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

Przemysław Jeziorski*

August 1, 2015

1 Static payoffs

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

1.1 Static supply and demand

The single-period profit function is identical to one used in Jeziorski (2014a) with the exception that for this study I employ three meta-formats, that is $F = 3$, instead of eight. Below I describe the parametrization in order to keep the paper self-contained; however, the discussion is also kept rather brief to avoid duplications.

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

$$\pi_k(\mathcal{J}^t) = \sum_{\substack{\text{j owned by k} \\ \text{in market m}}} \left(p_j(\bar{q}^t_j|\mathcal{J}^t) r_j(\bar{q}^t|\mathcal{J}^t) - \mathrm{MC}_j(\mathcal{J}^t) \right) \bar{q}^t_j, \tag{1.1}$$

*Haas School of Business, UC Berkeley

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

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

$$u_{ij} = \theta_{1i}^L \iota_j - \theta_{2i}^L q_j + \theta_3^L \mathrm{FM}_j + \xi_j + \epsilon_{ji}, \tag{1.2}$$

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

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

and

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

where Σ_1 is a diagonal matrix, $F_m(D_i|d)$ is an empirical distribution of demographic characteristics, ν_i is an unobserved taste shock, and Π is the matrix representing the correlation between demographic characteristics and format preferences. I assume draws for ν_i are uncorrelated across time and markets. The term ξ_j represents the unobserved quality of station j. The assumptions on ξ_j are equivalent to those in Berry et al. (1995).[1] The model allows for an outside option of not listening to radio u_{i0}, which is normalized to zero in the years 1996 and 1997. For subsequent years, u_{i0} contains time dummies to control for the influx on new broadcasting technologies such as satellite radio and internet.

The market share of the station j is given by

$$r_j(q|\mathcal{J}^t) = \mathrm{Prob}\big(\{(\nu_i, D_i, \epsilon_{ij}) : u_{ij} \geq u_{ij'}, \text{for } j' = 1, \ldots, J\}\big| q, \mathcal{J}^t\big). \tag{1.3}$$

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

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

$$p_j = \theta_1^A \left(1 - \theta_2^A \sum_{f' \in \mathbb{F}} w_{ff'}^m q_{f'} \right), \tag{1.4}$$

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

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

$$\mathrm{MC}_{jmt}(\theta^A, \theta^C) = \theta_1^{Am}[\theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Cm}\xi_{jt} + \theta_3^{Cm}\mathrm{SYN}_{jt} + \epsilon_{jt}^C]q_{jt}, \tag{1.5}$$

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

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

1.2 Data used to esimate the static model

The data come from four main sources: two consulting companies, BIA Inc and SQAD, a Common Population Survey, and Radio Today publications by Arbitron. BIA provides two comprehensive data sets on the vast majority of U.S. radio broadcasting firms. The first data set covers the years 1996-2001 and the second, 2002-2006. I combined these two data set to form a large panel for

1996-2006. SQAD provides a data set on average prices per rating point (cost-per-point or CPP) for each market and half year, grouped by demographics and time of the day. Unfortunately, SQAD does not provide data on station-level per-listener pricing. However, because the pricing is done on a per-listener basis, one can still compute a station-level price of an advertising slot by multiplying the CPP by the station rating. According to anecdotal evidence, many advertisers follow this procedure to figure out the prices they are likely to pay. This procedure does not account for the fact that stations may have different listenership pools and, consequently, the CPPs for different stations can vary. I alleviate this concern by computing a proxy for a station-level CPP. In particular, I take a weighted average of prices by demographics and time of the day, where the weights are the relevant ratings of the station. In so doing, I assume that stations with most of their listenership during a particular time of the day set a price that is close to the market average for that time. Although this estimate of station-level prices is not perfect, it produces a considerable amount of variation within the market. Subsequently, I use these price proxies to compute station-level advertising quantities by dividing estimates of station revenues (provided by BIA) by a product of prices and ratings. Note that ad quantity computed in this manner may carry some measurement error because it is a function of two estimates. However, if this measurement error is not endogenous – for example, if it only introduces error to an overall level of advertising in each market – it would not affect the results.

