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

Przemysław Jeziorski[†]

April 15, 2022

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, enabling the regulator to obtain dynamically robust welfare comparisons. The paper demonstrates the framework's application to regulate the U.S. radio broadcasting industry. In particular, it investigates the long-run efficacy of two commonly used merger heuristics: radio station ownership caps (see Telecom Act (1996)) and static merger simulations. The paper finds that raising the ownership cap results in higher total welfare and demonstrates varying long-run effectiveness of merger simulations.

[*]I would like to thank: Lanier Benkard, Luis Cabral, 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

The antitrust law is set to protect the market economy by promoting competition. An essential part of the law are horizontal merger enforcement rules. The landmark legislation, the Clayton Act of 1936, prohibits mergers and acquisitions whose effects "may be substantially to lessen competition." Applying this rule involves resolving a trade-off between market power and cost efficiencies resulting from consolidation (see Williamson (1968)). A potential horizontal merger is evaluated against the law before it can be executed. This evaluation has several implications. First, no mergers can be finally implemented that violate the rules. Thus, in the short run, changes in the rules directly impact the market structure by altering the set of executed mergers. Second, firms have incentives to alter their conduct in response to changes to the merger rules. Because regulatory challenges to the mergers are costly, firms are likely to propose mergers strategically, balancing the profit and the likelihood of clearing regulatory hurdles. Firms are also likely to reposition their product portfolio to amplify the future benefits from ownership consolidation. Thus, in the longer run, changes in the rules are likely to indirectly impact the market structure via an altered set of proposed mergers and subsequent changes in product characteristics. In this paper, I propose a framework to obtain a forward-looking assessment of mergers enforcement rules that incorporates changes in firms' conduct by endogenizing mergers and product characteristics.[1] The framework is applied to evaluate the efficiency of merger enforcement in the U.S. radio industry after enacting the 1996 Telecom Act. The modeling approach is general and applicable to most industries with multiproduct firms.

The 1996 Telecom Act was a major change in merger enforcement rules in the radio industry. The Act abolished a national ownership cap of 45 radio stations and nearly doubled market-level ownership caps. This change was followed by thousands of mergers and resulted in a rapid ownership consolidation. As a result, between 1996 and 2003, the number of radio station owners dropped by 25%. In 2003, the largest player in the market, Clear Channel, owned more than 1250 radio stations nationwide, increasing from only 45 stations in 1996. The vast number of mergers is not the only aftermath of the Act. Between 1996 and 2006, nearly 10% of radio stations per

[1] Lyons (2002) notes that the endogeneity of mergers affects the choice between enforcement rules. Lyons shows that if the mergers are endogenous, the regulator may want to commit to a consumer surplus approval criterion instead of a total surplus criterion, even if the goal is to maximize the total surplus.

year significantly altered the type of their programming, further reshaping the competitive landscape for listeners and advertisers (see Leeper (1999)). The Act aimed to enable the realization of cost efficiencies created by a joint operation of multiple radio stations. However, the legislation 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 complicates an assessment of whether further deregulation would be socially beneficial. Moreover, the stations acquired in the previous year are twice as likely to change their programming compared to stations that did not change their owner. This gap in repositioning frequency suggests a direct relationship between ownership consolidation and repositioning, implying that mergers and repositioning should be studied jointly. This relationship has been documented by Berry and Waldfogel (2001), and Sweeting (2009), who find that post-1996 radio mergers increased product variety.

Jeziorski (2014a) examines the impact of post-1996 mergers and repositioning on consumer surplus within the static lens. He shows that the changes in market structure following the Act increased listener surplus and decreased advertiser surplus. Jeziorski (2014b) supplements these results with estimates of fixed cost synergies from mergers to show that the post-1996 consolidation raises total surplus. The estimates are obtained by endogenizing mergers and repositioning and obtaining the level of fixed cost synergies that rationalize the firms' actions observed in the data. Despite its usefulness for the estimation of fixed cost synergies, the model postulated by Jeziorski (2014b) is not computable. As a result, the mergers and repositioning are assumed to be exogenous when evaluating counterfactual merger enforcement policies. This paper relaxes this assumption and develops a computable model that allows mergers and repositioning to be endogenous to changes in the regulatory regime. The paper asks two questions. First, it aims at a fuller understanding of the implication of the 1996 Act by comparing the factual industry path under the Act with the hypothetical path that would occur without the Act. Second, it aims to evaluate merger policies that are not observed in the data. Again, the policy prescription compares the counterfactual industry path for a given novel policy with the factual path. However, since the hypothetical path under the novel policy is unobserved, the static merger simulations are not directly applicable. A computable model in which market structure is endogenous becomes necessary.

This paper confirms the earlier results that post 1996 deregulation increases total surplus.

In particular, in the long run, under the factual industry path with the 1996 Act, the producer surplus is greater by 10.1%, listener surplus is greater by 0.07%, and advertiser surplus is lower by 1.7%. The impact on advertiser welfare is significantly smaller than Jeziorski (2014a), which reports a 21% drop. This gap is caused by the difference in assumptions between the two papers. Approximately half of the gap can be explained by a more granular treatment of programming formats in Jeziorski (2014a). The remaining half of the gap is caused by a more complete treatment of dynamics in this paper. For example, the paper demonstrates that in 72% of markets, the mergers are followed by product competitors' repositioning into higher-margin product space. This post-merger repositioning intensifies the competition mitigating the impact of the merger. Further, it is demonstrated that some of the extra product variety observed after 1996 would have occurred in the hypothetical world without the Act. Thus, the static analysis overestimates the Act's impact on consumer welfare via a variety provision.

Next, the paper evaluates the efficiency of counterfactual regulatory regimes. It explores raising the local ownership cap from between three to five FM stations (depending on the market size), as specified by the Act, to a uniform cap of seven FM stations. This change would lead to an additional 4% increase in producer surplus, 0.01% increase in listener surplus, and 1% decrease in advertisers surplus. This impact of this further deregulation is in the same direction but smaller than the impact of the Telecom Act.

The policy changes based on increasing local ownership caps increase total surplus; however, the increase is predominantly from greater profits to radio station owners. Thus, a regulatory agency could consider complementing the increase in the ownership caps with measures that directly focus on consumer surplus. I consider two policies that directly target consumer surplus and rely on merger simulations, following Nevo (2000). In the first policy, the agency raises the ownership cap to seven FM stations and rejects all mergers resulting in a lower static listener surplus. This policy leads to a 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 resulting in a lower advertiser surplus. Such policy presents only a short-term solution to the decrease in advertiser surplus. 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. The result is nearly the same as using the policy without any advertiser surplus provisions. This counterintuitive result occurs because the policy based on static merger simulations presumes that the product characteristics are perpetually constant. As a result, the companies, being forward-looking, strategically propose acceptable mergers to the regulator today and reposition the products after the merger is approved. As a result, the companies can extract advertiser surplus, circumventing the simulation-based antitrust criteria in the long run.

The closest empirical work to this paper is Benkard et al. (2010), which studies 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. Finally, in another paper, Mazzeo et al. (2012) argues that post-merger product repositioning may significantly alter the welfare assessment of mergers if competitors are likely to reposition into the profitable niches. This concern is also contained in Section 9 of Horizontal Merger Guidelines (2010).

