# Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry

Andrew Sweeting*

Duke University and NBER

July 2012

**Abstract**

This article predicts how radio station formats would change if, as was recently proposed, music stations were made to pay fees for musical performance rights. It does so by estimating and solving, using parametric approximations to firms' value functions, a dynamic model which captures important features of the industry such as vertical and horizontal product differentiation, demographic variation in programming tastes and multi-station ownership. The estimated model predicts that high fees would cause the number of music stations to fall significantly and quite quickly. For example, a fee equal to 10% of revenues would cause a 4.6% drop in the number of music stations within $2\frac{1}{2}$ years, and a 9.4% drop in the long-run. The size of the change is limited, however, by the fact that many listeners, particularly in demographics that are valued by advertisers, have strong preferences for music programming.

*Please send comments to atsweet@duke.edu. I would like to thank Jerry Hausman, Igal Hendel, Aviv Nevo, Amil Petrin, Ariel Pakes, Rob Porter, Steve Berry, Pat Bayer, Peter Arcidiacono, Kate Ho, Allan Collard-Wexler, Paul Ellickson, Arie Beresteanu, three referees and seminar participants for valuable comments. I would like to thank the National Association of Broadcasters and the Center for the Study of Industrial Organization at Northwestern University for financial support. All errors are my own.

# 1 Introduction

This article develops and estimates a dynamic oligopoly model to predict the effect on product variety of legislation (the Performance Rights Act[1]), introduced into Congress in 2009 with the support of the Obama Administration and members of both parties, requiring music radio stations to pay fees for musical performance rights. This legislation provides one particularly clear example, among many, of a policy which, intentionally or unintentionally, changes the incentives of firms to offer particular types of product.[2] In the long-run, the changes in product variety that these policies cause may have much larger effects on welfare than changes in the prices of existing products.

An analysis of variety requires a model with differentiated products and when firms have to pay significant sunk costs to develop new products or reposition existing ones, a dynamic model is required to predict how a policy change will affect an industry's evolution.[3] Solving and estimating dynamic models with differentiated products is not straightforward because a large state space is required to capture the features of the market that are likely to affect how the industry evolves. For example, in my setting it is necessary to allow for at least 8 types of product (programming formats), a large number of multi-product firms (up to 18 firms per market in my data, owning up to 8 stations each), and vertical differentiation (on average, the largest station in a market has more than 60 times as many listeners as the smallest station). I address the problem of the large state space by using parametric approximations to the value function when estimating and solving the model.[4] The use of value function approximation to solve dynamic models has been suggested by Benitez-Silva et al. (2000), Farias et al. (2012) and Arcidiacono et al. (2012)[5] and used in applied contexts by Hendel and Nevo (2006), Fowlie et al. (2011) and Barwick and Pathak (2012). I incorporate value function approximation into estimation procedures that build off the methods proposed by Aguirregabiria and Mira (AM hereafter) (2007, 2010) and Pakes et al. (POB) (2007). These procedures give quite

[1] H.R. 848 and S. 379, $111^{th}$ Congress.

[2] Other examples include gasoline taxes, fuel efficiency standards and trade policies that affect the incentives of domestic automakers to produce particular types of vehicles (Berry et al. (1993)) and taxes, labelling and advertising restrictions that increasingly affect the incentives of food and beverage manufacturers to make healthier products.

[3] Static models may be appropriate for settings where firms can change their assortments quickly and easily. For example, Draganaska et al. (2009) use a static model to predict how the set of ice cream flavors that firms sell in particular markets would change after a merger, out of the set of flavors that firms already produce.

[4] In an earlier version of this paper, Sweeting (2011), I presented Monte Carlo results for my estimation and solution procedures based on a much-simplified version of my model.

[5] Outside of the economics literature, Bertsekas and Yoffe (1996), Bertsekas and Yu (2007), Bertsekas (2010) and Ma and Powell (2009) investigate the effectiveness of approximate dynamic programming approaches.

similar estimates.[6] To provide additional confidence in my methods and results, I also compute estimates based on procedures that approximate value functions using forward simulation. I show that my preferred estimates lie within the bounds on the structural parameters computed using forward simulation and the moment inequality approach proposed by Pakes et al. (PPHI) (2011).

The Performance Rights Act was motivated by the declining revenues of the recording industry. The legislation proposed that commercial, broadcast radio stations, whose primary programming is music, should pay for musical performance rights, which are owned by musicians, performers and record labels, in addition to the fees that they currently pay for composition rights.[7] Requiring stations to pay fees for performance rights would bring the broadcast industry in the US into line with the broadcast industries in most other countries, and the cable, satellite and internet radio industries in the US, where stations already pay fees. The legislation proposed that for music stations with revenues above some cap, fees would be determined as a % of advertising revenues, and would not depend on the exact amount of music that the station played. Non-commercial stations and stations that provide primarily talk programming were exempt. The legislation did not specify how high the fees should be, so that in the absence of an agreement between the radio and recording industries, they would be determined by copyright judges. Media law experts have argued that existing case law might justify performance rights fees as high as 25% of advertising revenues, which would have much larger effects on stations' incentives to play music than the 2-3% fees that music stations currently pay for composition rights.[8]

[6] AM (2007) propose estimating dynamic models using an iterated nested pseudo-likelihood procedure where players' conditional choice probabilities are updated during estimation. This procedure may be feasible when a nested fixed point procedure would not be feasible. AM (2010) consider a modified version of this procedure where other players' choice probabilities are not updated during estimation. I present some results using estimators based on both of these approaches. POB present evidence that a moment-based objective function, which matches entry rates that are averaged across states, may be superior to likelihood-based ones in small samples. I therefore also consider a version of the modified AM (2010) estimator where this type of objective function is used, rather than one based on the pseudo-likelihood.

[7] For more details, see US GAO (2010).

[8] See for example, http://www.broadcastlawblog.com/2010/03/articles/music-rights/copyright-royalty-board-approves-settlement-for-sound-recording-royalty-rates-for-new-subscription-services-any-hints-as-to-what-a-broadcast-performance-royalty-would-be/ (accessed December 5, 2010). XM Sirius paid 8% of its subscription revenues for performance rights in 2010-12, even though some of its programming is not musical, and this fee included a discount recognizing that satellite radio was struggling to become established (Federal Register vol. 75, p. 5513 (2010-02-03)). Companies providing audio programming on cable pay 15% of revenues (Federal Register vol. 75, p. 14075 (2010-03-24)). The legislation was not passed, but in June 2012 Clear Channel struck a landmark detail with the Big Machine record label in which it agreed to pay a fee when it played Big Machine's songs on its broadcast stations in exchange for a rebate when it broadcasts the label's songs on the internet where performance fees are levied, although no more details of the deal were released (Wall Street Journal, June 6, 2012, accessed on that date). The same article noted that Pandora pays approximately 60% of its revenues in fees.

Not surprisingly, broadcasters argued that fees at this level would make it unprofitable for them to play music, and they predicted that many music stations would switch to non-music programming.[9] This claim cannot be evaluated using existing data because this level of fee has never been levied on broadcast stations in the US, and in countries where performance rights fees are levied programming is often affected by content regulation or the presence of state-owned broadcasters. For this reason, I develop, estimate and solve a discrete-choice model of format choice to predict how much format variety and music listenership would change in a large sample of markets if fees equal to 10% or 20% of music station revenues were introduced. The model captures important features of the industry that should affect the response to fees. In particular, I let different demographic groups have heterogeneous programming tastes, and for advertisers (broadcast stations' primary source of revenue) to place different values on listeners with different demographics. As I illustrate using counterfactuals, taste heterogeneity has important effects because the listeners most valued by advertisers (white listeners aged 25-49, especially women) prefer stations in music formats (such as Adult Contemporary, Country and Rock). A station that switches to a non-music format (such as News, Easy Listening, Religion or Spanish-language programming) will therefore tend to attract less valuable listeners, while the remaining music stations will see their audiences and revenues increase offsetting the effects of fees. The model also allows for the fact that format switching can be costly as a station may lose the value in its established relationships with listeners and advertisers, as well as having to replace all of its on-air staff, its programming director and many of its advertising sales staff.[10] I use counterfactuals to investigate how these costs affect how many stations switch formats when fees are introduced and the speed of adjustment.

While one could study the effects of policies favoring particular types of product in many industries, several features of the radio industry make it ideal for estimating this type of model. First, the industry has many local markets, each with its own local stations. This provides thousands of

[9] For example, the National Association of Broadcasters Radio Board Chairman, Steven Newberry, stated before the House Judiciary committee that "the number of stations playing music would dramatically decrease" because of the Act (The Performance Rights Act, Hearing Before the Committee on the Judiciary, House of Representatives, $111^{th}$ Congress, $1^{st}$ session on HR 848, March 10, 2009, Oral Testimony of Steven Newberry).

[10] The manager of a station in the Raleigh-Durham market that moved from Country to Sports programming in 2007 described how the station replaced all of its on-air staff and all of its advertising sales staff, and how it had chosen to play Rock music for a couple of months before moving to Sports to "kill off" its original audience so that it would face fewer complaints when it finally started its Sports programming. The switch also required an extensive 9 month planning process which involved hiring format consultancies to advise the station's owners on their strategic options. At the time, the owners predicted that the move would take at least two years to pay for itself.

station-level observations with which to estimate the main parameters of the model, with exogenous variation in market demographics, as well as station characteristics, helping to provide identification. These sources of identification are explicitly used in constructing the type of moment-based objective function suggested by POB. Second, widely-accepted programming categories (formats) facilitate the estimation of a model of product positioning where firms make discrete choices. Third, there is excess demand for station licences in most markets because of spectrum constraints. As a result it is possible to abstract away from station entry and exit decisions to focus on repositioning across formats.[11]

Because of the desirability of reaching the types of listeners who like music, fees are predicted to have only moderate, but still significant, long-run effects on the number of music stations. For example, 10% fees are predicted to reduce the number of music stations after 20 years by 9.4%, with music listening falling by 6.3%. Fees of 20% would produce changes that are approximately twice as large, and all of these effects would be much larger if there was less taste heterogeneity. My preferred estimates imply that the transition happens quite quickly with the majority of these long-run changes completed within 5 years. These counterfactual results contribute to the small literature that seeks to understand what determines product variety (examples include Borenstein and Netz (2002), George and Waldfogel (2003), Watson (2009), and also looking at radio Berry and Waldfogel (2001) and Sweeting (2010)), as well as the recent literature using static structural models to predict how specific policy-interventions or mergers would affect product characteristics (Fan (2012), Nishida (2012) and Datta and Sudhir (2012)). Jeziorski (2012) estimates a closely-related dynamic model to understand the benefits and costs of radio mergers.[12] While I do allow for multiple station ownership to affect format choices, the model presented here is simpler than Jeziorski's, which allows for endogenous mergers whereas I assume that firms expect the current ownership structure of the industry to persist. For this reason I focus on a later data period (2002-2005) when there was substantially less merger activity than immediately following the 1996 Telecommunications Act.[13]

[11] When evaluating the Performance Rights Act, the US GAO (2010), p. 25, argued that a reduction in the number of stations was unlikely given that the FCC reported substantial excess demand for broadcast licenses even when advertising revenues sank substantially in 2008. Of course, exceptionally high fees would likely induce exit.

