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

Przemysław Jeziorski†

May 4, 2014

Abstract

The paper introduces a method to conduct a forward-looking antitrust review of horizontal mergers. The method utilizes a dynamic oligopoly model in which mergers, entry/exit and product repositioning are endogenous. The model provides long-run industry trajectories with and without the merger under review, which enables the regulator to obtain dynamically robust welfare comparison. The paper demonstrates the application of the framework to regulate U.S. radio broadcasting industry. In particular, it investigates long-run efficacy of two commonly used merger heuristics: radio station ownership caps (see Telecom Act (1996)) and static merger simulations (see Nevo (2000)). The paper finds that raising the ownership cap results in higher total welfare, and demonstrates that myopic merger simulations may be ineffective in preventing the losses to consumer welfare.

*I would like to thank: Lanier Benkard, Paul Ellickson, Teck Ho, Carl Shapiro, Andrew Sweeting, Miguel Villas-Boas, as well as participants of the seminars at Stanford, NYU, U Chicago, UCLA, U of Michigan, U of Toronto, U of Virginia, and numerous conferences.

†Haas School of Business, UC Berkeley

1 Introduction

Enforcement of horizontal merger policy involves assessing the trade-off between market power and cost synergies resulting from mergers (see Williamson (1968)). Traditionally, the economic literature on antitrust policy studies this trade-off in a static environment, assuming that the set of competing firms and their product characteristics, other than prices, remain perpetually constant (see structure-conduct-performance, as in Bain (1968), equilibrium analysis, as in Farrell and Shapiro (1990), and more recently, merger simulations, as in Nevo (2000)). In practice, the above assumptions rarely hold, as markets experience merger waves as well as entry, exit and product repositioning. In case these assumptions are indeed not satisfied, the static merger enforcement heuristics may lead to inefficient or erroneous regulation. The inefficiency is hard to quantify theoretically and can manifest itself as Type I error, in which the regulator rejects the merger that should be approved, or as Type II error, in which the regulator approves the merger that should be rejected.

An instance of Type I error (merger overblocking) is recognized in the Horizontal Merger Guidelines published by the United States Department of Justice. Specifically, the competitive pressure generated by an effective entry of new competitors or product repositioning by existing ones can mitigate the effects of the merger by lowering the post-merger market power to the pre-merger level. In such cases, a myopic enforcement agency may overestimate market power and reject an efficient merger that should be approved. In more subtle cases, myopic regulation can lead to Type II error (merger underblocking). For example, an efficiency enhancing merger can lead to future monopolization by forcing less efficient competitors to exit or reposition. A myopic regulator who does not account for such exit will underestimate the post-merger market power and may approve a merger that should be rejected. Another source of Type II error is a selection bias that can arise if forward-looking companies attempt to circumvent the myopic regulations by proposing mergers strategically. For example, companies may attempt a merger which leads to excessive market power only if followed by the repositioning of consolidated products. In such case a myopic regulator, who presumes that product characteristics are perpetually constant, would allow a merger that should be blocked.

The above examples suggest that in order to make a robust decision, the regulator should employ a forward-looking approach and compare the trajectory of the industry in the presence of the merger in question to the counterfactual trajectory in the absence of that merger. Computing such industry trajectories requires solving and estimating a forward-looking model in which mergers, entry/exit and product characteristics are endogenous. Such approach is scarce in the literature because dynamic models of endogenous mergers and product characteristics pose theoretical and econometric challenges. In the general dynamic model of mergers, profit-maximizing players consider every feasible merger possibility, as well as predict which mergers will be executed by competitors in the future. However, the number of merger possibilities, among even a handful of active players, grows exponentially, making the computation of Nash equilibrium extremely burdensome. The model becomes even more complex when allowing for differentiated products, as it needs to keep track of the product portfolios of all active players. Past attempts to solve the aforementioned issues resulted in complicated frameworks that are impractical for estimation and counterfactual predictions (see Gowrisankaran (1999), Stahl (2011) and Jeziorski (2014b)). Thus, designing a simple, but empirically relevant, endogenous merger model seems necessary.

This paper proposes a general empirical framework that can be used to conduct a dynamically robust antitrust merger review. The framework is easily estimable using commonly obtainable panel data on mergers, and the associated prices, market shares and product characteristics. It is also easily computable which enables experiments evaluating counterfactual merger enforcement policies. The key component of the proposed methodology is a dynamic model with endogenous mergers, entry/exit and product repositioning. The model is an extension of static merger simulations and aims at three objectives: (i) allow the regulator to be forward looking by providing industry trajectories with and without the merger in question, (ii) minimize Type I and Type II error by allowing the set of competitors and characteristics of all products to change as a result of the merger, and (iii) limit the Type II error from merger selection bias by allowing the attempted mergers to be endogenous.

The model is a natural continuous time generalization of Rubinstein's 1982 bargaining model (for earlier applications of continuous time to dynamic oligopoly, see Kryukov (2008), Arcidiacono et al. (2010), Doraszelski and Judd (2012)). The pivotal feature of reframing the bargaining process in continuous time is that the probability that more than one company would attempt a merger or repositioning at the same instant is negligible. Such simplification dramatically lowers the conceptual and computational burden of predicting which mergers and repositioning events occur in the equilibrium. Consequently, it enables a modeling compromise that maintains the necessary economic complexity that is required to conduct a robust dynamic merger review, while alleviating the computational obstacles.

I apply the proposed model to analyze retrospective and hypothetical antitrust policy changes in the U.S. radio broadcasting industry and demonstrate inefficiencies in myopic regulation. The radio industry presents a natural experiment for studying the consequences of changes in antitrust policy. The 1996 Telecom Act doubled local radio station ownership caps and abolished the national ownership cap resulting in the deregulation of the industry. These changes facilitated over 6,000 acquisitions within the period of 1996-2006, constituting a major structural change from a fragmented into a consolidated market (see Leeper (1999)). The Act aimed to enable realization of cost efficiencies created by joint operation of multiple radio stations. However, it generated controversy concerning the increase in radio owners' market power over listeners and advertisers (see Drushel (1998)). The coexistence of cost synergies and market power raises questions about the efficiency of the 1996 deregulation and it complicates an assessment of whether further deregulation would be socially beneficial. This paper answers these two questions by applying the proposed empirical framework to conduct antitrust policy counterfactuals.

First, I investigate the consequences of the doubling of local ownership caps mandated by the 1996 Telecom Act. For that purpose, I compute the hypothetical industry path under the pre-Telecom-Act local caps and compare it to the actual post-Telecom-Act industry path. I find that, in the long run, the 1996 deregulation increases total surplus. In particular, under the actual industry path, producer surplus is greater by 10.1%, listener surplus is greater by 0.07%, and advertiser surplus is lower by 1.7% as it compares to the hypothetical industry path. Next, I investigate the consequences of further deregulation. Specifically, I explore the possibility of raising the local ownership cap from between three to five FM stations (depending on the market size) to a uniform cap of seven FM stations. I find that this change would lead to an additional 4% increase in producer surplus, 0.01% increase in listener surplus and 1% decrease in advertiser surplus. I note that this policy would have a smaller impact on the industry than the Telecom Act.

The thus far considered policy changes, based on increasing local ownership caps, increase total surplus. However, the increase is predominantly in the form of greater profits to radio station owners. Thus, a regulation agency could consider complementing the increase in the ownership caps with measures that directly focus on consumer surplus. In particular, the recent version of Horizontal Merger Guidelines (2010) recommends using merger simulations as introduced by Nevo (2000), and Ivaldi and Verboven (2005). These simulations compute unilateral price effects of proposed mergers allowing the regulator to reject mergers that lead to substantial price increases. Nevertheless, existing versions of such simulations are inherently static. Thus, similarly to ownership caps, their efficiency is likely to be affected by industry dynamics.

To investigate the robustness of static merger simulations I consider two exemplary simulation-based policies. In the first policy, the agency raises the ownership cap to seven FM stations and additionally rejects all mergers that lower static listener surplus. In the long run, this leads to 2.24% increase in producer surplus, a 0.08% increase in listener surplus, and a 0.66% decrease in advertiser surplus. I note that using a combination of ownership caps and merger simulations results in a larger listener surplus than using ownership caps alone.