The model is an extension of Arcidiacono et al. (2016), which studies endogenous product characteristics. The extension consists of adding endogenous mergers.[2] The approach to modeling endogenous mergers is a natural continuous-time generalization of Rubinstein (1982) bargaining model. 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 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 required to conduct a robust dynamic merger review while alleviating the curse of dimensionality caused by many possible merger configurations.

[2]In the case of the radio industry, product repositioning occurs by horizontal changes in the type of programming, ex., changing from a "talk" format to a "rock" format, as in Sweeting (2013). However, for markets where vertical characteristics, such as product quality, are more critical, the endogenous mergers can be instead coupled with R&D models similar to Kryukov (2008), or Doraszelski and Judd (2012).

Gowrisankaran (1999) was one of the first papers that applied a dynamic oligopoly model with endogenous mergers modeling to obtain merger counterfactuals. Using a series of computational experiments, he demonstrates that static models of merger enforcement, which presume that mergers are exogenous, may lead to different conclusions than dynamic models with endogenous mergers. An article by Igami and Uetake (2020), developed concurrently with this paper, conducts an empirical evaluation of mergers in the Hard Drive industry. Both Gowrisankaran (1999) and Igami and Uetake (2020) employ a discrete-time approach. Discrete-time requires rules on move sequencing to break the curse of dimensionality caused by the multitude of possible merger configurations. The viability of such an approach depends on the applicability of the sequencing rules to describe the reality of a particular industry. For example, Gowrisankaran (1999) presumes that mergers occur from the largest to the smallest. This sequencing rule is a viable alternative for markets where large players have notable institutional advantages of executing mergers more rapidly. Igami and Uetake (2020) apply a different approach and permit only one merger per period while assuming that the right to merge is awarded randomly. Such a sequencing rule assumes that the probability of obtaining the ability to merge is always decreasing in the number of active players. This assumption is a somewhat less natural choice to model radio broadcasting and other markets that experienced a structural shift from dispersed to concentrated ownership. In radio, we observe a rapid industry consolidation in the first few years after the Act, when the number of players is large, and muted merger activity in later stages, as the number of players decreases. This result, if anything, suggests an opposite relationship between the number of active players and the ability to propose mergers. The continuous-time approach offers a mathematically elegant solution to the curse of dimensionality without imposing complicated sequencing rules and without tying move frequency to the number of active players.

This paper is related to a numerical study by Mermelstein et al. (2020), which analyzes a two-player endogenous merger game with R&D investment. Similarly to Gowrisankaran (1999), they use a discrete-time dynamic oligopoly model, but they opt for a co-operative Nash bargaining instead of an alternative-offer non-cooperative bargaining game. A vanilla Nash bargaining framework accommodates only two bargaining counterparts. Thus, an alternating-offer game is perhaps a more natural choice for the radio industry, with many active radio station owners who can become potential acquirers.

The empirical results in this paper build on a vast game theoretical literature on mergers pioneered by Farrell and Shapiro (1990). The most closely related studies examine mergers within a dynamic framework. For example, Nocke and Whinston (2010) show that myopic merger rules are dynamically optimal when applied in a simplified environment with homogeneous products.Cabral (2003) examines the impact of free entry on the outcomes of the merger between two firms in a spatially differentiated oligopoly. 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 the complete 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. The treatment of the regulator in this paper is similar to Motta and Vasconcelos (2005). Another related paper Marino and Zábojník (2006) studies the post-merger entry, which is similar to product repositioning introduced in this paper.

The proposed model is 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 complements. Another related study is Mazzeo et al. (2012), who numerically demonstrate that post-merger repositioning can significantly alter the welfare assessment of the merger.

2 Data and industry background

Between 1996 and 2006, the radio broadcasting industry in the United States underwent major structural changes. Before 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 1996 Telecom Act significantly relaxed these restrictions. 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 and 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.[3] The number of frequencies does not change significantly over time; thus, most entry is executed by acquiring 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 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, to model the demand side, the 300 local markets might be regarded as distinct, although the existence of some cross-market fixed cost synergies is possible.

Before 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 by the number of allocated frequencies, as described in Table 1. The 1996 Telecom Act abolished the national cap, nearly doubled the local cap. Over the following decade, the industry moved from fragmented local ownership (about 1.64 stations per owner in 1996) to a market in which fewer than 10 parent radio companies dominating two-thirds of both revenues and listeners nationwide (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 one can capture the first-order magnitude of market power and cost synergies by examining a handful of prominent owners.

[3]Radio is convenient to study mergers since its geographically separated markets result in little multimarket contact. There is some overlap between radio markets, but those were excluded using the procedure developed by Sweeting (2013).

The consolidation of ownership mentioned above triggered an extensive product repositioning. 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 biannual basis by Arbitron, a consulting company, and are frequently utilized as marketing tools for targeted advertising. Notably, the formats frequently change, with approximately 10% of all stations switching formats annually. 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 impact the antitrust policy evaluation significantly and should be incorporated into the analysis. This point is demonstrated numerically in the online appendix. The data used to estimate a dynamic model covers the period from 1996-to 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. To avoid modeling cross-market interactions, I drop the overlapping markets following the method of Sweeting (2013); that is, I drop markets "were more than 6% of listening was to stations based in other markets". I also drop markets that do not have data on advertising prices. The online appendix contains the details on the price/quantity data used for static estimation. The remainder 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, and Viacom, which own a complicated network of stations nationwide. I allow these companies to own multiple stations in local markets and 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 the companies composes 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.

I label three active companies with the largest national revenue share in 2006 (the last year of the data set) as dominant owners in each local market.[4] Consequently, each local market may 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 entered through acquisitions.

In addition to tracking dominant owners, I label 22 of the remaining radio stations with the highest local listenership share as local owners in each local market. 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. AM stations become important in rural markets with less than 15 active stations, so I allow both AM and FM stations to be outside of the fringe.

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 on 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 seldom acquired by larger owners, as indicated in the data. Thus, while increasing the complexity of the estimation, modeling the forward-looking decisions of every small owner has little benefit.

[4]Labeling a subset of firms as national using future growth introduces an implicit assumption that all players have correct expectations about which firms will become large. This distinction is quite evident in the radio industry because firms such as Clear Channel, ABC, or Viacom were expected to execute many merger transactions after the deregulation. In other industries, such distinction may be less clear.

Nevertheless, dropping the smallest owners altogether is unrealistic because they collectively affect markups in the pricing game. Therefore, a realistic compromise is fully tracking only large and medium owners and partially tracking the smallest firms.

One artifact of not allowing smaller players to make merger bids is the prohibition of spin-offs. I observe spin-offs in the data, but they are primarily 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 merged entity's market power. I use this convenient fact and ignore the acquisition of stations 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, Rock, and Country, (ii) "Hits Music," containing such formats as Contemporary Hit Radio, Urban, and Alternative, and (iii) "Non-music," containing such formats as Talk, News, Ethnic and Religious formats. 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 younger crowd. The aggregation trades off static realism for dynamic realism. Namely, I sacrifice 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 present the model of the radio industry and come back to this issue when discussing the results.

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 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.