[12] Mooney (2010a,b) and O'Gorman and Smith (2008) also estimate structural models to understand the effects of mergers in the radio industry.

[13] Relative to Jeziorski, my model is also simpler in assuming that I use a reduced form revenue function rather than structurally modeling the market for advertising, and I restrict the number of stations that a firm can move each period. With additional computation, both a structural model of advertising and endogenous mergers could be included in the current analysis.

The paper is structured as follows. Section 2 describes the data and Section 3 presents the model. Section 4 outlines the procedures used to solve and estimate the model, with full details in Appendices B and C. Section 5 presents the baseline coefficient estimates, and compares them with alternative estimates from other procedures, which are detailed in Appendix D. Section 6 presents the results of the counterfactual experiments to investigate the effects of different levels of fees. Section 7 concludes.

# 2 Data

I estimate the model using data from 102 local radio markets from 2002-2005, and the data comes from BIAfn's *MediaAccess Pro* database (BIAfn) unless otherwise noted. From the 274 markets that Arbitron surveyed throughout these years, the ten largest markets are dropped to reduce the computational burden, and I exclude 162 markets where more than 6% of listening was to stations based in other markets in order to avoid modeling cross-market interactions. By linking Arbitron's market definitions (which usually correspond to MSAs) to the US Census's *County Population Estimates*, I measure each market's population in 18 mutually exclusive demographic groups, which are the product of three age categories (12-24, 25-49 and 50 plus), three ethnic categories (black, white and Hispanic) and gender.[14] The sample markets range in population from Caspar, WY (with 55,000 people aged 12 and above in 2002) to Atlanta, GA (3.4 million).

The data contains Arbitron's estimates of stations' audience shares, which are based on listening by people aged 12 and above, in the Spring and Fall quarters each year. Potential listening by children younger than 12 is ignored. I use data from these quarters, so that there are two periods each year in my model, from Spring 2002 to Spring 2005. I exclude non-commercial stations (approximately 18% of all stations), as these stations lack audience data. I also drop 3% of stations that Arbitron did not report share estimates for in four or more of the sample quarters, because these estimates fell below its Minimum Reporting Standard.[15] I also drop AM stations that appear to have been simulcasting what was being broadcast on an FM station with the same owner. These deletions leave

[14] The *County Population Estimates* are calculated for July each year. I choose to assign these to the Spring quarter and use linear interpolation to find estimates for the Fall. My estimation of the process by which demographics evolve explicitly addresses the fact that I only observe demographics every other period (see Appendix C for details).

[15] The shares of stations that do not meet the MRS in a specific period are assumed to be 10% less than the smallest share that Arbitron does report the market in that quarter.

a sample of 2,375 stations and 16,566 station-market-period observations. I turn audience shares into market shares for estimating demand by assuming that every person in the market could listen to the radio for up to 6 hours per day (average listening is around $2\frac{1}{2}$ hours per day) and by using Arbitron's estimate of how many people listened to the radio in a given market-quarter.[16] BIAfn reports an estimate of advertising revenues, based on a proprietary formula, for 96% of station-years in the sample between 2002 and 2004. While it is impossible to know whether these estimates are accurate for individual stations, they are widely cited within the industry, so they should be useful for approximating how revenues per listener differ with demographics and across markets, which is how they are used here.

BIAfn lists a primary programming format for each station in each quarter, and categorizes these into 20 broader format categories.[17] I aggregate these categories into seven active formats, listed in Table 1, which pick up the main demographic variations in listener tastes. For example, News/Talk stations attract older, male listeners, Easy Listening and Religious stations attract an equal mix of men and women, and Urban and Spanish formatted stations attract Black and Hispanic audiences respectively. Comparing across markets, these differences in demographic tastes clearly affect format-switching, which will provide a source of identification when estimating the model. For example, ten times as many stations are moved to Spanish in markets with an above median proportion of Hispanics than in markets with a below median proportion. The table also shows that music stations tend to have higher audience shares, consistent with many listeners preferring music, and that AM band stations are primarily concentrated in non-music formats, particularly News/Talk. This reflects the fact that AM stations provide poorer audio quality for music programming, and the propensity of owners to switch AM stations to News/Talk will also provide a source of identification when the model is estimated. For the counterfactual, I assume that stations in the first four formats would have to pay fees.

[16] Arbitron's estimate of the proportion of time spent listening in each market (APR) is not reported in *MediaAccessPro*. However, I was able to collect APRs from M Street's STAR database for 2002 and Spring 2003 and BIAfn were able to provide me with these numbers from Fall 2004. The APRs for Fall 2003 and Spring 2004 were interpolated which is a reasonable approach as they change very slowly over time.

[17] A music station may play (say) a Country music song without being in the Country format, or have a personality-based morning show without being in the Talk format. Instead, the format reflects the most common type of programming on the station, and in the proposed legislation it was assumed that music stations would pay fees based on all of their advertising revenues without any attempt to allocate them between periods of music and talk programming. In the counterfactuals I ignore the fact that stations in non-music formats might choose to buy program-specific licences in order to play a few hours of music programming.

Table 1: Format Aggregation

| Aggregated Format (* = music format) | BIAfn Format Categories | Number of Station-Qtrs FM | Number of Station-Qtrs AM | Mean Audience Share | % of Listeners in Demographic: Male | 12-24 | 25-49 | Black | Hisp. |
|---|---|---|---|---|---|---|---|---|---|
| 1 AC/CHR* | Adult Contemporary, Contemporary Hit Radio, | 3,299 | 47 | 5.3% | 39.1 | 28.4 | 49.8 | 13.1 | 15.4 |
| 2. Rock, AOR/Classic Rock* | Rock, AOR/Classic Rock | 2,597 | 6 | 4.7% | 70.5 | 25.6 | 63.0 | 2.3 | 8.9 |
| 3. Country* | Country | 1,799 | 223 | 6.3% | 47.4 | 15.1 | 44.2 | 2.4 | 5.0 |
| 4. Urban* | Urban, Gospel | 1,172 | 609 | 4.5% | 43.5 | 29.3 | 50.2 | 83.6 | 5.0 |
| 5. News/Talk | News, Talk, Sports | 267 | 2,656 | 3.2% | 60.7 | 3.7 | 38.1 | 7.5 | 5.1 |
| 6. Spanish | Spanish-language | 562 | 460 | 2.2% | 51.4 | 20.0 | 59.9 | 0.8 | 91.3 |
| 7. Other Programming | Oldies, Easy Listening, Variety Classical, Jazz, Big Band, Religious (non-Gospel) | 1,873 | 911 | 3.2% | 47.1 | 6.8 | 36.4 | 9.9 | 10.2 |

Note: Demographic %s based on Arbitron's *Radio Today* publications 2003-2006 and author's calculations.

40 stations go temporarily off-air before returning to service, so I also include a non-active 'Dark' format in the choice set. On the other hand, only one station permanently closes during the sample, so I do not include an option of permanent exit in the model, and treat the exit of this station as an unanticipated shock.[18] I also do not model *de novo* entry into the industry, although 55 stations (almost all of which remain very small throughout the data) begin operating when the FCC grants new licenses. The entry of these stations is also treated as a unanticipated shock, and they are included in the Dark format in the period immediately before they start operating. A more complete model of entry and exit could be included with additional computation. BIAfn provides the ownership history of each station, including the month of each transaction, and in each period I assign ownership to the firm that owns the station at the end of that period.[19]

Stations in the same market-format can have quite different market shares. I allow for these differences to be explained by several observable variables, specifically AM band*format interactions, the proportion of the market's population covered by the station's signal (interacted with the station's band), an 'out of market' dummy for whether the station is based outside the geographic boundaries of the local market (e.g., in the surrounding countryside) and a dummy for whether the station has an imputed share in one or more periods. With the exception of the band-format interactions, firms treat these observed characteristics as permanent and fixed, which is a reasonable approximation as less than 1% of stations change signal strength or tower height between 2002 and 2006. The desirability of being a music station, in the absence of fees, is indicated by the fact that stations with the best signals tend to select into music formats: for example, on average the signals of AC/CHR, Country and Rock stations reach 85% of their market population, whereas the signals of Spanish or Other Programming stations only reach 70%.[20]

## 2.1 Summary Statistics

Table 2 contains summary statistics on the main market and station variables. On average, 3.2% of

[18] More generally, exit in the radio industry usually occurs for non-economic reasons such as the death or incarceration of the owner of a small station or the revocation of a license by the FCC due to breaches of FCC regulations.

[19] While I allow for firms to own multiple stations in the same market I ignore the fact that there are many large radio companies that own stations in multiple markets. Allowing for cross-market economies of scope complicates the model, and cannot be incorporated perfectly into the counterfactual without modelling the large number of markets that are not incorporated in the sample. Sweeting (2011) did allow for cross-market economies of scope from operating stations in different markets in the same format, and estimated them to be small.

[20] News station signals reach 85% of their market populations on average.