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

1.3 Static model estimation

The following section is a parsimonious description of the estimation procedure I use to recover the parameters of the static model (for the full description see Jeziorski (2014a)). I conduct the estimation of the model in two steps. In the first step, I estimate the demand model, which includes parameters of the consumer utility θ^L (see equation (1.2)), and in the second step I recover the parameters of the inverse demand for advertising θ^A, $w_{jj'}$ (see equation (1.4)) and marginal cost parameters θ^C (see equation (1.5))

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

$$E[\xi_{jt}|Z_1, \theta^L] = 0, \tag{1.6}$$

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

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

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

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

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

The estimator is based on the following supply conditions:

$$r_{jt} + \sum_{j' \in s_{kt}} q_{j't} \frac{\partial r_{j't}(q_t)}{\partial q_{jt}} =$$

$$\theta^{Cmt} + \theta_1^{Cm} + \theta_2^{Am} \left[r_{jt} v_j + \sum_{j' \in s_{kt}} \left(r_{j't}(q_t) \omega_{jj'}^m + v_{j'} \frac{\partial r_{j't}(q_t)}{\partial q_{jt}} \right) \right] + \theta_2^{Cm} \xi_{jt} + \theta_3^{Cm} \mathrm{SYN}_{jt} + \eta_{jt}, \tag{1.7}$$

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

Because the equation does not depend on θ_1^A, I can use it to estimate θ_2^A and θ^C. Two sources of heterogeneity in marginal cost and slope coefficients exist across markets. Effective marginal cost parameters for each station in market m are given by $\theta_1^{Am} \theta^{Cm}$, and θ_1^{Am} is allowed to differ across markets. Moreover, to control for potential heterogeneity that is not captured by a level of revenues, I allow for three different sets of values for all parameters in θ^{Cm}: for small (up to 500,000 people), medium (between 500,000 and 1,500,000), and large (more than 1,500,000) markets. To avoid having a full set of dummies and to facilitate identification, I set time dummies for years 1996 and 1997 to zero. Similar specification is true for the slope of the inverse demand for ads and its effective slope is given by $\theta_1^{Am} \theta_2^{Am}$. To control for the fact that stations might have different market power in the advertising market depending on size, I allow for four different values for the slope of inverse demand, depending on the population of the market (up to 500,000 people, between 500,000 and 1,500,000, between 1,500,000 and 4,500,000, and more than 4,500,000). Given the estimates of θ_2^{Am} and θ^C, I can back out θ_1^{Am} by equating the observed average revenue in each market with its predicted counterpart. To control for the fact that ratings depend on quantity, which is likely to be correlated with η, I estimate the model with a two-stage least squares procedure which employs the following instruments: the number of stations in the same format and the ad quantities of competitors. Additionally, the instruments were lagged one period to control for potential serial correlation in η.

1.4 Simplification of the state space

The profit function varies across markets because of market-specific parameters and because the demographic composition of listeners is heterogeneous. The estimation of the static profit function is explained in the Appendix 1. I note that the profit function is estimated under the assumption that the station ownership and formats are fixed within a give half-year time period. To relax this assumption one would need to estimate the profit function jointly with a dynamic model, which would render this analysis infeasible.

The static model is non-stationary because it contains time dummies in the demand and supply equation, and because the distribution of listeners' demographics varies from year to year. Consequently, I estimate the static model as non-stationary to obtain more robust measures of listener and advertiser price elasticity. However, after the estimation, I detrend the static profits to fit the static model into the dynamic framework. Specifically, I remove the trends in supply and demand by using an average value of the time dummies. Additionally, I draw listeners from a joint 1996-2006 market-specific empirical distribution instead of year-by-year distributions, which assumes that no demographic trends within local markets. The former assumption is likely to result the conservative estimates of cost synergies. For example, missing an possible long-run downward trend in listenership would overestimate the profitability of mergers and consequently the estimator would require smaller cost synergies to rationalize the consolidation. The latter assumption may possibly affect the estimates of repositioning strategy, which may be responsive to changes in demographic composition of the market. To alleviate this concern, I analyze goodness of fit of the model and show that adding non-stationarity does not significantly improve the predictive power of the model. Consequently, the impact of the changes in demographics on repositioning is likely to be dwarfed by the impact of own and competitors' portfolios, so I ignore the former and analyze the latter.