The static profit function π is obtained by solving an advertising pricing game that follows Jeziorski (2014a). The game is fully described in the online appendix, and I include its main features here for completeness. The radio station owners choose advertising quantities for each station to maximize the joint profit from selling advertising in the advertising game. The advertising quantities and programming formats $\mathcal{J}^t$ result in endogenous listenership shares. The listeners are strategic when choosing the radio stations to listen to and have preferences over formats of programming, as well as the amount of advertising. Listenership shares and advertising quantities are supported by advertising prices that clear the advertising spot market. The advertising game is assumed to be in Nash equilibrium, in which radio station owners choose advertising quantities simultaneously.

The model is composed of two parts: (i) a model of endogenous mergers in which acquisition prices are determined by non-cooperative bargaining game, and (ii) a model of product repositioning, which is a continuous time version of Sweeting (2013). Both of these models are versions of dynamic discrete choice (see Miller (1984), Pakes (1986), Rust (1987), and Wolpin (1984)). Mergers and product repositioning can happen concurrently, with mergers leading to subsequent repositioning and *vice versa*. This feature makes this model distinct from alternative static models and from two-period models in which a repositioning stage follows the merger stage. To simplify the exposition, I discuss mergers and repositioning separately.

3.1 Mergers

The merger model can be summarized as a multi-agent version of a Rubinstein (1982) model with a random right to move. Each company is associated with a Poisson process that determines the arrival of the right to acquire a competitor. These Poisson processes are independent and share a common arrival rate λ^A. Once the company obtains the right to acquire, it chooses an acquisition target that maximizes its discounted stream of profits. Before making this choice an acquirer observes a vector of stochastic merger costs denoted by $\zeta^{A,t}(\cdot)$, where $\zeta^{A,t}(k')$ is the cost of acquiring competitor k'. I consider a flexible form of merger costs:

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

The term $\mu_k^A(\mathcal{J}^t, k')$ is a persistent part of the cost, which is a deterministic function of the industry state and is commonly known. On the other hand, the term $\epsilon_k^{A,t}(k')$ is an idiosyncratic stochastic shock to the cost and is private information of the acquirer.

The acquirer makes a single take-it-or-leave-it merger proposal that maximizes the continuation value or offers no merger. The potential acquiree instantaneously accepts or rejects the offer taking into account his opportunity cost. However, once the offer is rejected, it cannot be reconsidered unless a similar offer is extended in a future period.[5] If the merger offer is accepted, it is reviewed by an antitrust authority. During the review, the regulator blocks the merger of k and k' with a commonly known probability $\mathbf{G}(\mathcal{J}^t, k, k')$.[6] If the merger fails the review, it is not implemented, and neither party incurs merger costs. If the merger passes the review, the merger offer is implemented; that is, the following 3 things co-occur: (i) the acquirer incurs merger cost $\zeta^A(k')$, (ii) the merger bid P are transferred to the acquiree, and (iii) the companies merge their portfolios of products. The industry continues with a new market structure.

A significant advantage of a continuous-time bargaining model is that it resolves the issue of merger conflicts arising in discrete-time models. 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 of the deal k execution. Over a short period of time Δ, the probability of

[5]This assumption is challenging to relax because withholding offers would require a model in which multiple offers can be made at the same time. This assumption should not be pivotal if the frequency of making offers/counteroffers is assumed to be sufficiently high.

[6]I focus on the commitment of the regulator to the policies, that is, the regulator cannot renege on the announced policy at any point. I also focus on Markov policies of the regulator, which is without any loss to efficiency if the firms play a Markov Perfect Equilibrium.

execution of an attempt a_k is equal to:

$$\lambda^A\Delta(1-\lambda^A)\Delta + \text{CON}_{k,k'}\lambda^A\Delta\lambda^A\Delta + O(\Delta^2) = \lambda^A\Delta + O(\Delta^2)$$

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 model component 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)).

The model of mergers allows for entry through the acquisition of other active firms. Potential entrants are firms that hold empty portfolios of stations; that is, $\omega_k^t = \overrightarrow{0}$. [7] Modeling of entry and exit through acquisitions 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. Due to 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 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 and re-entry. Similar to entry, exit is modeled as selling off all owned stations.

Bargaining with take-it-or-leave-it offers determines the split of bargaining power depending on the outside option of the acquiree. This outside option is determined by many dynamic factors, which include: (i) the value of waiting until the market becomes more concentrated, (ii) the value of making a counteroffer in the future, and (iii) the value of repositioning. The value of the outside option inherently affects bargaining power. The larger the acquiree's outside option, the more significant part of the merger surplus they receive. For the industries with few large companies and many relatively homogenous smaller companies (radio industry included), the model is likely

[7] One way to allow the possibility of a more traditional entry is to endow the type space $\mathcal{F}$ with an inactive state. This extension is possible but is not implemented because traditional entry is insignificant in the radio industry.

to allocate most of the surplus to the large acquirers. However, in the industries where the number of firms is small, and the frequency of offers is large, the model can allocate a significant amount of the bargaining power to the acquirees. In this sense, a cooperative alternative, such as the Nash Bargaining Solution, should be regarded as a reduced form of this model (see (Collard-Wexler et al., 2019)).

The above model of mergers is versatile and can incorporate alternative methods of soliciting merger offers. One alternative is a model in which nature chooses an acquisition target. After being selected, the target is auctioned off and acquired by the highest bidder. Such an auction process is relatively straightforward to incorporate within the current framework; however, it may be inappropriate to analyze the radio industry. There are many potential acquisition targets in the radio industry, and acquisition bidding wars are rare. Consequently, if the move (auction) frequency λ^A is large enough, a large proportion of auctions will have only one bidder. In such a case, the auction would be equivalent to making a take-it-or-leave-it offer.

Another possible extension of the model is allowing the move arrival rate λ^A to be different (and possibly larger) when making a counteroffer. Such extension is theoretically possible but has insignificant consequences if the overall frequency of moves is large.

3.2 Repositioning actions

Repositioning is modeled as a dynamic discrete choice with random arrival of a right to reposition. In addition to a Poisson process gathering merger actions, each company is endowed with a Poisson process with an arrival rate λ^R, which determines the timing of possible repositioning actions. Upon the arrival of the right to reposition, the company observes a vector of repositioning costs

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

where $r = (f, f')$ and represents repositioning a product from type f to type f'. Similarly to merger costs, μ_k^R is a deterministic function of the state and is common knowledge, while $\epsilon_k^{R,t}$ is a private, idiosyncratic shock. After observing the cost, the company takes a single repositioning action r or no repositioning action. The repositioning is instantaneous, and the new state of the industry becomes common knowledge. One cannot delay a repositioning action.

3.3 Timing

As mentioned before, the merger and repositioning actions are interrelated and are executed concurrently. The joint model of merger and repositioning 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.
- (iv) In case of acceptance, the merger review is conducted, and the new market structure is determined. The flow returns to (1).

(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 immediately decides 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 the stage (1).

3.4 Strategies and equilibrium

This section contains a formal definition of the strategies and defines an 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: $\mathbf{a}_k(\mathcal{J}^t, \zeta^{A,t}) \in \{0, \dots, K\}$. 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 $\zeta_k^{R,(m)}$ cost 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 following equation gives the value function for company k (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 the strategies maximize a stream of discounted profits at any state,*

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

The above equation states that each player best responds to the opponents' strategies and a pre-announced enforcement rule. It includes a requirement that 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, that is,

$$\mathbf{P}^*_k(\mathcal{J}, \zeta^A_k, k') = V_{k'}(\mathcal{J}; \mathbf{g}^*, \mathbf{G}); \quad \forall k' > 0, k, \mathcal{J}, \zeta^A_k.$$

Under this condition acquisition events do not affect the acquiree's value function and 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 when receiving a merger bid, and (ii) acquirees receive only one merger offer at a time, almost surely. In such a case, knowing there are no other outstanding offers, 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.