Table 2: Summary Statistics

| Market Characteristics | | | | | |
|---|---|---|---|---|---|
| Variable | Observations (market-periods) | Mean | Std. Dev. | Min. | Max. |
| Number of stations in market | 714 | 23.2 | 5.8 | 11 | 38 |
| Number of different station owners in market | 714 | 9.4 | 3.2 | 3 | 18 |
| Population (12 and above, in millions) | 714 | 0.73 | 0.75 | 0.06 | 3.62 |
| Proportion of population black | 714 | 0.11 | 0.12 | 0.002 | 0.340 |
| Proportion of population Hispanic | 714 | 0.09 | 0.10 | 0.005 | 0.583 |
| Combined listening to sample stations (% of total market[21]) | 714 | 35.0 | 3.3 | 22.8 | 45.9 |
| BIAfn Estimated Annual Market Advertising Revenues (2002-2005, $m.) | 408 | 53.5 | 66.0 | 3.9 | 403.7 |

| Station Characteristics (note: statistics exclude Dark stations) | | | | | |
|---|---|---|---|---|---|
| Variable | Observations (station-periods) | Mean | Std. Dev. | Min. | Max. |
| Market share | 16,481 | 1.5% | 1.3% | 0.0% | 10.5% |
| BIAfn Estimated Advertising revenues (station-year 2002-2004, $m) | 6,413 | 2.4 | 3.7 | 0.003 | 45.6 |
| Dummy for station located outside market boundaries | 16,481 | 0.06 | 0.24 | 0 | 1 |
| Dummy for AM band | 16,481 | 0.30 | 0.46 | 0 | 1 |
| Proportion of market covered by signal (out of market stations excluded)[22] | 15,464 | 0.79 | 0.36 | 0.002 | 1.1 |
| Dummy for station that has imputed market share in at least one quarter | 16,481 | 0.14 | 0.35 | 0 | 1 |
| Dummy for station that switches formats before next quarter | 14,120 | 0.032 | 0.18 | 0 | 1 |
| Market share of a switching station | 454 | 1.0% | 0.9% | 0.0% | 6.3% |

[21] The total market includes the outside good of not listening to radio and listening to non-commercial and commercial non-sample stations. Market definition allows for each individual to spend up to 6 hours listening to the radio between 6am and midnight.

[22] Signal coverage is defined relative to the market population and I cap it at 1.1 to address outliers that appeared to distort the demand estimates.

stations switch between active programming formats each period. This rate is quite similar across markets of different sizes (2.9% in markets with more than 1 million people, and 3.3% in the remaining markets, and the difference is not statistically significant) even though station revenues tend to be much higher in larger markets. In order to allow some of the parameters to vary across markets I define three "market-size groups" based on the 12+ population: above 1 million (26 markets), 0.25 million-1 million (43) and less than 0.25 million (33). Switching stations have lower market shares than other stations, and the market share of just under 60% of switching stations increases in the year after they switch. Markets are heterogenous in size, ethnic composition and total advertising revenues. The average owner in a market (a firm in the model) operates 2.5 stations, with the number varying from 1 to 8. There is a slight tendency for firms to cluster their stations in the same format (when two local stations with the same owner are drawn at random the probability they are in the same format is 0.195, compared with 0.160 for stations with different owners). However, multi-station firms are observed moving stations to formats where they own fewer stations almost as frequently as they are observed moving stations to formats where they own more stations (111 vs. 115 moves in the data). This balance explains why I find only slight evidence of economies of scope from operating stations in the same format when I estimate the model.

# 3 Model

## 3.1 Overview and Notation

Radio station owners (firms) $o = 1, ..., O_m$ in each market $m$ play an infinite horizon discrete time game with periods $t = 1, ..., \infty$. Markets are assumed to be completely independent of each other, and I ignore the effects of common ownership across markets. The exogenous characteristics of market $m$ ($X_{mt}$) are its population size, the proportion of its population in each of 18 age-gender-race/ethnicity groups, the growth rates of its white, black and Hispanic populations and the value of each listener to advertisers. Market demographics change over time due to these growth rates, and the growth rates evolve exogenously from period to period according to an AR(1) process, with normally distributed innovations.[23] For the rest of the presentation I suppress the market index,

[23] This means that the 6 white (black/Hispanic) population sub-groups grow at the rate of the white (black/Hispanic) population, which implies that the relative age and gender balance of the six sub-groups remains the same over time which is a close reflection of what happens in the data.

except where absolutely necessary. A firm $o$ owns a set of stations $S^o$. The set of players and stations are assumed to remain the same over time, so that in the model firms do not expect new entry, permanent exit or changes in ownership. Each station's quality, which affects its audience, consists of three components. The first component depends on observed characteristics, such as signal coverage and whether the station is located inside the market, and has a common effect across formats and is assumed to be fixed over time. The second component reflects a format-specific quality effect for AM stations, as this band provides relatively poor audio quality in music formats. The first two components are denoted $X_{st}\gamma^S$. The final component is a one-dimensional level of time-varying quality $\xi_{st}$ that is assumed to be independent of observed station characteristics, and which evolves according to an AR(1) process that is not controlled by the firm, except that $\xi_{st}$ may change discretely when $s$ changes formats. There are $F = 0, 1, ..., 7$ discrete formats, where format 0 is the Dark (temporarily off-air) format, and each station is in exactly one format each period. $F_{st}$ is a vector which indicates the format of station $s$ in period $t$.

Each period local firms generate revenues by selling their stations' audiences to advertisers. Station audiences are determined by the demographic make-up of the market, as demographics affect programming tastes, station quality characteristics and station programming formats. Revenues are then determined by the market price for listeners and the relative price of listeners in each demographic group. Demographic tastes for programming, market prices and relative prices are assumed to be fixed over time. Stations formats and characteristics, market prices and demographics are all publicly observed at time $t$, and in describing the model I use $\mathcal{M}_{j,o,t}$ to denote the collection of all of these publicly observed variables in a firm($o$)-specific state $j$ in period $t$. For ease of reference, Appendix A contains a complete list of the state variables. All state variables

In the game, owners choose the formats of their stations for the next period. $A_o(\mathcal{M}_{j,o,t})$ denotes the discrete set of actions (next period choices) available to $o$. The choice set is state-dependent because I assume that each firm can move at most one station per period, which limits the choice set of multi-station firms.[24] Each possible action is associated with a private information, iid payoff

[24] While 99.5% of firm-period observations satisfy this constraint, there are 28 observations where firms move two stations and 1 observation where a firm moves 3 stations. These observations are ignored when calculating the pseudo-likelihood as part of the estimation process, and it is assumed that all other local firms optimize assuming that other firms can only move one station each period. Relaxing this restriction in a dynamic model with eight formats would be burdensome, but I have investigated how allowing each firm to make two moves, rather than one move, affects the results in a two-period version of the model where firms only care about their revenues in the following period. This results in positive, but still fairly small, estimates of economies of scope which is sensible as 9 out of the 28 two move

shock $\varepsilon_{ot}(a)$ which is received when action $a$ is chosen. These shocks will be distributed Type 1 extreme value, scaled by a parameter $\theta^{\varepsilon}$ which can vary with market size.

## 3.2 Timing

State is market

Within each period $t$, the timing of the game is as follows:

1. each firm $o$ observes the current state, $\mathcal{M}_{j,o,t}$;

2. each firm $o$ pays fixed costs for each of its active stations. The cost of operating a station is reduced by $\theta^C$ when a station operates another station in the same format, creating a total cost saving of $C(\mathcal{M}_{j,o,t})\theta^C$ where $C(\mathcal{M}_{j,o,t})$ is simply a count of how many stations it operates in formats where it has multiple stations. Given my specification of repositioning costs, only $\theta^C$, and not the level of fixed costs, is identified, so I will proceed treating fixed costs as equal to zero;

3. each firm $o$ observes the private information shocks $\varepsilon_{ot}$ to its format choices, and makes its format choice $a_{ot}$;

4. each firm receives advertising revenues $\sum_{s \in S^o} R_s(\mathcal{M}_{j,o,t}|\gamma)$, where $\gamma$ are the parameters of the demand and revenue models, pays repositioning costs $W_o(a_{ot})\theta^W$ and receives the payoff shock $\varepsilon_{ot}(a_{ot})$;

5. $\mathcal{M}_{j,o,t}$ evolves to the state in the next period, reflecting firms' format choices, and the stochastic evolution of station qualities and the growth rates of the white, Black and Hispanic populations.

For the purposes of solving and estimating the model it is useful to define the firm's flow profit function $\pi_{ot}$ as including payoffs accruing from point 3 in the current period to point 2 in the next period, as next period's fixed cost savings will depend deterministically on the action chosen in the

observations involve a firm moving two stations to the same format at the same time, which would be unlikely without scope economies.

current period

format change

$$\pi_{ot}(a_{ot}, \mathcal{M}_{j,o,t}, \theta, \gamma) + \theta^{\varepsilon}\varepsilon_{ot}(a_{ot}) \tag{1}$$
$$= \underbrace{\sum_{s \in S^o} R_s(\mathcal{M}_{j,o,t}, \gamma)}_{\text{Period } t \text{ Advertising Revenues}} + \underbrace{\beta C_o(a_{ot})\theta^C}_{\text{Fixed Cost Savings}} - \underbrace{W_o(a_{ot})\theta^W}_{\text{Repositioning Costs}} + \underbrace{\theta^{\varepsilon}\varepsilon_{ot}(a_{ot})}_{\text{Payoff Shock For } a_{ot}}$$

where the $\theta$ parameters will be estimated using the dynamic model. $\beta = 0.95$, implying an annual discount factor of just above 0.9. This is lower than factors that are often used in dynamic models, but it is higher than the ones typically used to value radio stations (Albarran and Patrick (2005)).

## 3.3 Components of the Per-Period Payoff Function

I now describe each component of the payoff function.