In the second policy, the regulator raises the ownership cap to seven FM stations and rejects the mergers that lower advertiser surplus. This policy aims to prevent the losses to advertiser surplus that result from all thus far considered deregulatory measures. Unfortunately, it presents only a short term solution. Specifically, in the first 5 years the policy leads to a 0.1% increase in advertiser surplus. However, in 20 years, this policy results in a 0.79% decrease in advertiser surplus, which is the same as in the policy without any merger simulations. [1] This counterintuitive result is a direct consequence of the myopic nature of merger simulations and exemplifies the Type II error mechanism described earlier. The companies, being forward looking, strategically propose mergers that are acceptable to the regulator and reposition the products after the merger is approved. As a result, in the long run, the companies are able to extract advertiser surplus circumventing the simulation-based antitrust criteria. This example of the failure of merger simulations shows that static heuristics should be applied with caution in the industries that experience considerable amount of post-merger activity, such as product repositioning.

[1] This result is consistent with the fact, that according to judicial documents, a version of advertiser-centered policy was sporadically applied between 1996 and 2000 but did not prevent losses to advertiser surplus, as documented by Jeziorski (2014a).

The closest empirical work to this paper is that of Benkard et al. (2010), who study the long run effects of mergers in the airline industry using a dynamic model of entry and exit with exogenous mergers. They find that the mergers between major airlines lead to increased entry by low cost carriers. Similarly, Collard-Wexler (2014) examines the duration of the merger effects in the ready-mix concrete industry and finds that a merger from duopoly to monopoly generates between 9 and 10 years of monopoly.

This paper extends the theoretical analysis by Nocke and Whinston (2010), who show that myopic merger rules are dynamically optimal when applied in a simplified environment with homogenous products and when possible mergers are disjoint (no firm has the possibility of being part of more than one merger). However, their paper does not cover markets with differentiated products and overlapping mergers, such the ones studied in this paper. Similarly, this paper extends Cabral (2003) who examines impact of free entry on the outcomes of the merger between two firms in a spatially differentiated oligopoly. Cabral shows that post-merger cost efficiencies can have detrimental effects on consumers by lowering entry. The possibility of such effects is allowed within the framework used in this paper.

The model utilized in this paper has features in common with the theoretical model by Armstrong and Vickers (2010), who examine a static principal and agent problem. The agent proposes projects to a principal, but 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. Furthermore, the present study extends the arguments of Lyons (2002), who emphasizes that regulators cannot choose which mergers to execute, but rather they can only approve or reject mergers from the set chosen by strategic players. Lyons offers 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 model is also compatible with a vast theoretical and numerical literature on endogenous mergers and product repositioning. For example, as in Kamien and Zang (1990), Rodrigues (2001), and Gowrisankaran and Holmes (2004), the model assumes that both sellers and buyers are fully forward looking. This work is also related to the literature on merger waves (see Harford (2005) and Qiu and Zhou (2007)) allowing mergers to be strategic compliments.

Another related study is Mazzeo et al. (2012), who numerically demonstrate, through the use of a static model, that post-merger repositioning can significantly alter the welfare assessment of the merger. In my present effort, I provide an empirical demonstration of a similar phenomena utilizing a dynamic model. Finally, the model is this paper builds on Sweeting (2013), who considers endogenous repositioning without considering mergers.

The study is organized as follows. In the next section, I provide a description of the U.S. radio broadcasting industry. In the third section, I present the dynamic model. The fourth section provides a description of the data, the fifth section describes the estimation algorithm, the sixth section contains results, the seventh section presents counterfactual, and I conclude in the eighth section.

2 Industry background

Between 1996 and 2006 the radio broadcasting industry in the United States underwent major structural changes. Prior to 1996 the industry was heavily regulated. In particular, Federal Communication Commission limited the joint ownership of radio stations through the imposition of national and local ownership caps. The restrictions were significantly relaxed with the enactment of the 1996 Telecom Act. This legislative change spurred a massive wave of ownership consolidation and product repositioning in which about a half of the 12,000 active radio stations changed ownership and a similar fraction of radio stations changed the type of broadcast content. These changes significantly affected both the demand and supply side of the radio industry by creating market power over advertisers and listeners as well as fixed and marginal cost efficiencies. Because of these developments, radio has become a viable case study for evaluating the consequences of antitrust policy changes.

The American radio broadcasting industry is composed of more than 300 relatively separated geographical markets. The broadcast spectrum in each market is partitioned into a set of discrete frequencies each hosting a single radio station. The number of frequencies does not change significantly over time; thus, most of the entry is executed though the acquisition of one of the assigned frequencies (between 1996 and 2006 FCC granted less than 60 new licenses nationwide which constitutes approximately 0.5% of active radio stations). Due to the geographical fragmentation the competition between radio stations is localized. For example, according to the Radio Advertising Bureau national advertising contributed only 15% of overall advertising revenue in 2009. As a result of this convenient fact, restricting attention to the competition for local listeners and advertisers captures first order revenue sources in this industry. In particular, for the purpose of modeling the demand side, the 300 local markets might be regarded as distinct, although the existence of some cross-market fixed cost synergies is possible.

Prior to the 1996 Telecom Act the industry was limited by national and local ownership caps. The national cap prevented any company from owning over 45 stations which effectively prevented the formation of large cross-market chains. The local cap was determined on the number of allocated frequencies, as described in Table 1. The 1996 Telecom Act abolished the national cap, nearly doubled the local cap, and over the following decade the industry moved from fragmented local ownership (about 1.64 station per owner in 1996) to a market in which fewer than 10 parent radio companies dominate two-thirds of both revenues and listeners nation wide (as of 2010). Furthermore, according to the 2010 data by BIA/Kelsey, the two largest companies, Clear Channel and Viacom, account for about 42% of listeners and 45% of advertising revenues. In particular, Clear Channel grew from 40 stations to over 1,200, executing about 15% of all 1996-2006 industry acquisitions. Thus, it is reasonable to expect that the first order magnitude of market power and cost synergies can be captured from an examination of a handful of large owners.

An extensive product repositioning followed the aforementioned consolidation of ownership. In the case of radio market such repositioning is relatively easy to identify because each radio station is uniquely categorized into a distinct programming format. Each format describes the overall type of programming and is directly related to the demographics of the potential listenership base. Thus, local radio markets can be categorized as differentiated product oligopolies in which the degree of differentiation is endogenous and is measured by the variety of supplied formats (see Berry and Waldfogel (2001)). The formats are announced on a biannually basis by Arbitron, a consulting company, and are frequently utilized as marketing tools for targeting advertising. Importantly, the formats frequently change, with approximately 10% of all stations switching formats annually, and these changes can be regarded as a generalized version of entry and exit into and out of particular industry niches; therefore, they affect the amount of market power and cost synergies. Thus, format changes are likely to have first order impact on the antitrust policy evaluation and should be incorporated in the analysis. This point is demonstrated numerically in the online appendix.

The next section contains the model of the industry that captures the above features.

3 Model

Consider a market over an infinite continuous time horizon. The market consists of a maximum of 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 programming format 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.

3.1 Actions

The firms' actions are either mergers or repositionings with entry and exit as special cases. 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 the curse of dimensionality, I assume these Poisson processes to be independent, across firms and event types, conditional on $\mathcal{J}^t$. Doraszelski and Judd (2012) demonstrated that such assumptions parallel the independent state transitions frequently used in discrete-time games.

Conditional on the arrival of an opportunity to merge, a player may choose one company to acquire. [2] During the acquisition the acquirer makes one take-it-or-leave-it offer P to the chosen

[2]Because the model is written in continuous time, it can still generate an arbitrary number of mergers within any fixed time period; and moreover, these multiple merger events are correlated endogenously. By contrast, the consequences are more restrictive for an equivalent assumption in a discrete-time model, which would effectively

acquiree. The acquiree accepts or rejects the offer, taking into account the dynamic opportunity cost. The cost of executing a merger[3] with k' is given by $\zeta^{A,t}(\mathcal{J}^t, k')$, which includes the 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 may choose a repositioning action $r = (f, f')$, which changes the format of an owned station from f to f', and involves 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 also arrive as a Poisson process (see Arcidiacono et al. (2010)).