The discussion of the existence of an equilibrium and computational strategy is contained in the Appendix B. In summary, similarly to dynamic discrete choice in discrete time, one can express the merger and repositioning strategies in terms of instantaneous conditional choice probabilities, denoted by $\text{CCP}^A(\mathcal{J})$, and $\text{CCP}^R(\mathcal{J})$. These probabilities gather merger and repositioning actions conditional on the arrival of the right to move. After this reformulation, one can directly apply the existence result and computational tools from Doraszelski and Judd (2012).

Section 2 introduces three types of firms: national firms, local firms, and fringe firms. These firm types can be accommodated as restrictions on model parameters. In particular, setting large enough acquisition costs for local firms would preclude them from acquiring other firms. Similarly, fringe firms cannot acquire and reposition, which one can accommodate by assuming large enough acquisition and repositioning costs.

Dividing owners into the groups mentioned above has some significant consequences. The upside is that it captures the essential features of the radio market and reduces the complexity of the estimation. In particular, it enables estimating acquisition price (value function of the acquiree) without tracking the possibility that the acquiree can make merger offers himself. Consequently, 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.[8]

4 Estimation

The road map of the estimation is as follows. First, I pre-estimate the static profit function using the data on radio station market shares, formats, ownership, and advertising quantities and prices. The procedure closely follows Jeziorski (2014a) and delivers a static profit function $\pi(\cot)$ for any state of the industry. The profit function is obtained by solving a Nash equilibrium of the static advertising game. Second, I use the static profit function to estimate the dynamic model.

The estimator used to estimate the dynamic model 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 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 described by equations (B.2) and (B.3) in the Appendix B. An IPML is a modification 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. (2016), 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 $\{(g_h, \mathcal{J}_h) : h = 1, \dots, H\}$, 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 (B.2) and (B.3); that is,

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

[8] 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.

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 may 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

$$Q_H(\theta) = \sum_{h=1}^{H} \log \mathrm{CCP}(g_h | \mathcal{J}_h; \hat{V}(\cdot;\theta), \theta). \tag{4.1}$$

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 four remaining parts of this section. The first part describes the pre-estimation of a one-shot profit function; the second contains the description of the state space of the model; the third explains the estimation of acquisition and repositioning strategies; and the fourth describes the simulated pseudo-likelihood estimation of structural parameters.

In the remainder of the paper, I presume that the static profits $\pi(\mathcal{J})$ are known. The estimation procedure to obtain $\pi(\mathcal{J})$ is very similar to the one employed by Jeziorski (2014b). The details of this procedure are contained in the Online Appendix.

4.1 description of the state space

After the simplifications, the state variables are the market's identity (because the profit functions vary from market to market), station portfolios of 3 dominant owners, and station portfolio of fringe owners. I solve the game market by market because the markets are assumed to have no interactions. Table 3 shows the number of states of the game with post-1996 ownership caps and in the game with a uniform ownership cap of 7 stations. The largest market in the game with post-1996 ownership caps has 21 million states, and the smallest has about 700,000 states. The number of states increases by order of magnitude when the ownership caps increase to 7. In particular, the largest market has nearly 100 million states.

I find that solving the game market by market does not provide direct savings in terms of computation time because evaluating the Bellman equation is relatively cheap and the number of these evaluations grows linearly in the overall number of states for all markets. However, separating markets is still crucial for the feasibility of the exercise because it enables significant memory savings. In particular, memorizing a value function for one market at a time, lowers the requirements from approximately 800GB to approximately 17GB of Random Access Memory (RAM).

Overall computation time is determined by the total number of states in all markets, which amounts to 4.5 billion. Consequently, despite heavy optimizations, such as using compiled C-code, precomputing profits for all states, and using look-up tables for state transitions, the counterfactual with a cap of 7 requires more than 6,000 Central-Processing-Unit-Hours (CPU-hours) to compute. The algorithm is "embarrassingly parallelizable"; that is, it delivers linear efficiency gain with the number of applied CPU-hours. Consequently, the counterfactual with a cap of 7 requires about a week to compute on a 48 CPUs server.[9] Increasing the number of product types (formats) in my application is infeasible with today's computers; however, it is likely possible in the near future. Also, analyzing more product types for other industries with less active firms than radio should be possible.

Both stages of dynamic estimation involve the imposition of current FCC ownership caps while assuming that there is no other and meaningful antitrust scrutiny. Such assumption largely reflects the reality of the radio industry but can be easily relaxed if needed. Having more antitrust restrictions makes the analysis easier by lowering the set of possible industry configurations (for example, larger ownership caps have higher computational complexity that smaller ownership caps).

4.2 Estimation of acquisition and repositioning strategies

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

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

[9]The server setup used in this paper consists of 48 CPUs sharing the same RAM. Sharing the RAM is crucial because when running on non-shared RAM setups (such as clusters or cloud) the algorithm is not embarrassingly parallelizable. In particular, one needs to accommodate an extra overhead from communicating the new value function to the nodes in each iteration. For this reason, comparable performance (in the year 2014) would require hundreds of cloud nodes instead of 48 units.

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 and 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 with multiple mergers within the same month, the state space when taking 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, while mergers are observed monthly. Therefore, multiple mergers and repositioning actions during the same semi-yearly 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.

Equilibrium CCPs, given by equations B.2 and B.3, depend on the state through unknown value functions requiring semi-parametric estimation. In particular, the acquisition CCPs are given by

$$\widehat{\mathrm{CCP}}^A(k'|k, \mathcal{J}, \theta^A) = \frac{\exp\left\{\Upsilon^A(k, k', \mathcal{J})\right\}}{\sum_{k''} \exp\left\{\Upsilon^A(k, k'', \mathcal{J})\right\}}, \tag{4.2}$$

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 helps express 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 particular 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. But unfortunately, in practice, even though 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, I approximate reposition action with

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

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 must 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, 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, preventing me from computing the full information likelihood function. Instead, one could use a simulated likelihood or a generalized method of moments estimators. 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 of efficiency. Another option is to perform an analytical integration of unobservables using Chapman-Kolmogorov equations describing state transitions, as suggested by Arcidiacono et al. (2016). 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 relatively sparse, it would not be sufficiently sparse to store in the computer memory or recompute "on the fly." So instead, I developed a method of integrating the likelihood based on partial Chapman-Kolmogorov equations. These partial equations take advantage of the fact that only a tiny subset of feasible latent industry states is relevant for the estimation. However, this simplification relies on two facts: (i) I observe all merger events; thus, no latent mergers have to be integrated out, (ii) two repositioning events by the same station are improbable and are assumed away.[10] 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 can only identify the product of the move arrival rate λ and CCPs. However, as long as the actual move arrival rate is not excessively small, I can estimate the first stage by choosing a reference value of λ set to 1. Then, relevant CCPs for the desired value of an arrival rate could be obtained by dividing the estimates by λ.