### 3.3.1 Station Revenues ($R_s(\mathcal{M}_{j,o,t}|\gamma)$)

A station's revenues are the sum of the number of listeners that a station has in each demographic group, determined by a listener demand model, multiplied by the market price of each of these listeners, determined by a revenue function.

listener demand model

**Listener Demand** A station's audience in each of 18 demographic groups is determined by a static, discrete choice random coefficients logit model as a function of the state variables in its own market. Each consumer $i$ in this market chooses at most one station to listen to, and $i$'s utility if she listens to non-Dark station $s$ is

reference

The random coeff. on format preferences

$$u_{ist} = \gamma_i^R + X_{st}\gamma^S + F_{st}(\overline{\gamma^F} + \gamma_D^F D_i) + \xi_{st} + \varepsilon_{ist}^L \tag{2}$$
$$= \delta_{st}(F_{st}, X_{st}, \gamma^S, \overline{\gamma^F}, \xi_{st}) + \gamma_i^R + F_{st}\gamma_D^F D_i + \varepsilon_{ist}^L \tag{3}$$

where $\delta_{st}$ is the 'mean utility' of the station for a consumer with baseline demographics (white, male, aged 12-24) and $\varepsilon_{ist}^L$ is an iid logit shock to individual preferences. $X_{st}$ and $\xi_{st}$ (unobserved quality, which can be inferred from the estimated demand model) are assumed to be valued by all consumers in the same way. $\overline{\gamma^F}$ are the format preferences of baseline demographic consumers, while $\gamma_D^F D_i$

allows format preferences to vary with age, gender and race. $\gamma_i^R$, assumed to be distributed normal with mean zero and variance $\gamma^\sigma$, allows for heterogeneity in how much consumers value listening to the radio. Choice of the outside good, defined as not listening to one of the commercial stations included in the model, is assumed to give utility of $\varepsilon_{i0t}^L$.

This is a rich specification, but it makes two significant simplifications. First, consumers are assumed to choose at most one station, whereas in reality people listen to several stations for different lengths of time during a period (ratings quarter). This is a standard simplification when using aggregate data (e.g., Nevo (2001)).[25] Second, the model is entirely static, whereas listening habits might make shares adjust slowly to changes in formats. This simplification seems reasonable as I consider only major programming changes, that should be obvious to all consumers, and six-month time periods, which are likely to be longer than the time required for listeners to adjust.[26]

**Revenues Per Listener** The advertising revenue that a station $s$ receives for a listener with demographics $D_d$ is determined by a parametric function

? why ...

$$r_{st}(D_d) = \gamma_{my}(1 + Y_{st}\gamma^Y)(1 + D_d\gamma_d) \tag{4}$$

where $\gamma_{my}$ is a market-year fixed effect and $\gamma_d$ allows advertiser valuations to vary with listener demographics.[27] In order to allow for format switching and market structure to affect revenues, the variables in $Y$ include the number of other stations that the firm has in the format, the number of other stations in the format and a dummy variable for whether the station switched formats in the previous period. Total station revenues $R_{st}(\mathcal{M}_{j,o,t}|\gamma)$ are the sum, over the 18 demographic groups, of the number of listeners multiplied by these prices. Total station revenues (at the annual level), aggregate share data and the predictions of the listener demand model about how many listeners the station will have in each demographic group are used to estimate these parameters, as station-level prices per listener by demographic are not available.

? used to estimate

The listener utility and revenue specifications (2) and (4) do not contain time effects, even though

[25] It can be rationalized as a representation of consumers' preferences during shorter time periods, which are aggregated to give period market shares. This representation, which assumes iid draws in preferences across the shorter time periods, is adequate as long as stations and advertisers are indifferent between audiences of the same size made up of either a few people who listen a lot, or a lot of people who listen a little.

[26] The example in footnote 10 suggests that stations that make major programming changes may take actions to encourage their old listeners to move to other stations, in order to avoid a protracted period of complaints.

[27] The coefficients of this reduced-form function are assumed to stay the same when performance fees are introduced.

So essentially no time fixed effect

broadcast radio audiences have been falling slowly since the late 1980s and revenues per listener were rising during the sample period. When I estimate these models I do include time effects to avoid biasing other coefficients, but when estimating the dynamic model I assume that all firms expect the current values of these time effects to remain fixed into the future, as methods for solving and estimating dynamic games assume stationarity. This is a simplification but it is consistent with the fact that real revenues, which are what firms ultimately care about, changed very little during the sample, as the trends in listenership and revenues per listener approximately cancelled out.[28,29]

?

### 3.3.2 Repositioning Costs ($W_o(a_{ot})\theta^W$)

A firm pays a repositioning cost when it moves a station to a new format. I allow repositioning costs to vary across the market size groups, with the revenues of the station being moved and whether the station was moved in the previous period, and to depend on whether the switch is between active formats or to or from Dark.

switching cost

### 3.3.3 Fixed Costs and Economies of Scope ($\beta C_o(a_{ot})\theta^C$)

fixed cost of being active
reference point

Because I estimate the value of repositioning costs to and from the Dark format, the common fixed cost incurred by all active stations is not identified, but any reduction in fixed costs from operating multiple stations in the same format is identified from the format choices of multi-station firms who must trade off this efficiency against audience cannibalization. As a simple specification for these economies of scope, I assume that the fixed costs of any station operating in the same format as one of its sister stations are reduced by $\theta^C$, which can vary across market size groups.

[28] The Radio Advertising Bureau estimates annual industry revenues from 2002 to 2006 of \$19.4 bn., \$19.6 bn., \$20.0 bn. and \$20.1 bn. (personal correspondance, November 29, 2010).

[29] Given my counterfactual, it is important that these trends are common across formats. This appears to be approximately the case in the data. For example, based on Arbitron's *Radio Today* reports, time spent listening between 2002 and 2005 fell by 0.45% per period for the population as a whole (implying a drop in listening of 6.5 minutes per day over the entire sample period), 0.47% per period for blacks and 0.35% for Hispanics who were being served by more Spanish language stations over this period. When I regress station revenues per listener on format dummies, market dummies, year dummies and year*format interactions, the coefficients on the year*format interactions are jointly insignificant (p-value 0.3142) which also suggests that revenues per listener were changing in a similar way across formats over time.

what's $\theta^\varepsilon$

### 3.3.4 Payoff Shocks ($\theta^\varepsilon \varepsilon_{ot}(a_{ot})$)

Firms receive iid (across firms and over time) private information shocks to their payoffs from each possible format choice, including keeping stations in the same format. These shocks are drawn from a Type 1 extreme value distribution, scaled by a parameter $\theta^\varepsilon$, which can vary across the market-size groups. These scale parameters are identified because revenues are treated as observed when estimating the dynamic model. These shocks should capture all factors that affect a firm's format choices but are not captured in the expected revenue or cost functions assumed by the model. For example, the owner of the station that switched to Sports programming discussed in footnote 10 had an existing business that sold advertising for local sports facilities, leading to a possible synergy that is not captured by the model. In the same example, the firm tried to keep its plan to move to Sports programming secret until the move was made, which helps to rationalize the private information assumption.[30] The strong assumption is that the $\varepsilon$s are serially uncorrelated, and this is required for tractability.

Not latent utility model

So $\theta^\varepsilon$ is observed

## 3.4 Evolution of the State Variables

At the end of period $t$, the state variables evolve for the following period. Station formats change deterministically with firms' choices. Unobserved station quality is assumed to evolve according to an AR(1) process with normally distributed innovations that are iid across stations

$$\xi_{st} = \rho^\xi \xi_{st-1} + \nu^\xi_{st} \text{ if } F_{st} = F_{st+1} \quad (5)$$

$$= \rho^\xi \xi_{st-1} + \nu^\xi_{st} + \gamma^\xi \text{ if } F_{st} \neq F_{st+1} \quad (6)$$

fixed shift

where $\nu^\xi_{st} \sim N(0, \sigma^2_{v_\xi})$.[31] The $\gamma^\xi$ term allows for a fixed shift in quality when the firm changes format. I assume that firms do not know the innovations $v^\xi_{st+1}$ innovations when they make format choices

[30] A standard objection to the private information assumption in static models (e.g., Seim (2006)) is that it can lead to firms experiencing ex-post regret, because, for example, more firms choose the same location than was expected. However, in my model the rate of switching is relatively low and in the data it is relatively rare for two firms in the same market to make switches that would have a large impact on the expected profitability of each other's switch. In a dynamic model firms are also able to quickly reverse choices that turn out to be particularly sub-optimal.

[31] It is not necessary to assume that the innovations are normal and the process is estimated without imposing normality. However, a normal probability plot indicates that the implied distribution of innovations matches a normal distribution very well except at percentiles below 5% and above 95%, and drawing from the empirical distribution of innovations gives very similar results.

$\xi_{ct}$ & $X_{sc}$

in period $t$, which allows me to form moments for consistently estimating demand based on these innovations which should be valid even if the levels of $\xi$ affect format choices. By assumption, this rules out some forms of selection that might drive format repositioning. However, I provide evidence that my modeling assumptions are consistent with the data by showing that I can closely match the empirical distribution of share changes for stations that switch formats, even when I estimate the parameters $\rho^{\xi}$ and $\sigma^2_{v_\xi}$ using only stations that do not switch.

u changes

While listener demand depends on 18 mutually exclusive age-gender-ethnic/racial groups, it would cumbersome to model the evolution of the population in so many groups. Instead, I model the growth rate for each ethnic/racial group (white, black and Hispanic) and assume that the same growth rate applies to each of the associated age-gender groups. I assume that for ethnic group $e$

$$\log\left(\frac{pop_{met}}{pop_{met-1}}\right) = \tau_0 + \tau_1 \log\left(\frac{pop_{met}}{pop_{met-1}}\right) + u_{met} \tag{7}$$

population

which allows population growth for particular groups to have the type of serial correlation that is observed in the data.[32] This particular specification also allows me to address the problem that population estimates are annual (see Appendix C for a detailed discussion).

what's the purpose! this

For the firm

### 3.5 Value Functions and Equilibrium Concept

As in almost all of the literature following Ericson and Pakes (1995), I assume that firms use stationary Markov Perfect Nash Equilibrium (MPNE) pure strategies.[33] A stationary Markov Perfect strategy for a firm $o$, $\Gamma_o$, is a mapping from any state $(\mathcal{M}_{j,o,t}, \varepsilon_{ot})$ to an action $a_{ot}$ that does not depend on $t$. I use $\Gamma$ to denote the strategies of all firms. $V_o^{\Gamma}(\mathcal{M}_{j,o,t}, \varepsilon_{ot})$ defines a firm's value in a particular state when it uses an optimal strategy and other firms use strategies defined in $\Gamma$. By Bellman's

[32] Alho and Spencer (2005), Chapter 7 discuss the application of time series models, including AR(1) to demographic growth rates. Models with additional lag terms would complicate the state space of the dynamic model.