Because the number of active broadcast frequencies is limited, entry is possible only through the acquisition of other active firms. Similarly, an exit is modeled as selling off all owned stations. Potential entrants are firms that hold empty portfolios of stations; that is, $\omega_k^t = \overrightarrow{0}$. [4] 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 entry into a particular local market waiting for favorable market conditions. Consequently, the common assumption that potential entrants are short-lived needs to be modified. Instead, I assume 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

$$\zeta_k^{A,t}(\mathcal{J}^t, k') = \mu_k^A(\mathcal{J}^t, k') + \sigma_k^A(\mathcal{J}^t, k')\epsilon_k^{A,t}(k'). \tag{3.1}$$

cap the number of acquisitions by a potential firm to one (or a finite number) per period.

[3] 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; thus, these costs coincide on the equilibrium path.

[4] 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.

Similarly, the repositioning payoff shock is given by

$$\zeta_k^{R,t}(\mathcal{J}^t, r) = \mu_k^R(\mathcal{J}^t, r) + \sigma_k^R(\mathcal{J}^t, r)\epsilon_k^{R,t}(r). \tag{3.2}$$

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 hand, 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, while ϵ picks up remaining local fluctuations in actions.

A major advantage of a continuous time model is that the method to resolve merger conflicts does not need to be specified. Consider the 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 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 component of the model would be difficult. In practice, it would force the modeler to assume such events away, for example, by putting a structure on a sequence of moves (see Gowrisankaran (1999), Gowrisankaran and Holmes (2004) and Jeziorski (2014b)).

3.2 Timing

The model includes the following sequence of events:

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

(2) Players collect the payoff $\pi_k(\mathcal{J}^t) - F_k(\mathcal{J}^t)$ 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 to make a merger bid. If he chooses to make the 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 stage (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 moves to stage (1) without any intermediary step.

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

3.3 Strategies and equilibrium

A strategy consists of four components: a merger strategy, a pricing strategy, a strategy to accept or reject the merger bid, and a repositioning strategy. A merger strategy has the following form: $\mathbf{a}_k(\mathcal{J}^t, \zeta^{A,t}) \in \{0, \ldots, K\}$. This formula 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 the accept/reject function is constant for all players and can be omitted (I relax the assumption later in this study). 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 remaining idle or for 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 arrivals of the l-th merger and the m-th repositioning opportunity for player k, respectively. With some abuse of notation, denote by $\zeta_k^{A,(l)}$ and $\zeta_k^{R,(m)}$, respectively, new private information shocks revealed at $\tau_k^{A,(l)}$ and $\tau_k^{R,(m)}$. 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 (I temporarily ignore the events by which company k is acquired):

$$V_k(\mathcal{J}^0; \mathbf{g}_k, \mathbf{g}_{-k}, \mathbf{G}) = E_{\mathbf{g}} \Bigg\{ \int_0^\infty e^{-\rho t} \left[\pi_k(\mathcal{J}^t) - F_k(\mathcal{J}^t)\right] dt + \sum_{l=1}^{\infty} e^{-\rho \tau_k^{A,(l)}} \left[\zeta_k^{A,(l)} \left(a_k^{(l)}\right) - G_k^{(l)} P_k^{(l)}\right] + \sum_{m=1}^{\infty} e^{-\rho \tau_k^{R,(m)}} \zeta_k^{R,(m)} \left(r_k^{(m)}\right) \Bigg\}. \tag{3.3}$$

The equilibrium of the game is defined as follows.

Definition 3.1 (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,

$$\mathbf{g}_k^*(\mathcal{J}, \zeta_k) \in \arg\max_{\mathbf{g}_k} V_k(\mathcal{J}; \mathbf{g}_k, \mathbf{g}_{-k}^*, \mathbf{G}); \quad \forall k, \mathcal{J}, \zeta_k. \tag{3.4}$$

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

$$\mathbf{P}_k^*(\mathcal{J}, \zeta_k^A, k') \geq V_{k'}(\mathcal{J}; \mathbf{g}^*, \mathbf{G}); \quad \forall k, \mathcal{J}, \zeta_k^A. \tag{3.5}$$

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

In this study, I examine only equilibria in which the acquisition price is equal to the acquiree's value function, thus acquisition events 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. In such a case, knowing there is no other outstanding offers at each particular instant, the acquirer would propose the acquisition price equal to the reservation value of the acquiree (value function) or would not make an offer if this reservation value is too large. Such a 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.

3.4 Existence

In order 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 the 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 of 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 occurs. The key fact is that these expectations can be expressed as functions of CCPs. Hotz and Miller (1993) established this 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:

$$\begin{aligned}
\rho V_k(\mathcal{J}) = \max_{\mathrm{CCP}_k^A, \mathrm{CCP}_k^R} \Bigg\{ & \pi_k(\mathcal{J}) - F_k(\mathcal{J}) - \left(\lambda_k^A(\mathcal{J}) + \sum_{k=1}^{K} \lambda_k^R(\mathcal{J}) \right) V_k(\mathcal{J}) - \\
& \lambda_k^A(\mathcal{J}) \left[\sum_{a \in \Gamma_k^A(\mathcal{J})} \mathrm{CCP}_k^A(a) \Big(V_k(\mathcal{J}'(k,a)) - V_a(\mathcal{J}) + W_a^A(\mathrm{CCP}_k^A, \mathcal{J}) \Big) \right] + \\
& \lambda_k^R(\mathcal{J}) \left[\sum_{r \in \Gamma_k^R(\mathcal{J})} \mathrm{CCP}_k^R(r) \Big(V_k(\mathcal{J}'(k,r)) - V_{a'}(\mathcal{J}) + W_r^R(\mathrm{CCP}_k^R, \mathcal{J}) \Big) \right] + \\
& \sum_{k' \neq k} \lambda_{k'}^A(\mathcal{J}) \sum_{a \in \Gamma_{k'}^A(\mathcal{J})} \mathrm{CCP}_{k'}^A(a) V_k(\mathcal{J}'(k',a)) + \\
& \sum_{k' \neq k} \lambda_{k'}^R(\mathcal{J}) \sum_{r \in \Gamma_{k'}^R(\mathcal{J})} \mathrm{CCP}_{k'}^R(r) V_k(\mathcal{J}'(k',r)) \Bigg\},
\end{aligned} \tag{3.6}$$

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 from Doraszelski and Judd (2012).

3.5 Computational strategy

Equation 3.6 can be used to compute the equilibrium of the game, and the algorithm has relatively low computational requirements. Suppose the idiosyncratic parts of the payoff shocks, as defined in equations (3.1) and (3.2), 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 are given by a closed-form formula,

$$\mathrm{CCP}_k^A(a|\mathcal{J}) = \frac{\exp\left\{\sigma_k^A(\mathcal{J},a)^{-1}\left[V_k(\mathcal{J}'(k,a)) - V_a(\mathcal{J}) + \mu_k^A(\mathcal{J},a)\right]\right\}}{\sum_{a'\in\Gamma_k^A(\mathcal{J})}\exp\left\{\sigma_k^A(\mathcal{J},a')^{-1}\left[V_k(\mathcal{J}'(k,a')) - V_{a'}(\mathcal{J}) + \mu_k^A(\mathcal{J},a')\right]\right\}}, \tag{3.7}$$

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 (3.1). Repositioning CCPs are given by the following formula:

$$\mathrm{CCP}_k^R(r|\mathcal{J}) = \frac{\exp\left\{\sigma_k^R(\mathcal{J},r)^{-1}\left[V_k(\mathcal{J}'(k,r)) + \mu_k^R(\mathcal{J},r)\right]\right\}}{\sum_{r'\in\Gamma_k^R(\mathcal{J})}\exp\left\{\sigma_k^R(\mathcal{J},r')^{-1}\left[V_k(\mathcal{J}'(k,r')) + \mu_k^R(\mathcal{J},r')\right]\right\}}. \tag{3.8}$$

The computational algorithm involves iterating on the value function using a Bellman equation (3.6) and equations (3.7) and (3.8). The procedure can be summarized as follows:

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

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

  (i) use $V^{(j)}$ to compute the CCPs of all players at $\mathcal{J}$, given by equations (3.7) and (3.8),

  (ii) use the CCPs from (i) to obtain a new value function $V^{(j+1)}(\mathcal{J})$ by iterating a Bellman equation (3.6).