The state space of the market needs to be further simplified in order to make the dynamics manageable. In particular, I compute an average station quality ξ and tower power for each market-format combination and endow each station with these averages depending on the station's market and format. Additionally, to control for the differences in the radio stations' sizes between fringe and non-fringe owners, I allow the averages of ξ and tower power to be different for these two groups of radio stations. Such procedure further standardizes the profit function and makes it consistent with the model laid out in the model section of the paper. In particular, this standardization makes stations homogenous with an exception of the meta-format and fringe designation. Note this assumption is stronger than the one in Jeziorski (2014b), which allows for station-level persistent but exogenous heterogeneity. If the radio stations are indeed heterogeneous in more dimensions and such heterogeneity creates a large amount of extra market power, this procedure could pollute the estimates of cost synergies as well as the counterfactuals. To alleviate some of these concerns, I compare cost synergy estimates to those in Jeziorski (2014b); however, all results in the present study should be interpreted keeping this standardization in mind. Further, I note that the assumption of fixed station quality is testable. Indeed, Jeziorski (2014b) contains such tests and shows that mergers do not affect quality in an economically meaningful way, and that the quality itself is stationary. I do not repeat these tests in paper because the industry and time span of the data analyzed in this paper is the same. Two other studies of the same industry, that is Sweeting (2013) and Jeziorski (2014a), allow for AR(1) process for an unobserved quality ξ, estimating that the majority of 85% and 75% of the ξ is persistent, respectively. I note that Sweeting (2013) finds economically small, but statistically significant drop in quality after switching a format. This drop is unlikely to influence my results, but would require many additional assumptions to be incorporated in the structural model. In the industries where the interaction of quality and mergers is important appropriate modifications to the model may be necessary.

1.5 Estimation results of a static model

This subsection contains a brief description of the static profit function estimates. The model of a profit function is a simplified version of Jeziorski (2014a), and to minimize the duplication I provide only a brief description of the parameters.

Table 1 presents the estimates of listenership demand. The first and second columns contain the mean and the standard deviation of the random coefficients. I find that advertising has a negative effect on listenership, and the effect is fairly homogenous among listeners. I also find that listeners prefer FM to AM stations, and that greater transmitting power corresponds to higher listenership. Format dummies are negative; however, by construction, they capture preferences of a specific demographic group: male, uneducated, low income, white, non-Hispanic teenagers. To obtain preferences of other demographic groups, one has to add appropriate demographic interactions, which are presented in Table 2. The first-meta format, which delivers adult-oriented music, such as, classic rock, country, or jazz, is most popular among middle-aged listeners, with women constituting a slight majority of the demographics. The second meta-format, which delivers contemporary pop and alternative music, appeals to younger people. This meta-format contains popular pop and urban formats, as well as hip-hop, which explains the large and highly significant African-American fixed effect. The last meta-format, which contains talk radio and ethnic or religious stations, is popular among older listeners, mainly higher-educated females with relatively large incomes. A positive Hispanic dummy is related to Hispanic and religious stations in this meta-format.

Table 3 describes the national trends in radio listenership. The values represent the residual trend in radio listenership beyond the changes in the demographic composition. In general, I find the trend in the utility of the outside option to be non-monotonic. I use these numbers to detrend the profit function in the dynamic estimation.

Table 4 presents the coefficients of the inverse demand for advertising. The inverse demand curve is downward sloping, indicating the radio stations' direct market power over advertisers. The inverse demand is steeper in smaller markets.

Tables 5 and 6 contain the marginal cost estimates. The marginal cost is larger in smaller markets. I find evidence of marginal cost synergies in small and large markets, but not in medium markets.

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

Standard errors in parentheses

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

Table 1: Estimates of the random-coefficients logit model of radio listeners' demand. The first column consists of the mean values of parameters in the utility function. The second row consists of the standard deviations of a random effect ν.

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

Standard errors in parentheses

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

Table 2: The table presents estimates of covariances in the random-coefficients logit model of radio listeners' demand. Each cell represents a covariance between specific demographic characteristics and listening to a particular radio =station format.

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

Standard errors in parentheses

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

Table 3: Estimates of utility (exponentiated) of not listening to radio. Value for 1996 is normalized to 1.

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

Standard errors (corrected for the first stage) in parentheses

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

Table 4: The slope of advertising price per rating point (CPP). Intercept is set to 1. Units are standard deviations of quantity supplied on a station level.

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

Standard errors (corrected for the first stage) in parentheses

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

Table 5: The marginal cost per minute of advertising sold. The intercept of the advertising price per rating point is set to 1. Note that these numbers may be higher than 1 because the final price of advertising is the CPP times the station rating in per cent. Units for quality are standard deviations of quality in the sample.

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

Standard errors (corrected for the first stage) in parentheses

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

Table 6: The marginal cost synergies of owning multiple stations in the same format.

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

Table 7: National owner repositioning CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages.