[33] With continuous states it is an assumption that a pure strategy MPNE exists. Dorazelski and Satterthwaite (2010) prove the existence of a pure strategy MPNE for a model with discrete states when the random component of payoffs has unbounded support. Conceptually it would be possible to convert my model into one where existence was guaranteed by using an arbitrarily fine discretization of the continuous states. A similar argument is made by Jenkins et al. (2004).

Markov Perfect Nash Equilibrium stationary?

Dynamic IO chu

三个项课程 8

optimality principle

$$\boxed{V_o^{\Gamma}(\mathcal{M}_{j,o,t}, \varepsilon_{ot})} = \max_{a \in A_o(\mathcal{M}_{j,o,t})} \left[ \begin{array}{c} \pi_{ot}(a, \mathcal{M}_{j,o,t}) + \theta^{\varepsilon}\varepsilon_{ot}(a) + \\ \beta \int V_o^{\Gamma}(\mathcal{M}_{h,o,t+1}) g(\mathcal{M}_{h,o,t+1}|a, \Gamma_{-o}, \mathcal{M}_{j,o,t}) d\mathcal{M}_{h,o,t+1} \end{array} \right] \tag{8}$$

where $g()$ is the transition density when $o$ chooses $a$ and other firms use strategies $\Gamma_{-o}$. Given the distribution of the payoff shocks, an optimal strategy for firm $o$ will map into conditional choice probabilities

$$P^{\Gamma_o}(a, \mathcal{M}_{,o,t}, \Gamma_{-o}) = \frac{\exp\left(\frac{v_o^{\Gamma}(a, \mathcal{M}_{j,o,t}, \Gamma_{-o})}{\theta^{\varepsilon}}\right)}{\sum_{a' \in A_o(\mathcal{M}_{j,o,t})} \exp\left(\frac{v_o^{\Gamma}(a', \mathcal{M}_{j,o,t}, \Gamma_{-o})}{\theta^{\varepsilon}}\right)} \tag{9}$$

where $v_o^{\Gamma}(a, \mathcal{M}_{j,o,t}, \Gamma_{-o})$ is a choice-specific value function which excludes the current payoff shock

$$\boxed{v_o^{\Gamma}(a, \mathcal{M}_{j,o,t}, \Gamma_{-o})} = \pi(a, \mathcal{M}_{j,o,t}) + \beta \int V_o^{\Gamma}(\mathcal{M}_{h,o,t+1}) g(\mathcal{M}_{h,o,t+1}|a, \Gamma_{-o}, \mathcal{M}_{j,o,t}) d\mathcal{M}_{h,o,t+1} \tag{10}$$

and $\boxed{V_o^{\Gamma}(\mathcal{M}_{h,o,t+1})}$ is the firm's value in state $\mathcal{M}_{h,o,t+1}$ before that period's payoff shocks are realized.

Intermediate value function.

# 4 Methods Used to Solve and Estimate the Model

In this Section I outline the methods used to solve and estimate the model, focusing on how I incorporate parametric approximations to the value function into existing procedures that have been applied to dynamic models with smaller state spaces. Complete details are given in Appendices B and C. I begin by discussing how the model is solved, as estimation of the dynamic model involves some simple changes to the solution procedure.

Now we have large state spaces

## 4.1 Solution Method

For exposition it is useful to define a firm's expected flow profits (prior to the realization of the $\varepsilon$s) in a particular state when it uses optimal choice probabilities $P_o$ as

$$\widetilde{\pi}(P_o(\mathcal{M}_{j,o,t})) = \sum_{s \in S^o} R_s(\mathcal{M}_{j,o,t}|\gamma) + \sum_{a \in A_o(\mathcal{M}_{j,o,t})} P_o(a|\mathcal{M}_{j,o,t}) \left( \begin{array}{c} \beta C_o(a)\theta^C - W_o(a)\theta^W + \\ +\theta^{\varepsilon}(\varkappa - \log(P_o(a|\mathcal{M}_{j,o,t})) \end{array} \right) \tag{11}$$

where $\varkappa$ is Euler's constant. If we were considering a finite set of states, then we could express value functions given a set of choice probabilities $P$ (for all firms in all states) as a set of equations

$$V^P = [I - \beta E_P]^{-1} \widetilde{\pi}(P) \tag{12}$$

where $E_P$ is the Markov operator corresponding to policies $P$, so that the $(i,j)^{th}$ element is the probability of moving from state $i$ to state $j$ given strategies $P$.

A standard way of trying to solving a dynamic model is to use policy iteration (Judd (1998), Rust (2000)), which involves repeatedly iterating two steps. At iteration $i$, in the first step (*policy valuation*), (11) and (12) are applied to calculate values $V^{P^i}$ associated with choice probabilities $P^i$ that may not be optimal. In the second step (*policy improvement*), these values $V^{P^i}$ are used to update $P^i$ by computing choice-specific value functions

$$v_o^{P^i}(a, \mathcal{M}_{j,o,t}) = \pi(a, \mathcal{M}_{j,o,t}) + \beta \int V_o^{P^i}(\mathcal{M}_{h,o,t+1}) g(\mathcal{M}_{h,o,t+1} | P^i_{-o}, \mathcal{M}_{j,o,t}, a) d\mathcal{M}_{h,o,t+1} \tag{13}$$

where $g()$ is now the transition density as a function of choice probabilities, and applying formula (9). Iteration continues until both values and choice probabilities converge, up to a pre-specified numerical tolerance.

As the state space is exceptionally large and some state variables are continuous it is impossible to solve for values at all states. Instead I choose a specific set of $N$ states, and assume that the value function can be approximated by a parametric linear function of $K$ functions ($\phi$) of the state variables i.e.,

$$V_o^{P^*}(\mathcal{M}_{j,o,t}) \simeq \sum_{k=1}^{K} \lambda_k \phi_{ko}(\mathcal{M}_{j,o,t}) \tag{14}$$

Solving the value function now requires finding $K$ $\lambda$ coefficients rather than $N$ values. Stacking the equations for each of the $N$ selected states in matrix form, the following equations should hold for equilibrium strategies $P^*$

$$\Phi\lambda = \widetilde{\pi}(P^*) + \beta E_{P^*} \Phi\lambda \tag{15}$$

where $\Phi$ is the matrix of the functions of the state variables and $E_{P^*}\Phi$ is a matrix with element $(j,k)$

$$E_{P^*}\Phi_{j,k} = \int \phi_{ko}(\mathcal{M}_{h,o,t+1}) g(\mathcal{M}_{h,o,t+1} | P^*, \mathcal{M}_{j,o,t}) d\mathcal{M}_{h,o,t+1} \tag{16}$$

where row $j$ is associated with state $\mathcal{M}_{j,o,t}$. The choice of states $\mathcal{M}_{h,o,t+1}$ to approximate the integral is described in Appendix B.

For the overidentified case ($N > K$), $\widehat{\lambda^P}$ can be found using an OLS estimator

$$\widehat{\lambda^P} = ((\Phi - \beta E_P \Phi)'(\Phi - \beta E_P \Phi))^{-1}(\Phi - \beta E_P \Phi)'\widetilde{\pi}(P) \tag{17}$$

The *parametric policy iteration* procedure (Benitez-Silva et al. (2000)) now consists of iterating several steps. Before the procedure begins, I calculate $\Phi$ for the $N$ selected states, which include the observed states and a 499 duplicates of these states where the features that can vary over time are perturbed. The variables in the approximation include measures of firm revenues, opportunities to increase revenues by switching formats (given the current configuration of competitors) and firm and rival characteristics. For the counterfactual the model is solved market-by-market, so no restrictions that the approximating function has to be the same across markets are imposed. For an iteration $i$, the following steps are used:

1. calculate $\widetilde{\pi}(P^i)$ and $E_{P^i}\Phi$.
2. create matrices $(\Phi - \beta E_{P^i}\Phi)$ and use (17) to calculate $\widehat{\lambda^{P^i}}$
3. use $\widehat{\lambda^{P^i}}$ to calculate the choice-specific value functions for each choice for each firm and the multinomial logit formula (9) to calculate updated probabilities $P'$
4. if the maximum absolute difference between $P'$ and $P^i$ is sufficiently small (I use a tolerance of $1e-5$) the procedure stops and $\widehat{\lambda^{P^i}}$ is saved as $\lambda^*$; otherwise $P^{i+1} = \psi P' + (1-\psi)P^i$, and iteration $i+1$ begins at step 1. I use $\psi = 0.1$.

This procedure solves for the conditional choice probabilities in the observed states. However, to perform the counterfactual, I need to simulate the model forward to states that which will not have been included in the initial selection of $N$ states. Therefore, in each future period, $\lambda^*$ is used to solve for equilibrium choice probabilities. Full details of this procedure are in Appendix B.

## 4.2 Estimation

Estimation of the model proceeds in two main stages. The first stage involves estimation of (i) the listener demand model and the process for the unobserved ($\xi$) component of station quality; (ii) the why can you do that?

revenue function; (iii) an initial guess of firms' conditional choice probabilities; and (iv) the transition process governing demographics. The methods used are based on existing literature, and complete details are in Appendix C. The main innovation is that I formulate demand moments that are based on *innovations* in unobserved quality. Under the timing assumption that firms have no knowledge of these innovations when they make format choices, this allows for consistent estimation of demand even when format choices are affected by the level of quality. The second stage involves estimation of repositioning costs and economies of scope using the dynamic model. As the particular method I use is new, I discuss the procedure in the text, but Appendix C provides additional detail. Appendix D describes the implementation of methods based on forward simulation that are used as robustness checks.

fixed cost saving

### 4.2.1 Estimation of the Dynamic Model

My preferred estimates come from an estimator which combines parametric approximation of the value function with a pseudo-likelihood procedure that follows AM (2010) (discussed in Aguirregabiria and Nevo (2012)). This procedure follows a similar iterative procedure to the one used to solve the model, with an added estimation step, but the choice probabilities of other firms ($P_{-o}$) are held constant at initial first-stage estimates.[34]

Before the procedure begins, I calculate $\Phi$ for the $N$ selected states, which include the observed states and 9 duplicates of these states ($N = 60,610$) where the features that can vary over time are perturbed. The variables in the approximation include measures of firm revenues, opportunities to increase revenues by switching formats (given the current configuration of competitors), and firm and rival characteristics. For estimation it is necessary to pool markets together, so the variables in the approximating function are constrained to be the same across markets, although market-quarter fixed effects and interactions with market characteristics, provide flexibility.