4.3 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:

$$F_{kj}^{m}(\mathcal{J}^t|\theta^F) = \bar{F}_f^m \times F^S(\omega_{kf}^t, z_k|\theta^F) \times F^E(n_k^t, z_k|\theta^E). \tag{4.3}$$

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.[11] Note that for local owners, F^E and F^S are equal to 1.

[10] I verify this assumption using a subset of the data for which I observe formats quarterly. Notably, none of the stations in this subsample changed format twice or more times in the same half-year.

[11] I include cross-market economies of scale, but I do not model the joint decision to acquire multiple stations across markets. The implications of this assumption depend on the form of cross-market synergies. In my particular case, I use a simple specification in which large owners are assumed to have a cost discount that does not increase in the number of stations owned in other markets. Without loss of generality, one can disaggregate joint decisions to acquire multiple stations into market-by-market acquisition choices. The disaggregation may matter if the multiple

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 it 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[12]

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

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 where a single station has average variable profits greater than \$150,000. The second category has half-yearly variable profits in the range of \$150,000-\$60,000. The third is the \$60,000-\$20,000 range. Finally,

stations are acquired from the same owner and if there are economies of scale while bargaining.

[12] 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.

the fourth is less than \$20,000.[13] Additionally, $\theta^{A,m}$ may vary across formats because of 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 minor (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 temporarily allow the dominant owners to bring announcers from other markets to streamline format switching. However, local owners may have better access to 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 more significant capital investments and access to specialized production factors. Thus I expect dominant owners to have a 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 can depend on the source and target format and vary by 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.

[13] The variable profits are small for two reasons. First, they are an average across a relatively small number of large stations and a more significant number of fringe stations. Second, the variable profits are net of variable costs. Overall evidence on radio station profitability is scarce; According to Federal Communications Commission (2003), Figure VI, the net profit margin of the radio industry was negative in 33% of quarters from 1995 to 2003. Moreover, the radio industry underperforms relative to S&P-500 companies. Small average profits are also consistent with Jeziorski (2014b).

The above specification is used to simulate the value function

$$\begin{aligned}V_k(\mathcal{J}^t|\theta) = \int_{s=t}^{\infty} e^{-\rho s}\pi_k(\mathcal{J}^s)ds - \int_{s=t}^{\infty} e^{-\rho s}F_k(\mathcal{J}^s|\theta)ds + \sum_{l=1}^{\infty} e^{-\rho\tau_k^{A,(l)}} P(a_k^{(l)}, \mathcal{J}^{\tau_k^{A,(l)}}|\theta) + \\ \sum_{l=1}^{\infty} e^{-\rho\tau_k^{A,(l)}} W^A_{a_k^{(l)}}(\text{CCP}^A_k, \mathcal{J}^{\tau_k^{A,(l)}}|\theta) + \sum_{m=1}^{\infty} e^{-\rho\tau_k^{R,(m)}} W^R_{r_k^{(m)}}(\text{CCP}^R_k, \mathcal{J}^{\tau_k^{R,(m)}}|\theta).\end{aligned} \tag{4.4}$$

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:

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

The closed-form solution for the conditional expected value of shocks W is unknown for the number of alternatives 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 (4.1)), obtained using the procedure described in Appendix C. Note that the value function must be simulated for every potentially feasible latent state and 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 comprises $1,000$ 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 a 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.

4.4 Identification

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 a pre-estimated static model can predict the revenues of radio owners in each industry state. 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. Similarly, I identify the merger cost and cost efficiencies from consolidation; I use the convenient fact that entry is possible only through acquisition and choose the merger cost that rationalizes the observed timing of such entry decisions. Similarly, the estimates of the cost efficiencies have to rationalize subsequent acquisitions. The fixed cost is bounded from above, stemming from the fact that, by revealed preference, entry is profitable in all markets. The cost is also bounded from below by the cost discount level required to justify the mergers.

Identifying rate λ separately from other structural parameters is difficult without observing all move opportunities. However, such identification may still be possible with an exclusion restriction that shifts the continuation value but does not affect current payoffs. One candidate for such exclusion is ownership caps. If 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. 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 qualitative and quantitative results.

5 Results

This section reports and discusses the estimates of the model's structural parameters. The static part and first stage estimates are discussed in the Online Appendix. I also discuss the goodness of fit metrics for the first stage estimation.

5.1 First stage goodness of fit

To evaluate goodness of fit of the first-stage estimation I use a pseudo R^2 measures proposed by McFadden (1973). In particular, I compute $R^2 = 1 - \frac{\log(\text{likelihood of a full model})}{\log(\text{likelihood of a benchmark model})}$. This statistic delivers a test for the significance of the first-stage estimation equation, when a benchmark model has constant merger and repositioning propensities, and the full model has the logistic structure (see equation (4.2)). Alternatively, the statistic can be interpreted as a "mean square error" of the "variance" explained by the model, analogous to a Multiple Correlation Coefficient (R^2) in the linear regression. The model used as the first stage in this study can be regarded as a generalization of the static logit model; it comprises discrete choice problems arriving with exponential arrival times. Thus, McFadden's R^2 should be informative about the goodness of fit. Since the merger and repositioning strategies are estimated jointly, I compute a joint R^2 measure using a model with two constants (one for merger strategy and one for repositioning strategy) as a benchmark. The statistics are presented in Table 9. I obtain $R^2 = 0.1$, which indicates that the model can explain about 10% of variation in merger and repositioning activity in the data. The unit of observation is a merger month; thus, it is expected to have a reasonable amount of the residual variance because it is difficult to predict the exact month the merger takes place. Further information can be obtained by looking at the model's fit over time. For example, I observe a sharp decrease in the merger activity after 2000, and it is crucial to verify if the first stage model can reasonably predict this qualitative change. For this purpose, I report the R^2 statistic for each year in the data. The R^2 statistic is always above 0.07 and stays above 0.1 for the pivotal years of 2000 and 2001.

The original McFadden's R^2 is designed to compare nested models; however, it can be extended to compare non-nested models. Such analysis aims to investigate if relaxing certain assumptions of the model would lead to a better fit. A particularly relevant non-nested model is a non-stationary model that allows for time-varying merger and repositioning propensity. Using this model as a benchmark demonstrates how much explanatory power is gained when allowing the primary model to be non-stationary. The second row of Table 9 presents the results of using a model with year dummies for merger and repositioning propensities (20 parameters)[14] instead of using a model with constants as a benchmark. The results suggest that allowing for non-stationarity does not add a significant explanatory power. In particular, The main (stationary) model has more explanatory power than the alternative non-stationary model. Moreover, since the R^2 computed using the model with the constant probabilities is quite similar to the one computed using the non-stationary model, I conclude that allowing a simple time trend does not capture a significant amount of variation in data.

Another way to judge the model's fit is to compare the consumer surplus observed in the data to that predicted by first-stage estimates. The largest difference is observed at the end of the data, 2006. The simulated advertiser surplus in that year is within 2% of the observed listener surplus. The simulated listener surplus is within 0.4% of observed surplus.

5.2 Structural parameters

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

According to equation (4.3), 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

[14]This alternative formulation allows both the arrival of moves λ and move execution probabilities to be non-stationary.

owning multiple stations in the same format $F^S(\mathcal{J}^t_{kf}|\theta^F)$, and a multiplier represents the economies of scale for owning multiple stations of any format $F^E(n^t_k|\theta^E)$.