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

Table 8: National owner repositioning CCP: Coefficients on the covariates related to the current format.

2 First stage results – suplement

This section contains unreported tables from the estimation of the dynamic model.

2.0.1 Format-switching strategy

Table 7 contains from-to format-switching dummies, as well as, interactions between demographics and target formats. As with acquisition strategy, I find that format switching largely reflects listeners' tastes. Tables 8 and 9 describe the impact of the industry state on format switching. In general, the number of stations owned in a particular format correlates positively with switching to that format, while the opposite holds for the number of stations owned by competitors.

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

Table 9: National owner repositioning CCP: Coefficients on the covariates related to the target format.

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

Table 10: National owner repositioning CCP: Dummies for proximity to the ownership cap.

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

Table 11: Local owner repositioning CCP: Format dummies and format-demographics interactions; demographics variables are 1996-2006 market-level averages.

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

Table 12: Local owner repositioning CCP: Coefficients on the covariates related to the current format.

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

Table 13: Local owner repositioning CCP: Coefficients on the covariates related to the target format.

<table>
<tr><th rowspan="3">Owner</th><th rowspan="3">Source format</th><th colspan="4">Mean</th><th colspan="2">Std. deviation</th></tr>
<tr><th colspan="3">Target format</th><th rowspan="2">Variable profit</th><th rowspan="2">Intercept</th><th rowspan="2">Variable profit</th></tr>
<tr><th>Adult Music</th><th>Hits Music</th><th>Non-Music</th></tr>
<tr><td rowspan="3">National</td><td>Adult Music</td><td>-</td><td>8.873***
(0.999)</td><td>8.538***
(0.922)</td><td rowspan="6">2.084
(1.842)</td><td rowspan="6">1.713***
(0.180)</td><td rowspan="6">0.069
(0.329)</td></tr>
<tr><td>Hits Music</td><td>6.949***
(0.773)</td><td>-</td><td>8.381***
(0.926)</td></tr>
<tr><td>Non-Music</td><td>7.611***
(0.858)</td><td>10.004***
(1.145)</td><td>-</td></tr>
<tr><td rowspan="3">Local</td><td>Adult Music</td><td>-</td><td>8.887***
(0.963)</td><td>6.450***
(0.701)</td></tr>
<tr><td>Hits Music</td><td>5.093***
(0.616)</td><td>-</td><td>4.929***
(0.593)</td></tr>
<tr><td>Non-Music</td><td>8.100***
(0.876)</td><td>10.420***
(1.129)</td><td>-</td></tr>
</table>

Standard errors in parentheses

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

Table 14: Market category 2: Estimates of format-switching costs (in millions USD). The table contains to-from format fixed effects for local and national owners.

Tables 11 through 13 present the covariates of format switching by local owners. The numbers are similar to those of national owners, so I omit the discussion.

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	-	6.531*** (0.892)	6.284*** (0.838)	2.084 (1.842)	1.226*** (0.163)	0.069 (0.329)
	Hits Music	5.114*** (0.711)	-	6.169*** (0.831)			
	Non-Music	5.602*** (0.775)	7.363*** (1.036)	-			
Local	Adult Music	-	6.541*** (0.881)	4.747*** (0.641)			
	Hits Music	3.749*** (0.537)	-	3.628*** (0.525)			
	Non-Music	5.961*** (0.797)	7.669*** (1.033)	-			

Standard errors in parentheses

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

Table 15: Market category 3: Estimates of format-switching costs (in millions USD). The table contains to-from format fixed effects for local and national owners.

3 Numerical examples

In this appendix, I use numerical examples to show how various long-run processes can affect the dynamics of merger bids and the efficacy of the merger enforcement process. The first two examples demonstrate that myopic regulation could be suboptimal. Specifically, I identify cases in which, relative to a dynamic optimum, a myopic antitrust agency blocks mergers exessively (Example 1) or not often enough (Example 2). Additionally, I demonstrate that the efficacy of predicting waves of mergers using a myopic model might be limited (Example 3).

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

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

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

Standard errors in parentheses

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

Table 16: Market category 4: Estimates of format-switching costs (in millions USD). The table contains to-from format fixed effects for local and national owners.

of listening to station j of type f is given by

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

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

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

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

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

In the remainder of this section, I assume, for similiticy, that $C = F = 2$ and that each consumer type has a favorite format, that is, $\alpha_{cc} = \gamma_1$ and $\alpha_{fc} = \gamma_2$ if $c \neq f$, where $\gamma_1 > \gamma_2$. The difference between γ_1 and γ_2 measures a consumer's willingness to switch to a format that is not their favorite and determines the degree of within-format market power.

