error is not endogenous within markets – for example, if it only introduces error to an overall level of advertising in each market – it would not affect the results.

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

B.2 Static estimation

The following section is a parsimonious description of the estimation procedure I use to recover the parameters of the static model (for the full description see Jeziorski (2012)). I conduct the estimation of the model in two steps. In the first step, I estimate the demand model that includes parameters of the consumer utility θ^L (see equation (5.2)). In the second step, I recover parameters of the inverse demand for advertising θ^A, $w_{jj'}$ (see equation (5.4)) and marginal cost parameters θ^C (see equation (5.5))

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

$$E[\xi_{jt}|Z_1, \theta^L] = 0, \tag{B.1}$$

This moment condition follows Berry, Levinsohn, and Pakes (1995). I use instruments for advertising quantities because these quantities are likely to be correlated with unobserved station quality. My instruments include lagged mean and second central moment of competitors' advertising quantity, lagged market HHIs and lagged number and cumulative market share of other stations in the same format. These instruments are valid under the following assumptions: (i) ξ_t is independent across time and radio stations, and (ii) decisions about portfolio selection are made before decisions about advertising.

A second set of moment conditions is based on demographic listenership data. Namely, I equate a national share R_{fc} of format f among listeners possessing certain demographic characteristics c to it's predicted empirical counterpart $\hat{R}_{fc}$. Formally, I use an unconditional moment $E[\hat{R}_{fc} - R_{fc}|\theta^L] = 0$. I obtain the conditional empirical moments $\hat{R}_{fc}$ by drawing listeners of characteristic c from the conditional national empirical distribution (based on Common Population Survey) and averaging their format choice probabilities implied by the model.

The second stage of the estimation obtains the competition matrix Ω, the parameters of demand for advertising θ^A, and marginal cost θ^C. The elements of the matrix Ω are postulated to take the following form:

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

where $r_{f|a}$ is a nationally aggregated probability that the advertiser of type a chooses format f ($r_{a|f}$ can be obtained by Bayes' separately for each market, knowing the market proportion of types).

The estimator is based on the following supply conditions:

$$\begin{aligned} r_{jt} + \sum_{j'\in s_{kt}} q_{j't}\frac{\partial r_{j't}(q_t)}{\partial q_{jt}} = & \\ & \theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Am}\left[r_{jt}v_j + \sum_{j'\in s_{kt}}\left(r_{j't}(q_t)\omega_{jj'}^m + v_{j'}\frac{\partial r_{j't}(q_t)}{\partial q_{jt}}\right)\right] + \theta_2^{Cm}\xi_{jt} + \theta_3^{Cm}\mathrm{SYN}_{jt} + \eta_{jt}, \end{aligned} \tag{B.2}$$

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

Because the equation does not depend on θ_1^A, I can use it to estimate θ_2^A and θ^C. Two sources of heterogeneity in marginal cost and slope coefficients exist across markets. Effective marginal cost parameters for each station in market m are given by $\theta_1^{Am}\theta^{Cm}$, and θ_1^{Am} is allowed to be different across markets. Moreover, to control for potential heterogeneity that is not captured by a level of revenues, I allow for three different sets of values of all parameters in θ^{Cm}: for small (up to 500 people), medium (between 500 and 1500), and large (more than 1500) markets. To avoid having a full set of dummies and to facilitate identification, I set time dummies for years 1996 and 1997 to zero. Similar specification is true for the slope of the inverse demand for ads and its effective slope is given by $\theta_1^{Am}\theta_2^{Am}$. To control for the fact that stations might have different market power in the advertising market depending on its size, I allow for four different values for the slope of inverse demand, depending on the population of the market (up to 500 people, between 500 and 1500, between 1500 and 4500, and more than 4500). Given the estimates of θ_2^{Am} and θ^C, I can back out θ_1^{Am} by equating the observed average revenue in each market with its predicted counterpart. To control for the fact that ratings depend on quantity, which is likely to be correlated with η, I estimate the model with two-stage least squares using the following instruments: number of stations in the same format and ad quantities of competitors. Additionally, the instruments were lagged one period to control for potential serial correlation in η.