(2) Stop if $\|V^{(j)} - V^{(j+1)}\| <$ tolerance; otherwise, go to stage (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 on a player by player basis, 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 a 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 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 $W_a^A(\mathrm{CCP}_k^A, \mathcal{J}^t)$ and $W_r^R(\mathrm{CCP}_k^R, \mathcal{J}^t)$ for more than two feasible actions is unknown. Instead, these expectations must be simulated (see section 5 for details).

4 Data

The data used to estimate a dynamic model covers the period of 1996-2006 and consists of (i) a complete set of radio station acquisition transactions with monthly time stamps, and (ii) formats of every radio station in the United States with half-year time stamps. Additionally, the study 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. In order to avoid modeling cross-market interactions I drop the overlapping markets in a way following the method of Sweeting (2013); that is, I drop markets "where more than 6% of listening was to stations based in other markets". I also drop markets that do not have data on advertising prices. Appendix B contains the details on the price/quantity data 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 described in section 2. Primarily, I divide the set of players into three groups: dominant owners, local owners, and fringe. Dominant owners include companies such as Clear Channel, ABC, or Viacom, which own 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 dominant owners to reposition stations within their portfolios. The second group of companies consists of local owners. These companies are not allowed to own multiple stations; however, they are allowed to reposition. Both dominant 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 group are myopic and cannot reposition or be acquired, but they do participate in the static competition for advertisers and listeners.

In each local market, I label three active companies with the largest national revenue share in 2006 (the last year of the data set) as dominant owners. Consequently, each local market might have a different set of potentially active dominant owners, however this set almost always contains Clear Channel, accompanied by ABC, Viacom, Citadel or Cumulus (see Table 2). According to the data, during and immediately after 1996, when the ownership was still fragmented, the dominant owners were initially inactive in many local markets, and then subsequently entered through acquisitions.

In addition to tracking dominant owners, in each local market, I label 22 of the remaining radio stations with the highest local listenership share as local owners. All other radio stations, which are small and usually have less than 0.5% listenership, are labeled as fringe stations. An exception to the above rule consists the markets with more than 15 active stations, where I label all AM stations as fringe because in such markets FM stations generate a dominant part of total revenues. In rural markets with less than 15 active stations, AM stations become important, so I allow both AM and FM stations to be outside of the fringe.

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. Relaxation is also possible when computing counterfactuals and when using a nested-fixed-point estimator; but, given my data, these extensions would require further assumptions and have not been implemented.[5]

Modeling dominant 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 the online appendix, the regulator must track the product repositioning of smaller players in response to the merger. Also, as shown in example 3 in the online appendix, the acquirees should be modeled forward-looking. However, at the same time, the smallest stations rarely change formats and are almost never acquired by larger owners as indicated in the data. Thus, in practice, while increasing

[5]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.

the complexity of the estimation, modeling the forward-looking decisions of every small owner has little benefit. Nevertheless, dropping the smallest owners altogether is unrealistic because they collectively affect markups in the pricing game. Therefore, fully tracking only large and medium owners, and partially tracking smallest firms is a realistic compromise.

One artifact of not allowing smaller players to make merger bids is the prohibition of spin offs. I do observe spin offs in the data, but they are mostly a consequence of lumpy cross-market mergers that violate local ownership caps. Consequently, the owners must 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 into the new owner's portfolio in the first place. Thus, counting spun off stations 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 subsequently.

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, ock, 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. This choice is dictated by the consideration of the substitution patterns described by Jeziorski (2014a). The "Adult Music" format caters to a more mature population of listeners, while the "Hits Music" attracts a mostly younger crowd. The details on the substitution patterns can be found in section 6.1. The aggregation trades 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 greater 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 the estimation of large dynamic systems without re-solving for an equilibrium at each parameter value. Hotz and Miller (1993) developed their estimator for discrete-time single-agent problems, and many studies have extended their method to discrete-time dynamic games (see Pakes et al. (2007), Bajari et al. (2007), Aguirregabiria and Mira (2007) and Arcidiacono and Miller (2011)). The present paper develops a new two-step Instantaneous Pseudo Maximum Likelihood (IPML) estimator that maximizes an objective function based on the instantaneous choice probabilities from equations (3.7) and (3.8). An IPML is can be regarded as an extension of the 0-iteration Pseudo Maximum Likelihood introduced in Aguirregabiria and Mira (2007) for discrete time games. The estimation procedure builds on the previous work by Arcidiacono et al. (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.

Suppose one has the data on H players' actions and states of the game an instant prior to taking these actions $\{(g_h, \mathcal{J}_h) : h = 1, \ldots, H\}$, where g_h is either a merger or repositioning action. If the full solution to the game is available, then the game can be estimated using a full information maximum log-likelihood (FMLE). Because state transitions conditional on actions are deterministic an FMLE is obtained by plugging a computed value function $V(\cdot;\theta)$ into the equations (3.7) and (3.8); that is,

$$L_H(\theta) = \sum_{h=1}^{H} \log \mathrm{CCP}(g_h | \mathcal{J}_h; V(\cdot;\theta), \theta).$$

However, there are two reasons why $V(\cdot;\theta)$ is difficult to obtain: (i) the state space of the game is large, so it is infeasible to recompute the value function for many candidate values of θ; and, (ii) the game is likely to have multiple equilibria so obtaining $V(\cdot;\theta)$ for every possible equilibrium might be necessary. The IPML estimator is designed to solve these issues. It replaces the value function $V(\cdot;\theta)$ with its uniformly consistent estimator $\hat{V}(\cdot;\theta)$, and maximizes the instantaneous pseudo likelihood

$$Q_H(\theta) = \sum_{h=1}^{H} \log \mathrm{CCP}(g_h | \mathcal{J}_h; \hat{V}(\cdot;\theta), \theta). \tag{5.1}$$

I follow the usual way of obtaining a uniformly consistent estimator of the value function; that is, I first pre-estimate CCPs and subsequently simulate the value function using equation (3.3). The details are presented in the three remaining parts of this section. The first part describes the pre-estimation of a one-shot profit function; the second explains the estimation of acquisition and repositioning strategies; and the third describes the 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 (2014a) with the exception that for this study I employ three meta-formats, that is $F = 3$, instead of eight. Below I describe the parametrization in order to keep the paper self-contained; however, the discussion is also kept 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:

$$\pi_k(\mathcal{J}^t) = \sum_{\substack{\text{j owned by k} \\ \text{in market m}}} \Big(p_j(\bar{q}_j^t|\mathcal{J}^t) r_j(\bar{q}^t|\mathcal{J}^t) - \mathrm{MC}_j(\mathcal{J}^t) \Big) \bar{q}_j^t, \tag{5.2}$$

where $p_j(\cdot)$ is the 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 et al. (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 broadcast advertising minutes for station j as q_j. For a given consumer i, the utility from listening to a station j is given by

$$u_{ij} = \theta_{1i}^L \iota_j - \theta_{2i}^L q_j + \theta_3^L \mathrm{FM}_j + \xi_j + \epsilon_{ji}, \tag{5.3}$$

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 equivalent to those in Berry et al. (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

$$r_j(q|\mathcal{J}^t) = \text{Prob}\big(\{(\nu_i, D_i, \epsilon_{ij}) : u_{ij} \geq u_{ij'}, \text{for } j' = 1, \ldots, J\}\big|q, \mathcal{J}^t\big). \tag{5.4}$$

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 by advertisers, as well as advertising congestion. The simplest reduced-form model that captures these features is a linear inverse demand for advertising, such as

$$p_j = \theta_1^A \left(1 - \theta_2^A \sum_{f' \in \mathbb{F}} w_{ff'}^m q_{f'}\right), \tag{5.5}$$

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

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

$$\text{C}_{jmt}(\theta^A, \theta^C) = \theta_1^{Am}[\theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Cm}\xi_{jt} + \theta_3^{Cm}\text{SYN}_{jt} + \epsilon_{jt}^C]q_{jt}, \tag{5.6}$$

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 differ for each market, and θ^{Cm} is allowed to vary

[6] The assumptions on ξ are a simplification compared to the specification used by Jeziorski (2014b) and Sweeting (2013), who both assume ξ_j follows an AR(1) process. This decision was made to keep the dynamic model computable.