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

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

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

I analyze the aforementioned situation numerically in the context of both the total-surplus and consumer-surplus criteria. Table 17 presents the results of computational experiments for an industry with fixed cost synergies and a regulator using a total-surplus criterion. The static dead-weight loss varies between 0.6% (of the total surplus) and 4.2%, respectively, for high and low levels of fixed cost synergies. Thus the myopic total-surplus maximizer would reject this merger. However, depending on the repositioning cost and the size of cost synergies, a forward-looking regulator might approve the merger. For the highest level of fixed cost synergies, the merger should be approved even for large values of repositioning cost, bringing between a 0.7% and 2.2% increase in total welfare. Note that if the cost synergies are smaller, the regulator should approve the merger only if the entry is likely and timely, that is, when the repositioning cost is low. Thus, to make a correct decision, the regulator should have estimates of both cost synergies and repositioning cost.

In Table 18, I present a similar analysis for the industry with marginal cost efficiencies and for the regulator who enforces a consumer-surplus criterion. I consider the situation in which the corporation has lower marginal cost than the local firm. The top of the table contains a case in which marginal cost efficiencies from the merger are large enough to counteract a static increase in market power. Consequently, both myopic and forward-looking regulators would make the same decision and approve the merger. However, if the cost synergies are smaller, the static impact of the merger on consumer surplus is negative and a myopic regular rejects the merger. At the same time, the dynamic regulator approves the merger if the repositioning cost is low enough.

Example 2: Predatory mergers. Consider a situation in which a merger results in substantial marginal cost synergies. These synergies would benefit the conglomerate but might hurt the smaller competitors and prompt their exit. Following this logic, I call a merger predatory if it is aimed at driving competitors out of the market by using substantial marginal cost advantages. This example illustrates how the actions of myopic regulators may be suboptimal when approve rather than reject such predatory mergers.

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

Table 18 presents the results of multiple computational experiments exploring this dynamic. A corporation with two products has a marginal cost of 0.75, whereas a corporation with one product has a marginal cost of 1, and the same true for any other other company. The tables

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

Table 17: The table illustrates the difference between the myopic and long-term impact of the merger described in Example 1a on total and consumer surplus. Additionally, I report pre-merger and post-merger probabilities of repositioning that increases the number of competitors in the relevant format. I have underlined the relevant events, that would lead a myopic regulator to block the merger. However, the same events would lead a forward-looking regulator to allow the merger.

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

Table 18: The table illustrates the difference between the myopic and long-term impact on total and consumer surplus of the merger described in Example 1b. Additionally, I report pre-merger and post-merger probabilities of repositioning that increases the number of competitors in the relevant format. I have underlined the relevant events, that would lead a myopic regulator to block the merger. However, the same events would lead a forward-looking regulator to allow the merger.

contain percentage changes in consumer surplus caused by the merger and the probability that at least one firm will exit the first format during a unit interval following the merger. I demonstrate that the myopic regulator consistently underestimates the impact of mergers on consumer surplus and in many cases this regulator would approve too many mergers. In the most notable case, the myopic regulator would predict a 1.3% increase in consumer surplus, but in reality, the merger would result in a drop in consumer surplus of more than 3.6%.

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

Table 20 presents equilibrium probabilities of mergers for different primitives and levels of player sophistication. Intuitively, the likelihood of the merger is a decreasing function of a merger execution cost and an increasing function of marginal cost (because of the marginal cost efficiencies). First, I note that a probability of a merger in a myopic model is usually higher than in a dynamic model. Looking at the three-to-two merger first is helpful in explaining this higher probability. In this case, no further mergers are possible, so the marginal market power of merging in a myopic world is equal to the marginal market power of merging in a forward-looking world. However, the same is not true for acquisition prices, and Figure 2 depicts the difference between the myopic and forward-looking acquisition prices. In particular, I find the acquisition price is substantially higher in the forward-looking case resulting in a *hold out*. The reason is that a merger increases the markups of competitors that are not acquired, so a forward-looking acquiree has an incentive to reject a static merger bid and internalize these spillovers. Note that these incentives disappear once the economies of scale increase and competition with a conglomerate becomes more difficult.