Table 4 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 6 contains the estimates of the within-market economies of scale resulting from owning multiple stations. 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 insignificant. I also document further cost synergies of operating stations in the same format. According to Table 5, the operation of stations in the same format is additionally 14% cheaper on top of the economies of scale indicated in Table 6.

Table 7 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 relatively high volatility, which is homoscedastic. 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 costs for realized mergers.

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 the market category, source-target format, and being a national owner. Estimates of repositioning for Category 1 are reported in Table 8, and the remaining categories are contained in the Online Appendix. I find that the repositioning cost varies considerably across market categories. In particular, I find higher switching costs 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 patterns 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.

Similarly to the first-stage analysis, we can compare the consumer surplus that is observed in the data to that predicted by second-stage equilibrium CCPs. This should inform us how well the functional form of the cost structure can replicate the patterns in the data. Ten years forward, the simulated advertiser surplus is within 2.6% of the observed listener surplus. The simulated advertiser surplus is within 0.6% of observed surplus.

6 Counterfactuals

Using the estimates of the model's structural parameters, I perform several counterfactuals, which study alternative merger enforcement policies.[15]. To minimize the impact of the estimation error, I compare the simulations of counterfactual policies to a simulation of the current policy (not the actual outcomes).

6.1 Impact of the 1996 Telecom Act

The first set of experiments investigates the impact of the 1996 Telecom Act on producer, listener, and advertiser surplus. Table 10 presents counterfactuals that evaluate the impact of the looser post-1996 local ownership caps. In particular, I recompute the equilibrium merger and repositioning strategies for the pre-1996 ownership caps and compute 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 the producer surplus is 10% lower under the old caps. Approximately 6% of this decrease results from loss of market power and 4% results from lost cost efficiencies. Moreover, deregulation leads to lower

[15] The model of mergers is likely to have multiple equilibria. I use a selection rule that involves starting the best response dynamics algorithm from the current equilibrium in the data. While I cannot exclude the existence of other stable equilibria, I tried many other starting points, including a value function equal to zero, and could not obtain any other equilibrium. Keeping this caveat in mind, all results in the counterfactual section should be interpreted as "there exists an equilibrium such that...", and should not be interpreted as "in all equilibria it holds that...".

advertising supply 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. As a consequence, the realized mergers generate smaller gains in producer surplus. The overall effect is presented in the bottom three rows of Table 10.

It is helpful to compare the findings of dynamic analysis with those obtained using static analysis. In particular, Jeziorski (2014a) finds that for the decade between 1996 and 2006, ownership consolidation decreased advertiser welfare by 21%, whereas the present study determines that decrease stands at 1.7%. Approximately half of the gap can be explained by a more granular treatment of programming formats in Jeziorski (2014a). I investigate two factors that are mismeasured or omitted in static analysis and contribute to the remainder of the difference: (i) post-merger product repositioning, which could correct the negative effect of the mergers; and (ii) equilibrium provision of variety with and without the Telecom Act.

When assessing market power after acquiring a station in a particular format, it is particularly relevant to investigate competitors' incentives to reposition into that format. Namely, because the acquisition generates higher rents for all stations of the focal format (as demonstrated numerically in the Online Appendix), it invites repositioning by competitors. Such repositioning creates more competition and mitigates the effect of the merger. [16]

I measure the repositioning using *net entry rate* defined as the difference between the entry and exit rate of competitors in and out of the focal format. I demonstrate the impact of mergers on entry by comparing the net entry rates before and after acquiring a station of the focal format. I compare the net entry rates using two methodologies: (i) using raw data summarized by first-stage estimates (in the spirit of Benkard et al. (2010)), and (ii) using conditional choice probabilities obtained numerically from a model. These two methodologies of assessing post-merger entry are complementary. The model keeps the other market forces constant beyond the merger in question,

[16] This process is equivalent to the post-merger entry mentioned in Horizontal Merger Guidelines.

similar to running a field experiment. It is desirable when studying mergers because actual field experiments are tough to execute in such a context. However, the quasi-experimental properties come with the cost of relatively strong modeling assumptions.

I compute before and after net entry rates using the first-stage CCPs averaged across mergers observed in the data. 72% percent of markets have, on average, more net entry of stations into a format after the acquisition of a station in that format by a national owner. The average post-merger Poisson net entry rate equals 0.07 per month, which translates to about one year average waiting time for the repositioning of competitors.

I also compute net entry rates using the CCPs computed from the model. I compare average pre- and post-merger net entry on Figure 1 and I find that mergers lead to higher entry rates in all 88 markets. The average Poisson net entry rate predicted by the model amounts to 0.02 per month, translating to about a four-year average waiting time for a post-merger repositioning of competitors. Both results, the one obtained from the raw data and the one estimated model, are mutually consistent and suggest a significant increase in entry after the merger. The dynamic model accounts for such entry, whereas the static model does not.

Next, I examine the effects of the Telecom Act on the provision of variety. The reduced form evidence for the variety provision was provided by Berry and Waldfogel (2001), who found that mergers in radio markets generally lead to more variety. Jeziorski (2014a) further quantifies this effect and shows that the provision of more variety directly affects market power by increasing the surplus of the listeners and lowering the surplus of advertisers. These two papers compute the change in variety using temporal variation in formats before and after the Telecom Act. I supplement their analysis by comparing the provision of variety in the same year with and without the Telecom Act using counterfactual experiments.

I compute a variety index that measures the concentration of formats in the market, defined as

$$\sum_f \left(\frac{\sum_k \omega_{kf}}{N} \right)^2,$$

which resembles the Herfindahl index. If the variety index is 1, all the stations have the same format (no variety). If the variety index is $\frac{1}{F}$, the stations are distributed uniformly across formats (maximum variety). The variety index amounts to 0.45 in 1996 – the first year of my data. I replicate the result by Berry and Waldfogel (2001) and Jeziorski (2014a) by computing the change of variety index overtime under the post-1996 ownership caps. I find that post-1996 mergers lead to increased variety generating a variety index of 0.39 and 0.381 in 2001 and 2006, respectively. I recompute the variety index under the old ownership caps, and I find that the variety index amounted to 0.391 in 2001 and 0.388 in 2006. Relative to the changes in the variety index over time (comparing 0.45 in 1996 to 0.39 in 2001), counterfactual analysis reveals a negligible impact of the Telecom Act on variety 5 years after it was introduced (comparing 0.39 and 0.391) and a moderate impact ten years after it was introduced (comparing 0.381 to 0.388). Thus, the static analysis, which uses only temporal variation, overstates the impact of the Telecom Act on variety. Mismeasurement of the variety provision results in overstating antitrust implications on the Telecom Act.

6.2 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 welfare measures. 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)[17] and present the results in the first three rows of Table 11. 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. Market power is exercised predominantly on advertisers, which lose about 1% of their surplus in the long run. 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. In the first five years after moving to CAP7, the companies exercise market power on listeners. However, in 10 to 20 years, the firms will extract surplus from 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 align with the previous literature on retrospective post-merger repositioning in the radio industry. In

[17]Some results of the CAP7 counterfactuals rely on the functional form extrapolation of the cost synergies that are not observed in the data because they violate current ownership caps.

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, before this study, the applicability of these results to hypothetical and out-of-sample policies such as CAP7 had not been established.