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 it interacts 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 and supply equation, and because the distribution of listeners' demographics varies from year to year. Consequently, I estimate the model as non-stationary to obtain more robust measures of listener and advertiser price elasticity. However, after the estimation, I detrend the static profits to fit the static model into the dynamic framework presented in section 3. Specifically, I remove the trends in supply and demand by using an average value of the time dummies and I 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 results in the conservative estimates of cost synergies. For example, missing an possible long-run downward trend in listenership would overestimate the profitability of mergers and consequently the estimator would require smaller cost synergies to rationalize the consolidation. I return to this points when discussing the results and counterfactuals, however, I note that the existence of this kind of bias would make my final conclusions stronger.

The state space of the market needs to be further simplified in order to make the dynamics manageable. In particular, I compute an average station quality ξ and tower power for each market-format combination and endow each station with these averages depending on the station's market and format. Additionally, to control for the differences in the radio stations' sizes between fringe and non-fringe owners, I allow the averages of ξ and tower power to be different for these two groups of radio stations. Such procedure further standardizes the profit function and makes it consistent with the model laid out in section 3. 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 (2014b), which allows for station-level 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 the counterfactuals. To alleviate some of the concerns, I compare cost synergy estimates to those in Jeziorski (2014b); however, all results in the present study should be interpreted keeping this 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 within the same month, the state space at the time of taking an action is only partially observed.

The second set describes repositioning decisions:

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

where b^{mhd} is the observed set of repositioning events during half-year h. The formats are observed once every half a year, as opposed to the mergers, which are observed monthly. Therefore, 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 a repositioning action. Furthermore, I do not know how many active players were present during a repositioning action. For these reasons, the state, set of players, and players' actions are only partially observed, which must be taken into account during the estimation.

Equilibrium CCPs, given by equations 3.7 and 3.8, depend on the state through unknown value functions and semi-parametric estimation. In particular, the acquisition CCPs are given by

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

where $\Upsilon^A(k, k'\mathcal{J})$ are unknown functions of the state $\mathcal{J}$. The unknown functions Υ^A and Υ^R are approximated through the use of polynomial sieves (see Ai and Chen (2003)) described below.

I 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 dominant 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. I postulate that 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 dominant owners. With additional 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 dominant 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 indices 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 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

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

Note that dominant and local owners have different primitives for the model (possibly different fixed cost structures and local companies cannot acquire other firms), thus, equilibrium strategies have to be allowed to differ across these two types of players. For this reason, I utilize a different polynomial Υ^R to approximate the repositioning strategies of dominant and local firms. Despite the use of two different Υ^R, the coefficients of all polynomials are interrelated and require joint estimation because, as mentioned before, the state is only partially observed.

If the industry states were perfectly observable at the instant of the actions were takes, the sieves estimator would choose the θ^A and θ^R that maximize the data's pseudo-likelihood. However, as explained earlier, the data is imperfect, and full information likelihood cannot be computed. Instead one could use a simulated likelihood or a 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 a substantial loss in efficiency. Another option is to perform an analytical integration of unobsevables using Chapman-Kolmogorov equations describing state transitions, as suggested by Arcidiacono et al. (2010). However, this method cannot be applied directly, because the full intensity matrix (a continuous time equivalent of a transition matrix) for my largest markets can contain up to 4 million by 4 million entries. Although this matrix is quite sparse, it would not be sufficiently sparse to either store in the computer memory, or to recompute "on the fly." Instead, I develop a method of integrating the likelihood based on partial Chapman-Kolmogorov equations. These partial equations take advantage the fact that only a small subset of feasible latent industry states are relevant for the estimation. The details of this method are presented in the Appendix C.

I do not observe events in which players take no action, therefore, in the first stage, I am only able to identify the product of the move arrival rate λ and CCPs. However, as long as the true move arrival rate is not excessively small, 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 details the remaining parts of the estimation; namely, the parametrization of the model and the simulation of the value function.

Fixed cost. The fixed cost of player k to operate 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 dominant 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 as follows[7]

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

where θ_N^F is a discount for being a dominant 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 as follows:

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

The acquisition cost may 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 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 to be heterogeneous across markets, I allow the intercepts $\theta^{A,m}$ and $\theta_\sigma^{A,m}$ to vary across four market categories. The first category consists of markets in which a

[7] 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.

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) may vary with the type of programming. I try this specification and find that 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 dominant owners face different repositioning costs than local owners. For example, voice-tracking technology can allow the dominant owners to temporarily bring announcers from other markets to streamline format switching. However, local owners might have better access to local labor markets and a more flexible workforce. Differences may additionally vary by format. One example is the Hits Music format, which requires a large tower and costly marketing to gain sufficient listenership. For this reason, switching to the Hits format is likely to require greater capital investments and access to specialized production factors. Thus I expect dominant owners to have lower cost when switching 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 vary by the company type. In addition, I allow for the mean shifts and heteroscedasticity by size, which are 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 mean $\theta^{R,m}$, and the variance $\theta^R_{\sigma,m}$. I find that allowing for such flexibility in the specification is critical in fitting the model to the data.

The above specification is used to simulate the value function

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

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

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

The closed-form solution for the conditional expected value of shocks W is unknown for the number of alternatives larger than 1 (not including empty actions), which is a consequence of the fact that idiosyncratic shocks are not distributed as type-1 extreme value random variables. Instead, I simulate the idiosyncratic part of W on the grid of CCPs and fit the 4^{th}-degree complete Chebyshev polynomial. Likewise, 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.

Having the estimators of the value function $\hat{V}$ at hand, I maximize an expected version of the pseudo-likelihood (equation (5.1)), obtained using the procedure described in Appendix C. Note that the value function must be simulated for every potentially feasible latent state as well as for any state attainable by a single action from any feasible latent state. For example, in the case of Los Angeles, there are 46 feasible latent states, which generate 1,208 potentially accessible states. Overall, 88 markets contain 106,304 accessible states. Each simulation is composed of 1,000 draws, so the procedure involves obtaining 106,304,000 industry paths, which are assumed to evolve for 40 years and are kept constant thereafter. Because the number of industry paths is large 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 Appendix D). In such a case the computation of the pseudo-likelihood takes about as much as computing first-stage likelihood. (ii) Continuous time enables updating the industry state only at the arrival of the executed move, which saves computing power when simulating relatively infrequent actions such as mergers and product repositioning. In an extreme case, when the draw of the waiting time for the first executed move exceeds 40 years, the state is never updated and the draw of the value function collapses to perpetual static profits.

Several components of the model must be identified: (i) the 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. The identification strategy relies on the fact that revenues of radio owners at each industry state can be predicted by a pre-estimated static model. As in Sweeting (2013) and Jeziorski (2014b), the repositioning cost is identified as the residual from endogenizing format repositioning. In other words, pre-estimated revenue predictions and the estimated repositioning cost must rationalize the repositioning actions observed in the data. I identify the merger cost and cost efficiencies from consolidation in a similar way. Specifically, I use the convenient fact that entry is possible only though acquisition and choose the merger cost that rationalizes observed timing of such entry decisions. Similarly, the estimates of the cost efficiencies have to rationalize subsequent acquisitions. 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 difficult without observing move opportunities that were not executed. Such 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 are ownership caps. If an 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 that 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 analogous to a period length in a discrete-time game. Identifying this period length is similar to identifying a discount factor, which is known to be difficult (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 similar results, both qualitatively and quantitatively.

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 the first- and the 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 (2014a), and to minimize the duplication I provide only a brief description of the parameters.