The case of a four-to-three merger is more complicated. Figure 2 depicts the difference in acquisition price between myopic and forward-looking cases. For small marginal cost synergies,

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

Table 19: The table illustrates the difference between the myopic and long-term impact of the merger described in Example 2 on total and consumer surplus. Additionally, I report pre-merger and post-merger probabilities of repositioning that increases the number of competitors in the relevant format. I have underlined the relevant events, that would lead a myopic regulator to block the merger. However, the same events would lead a forward-looking regulator to allow the merger.

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

Table 20: Probability of diffent types of mergers for myopic and forward-looking buyers and sellers for different levels of merger execution cost and marginal cost synergies.

the results for a four-to-three case are similar to those of a three-to-two case. However, for large values of the marginal cost the acquisition price in the myopic model is greater than in the forward-looking model. To understand that, consider the firm's incentives to reject a four-to-three merger bid. In the case of a high marginal cost, the potential acquiree recognizes that rejecting the bid would likely lead to a competition against a highly efficient conglomerate. As a result, the acquirer might credibly threaten the acquiree to make a bid to a competitor and subsequently come back with a lower offer leveraging on its size. Thus, in the extreme case, the acquiree agrees to a haircut, that is, selling out below current static profits. Such a haircut explains why a four-to-three merger is more probable than a three-to-two merger (see the last rows of Table 20).

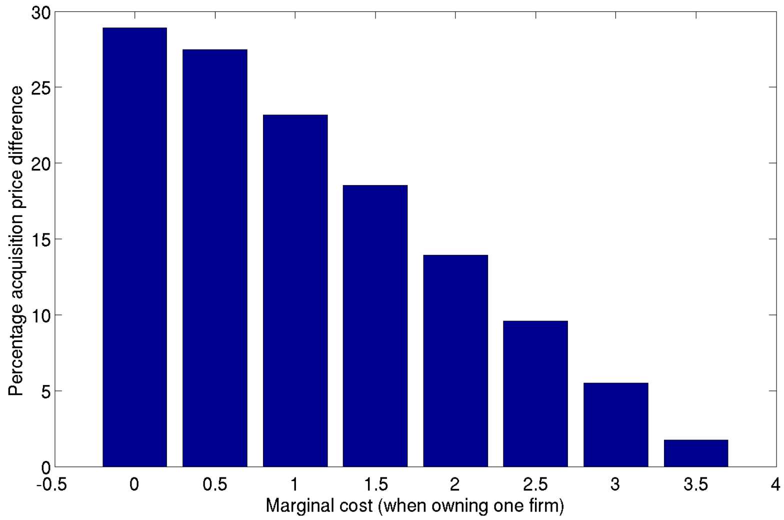

Figure 1: Percentage difference in the acquisition price of the three-to-two merger, between a forward-looking and a myopic model. The x-axis represents the different levels of marginal cost that determine the importance of marginal cost synergies. Note that a forward-looking acquisition price is greater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

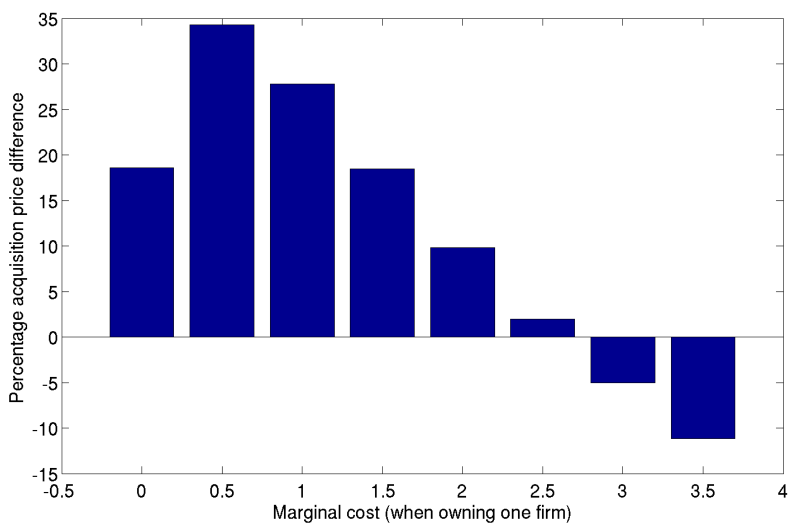

Figure 2: Percentage difference in acquisition price of the three-to-two merger, between a forward-looking and a myopic model. The x-axis represents different levels of marginal cost that determine the importance of marginal cost synergies. On the left side of the graph, the myopic price is smaller than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

References

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

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.

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

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