Next, I imposed additional antitrust criteria and recomputed the equilibrium of the dynamic game. Rows 4 to 6 of Table 11 present the results of experiments in which mergers that decrease static listener surplus are forbidden. On the one hand, the policy successfully selected 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 minor increase in producer surplus. However, because the executed mergers are generally more cost-efficient than 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 11. Contrary to the listener surplus policy, the advertiser surplus policy is unsuccessful in preventing mergers that harm advertisers in the long run. On the other hand, 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 accepted mergers. Moreover, contradictory to the static intuition, the static advertiser surplus criterion delivers a worse outcome for advertisers than the static listener surplus criterion. This difference demonstrates that myopic merger policy can be dynamically suboptimal and have somewhat counterintuitive long-run consequences.

7 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 result from a dynamic bargaining process. This model provides a tool to conduct an antitrust merger review that accommodates the distortions created by the endogeneity of mergers, entry/exit, and product repositioning.

I demonstrate the applicability of the procedure by examining the wave of consolidation in the U.S. radio broadcasting industry from 1996 to 2006. I find substantial fixed and marginal cost synergies from joint ownership. Namely, operating multiple stations within the local markets is cheaper than operating them individually. Additionally, significant cost synergies arise from operating multiple similar stations within a local market. Specifically, the operation of two stations by the same owner in a given local market is up to 14% cheaper than operating the same stations by different owners. Moreover, if the two stations 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 the total surplus by raising producer surplus and generating a negligible impact on listener and advertiser surplus. The small impact of the Act on advertisers contrasts with the significant 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, increasing ownership caps to seven stations increases the 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 do not sufficiently safeguard against significant 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 result from the fact that companies can circumvent the welfare rule by proposing a merger acceptable to the myopic regulator. Also, the companies alter the product characteristics to extract advertiser surplus after the merger. More generally, I show that in the industries where dynamic processes such as entry/exit or product repositioning are prevalent, the companies can 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 the merger this rule allows and less by the exact types of the merger this rule permits.

References

Aguirregabiria, V. and A. Magesan (2013): "Euler equations for the estimation of dynamic discrete choice structural models," in *Structural Econometric Models*, Emerald Group Publishing Limited.

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. Bayer, J. R. Blevins, and P. B. Ellickson (2016): "Estimation of dynamic discrete choice models in continuous time with an application to retail competition," *The Review of Economic Studies*, 83, 889–931.

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.

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. 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*.

Collard-Wexler, A., G. Gowrisankaran, and R. S. Lee (2019): ""Nash-in-Nash" bargaining: A microfoundation for applied work," *Journal of Political Economy*, 127, 163–195.

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.

Federal Communications Commission (2003): "Review of the Radio Industry," Tech. rep.

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.

Igami, M. and K. Uetake (2020): "Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996–2016," *The Review of Economic Studies*, 87, 2672–2702.

Jeziorski, P. (2014a): "Effects of Mergers in Two-Sided Markets: The US Radio Industry," *American Economic Journal: Microeconomics*, 6, 35–73.

——— (2014b): "Estimation of Cost Efficiencies from Mergers: Application to US Radio," *The RAND Journal of Economics*, 45, 816–846.

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.

Marino, A. M. and J. Zábojník (2006): "Merger, Ease of Entry and Entry Deterrence in a Dynamic Model," *The Journal of Industrial Economics*, 54, 397–423.

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

McFadden, D. (1973): "Conditional logit analysis of qualitative choice behavior," *Frontiers in Econometrics*, 105–142.

Mermelstein, B., V. Nocke, M. A. Satterthwaite, and M. D. Whinston (2020): "Internal versus external growth in industries with scale economies: A computational model of optimal merger policy," *Journal of Political Economy*, 128, 301–341.

Miller, R. A. (1984): "Job matching and occupational choice," *The Journal of Political Economy*, 1086–1120.

Motta, M. and H. Vasconcelos (2005): "Efficiency gains and myopic antitrust authority in a dynamic merger game," *International Journal of Industrial Organization*, 23, 777–801, merger Control in International Markets.

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. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54, 755–84.

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. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, 999–1033.

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

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.

Wolpin, K. I. (1984): "An estimable dynamic stochastic model of fertility and child mortality," *The Journal of Political Economy*, 852–874.

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.

	Number of markets having states within							Total states
	0-1M	1M-5M	5M-10M	10M-20M	20M-50M	50M-90M	90M-100M	
Post 1996	5	35	14	12	22	0	0	792,557,073
Cap 7	3	7	8	11	14	18	27	4,577,183,777
	Average computation time in CPU-hours							Total time
States	0-1M	1M-5M	5M-10M	10M-20M	20M-50M	50M-90M	90M-100M	
CPUh	1.4	5.7	10.7	19.3	47.7	88.6	138.5	Post 1996: 1,098 Cap7: 6,342

Table 3: The table shows the number of markets having number states in a particular interval, the total number of states for all markets, and average equilibrium computation time.

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 4: Fixed cost of owning one station in each market (half-yearly, millions USD).

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 5: 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 6: 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 7: Estimates of the acquisition cost (in millions USD). The table contains an intercept of the mean and standard deviation of the acqusition distribution. It includes market-size fixed effects (relative to the smallest category) and a coefficient on the static variable profits of an acquistion target.

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

Standard errors in parentheses

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

Table 8: Market category 1: Estimates of format-switching costs (in millions USD). The table contains to-from format fixed effects for local and national owners.

	All years	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Pseudo R^2 constant probabilities	0.10	0.07	0.11	0.12	0.12	0.11	0.09	0.08	0.09	0.08	0.10
Pseudo R^2 year dummies	0.08	0.05	0.11	0.11	0.10	0.07	0.10	0.05	0.09	0.07	0.08

Table 9: Pseudo R^2 measures for the first stage estimation, defined as $1 - \frac{\log(\text{likelihood of a full model})}{\log(\text{likelihood of a benchmark model})}$. I use two benchmark models: (i) a model with a constant probability of acqisition and repositioning (2 parameters), (ii) a model with time dummies (20 parameters). The former R^2 measure avaluates a general goodness of fit, the latter investigates to which degree a non-stationary model would fit the data better.

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 10: Impact of different enforcement regimes on producer (half-yearly, millions USD), listener (day-minutes of advertising per listener; listener surplus cannot be easily expressed in dollars because there are no dollar transfers between listeners and radio station owners), and advertiser surplus (half-yearly, millions USD), cumulative in the 88 markets. 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.