Table 3 presents the estimates of listenership demand. The first and second columns contain the mean and the standard deviation of the random coefficients. I find that advertising has a negative effect on listenership, and the effect is fairly homogenous among listeners. I also find that listeners prefer FM to AM stations, and that greater transmitting power corresponds to higher listenership. Format dummies are negative; however, by construction, they capture preferences of a specific demographic group: 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 4. The first-meta format, which delivers adult-oriented music, such as, classic rock, country, or jazz, is most popular among middle-aged listeners, with women constituting a slight majority of the demographics. The second meta-format, which delivers contemporary pop and alternative music, appeals to younger people. This meta-format 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, mainly higher-educated females with relatively large incomes. A positive Hispanic dummy is related to Hispanic and religious stations in this meta-format.

Table 5 describes the national trends in radio listenership. The values represent the residual trend in radio listenership beyond the changes in the demographic composition. In general, I find the trend in the utility of the outside option to be non-monotonic. I use these numbers to detrend the profit function in the dynamic estimation.

Table 6 presents the coefficients of the inverse demand for advertising. The inverse demand curve is downward sloping, indicating the radio stations' direct market power over advertisers. The inverse demand is steeper in smaller markets.

Tables 7 and 8 contain the marginal cost estimates. The marginal cost is larger in smaller markets. 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 jointly as described in section 5.2. I group the results into multiple tables to facilitate the exposition. Tables 9 through 12 contain estimates of the acquisition strategy; Tables 13 through 16 present estimates of the repositioning strategy of national owners; and finally, Tables 17, through 19 contain estimates of the repositioning strategy of the local owners.

6.2.1 Acquisition strategy

Table 9 contains the estimates of format-acquisition dummies and interactions between format acquisition and demographic composition of the local market. The 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 other observable covariates explain most of the variation in acquisition across formats. I find that the interactions between the demographic composition and acquisition propensity represent the preferences for formats described in Table 4, 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 interacts positively with Hispanic.

Tables 10 and 11 present the coefficients describing the relationship between the industry state and a propensity to acquire a particular format f. Specifically, Table 10 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 (2014b) find that owners avoid acquiring stations similar to their current portfolios. The reason for this discrepancy is the broader definition of the format used in this study. For example, an owner of 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 indicates that the payoffs from mergers have decreasing returns.

The third column of Table 10 contains a coefficient that captures the impact of the number of national competitors in the same format as a potential acquiree on the propensity to acquire. I find that a larger number of national competitors in a particular format correlates with a higher propensity to acquire in this format; however, the result is not statistically significant. The fourth column reports a coefficient on ownership concentration (similar to the Herfindahl index) of stations in the acquiree's format. In general, the less concentrated the ownership in the acquiree's format, the greater the propensity to acquire. The last two columns describe the impact of the portfolios of the local owners. I find that, unlike the number of national competitors, the number of local competitors in the acquiree's format is negatively correlated with an acquisition.

Table 11 contains further state covariates explaining the acquisition decision. I find that larger numbers of owned stations correlate positively with a greater likelihood of acquisition. Also, the more stations are held by competing national owners, the less probable the acquisition. The last two numbers in the table contain higher-order terms representing concentration of ownership across competitors and the distribution across formats of nationally owned stations.

Table 12 contains dummies reflecting proximity to the ownership cap. All else being equal, if the acquisition should result in coming closer to the maximum set by the cap, that acquisition becomes less probable. This means that as owners approach the cap, they become more selective in order to maintain an option value.

6.2.2 Format-switching strategy

Table 13 contains from-to format-switching dummies, as well as, interactions between demographics and target formats. As with acquisition strategy, I find that format switching largely reflects listeners' tastes. Tables 14 and 15 describe the impact of the industry state on format switching. In general, the number of stations owned in a particular format correlates positively with switching to that format, while the opposite holds for the number of stations owned by competitors.

Estimates in Table 16 describe an impact of proximity to the ownership cap on format switching. Note that the closer the owners are to the limits imposed by the cap, the more probable format switching becomes. In this case, 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 may as well acquire.

Tables 17 through 19 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 estimates of the structural parameters, including fixed cost, as well as persistent and variable components of the acquisition and repositioning costs.

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 the economies of scale for owning multiple stations of any format $F^E(n_k^t|\theta^E)$.

Table 20 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 22 contains the estimates of the within-market economies of scale resulting from owning multiple station. I find that 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, a status that captures cross-market cost synergies. I find that 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 in the same format. According to Table 21, the operation of stations in the same format is additionally 14% cheaper on top of the economies of scale indicated in Table 22.

Table 23 presents the estimates of the merger cost. I find that the mean acquisition cost varies by market type and company size. Moreover, I find that this acquisition cost has a 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 large mean and high volatility usually leads to low cost for realized mergers.

Estimates of repositioning cost are contained in Tables 24 through 27. 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 fixed effect for market category, source-target format, and being a national owner. I find that the repositioning cost varies considerably across 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 formats, suggesting that switching is cheaper between some formats than between others. These differences in switching costs are driven by the format switching 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 explain the other switching-cost estimates similarly.

7 Counterfactuals

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

7.1 Impact of the 1996 Telecom Act

The first set of experiments is meant to investigate the impact of the 1996 Telecom Act on producer, listener, and advertiser surplus. Table 28 presents counterfactuals that evaluate the impact of the looser post-1996 local ownership caps. In particular, I recomputed the equilibrium merger and repositioning strategies for the pre-1996 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 is difficult to identify separately from the move arrival rate λ. I find that radio-station owners benefit from the deregulation because, under the old caps, the producer surplus is 10% lower. Roughly 6% of this decrease comes a loss of market power and 4% comes from lost cost efficiencies. Moreover, the deregulation leads to less advertising supplied and higher per-listener prices. Consequently, 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 (the post-1996 conditions). An additional exercise aims to partially address changes in the national cap by nullifying the cross-market cost benefits. In general, a lack of national synergies leads to fewer mergers, which lowers the negative impact of the deregulation on advertiser surplus. However, mergers are also less efficient because cross-market synergies are not internalized, thereby the realized mergers generate smaller gains in producer surplus. The overall effect is presented in the bottom three rows of Table 28.

It is useful to compare the findings from dynamic analysis with the findings obtained using a static analysis, which assumes that mergers are exogenous. In particular, Jeziorski (2014a) finds that over the course of the decade between 1996 and 2006, ownership consolidation decreased advertiser welfare by 21%, whereas the present study determines that decrease to be 1.7%. The following two factors contribute to this difference: (i) static analysis cannot account for long-run product repositioning, which could correct the negative effect of the mergers; and (ii) 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. I compare pre- and post-merger net entry on Figure 1 and I find that mergers lead to higher entry rates in all 88 markets. As documented numerically in this study's on-line numerical appendix, the positive impact of the mergers on net entry rates means that these mergers are likely to be self-correcting. Thus, a regulator that does account for entry, or uses undervalued pre-merger entry rates to evaluate the merger, may overestimate the negative impact of the merger on consumer surplus, which results in over-blocking.

7.2 Alternative merger policies

Ownership caps are rarely applied in markets other than radio broadcasting. Alternatively, the regulator applies policies based on concentration indices or direct measures of welfare. In the next experiment, I increase the ownership caps to seven FM stations, as before, but additionally I 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 29. I find that in the long run, the relaxation of the caps leads to about a 4.2% increase in producer surplus. Approximately one third of this gain comes from fixed cost efficiencies, and the remaining two thirds from market power. In the long run, market power is exercised predominantly on advertisers, which lose about 1% of their surplus. At the same time, the listeners gain 0.01%. Shorter-run analysis (over the first 5 to 10 years) 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 increase variety in the industry (see Berry and Waldfogel (1999), and Sweeting (2009)) benefiting listeners but hurting advertisers by thinning competition (see Jeziorski (2014a)). However, prior to this study, the applicability of these results to hypothetical and out-of-sample policies such as CAP7 has not been established.

Next, I impose additional antitrust criteria and recompute the equilibrium of the dynamic game. Rows 4 to 6 of Table 29 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 cost efficient compared to CAP7, much of the cost synergy is still realized.