In the iterated procedure, the following steps are followed in iteration $i$ where the current guess of the structural parameters is $\theta^i$ and the current guess of $o$'s choice probabilities is $P_o^i$ (see Appendix

[34] I describe an iterative procedure to solving the dynamic model where, the approximation to the value function is calculated each time the structural parameters $\theta$ are updated. This works well for the parametric functions that I consider. An alternative approach would be to estimate both the structural parameters and the parameters of the parametric approximation as part of a single maximization problem with a very large set of equality constraints describing the equilibrium conditions of the model (MPEC). Su and Judd (2011) argue that an MPEC approach has superior numerical properties.

C for all of the details):

1. calculate $\widetilde{\pi}(P_o^i, \theta^i)$ and $E_{P^i}\Phi$

2. create matrices $(\Phi - \beta E_{P^i}\Phi)$ and use (17) to calculate $\widehat{\lambda^i(\theta^i, P_o^i)}$

3. use $\widehat{\lambda^i(\theta^i, P_o^i)}$ to approximate the future value from making each choice $a$

$$FV(a,\mathcal{M}_{j,o,t}, P_o^i, P_{-o}) = \sum_{k=1}^{K} \left\{ \left( \int \phi_{ko}(\mathcal{M}_{h,o,t+1}) g(\mathcal{M}_{h,o,t+1}|P_o^i, P_{-o},\mathcal{M}_{j,o,t}, a) d\mathcal{M}_{h,o,t+1} \right) \widehat{\lambda_k^i(\theta^i, P_o^i)} \right\} \tag{18}$$

   (this is not quite the same as the choice-specific value function defined before which also included current revenue and repositioning costs associated with action $a$)

4. estimate the structural parameters $\theta'$ using a pseudo-likelihood estimator where the probability that an action $a$ is chosen is

$$\frac{\exp\left(\frac{FV(a,\mathcal{M}_{j,o,t},P_o^i,P_{-o}) - W_o(a)\theta^{W\prime} - \beta C_o(a)\theta^{C\prime}}{\theta^{\varepsilon\prime}}\right)}{\sum_{a' \in A_{ot}} \exp\left(\frac{FV(a,\mathcal{M}_{j,o,t},P_o^i,P_{-o}) - W_o(a')\theta^{W\prime} - \beta C_o(a)\theta^{C\prime}}{\theta^{\varepsilon\prime}}\right)} \tag{19}$$

   where current revenues drop out because they are common across choices. I also report results based on using a moment-based estimator in this step, where a set of switching rates, which should be informative about the parameters, are matched (e.g., the overall switching rate between active formats, and the rate at which stations move to Spanish in markets with low and high Hispanic populations). POB argue that estimators based on this type of objective function may suffer from less bias in finite-samples because of averaging across states. Observations for firms moving more than one station are excluded when calculating the objective function.

5. if the maximum absolute difference between $\theta'$ and $\theta^i$ and between $P_o'$ (based on (19)) and $P_o^i$ is less than 1e-4 the procedure stops. Otherwise the algorithm returns to step 1 using $P_o^{i+1} = \psi P_o' + (1-\psi)P_o^i$ and $\theta^{i+1} = \psi\theta' + (1-\psi)\theta^i$ where $\psi = 0.1$.

Standard errors are calculated using a bootstrap where markets are re-sampled.

I also report results using the AM (2007) NPL procedure where $P_{-o}$ is also updated during estimation. This could be advantageous if the initial estimates of choice probabilities are inaccurate, but there is no guarantee that this type of procedure will cause the probabilities to converge to their true equilibrium values unless the initial estimates were already good and the equilibrium satisfies a particular kind of stability condition (Aguirregabiria and Nevo (2012)). Pesendorfer and Schmidt-Dengler (2008, 2010) illustrate the problems that can result when NPL updating is applied to models with multiple equilibria, and multiple equilibria may certainly exist in my model. Partly for this reason, I view the NPL results primarily as evidence of the robustness of the estimates where $P_{-o}$ is held fixed.

# 5 Empirical Results

This section presents the coefficient estimates, including robustness checks on the estimation of the dynamic parameters.

## 5.1 Listener Demand Model

Tables 3 and 4 report the estimated parameters of the listener demand model. The very large value of $\gamma^{\sigma}$ implies that there is relatively little substitution between listening to commercial radio stations and the outside good.[35] The demographic*format coefficients show that there is a lot of demographic heterogeneity in programming tastes, and they are precisely estimated because demographic*format-specific moments are used following Petrin (2002). For example, older listeners value all radio programming more than other listeners, but they particularly like News, Country and Other Programming, while blacks and Hispanics prefer Urban and Spanish programming respectively. The important implication of these coefficients for the counterfactual is that when a music station switches to non-music programming it will lose most of its younger, or black, listeners to the remaining music stations.

As expected, the station characteristic coefficients indicate that the AM band stations are better suited to News/Talk than other formats, that greater signal coverage increases quality and that out-

[35] The coefficients on the time dummies, which are not reported, indicate that the utility from radio listening was decreasing slowly over time.

Table 3: Estimates of Format Taste Parameters

| | Mean Tastes | $\gamma^{\sigma}$ | Age 25-49 | Age 50 plus | Female | Black | Hispanic |
|---|---|---|---|---|---|---|---|
| Radio Listening | -15.8753 | 16.7749 | 2.1404 | 7.7839 | 2.8197 | 3.6171 | 2.8563 |
| | (0.9224) | (1.3946) | (0.2046) | (0.7247) | (0.2098) | (0.3303) | (0.4419) |
| Format Interactions (AC/CHR excluded) | | | | | | | |
| Rock | 0.6575 | - | 0.2813 | -0.4054 | -1.2581 | -1.8814 | -0.8157 |
| | (0.0580) | | (0.0087) | (0.0158) | (0.0043) | (0.0292) | (0.0402) |
| Country | -0.1187 | - | 0.4847 | 1.3039 | -0.4428 | -1.9726 | -1.2733 |
| | (0.0646) | | (0.0095) | (0.0175) | (0.0046) | (0.0330) | (0.0524) |
| Urban | -1.2040 | - | -0.7373 | -1.0188 | -0.3996 | 3.9375 | 0.6158 |
| | (0.0832) | | (0.0393) | (0.0497) | (0.0086) | (0.0406) | (0.0501) |
| News/Talk | -1.2918 | - | 1.6979 | 3.1485 | -1.1171 | -0.7935 | -1.1071 |
| | (0.1443) | | (0.0080) | (0.0144) | (0.0049) | (0.0275) | (0.0385) |
| Other Programming | -0.9883 | - | 1.0958 | 2.4600 | -0.5384 | -0.4204 | -0.2916 |
| | (0.0650) | | (0.0079) | (0.0150) | (0.0049) | (0.0275) | (0.0385) |
| Spanish | -2.7945 | - | 1.0300 | 1.1111 | -0.3649 | -0.5138 | 3.9489 |
| | (0.1955) | | (0.0264) | (0.0506) | (0.0163) | (0.1519) | (0.1694) |

Notes: 16,481 observations, GMM objective function 2.90e-12, standard errors in parentheses.
Mean tastes will reflect valuations of a white male aged 12-24.

Table 4: Estimates of Station Quality Parameters

| | | | |
|---|---|---|---|
| AM * AC/CHR or Rock | -0.7781 | FM * Signal Coverage | 0.6057 |
| | (0.3174) | | (0.1068) |
| AM * Country | -0.4538 | Small Station Dummy | -0.4277 |
| | (0.1776) | (shares imputed for some quarters) | (0.0580) |
| AM * Urban | -0.3523 | Out of Market Dummy | -0.5082 |
| | (0.1296) | | (0.0834) |
| AM * News/Talk | -0.0806 | TRANSITION PROCESS | |
| | (0.1658) | FOR UNOBSERVED QUALITY | |
| AM * Other | -0.3811 | $\rho^{\xi}$ | 0.8421 |
| | (0.1196) | | (0.0058) |
| AM * Spanish | -0.2714 | $\sigma_{v^{\xi}}$ | 0.3132 |
| | (0.1593) | | (0.0020) |
| Signal Coverage | 1.5938 | Effect of Format Switch | -0.0501 |
| (for stations located in the market) | (0.1004) | on Unobserved Quality | (0.0103) |

Notes: 16,481 observations, GMM objective function 2.90e-12. Time coefficients not reported. Std. errors in parentheses. AM*AC/CHR and *Rock combined due to small number of observations.

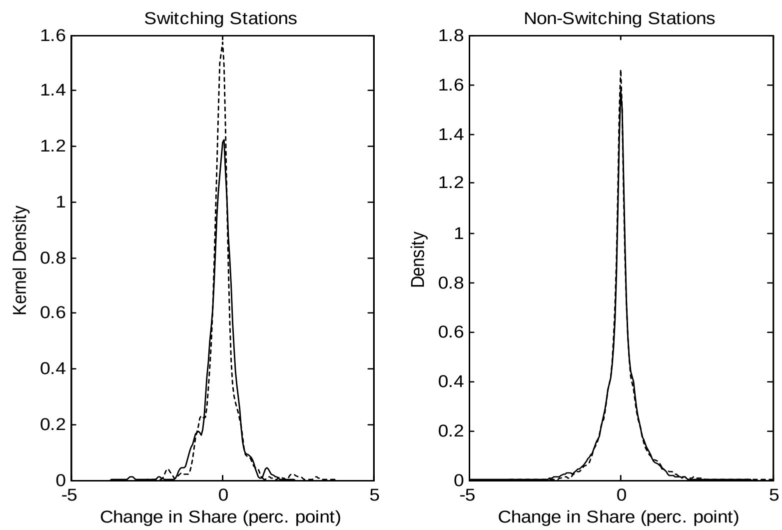
of-market stations, which likely only have partial coverage, have lower quality. Unobserved station quality is estimated to be quite persistent ($\rho^{\xi} = 0.84$), while a format switch results is estimated to cause a small but statistically significant *drop* in quality at the time of the switch.