C Value function simulation details

The value function at $\mathcal{J}^s$ can be decomposed into four components according to

$$V_k = V^{(\pi)} + V^{(P)} + V^{(F)} + V^{(A)} + V^{(R)},$$

where

$$\begin{aligned} V_k^{(\pi)} &= \int_{s=t}^{\infty} e^{-\rho s}\pi_k(\mathcal{J}^s)ds, \\ V_k^{(F)} &= -\int_{s=t}^{\infty} e^{-\rho s}F_k(\mathcal{J}^s|\theta)ds, \\ V_k^{(A)} &= \sum_{l=1}^{\infty} e^{-\rho\tau_k^{A,(l)}} W^A_{a_k^{(l)}}(\mathrm{CCP}_k^A, \mathcal{J}^{\tau_k^{A,(l)}}|\theta), \\ V_K^{(R)} &= \sum_{m=1}^{\infty} e^{-\rho\tau_k^{R,(m)}} W^R_{r_k^{(l)}}(\mathrm{CCP}_k^R, \mathcal{J}^{\tau_k^{R,(m)}}|\theta), \\ V_k^{(P)} &= \sum_{l=1}^{\infty} e^{-\rho\tau_k^{A,(l)}} - P(a_k^{(l)}, \mathcal{J}^{\tau_k^{A,(l)}}|\theta). \end{aligned}$$

Each of these components can be expressed as a linear function of parameters θ and sufficient statistics about the simulated industry paths $\hat{\mathcal{J}}^{s,r}$ for $r = 1, \ldots, 1000$. I discuss all components below.

The first component $V_k^{(\pi)}$ does not depend on dynamic parameters, so the sufficient statistic is just an average of all draws. The second component $V_k^{(F)}$ is a discounted sum of fixed costs. The sufficient statistic to compute this cost is a matrix

$$\mathrm{SIM}^F(f, x, y) = \int_{s=t}^{\infty} e^{-\rho s} \mathbf{1}(\omega_{kf}^s = x, n_k^s = y) ds.$$

Fixed cost can be obtained using

$$V_k^{(F)} = \sum_{f=1}^{F} \bar{F}_f^m \sum_{x,y} \mathrm{SIM}^F(f, x, y) F^S(x|\theta^F) F^E(y|\theta^E).$$

The third component consist of sum of discounted acquisition shocks. It can be decomposed into

$$V_k^{(A)} = \theta^{A,m} \sum_{l=1}^{\infty} e^{-\rho \tau_k^{A,(l)}} + \theta_\pi^A \sum_{l=1}^{\infty} e^{-\rho \tau_k^{A,(l)}} \pi_k(\mathcal{J}^{\tau_k^{A,(l)}}) + \\ \theta_\sigma^{A,m} \sum_{l=1}^{\infty} e^{-\rho \tau_k^{A,(l)}} E[\epsilon^A(k')|a_k^{(l)}] + \theta_{\sigma,\pi}^{A,m} \sum_{l=1}^{\infty} e^{-\rho \tau_k^{A,(l)}} \pi_k(\mathcal{J}^{\tau_k^{A,(l)}}) E[\epsilon^A(k')|a_k^{(l)}].$$

In such a case, I need four sufficient statistics to evaluate this part of the value function. One can similarly decompose the fourth component and obtain nine sufficient statistic. An extra five statics come from the fact that I allow six different means of repositioning cost depending on the source and target format.

The last component is the sum of discounted acquisition spending. To obtain it, I use the fact that the acquisition price $P(a_k^{(l)}|\mathcal{J}^t, \theta)$ is equal to the value function of the acquiree conditional on rejecting every equilibrium merger offer. Thus, obtaining an acquisition price is the same as simulating a value function for the fringe firm. Such a value function contains three of the above terms, namely, $V_k^{(\pi)}$, $V_k^{(F)}$, and $V_k^{(R)}$, which are simulated using the aforementioned sufficient statistics in the nested loop.