Lastly, I evaluate the policy based on advertiser surplus and present the results in rows 7 to 9 of Table 29. 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 an approximate 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. The reason for this reversal is that companies circumvent the regulation by proposing mergers that meet the static advertiser surplus criterion, and optimally repositioning to extract the advertiser surplus once the merger is approved. Consequently, in the long run, the impact of the enforcement policy on welfare depends more on how many mergers are approved and less on the type of mergers that are accepted. 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, entry/exit and product repositioning. The merger transfer prices are endogenous and are an outcome of a dynamic bargaining process. This model aims at providing a tool to conduct antitrust merger review that accommodates the distortions created by endogeneity of mergers, entry/exit and product repositioning.

I demonstrate the applicability if the procedure by examining the wave of consolidation in U.S. radio broadcasting industry that occurred from 1996 to 2006. I find substantial fixed and marginal cost synergies from joint ownership. Namely, operating multiple stations together within the local markets is cheaper than operating them individually. Additionally, significant cost synergies arise from the operation of multiple similar stations withing a local market. Specifically, the operating of two stations by the same owner in a given local market is up to 14% cheaper than the operation of similar stations by different owners. Moreover, if the two stations that are operated jointly have the same programming format, the fixed cost drops by an additional 14%.

After establishing the extent of cost synergies from mergers, I compute a merger retrospective that evaluates the 1996 Telecom Act. This retrospective compares the industry trajectory without the Act with the factual trajectory with the Act. I find that 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 ownership caps to seven stations increases total surplus. I demonstrate that the mergers in the radio industry are largely self-correcting because they invite a significant amount of repositioning, which mitigates the market power. However, I also demonstrate that static welfare criteria are not sufficiently safeguarding to all losses to advertiser surplus. Specifically, the criterion that rejects mergers lowering static advertiser surplus does not prevent long-run losses to that surplus. Such losses to the advertisers surplus are a consequence of the fact that companies can circumvent the welfare rule by proposing a merger that is acceptable to the myopic regulator and altering the product characteristics to extract advertiser surplus after the merger. Also, more generally, I show that in the industries where dynamic processes such as entry/exit or product repositioning are prevalent, the companies are able to extract consumer surplus despite the welfare criteria employed by the regulator. In such cases, the efficacy of a particular merger enforcement rule is predominantly driven by the raw number of merger this rule allows, and less by the exact types of merger this rule permits.

References

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

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

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

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

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

Bain, J. S. (1968): *Industrial organization*, Wiley New York.

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

Benkard, C. L., A. Bodoh-Creed, and J. Lazarev (2010): "Simulating the Dynamic Effects of Horizontal Mergers: U.S. Airlines," Tech. rep., Stanford University.

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

——— (2001): "Do Mergers Increase Product Variety? Evidence From Radio Broadcasting," *The Quarterly Journal of Economics*, 116, 1009–1025.

Cabral, L. (2003): "Horizontal mergers with free-entry: why cost efficiencies may be a weak defense and asset sales a poor remedy," *International Journal of Industrial Organization*, 21, 607–623.

Collard-Wexler, A. (2014): "Mergers and Sunk Costs: An application to the ready-mix concrete industry," *AEJ: Microeconomics*.

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

Drushel, B. E. (1998): "The Telecommunications Act of 1996 and Radio Market Structure," *Journal of Media Economics*, 11.

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

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

Gowrisankaran, G. (1999): "A Dynamic Model of Endogenous Horizontal Mergers," *The RAND Journal of Economics*, 30, 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, 561–582.

Harford, J. (2005): "What drives merger waves?" *Journal of Financial Economics*, 77, 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, 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, 669 – 691, ¡ce:title¿Merger Control in International Markets¡/ce:title¿.

Jeziorski, P. (2014a): “Effects of mergers in two-sided markets: Examination of the U.S. radio industry,” *AEJ: Microeconomics*, Forthcoming.

——— (2014b): “Estimation of cost synergies from mergers: Application to U.S. radio,” *The RAND Journal of Economics*, Forthcoming.

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

Kryukov, Y. (2008): “Dynamic R&D and the Effectiveness of Policy Intervention in the Pharmaceutical Industry,” Tech. rep.

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

Lyons, B. R. (2002): “Could politicians be more right than economists?: A theory of merger standards,” Tech. rep.

Magesan, A. and V. Aguirregabiria (2013): “Euler Equations for the Estimation of Dynamic Discrete Choice Structural Models,” Tech. rep.

Mazzeo, M., K. Seim, and M. Varela (2012): “The Welfare consequences of mergers with product repositioning,” Tech. rep.

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

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

——— (2013): “Merger Policy with Merger Choice,” *American Economic Review*, 103, 1006–33.

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, 373–399.

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

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

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

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

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.

——— (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, pp. 18–36.

Appendices

A Tables and Figures

# of active stations	Old ownership cap	New cap
45+	4	8
30-44	4	7
15-29	4	6
0-14	3	5

Table 1: Change in local ownership caps introduced by the 1996 Telecom Act.

Company	Number of active markets in 2006 (out of 88)
Clear Channel	79
ABC/Disney	35
Viacom Int'l Inc	23
Citadel Comm Corp	18
Cumulus Media Partners LLC	18
Forstmann Little	18
Davidson Media Group LLC	17
Family Stations Inc	16
Radio One Inc	16
Entercom	15
Cox Radio Inc	12
Multicultural Bcstg	12
CSN International	10
Crawford Broadcasting Company	10
Entravision Communications Corp	10
Univision	10

Table 2: Number of active local markets (out of 88) for the largest radio owners.

	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 3: 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.

	Age	Sex	Education	Income	Black	Spanish
	Demographics Characteritics					
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 4: 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 a particular radio =station format.

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

Standard errors in parentheses

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

Table 5: 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^{***} (0.00)	-0.04^{***} (0.00)	-0.05^{***} (0.00)	-0.03^{***} (0.00)
2SLS	-0.07^{***} (0.00)	-0.03^{***} (0.00)	-0.03^{***} (0.00)	-0.02^{***} (0.00)

Standard errors (corrected for the first stage) in parentheses

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

Table 6: 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^{***} (0.03)	2.16^{***} (0.03)	1.22^{***} (0.03)	0.22^{***} (0.00)	0.16^{***} (0.00)	0.08^{***} (0.00)
2SLS	2.99^{***} (0.04)	2.42^{***} (0.04)	1.67^{***} (0.05)	0.26^{***} (0.00)	0.18^{***} (0.00)	0.12^{***} (0.00)

Standard errors (corrected for the first stage) in parentheses

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

Table 7: The marginal cost per minute of advertising sold. The intercept of the advertising price per rating point is set to 1. Note that these numbers may be higher than 1 because the final price of advertising is the 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)	-0.02 (0.01)	-0.10^{***} (0.01)
2SLS	-0.21^{***} (0.02)	0.01 (0.01)	-0.05^{***} (0.01)

Standard errors (corrected for the first stage) in parentheses

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

Table 8: The marginal cost synergies of owning multiple stations in 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 9: Acquisition CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages.

$\eta_{f,k}$	$\eta^2_{f,k}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f} \right)^2$	$\sum_{k' \in \mathbf{K^L}} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^L}} \eta_{k',f} \right)^2$
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 10: 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)
$\left(\sum_{f' \neq f} \eta_{f',k}\right)^2$	$\sum_{k' \in \mathbf{K^N} \backslash k, f' \neq f} \eta_{k',f'}$
−0.353	−0.131
(5.247)	(4.912)
$\left(\sum_{k' \in \mathbf{K^N} \backslash k, f' \neq f} \eta_{k',f'}\right)^2$	$\sum_{k' \in \mathbf{K^N} \backslash k} \left(\sum_{f' \neq f} \eta_{k',f'}\right)^2$
−1.999	4.618
(3.845)	(0.940)
$\sum_{f' \neq f} \left(\sum_{k' \in \mathbf{K^N} \backslash k} \eta_{k',f'}\right)^2$	
−4.306	
(2.009)	

Table 11: Acquisition CCP: Coefficients on the covariates related to formats other-than-those of the acquiree.

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

Table 12: Acquisition CCP: Dummies for proximity 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 13: National owner repositioning CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages.