**Fit of the Listener Demand Model.** The listener demand model plays an important role in the counterfactual because it predicts a station's audience in different formats. Figure **??** shows that the estimated model does well at matching the distribution of period-to-period share changes observed in the data for stations that switch formats, as well as those that remain in the same format. The transition process for unobserved station quality is estimated (apart from the intercept term) using only stations that remain in the same format, and the fact that predicted and actual distributions fit well suggests that my assumption that the processes are the same is reasonable.[36,37]

[36]For example, if I assumed that a station received a new draw of $\xi$ when it switched formats then, because the std. deviation of the $\xi$s is much greater than the std. deviation of the innovations (0.86 vs. 0.31), the variance of share changes for switching stations would be much greater, whereas it is only slightly greater. It is harder to rule out the possibility that firms know something about the innovations that their stations will receive in different formats when choices are made, because of the complicated forms of selection that this might introduce. Allowing for this type of selection in entry models is a topic of on-going research (e.g., Roberts and Sweeting (2012)).

[37]One can also calculate the correlation between the share change predicted by the model (based on one simulation draw of $\nu$) and the share change observed in the data for switching stations. This correlation is 0.28, and it is statistically significant at any standard significance level.

Figure 1: Distribution of Share Changes: Data (solid line), Model Prediction (dashed line, one simulation draw)




## 5.2 Revenue Function

Table 5 shows the coefficient estimates from two specifications of the revenue function (market-year coefficients are not reported). The first specification assumes that only demographics affect a listener's value to a station, and the effects are measured relative to white, males aged 25-49, while the second specification, that will be used in the rest of the analysis, allows for format switching, common ownership and the number of stations in the same format (competition) to affect revenues. The demographic coefficients are similar in both specifications, and they indicate that advertisers place different values on listeners with different demographics. In particular, a female listener aged 25-49 is 17-18% more valuable than a male of the same age, a black or Hispanic is 20% less valuable than a white listener, and young and old listeners are worth less than those aged 25-49.[38] The fact that male, older and Hispanic listeners are less valuable will limit how many stations will want to switch from music programming when performance fees are introduced. In the second specification, a format switch is predicted to reduce station revenues by 10%, which may reflect the fact that switching stations discount the price of commercials while they develop new relationships, but the

[38] The age effects are particularly large. The discount for young listeners is consistent with the fact that these listeners are particularly likely to switch stations to avoid commercials (Speck and Elliot (1997)). The local advertising market for older consumers may be more competitive, because they tend to consume several local media. For example, the Radio Advertising Bureau estimates that people aged above 64 are twice as likely to read a local newspaper as those aged 18-34 (http://www.rab.com/public/mediafacts/details.cfm?id=8, accessed November 24 2010).

Table 5: Parameter Estimates for the Revenue Function

| | (1) | (2) |
|---|---|---|
| DEMOGRAPHICS | | |
| Female | 0.1797 | 0.1917 |
| | (0.0368) | (0.0374) |
| Age 12-24 | -0.5811 | -0.5883 |
| | (0.1075) | (0.1084) |
| Age 50+ | -0.4531 | -0.4572 |
| | (0.0577) | (0.0581) |
| Black | -0.1964 | -0.1961 |
| | (0.0148) | (0.0155) |
| Hispanic | -0.1593 | -0.1596 |
| | (0.0159) | (0.0159) |
| STATION CHARACTERISTICS AND COMPETITION | | |
| Number of stations with same owner in format | - | 0.0064 |
| | | (0.0047) |
| Number of other stations in format | - | -0.0019 |
| | | (0.0018) |
| Format switch in previous quarter | - | -0.1045 |
| | | (0.0279) |
| $R^2$ (compared to a model with only market-year fixed effects) | 0.3182 | 0.3208 |

4,483 annual station observations (observations with imputed shares excluded). Market-year coefficients not reported. Standard errors corrected for imprecision in the demand parameters.

other coefficients have no statistically significant effect.

The reported $R^2$s indicate that the model only explains some of the within-market variation in per listener revenues. However, the model does a reasonable job of predicting how station revenues change over time. For example, conditional on the observed change in audience, the correlation between observed year-to-year changes in station revenues and those predicted by the model for stations switching formats is 0.55.

## 5.3 Dynamic Parameters

Repositioning costs and economies of scope are estimated using the dynamic model. After presenting the results based on estimates that use parametric approximations to the value function, I present the results from the alternative estimators that approximate the value function using forward simulation.

### 5.3.1 Estimates Based on Parametric Approximations of the Value Function

Table 6 shows the coefficient estimates from three different estimators where the value function is approximated using the same parametric function. Specifications (1) and (3) are the modified and iterated estimators with a pseudo-likelihood objective function, while specification (2) is the modified estimator where the objective function is based on matching various switching rates following POB, and these estimates should be compared with the estimates in column (1). Standard errors are reported based on a bootstrap where markets are re-sampled with 20 replications.

As stations are rarely moved to Dark, the key parameters for the counterfactual are the costs of switching between active formats and the scale parameter of the $\varepsilon$s. All of the estimates indicate that these parameters vary systematically with market size, which allows the model to match the stylized fact that format switching rates are similar in small and large markets, even though average firm revenues (and hence values) are quite different.[39] One interpretation is that format switching is costly primarily because of the costs of marketing the station to new listeners (advertising will cost more in larger markets) and losing the goodwill in relationships with existing advertisers, as the value of goodwill may be proportional to revenues. The positive, but statistically insignificant, coefficients on the prior revenue of the station being moved (the mean of this variable is 0.92) is also consistent with this interpretation. Only one of the economies of scope coefficients is statistically significant at the 5% level, although 5 out of 6 coefficients are positive for markets with populations above 0.25m.. These weak results are consistent with the fact that in this sample of data, and using my format definitions, firms are not systematically making their stations more clustered in particular formats. However, because I am attempting only to estimate the size of incremental economies of scope from operating stations in the same format, these weak results are not inconsistent with Jeziorski's (2012) and O'Gorman and Smith's (2008) estimates of large economies of scope from owning multiple stations.

As suggested in Section 4, *a priori* arguments can be made for favoring each of these specifications. For example, the pseudo-likelihood estimators may be more efficient, but the moment-based estimator might be expected to be more robust in small samples and it does provide a closer match to some

[39] Mean annual (two period) station advertising revenues were \$4.8 m., \$1.4 m. and \$0.6 m. for the three market groups. Based on 10-K reports of publicly listed radio companies during my data period, operating income constituted between 20% and 40% of advertising revenues. However, it is likely that stations owned by smaller firms were less profitable.

Table 6: Parameter Estimates from the Dynamic Model

| Specification | (1) Modified Procedure P-Likelihood | (2) Modified Procedure Moments | (3) Iterated Procedure P-Likelihood | | (1) Modified Procedure P-Likelihood | (2) Modified Procedure Moments | (3) Iterated Procedure P-Likelihood |
|---|---|---|---|---|---|---|---|
| COSTS OF MOVE | | | | COST OF MOVING FROM | | | |
| TO ACTIVE FORMAT ($ M.) | | | | ACTIVE TO DARK ($ M.) | | | |
| Markets with pop. 1 m. + | 2.524 | 1.897 | 1.467 | Markets with pop. 1 m + | 2.704 | 3.126 | 1.572 |
| | (0.380) | (0.441) | (0.409) | | (0.744) | (0.964) | (0.503) |
| * Recent Format Switch | -0.001 | -0.001 | -0.000 | Markets with pop. 0.25- 1 m. | 0.636 | 1.636 | 0.369 |
| | (0.098) | (0.110) | (0.065) | | (0.230) | (0.582) | (0.130) |
| Markets with pop. 0.25-1 m. | 0.669 | 0.446 | 0.388 | Markets with pop. < 0.25 m. | 0.199 | 0.654 | 0.115 |
| | (0.157) | (0.167) | (0.102) | | (0.054) | (0.322) | (0.022) |
| * Recent Format Switch | 0.077 | -0.213 | -0.044 | Revenue of Switching | 1.108 | 3.105 | 0.642 |
| | (0.082) | (0.145) | (0.067) | Station | (1.330) | (1.506) | (1.003) |
| Markets with pop. < 0.25 m. | 0.233 | 0.069 | 0.135 | ECONOMIES OF SCOPE | | | |
| | (0.099) | (0.048) | (0.063) | Markets with pop. 1 m. + | 0.134 | -0.046 | 0.078 |
| * Recent Format Switch | 0.030 | -0.0310 | -0.017 | | (0.099) | (0.072) | (0.075) |
| | (0.023) | (0.105) | (0.054) | Markets with pop. 0.25-1 m. | 0.026 | 0.102 | 0.015 |
| *Revenue of Switching | 0.034 | 0.116 | 0.022 | | (0.018) | (0.021) | (0.010) |
| Station (all markets) | (0.168) | (0.211) | (0.118) | Markets with pop. < 0.25 m. | -0.006 | -0.014 | -0.033 |
| | | | | | (0.012) | (0.010) | (0.029) |
| ADDITIONAL COST OF MOVE | | | | | | | |
| TO ACTIVE FROM DARK ($ M.) | | | | SCALE OF $\varepsilon$s | | | |
| Markets with pop. 1 m + | -0.501 | -0.445 | -0.291 | Markets with pop. 1 m. + | 0.517 | 0.408 | 0.300 |
| | (0.220) | (0.183) | (0.095) | | (0.180) | (0.063) | (0.087) |
| Markets with pop. 0.25- 1 m. | -0.255 | -0.495 | -0.147 | Markets with pop. 0.25-1 m. | 0.144 | 0.091 | 0.083 |
| | (0.166) | (0.098) | (0.069) | | (0.055) | (0.037) | (0.026) |
| Markets with pop. < 0.25 m. | -0.061 | -0.078 | -0.035 | Markets with pop. < 0.25 m. | 0.050 | 0.014 | 0.028 |
| | (0.091) | (0.032) | (0.042) | | (0.014) | (0.007) | (0.009) |

of the main sources of variation in the data.[40] For example, in the data 10 times as many stations move to Spanish in markets with above median Hispanic populations as in markets with below median Hispanic populations. The moment based estimates predict that 7.2 times as many stations should move, whereas the estimates in column (1), which imply more random switching because the $\theta^{\varepsilon}$s are larger, only predict that 2.6 times as many stations should move. For blacks and the number of stations switching to Urban, the equivalent data, moment and modified pseudo-likelihood predictions are 5, 6.5 and 2.9 respectively. The iterated pseudo-likelihood based estimates in column (3) make predictions that are more like the moment-based estimates (7.0 for Hispanics/Spanish and 5.9 for blacks/Urban).