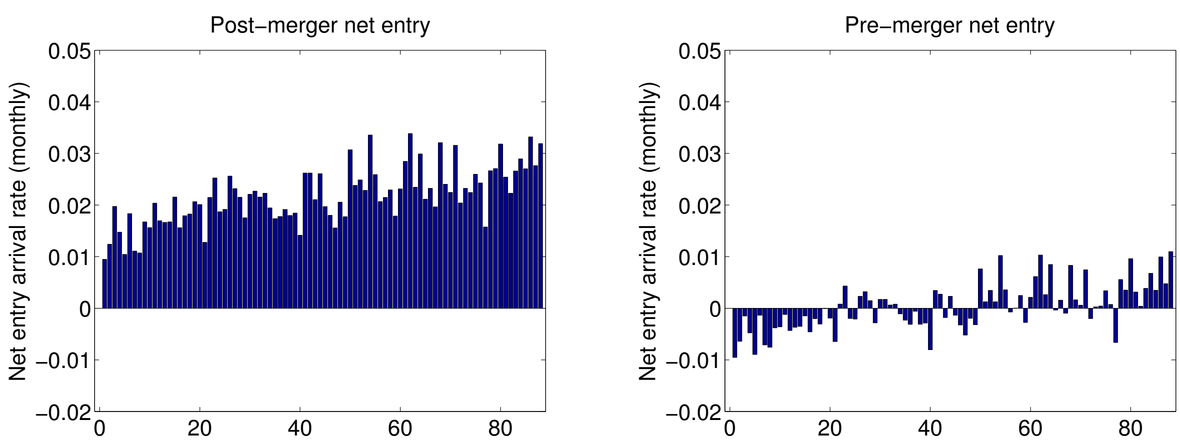

Figure 1: Post- and pre-merger net entry rates by market. Markets are sorted by population, from the smallest (Bismark) to the largest (Los Angeles).

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 11: Impact of increasing FM local ownership cap to 7 stations on producer (half-yearly, millions USD), listener (day-minutes of advertising per listener), and advertiser surplus (half-yearly, millions USD), cumulative in the 88 markets. 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 Formulation using conditional choice probabilities

To apply the 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 Aguirregabiria and Magesan (2013)).

Let $\text{CCP}_k^A(a|\mathcal{J})$ be an ex-ante probability of company k acquiring a company k' conditional on the arrival of a merger opportunity. Similarly, define $\text{CCP}_k^R(r|\mathcal{J})$ to be an ex-ante probability of repositioning from f to f'. After a minor adjustment to continuous time, the results contained in the proof of Theorem 1 from Hotz and Miller (1993) apply to 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

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

A similar expression can be written for the repositioning action:

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

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:

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

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).

Equation B.1 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,

$$\text{CCP}_k^A(a|\mathcal{J}) = \frac{\exp\left\{\sigma_k^A(\mathcal{J},a)^{-1}\left[V_k(\mathcal{J}'(k,a)) - V_a(\mathcal{J}) + \mu_k^A(\mathcal{J},a)\right]\right\}}{\sum_{a' \in \Gamma_k^A(\mathcal{J})} \exp\left\{\sigma_k^A(\mathcal{J},a')^{-1}\left[V_k(\mathcal{J}'(k,a')) - V_{a'}(\mathcal{J}) + \mu_k^A(\mathcal{J},a')\right]\right\}}, \tag{B.2}$$

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:

$$\text{CCP}_k^R(r|\mathcal{J}) = \frac{\exp\left\{\sigma_k^R(\mathcal{J},r)^{-1}\left[V_k(\mathcal{J}'(k,r)) + \mu_k^R(\mathcal{J},r)\right]\right\}}{\sum_{r' \in \Gamma_k^R(\mathcal{J})} \exp\left\{\sigma_k^R(\mathcal{J},r')^{-1}\left[V_k(\mathcal{J}'(k,r')) + \mu_k^R(\mathcal{J},r')\right]\right\}}. \tag{B.3}$$

The computational algorithm involves iterating on the value function using a Bellman equation (B.1) and equations (B.2) and (B.3). 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 (B.2) and (B.3),

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

(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 other players' strategies only through the value functions. In such a case, one does not need to remember a complete 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 4 for details).

C Time aggregation

The time aggregation is performed in steps. I will demonstrate the steps using a simple example contained at Figure 2. In this example,

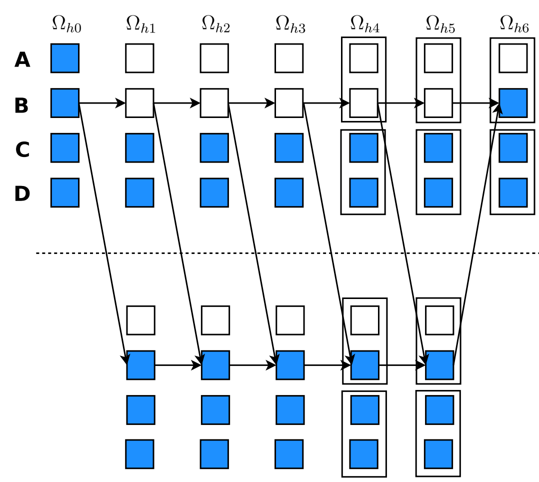

Figure 2: Feasible industry paths for a simple example with one repositioning event and two merger events.

there is one repositioning event and two merger events. The merger events happen during the month 4, and the repositioning event happens sometime between month 1 and month 6. So first, I construct a set of feasible intermediate 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 the month i by $\Omega^{hi} \subset \Omega^h$. States in Ω^{hi} incorporate all mergers that happened before 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. In the simple example, each end of the month has only two feasible latent states; however, the number of these sets proliferates as the number of repositioning events increases. 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. Furthermore, apart from the states in Ω^{hi}, the set Ω^h contains all transitory states between any states in Ω^{hi} and Ω^{hi+1}. For the simple example, the transitory states are presented in Figure 3. In particular, it is possible that upper merger happens before lower merger, and vice versa. It is also possible that repositioning happens at any point before, in between and after both mergers. The computation of the transitory states that

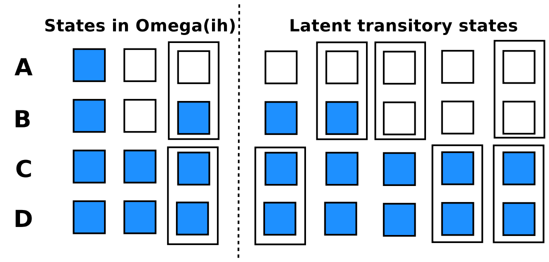

Figure 3: Feasible and transitory states for a simple example with one repositioning event and two merger events.

determine feasible paths of the industry becomes quite burdensome if many mergers and repositionings happen in a particular half-year. For example, the 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 computation was parallelized. The benefit of this process is that the final augmented state space, Ω^h, is thousands of times smaller than the entire state, which dramatically reduces the size of the intensity matrix.

Upon the arrival of the merger and repositioning actions at time t, the equilibrium strategies induce transitions according to instantaneous conditional choice probabilities of acquisition $\text{CCP}^A(\mathcal{J}^t)$ or repositioning $\text{CCP}^R(\mathcal{J}^t)$. Together with action arrival rates λ^A and λ^R, these CCPs generate an intensity matrix Q^h on the augmented state space Ω^h. The overall goal is to use the Markov process on Ω^h to compute the conditional likelihood of the data, that is, $L(\Omega^{h6}, \ldots, \Omega^{h1}|\Omega^{h0})$. The exact states from Ω^{hi} (except 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 $L(\Omega^{h2}|\Omega^{h1}, \Omega^{h0})$.[18] For this purpose, I compute $\iota^h(2)$ by solving equation (C.1) with

[18]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}.

$\iota^h(1)$ 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 $L(\Omega^{hi}|\Omega^{hi-1},\dots,\Omega^{h1},\Omega^{h0})$, and as a result, I get the joint likelihood $L(\Omega^{h6},\dots,\Omega^{h1}|\Omega^{h0})$ 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

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

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 $r = 1,\dots,1000$. 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

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

Fixed cost can be obtained using

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

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

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

I need four sufficient statistics to evaluate this part of the value function in such a case. 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 costs 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.