$\eta_{f,k}$	$\eta^2_{f,k}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}$	$\sum_{k' \in \mathbf{K^N} \setminus k} \eta^2_{k',f}$	$\left(\sum_{k' \in \mathbf{K^N} \setminus k} \eta_{k',f}\right)^2$	$\sum_{k' \in \mathbf{K^L}} \eta_{k',f}$	$\left(\sum_{k' \in \mathbf{K^L}} \eta_{k',f}\right)^2$
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 14: 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 15: 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 16: National owner repositioning CCP: Dummies for proximity 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 17: Local owner repositioning CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages.

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

Table 18: 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 19: 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 20: 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 21: Fixed cost: Table contains estimates of discounts to fixed cost resulting from local cost synergies of 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 22: 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 23: 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				Std. deviation	
		Target format			Variable	Intercept	Variable
		Adult Music	Hits Music	Non-Music	profit		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 24: Market category 1: Estimates of format-switching costs. The table contains to-from format fixed effects for local and national owners.

Owner	Source format	Mean				Std. deviation	
		Target format			Variable	Intercept	Variable
		Adult Music	Hits Music	Non-Music	profit		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 25: Market category 2: Estimates of format-switching costs. The table contains to-from format fixed effects for local and national owners.

Owner	Source format	Mean				Std. deviation	
		Target format			Variable	Intercept	Variable
		Adult Music	Hits Music	Non-Music	profit		profit
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 26: Market category 3: Estimates of format-switching costs. The table contains to-from format fixed effects for local and national owners.

Owner	Source format	Mean				Std. deviation	
		Target format			Variable	Intercept	Variable
		Adult Music	Hits Music	Non-Music	profit		profit
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 27: Market category 4: Estimates of format-switching costs. 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 28: 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 for the counterfactual and observed regimes. A positive number means that the counterfactual regime yields a higher value.

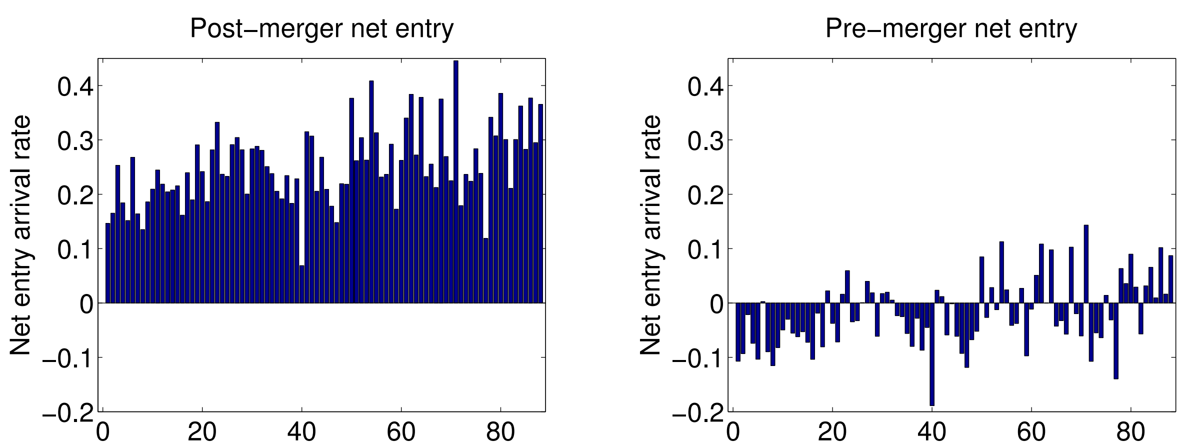

Figure 1: Post- and pre-merger net entry rates by market. Markets are sorted by population, from the smallest (Bismark) to the largest (Los Angeles).

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 29: 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 for the counterfactual and observed regimes. A positive number means that the counterfactual regime yields a higher value.

B Static payoffs

This section of the appendix contains a discussion of the data and estimation procedure used to obtain the 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 firms. The first data set covers the years 1996-2001 and the second, 2002-2006. I combined these two data set to form a large panel for 1996-2006. SQAD provides a data set on average prices per rating point (cost-per-point or CPP) for each market and half year, grouped by demographics and time of the day. Unfortunately, SQAD does not provide data on station-level per-listener pricing. However, because the pricing is done on a per-listener basis, one can still compute a station-level price of an advertising slot by multiplying the CPP by the station rating. According to 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 may have different listenership pools and, consequently, the CPPs for different stations can vary. I alleviate this concern by computing a proxy for a station-level CPP. In particular, I take a weighted average of prices by demographics and time of the day, where the weights are the relevant ratings of the station. In so doing, I assume that stations with most of their listenership during a particular time of the day set a price that is close to the market average for that time. Although this estimate of station-level prices is not perfect, it produces a considerable amount of variation within the market. Subsequently, I use these price proxies to compute station-level advertising quantities by dividing estimates of station revenues (provided by BIA) by a product of prices and ratings. Note that ad quantity computed in this manner may carry some measurement error because it is a function of two estimates. However, if this measurement error is not endogenous – 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 different demographic groups listening to a particular format, I use Radio Today publications, which 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 the 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 these years, I match these averages to data moments for 1996-2006.

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 (2014a)). I conduct the estimation of the model in two steps. In the first step, I estimate the demand model, which includes parameters of the consumer utility θ^L (see equation (5.3)), and in the second step I recover the parameters of the inverse demand for advertising θ^A, $w_{jj'}$ (see equation (5.5)) and marginal cost parameters θ^C (see equation (5.6))

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 of moment conditions 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 et al. (1995). I use instruments for advertising quantities because these quantities are likely to be correlated with unobserved station quality. My instruments include the lagged mean and second central moments of competitors' advertising quantity, lagged market HHIs and lagged numbers and cumulative market share for 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 its predicted empirical counterpart $\hat{R}_{fc}$. Formally, I use an unconditional moment $E[\hat{R}_{fc} - R_{fc}|\theta^L] = 0$, and 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) 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 the 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' theorem separately for each market, assuing knowledge of the market proportion of the 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 differ 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 for all parameters in θ^{Cm}: for small (up to 500 people), medium (between 500 and 1500), and large (more than 1500) markets. To avoid having a full set of dummies and to facilitate identification, I set time dummies for years 1996 and 1997 to zero. 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 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 a two-stage least squares procedure which employs the following instruments: the number of stations in the same format and the ad quantities of competitors. Additionally, the instruments were lagged one period to control for potential serial correlation in η.

C Time aggregation

The time aggregation is performed 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 the 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 $\bar{\omega}$, which encompasses all infeasible states. In practice, constructing the set of feasible states Ω^h might be computationally expensive. In this study, I employ a backward induction recursion that constructs and examines all feasible paths for the industry between Ω^{h0} and Ω^{h6}. For example, a computation of feasible paths for seven mergers and three repositioning events within a 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 study is feasible because the process was parallelized. Despite a long preparation time, this computation must be done only once for each data set. The final augmented state space is 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 $\mathrm{CCP}^A(\mathcal{J}^t)$ or repositioning $\mathrm{CCP}^R(\mathcal{J}^t)$. 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 the Ω^{hi}s are unobserved by the econometrician and must be integrated out.

Denote the time that passed since the beginning of h by $s \in [0,6]$, and 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 eliminating the infeasible paths. Note that $\iota^h(0)$ is a degenerate distribution at Ω^{h0}. First, I compute the distribution after the initial month, $\iota^h(1)$, by numerically solving a Chapman-Kolmogorov system of differential equations

$$\frac{d\iota^h(s)}{ds} = \iota^h(s)Q^h, \tag{C.1}$$

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 $L(\Omega^{h2}|\Omega^{h1}, \Omega^{h0})$.[8] For this purpose, I compute $\iota^h(2)$ by solving equation (C.1) 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. By repeating the procedure for every h the expected likelihood of the data is obtained.

D 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

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

[8]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}.

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 simply an average of all draws, and the second component $V_k^{(F)}$ is a discounted sum of fixed costs. The sufficient statistic to compute this cost is a matrix

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

Fixed cost can be obtained using

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

The third component consist of sum of discounted acquisition shocks, and it can be decomposed into

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

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 statistics. An extra five statistics 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 this figure, I make use of 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.