On the other hand, when the $\theta^{\varepsilon}$s are too small problems of multiple equilibria tend to become more acute, as each firm's choice probabilities become more sensitive to how they expect other stations should move. For example, the estimates from an iterated-version of the moment-based estimator appeared unstable[41] and so are not reported, and there was also some evidence of this type of instability performing the counterfactuals for smaller markets with the estimates in column (2). For this practical reason, I proceed using the estimates in column (1), drawing confidence from the fact that estimates of switching costs in the remaining columns are at least broadly similar.

### 5.3.2 Estimates Based on Forward-Simulation Approximations to the Value Function.

An alternative approach to estimating dynamic games with a rich state space involves approximating the value function via forward simulation based on initial first stage estimates of the conditional choice probabilities (e.g., BBL, and the applications by Ryan (2012), Snider (2009) and to radio by Jeziorski (2012)). Estimation is based on the inequalities, or moment inequalities, that are implied by the equilibrium assumption that each firm's actual strategy, reflected in its estimated conditional choice probabilities, should result in a higher value than any alternative, given the strategies of other players. Here I report estimates of repositioning costs, economies of scope and the scale of payoff shocks based on this type of approximation. This provides a robustness check on my coefficients,

[40] One reason why the moment-based and the pseudo-likelihood estimates are quite similar here may well be that the parametric approximation itself involves some averaging across states, which is one of the main advantages of the moment-based method.

[41] This instability may also arise from the fact that moment-based objective function is less well-behaved than the one based on the pseudo-likelihood, so an extensive search can be required to locate the global minimum. If the algorithm moves between local minima then this should also cause instability.

but it also more generally provides evidence on the performance of different approaches to estimating dynamic games. Here I give a very brief overview of the method, with full details in Appendix D.

As a firm's payoffs are linear in the parameters, a firm's value when it uses strategy $\Gamma_o$ and other firms use strategies $\Gamma^*_{-o}$ can be expressed as

$$V_o(\mathcal{M}_{j,o,t}|\Gamma_o, \Gamma^*_{-o}, \theta) = \mathbf{V}_{o,\Gamma_o,\Gamma^*_{-o}}\theta = \mathbf{R}_{o,\Gamma_o,\Gamma^*_{-o}} - \theta^W \mathbf{W}_{o,\Gamma_o,\Gamma^*_{-o}} + \theta^C \mathbf{C}_{o,\Gamma_o,\Gamma^*_{-o}} + \theta^{\varepsilon} \boldsymbol{\varepsilon}^F_{o,\Gamma_o,\Gamma^*_{-o}}$$

where $\mathbf{R}_{o,\Gamma_o,\Gamma^*_{-o}} = E_{o,\Gamma_o,\Gamma^*_{-o}} \sum_{t'=0}^{\infty} \beta^{t'} \sum_{s \in S^o} R_s(\mathcal{M}_{o,t+t'}|\gamma)$, the expected discounted sum of future revenues and $\mathcal{M}_{o,t+t'}$ is $o$'s state at time $t+t'$. $\mathbf{W}$ (expected discounted repositioning costs), $\mathbf{C}$ (economies of scope) and $\boldsymbol{\varepsilon}$ are defined similarly. The equilibrium restrictions used in estimation are that

$$V_o(\mathcal{M}_{j,o,t}|\Gamma^*_o, \Gamma^*_{-o}, \theta) - V_o(\mathcal{M}_{j,o,t}|\Gamma^a_o, \Gamma^*_{-o}, \theta) \geq 0 \ \forall \Gamma^a_o, \mathcal{M}_{j,o,t} \tag{20}$$

where $\Gamma^*$ are equilibrium strategies and $\Gamma^a_o$ is a particular alternative policy. Empirical implementation involves constructing estimates of $\mathbf{R}$, $\mathbf{W}$, $\mathbf{C}$ and $\boldsymbol{\varepsilon}$ based on observed policies (i.e., first-stage estimates of the conditional choice probabilities, and the demand and revenue models) and a finite number of alternatives (the ones used are detailed in the Appendix) using forward simulation. I consider two estimators of the parameters: the one proposed by BBL

$$\widehat{\theta^{BBL}} = \arg\min_{\theta} \sum_{o} \sum_{\forall a} \max\{(\mathbf{V}_{o,\Gamma^a_o,\Gamma^*_{-o}} - \mathbf{V}_{o,\Gamma^*_o,\Gamma^*_{-o}})\theta, 0\}^2 \tag{21}$$

and one which follows the moment-inequality approach of PPHI which finds the set of parameters that satisfy the following linear moment inequalities

$$\widehat{\theta^{PPHI}} \text{ is the set of } \theta \text{ where} \frac{1}{O} \sum_{o} (\mathbf{V}_{o,\Gamma^*_o,\Gamma^*_{-o}} - \mathbf{V}_{o,\Gamma^a_o,\Gamma^*_{-o}})\theta \geq 0 \ \forall \Gamma^a_o \tag{22}$$

The difference between these estimators is that the PPHI estimator uses the equilibrium implication that observed policies should do better on average (across states)[42], whereas the BBL estimator uses

[42] To be clear, this means that there is one inequality for a particular type of deviation $a$ (e.g., a policy which involves more switching than the estimated CCPs). One could create additional inequalities by interacting the difference between the optimal and alternative values with instruments (here, the instrument is a constant). This would shrink the size of the parameter set that satisfies all of the inequalities, sometimes to a point, but experimentation revealed that the estimates became sensitive to the alternatives and instruments used.

the fact that the equilibrium implies that they should do better in every state. In a similar way to POB's comparison between a moment-based entry rate matching estimator and a likelihood-based estimator, the PPHI estimator sacrifices information but, because of averaging, it may be more robust to approximation errors arising either from forward simulation or finite-sample or specification bias in the first-stage estimation of conditional choice probabilities.

As has been noted in the prior literature, it can be difficult to estimate more than a few parameters using these approaches. I therefore estimate the model separately for the three market-size groups, and for each group I only estimate a cost of switching to an active formats, the economy of scope parameter and the scale of the payoff shocks (this parameter is restricted to be non-negative).[43] The estimates, based on six alternative policies that are chosen to be intuitively informative about these parameters, are reported in Table 7 (note that exactly the same simulations and alternative policies are used for both types of estimator).

The PPHI estimates are sets because, for each group of markets, there is a convex set of parameters that satisfies all of the linear moment inequalities, and the reported estimates are the highest and lowest values of each of the parameters that satisfy each of the inequalities. The BBL numbers are point estimates, because no parameters satisfy all of the inequalities (the proportion violated is reported in the table).

Two features of these estimates deserve attention. First, the preferred estimates using value function approximation lie within or very close to the bounds implied by the PPHI estimator. In this sense, the preferred estimates in Table 6 are consistent with those of a quite different estimation methodology, while having the advantage of being point estimates that can be easily used in the counterfactuals. Second, with the possible exception of the second market group, the BBL estimates imply substantially higher repositioning costs and more volatile payoff shocks than either the value function approximation estimates or the upper bounds of the PPHI estimates. As these estimates of repositioning costs also look implausibly high, and we would not expect the BBL estimates to lie outside the PPHI bounds if averaging was only sacrificing information, rather than affecting bias, it seems appropriate to proceed using the estimates based on parametric value function approxima-

[43] In simplifying the model I assume that the cost of moving from Dark to an active format is the same as moving between a pair of active formats, and that there is no cost to moving to Dark. I have estimated specifications with separate coefficients for these costs, but without the imposition of additional and constraints, found that the estimates produced were often completely implausible (e.g., a cost of \$100 million for switching to Dark).

Table 7: Estimates Based on Forward Simulation

| | Repositioning Cost | Scope Economy | Scale of Payoff Shock |
|---|---|---|---|
| MARKETS WITH POP. >1 m. | | | |
| Moment inequality bounds (PPHI) | [2.194,11.031] | [-0.081,0.056] | [0.074,1.729] |
| 95% CIs | [0.652,13.877] | [-0.126,0.102] | [0.011,1.793] |
| BBL point estimate | 18.668 | 0.337 | 3.771 |
| std. error | (1.765) | (0.048) | (0.342) |
| proportion of inequalities violated | 24.4% | - | - |
| | | | |
| MARKETS WITH POP. 0.25m. -1m. | | | |
| Moment inequality bounds (PPHI) | [0.464,3.421] | [-0.071,0.031] | [0.015,0.549] |
| 95% CI | [0.232,4.035] | [-0.082,0.042] | [0,0.568] |
| BBL point estimate | 3.046 | 0.013 | 0.630 |
| std. error | (0.190) | (0.006) | (0.043) |
| proportion of inequalities violated | 10.0% | - | - |
| | | | |
| MARKETS WITH POP. <0.25m. | | | |
| Moment inequality bounds (PPHI) | [0.230,1.541] | [-0.022,0.008] | [0.005,0.251] |
| 95% CI | [0.081,1.690] | [-0.27,0.014] | [0,0.258] |
| BBL point estimate | 2.148 | 0.011 | 0.455 |
| std. error | (0.242) | (0.004) | (0.051) |
| % of BBL inequalities violated | 22.4% | - | - |

tion.[44] — station pay fee to music company

## 6 Counterfactual: The Effect of the Proposed Performance Rights Act on Format Choices

With the estimates in hand, I now use the model to predict how performance rights fees would affect format choices. I assume that these fees would be calculated as a percentage of advertising revenues for music stations, as assumed in US GAO (2010).[45] I assume that these fees were imposed as an unanticipated shock in Fall 2004, and, having solved the model, I simulate markets forward 40 periods from that date. The reported results are based on 51 of the 102 markets in my data.[46] Of course, fees were not imposed in 2004, but the industry is sufficiently similar now that the results should be informative about what would happen if they were introduced now.[47] I abstract away from other on-going demand changes by assuming that demographics remain fixed at their current levels.

I present two sets of results. The first set examines what would happen to the number of music stations and music audiences with no fees, 10% fees and 20% fees based on the estimates in column (1) of Table 6. I then examine how the effects of a 10% fee vary with some of the model's parameters, such as the level of repositioning costs and the degree of heterogeneity in listener tastes.

[44] Jeziorski (2012) also estimates repositioning costs using BBL. He estimates that repositioning costs in large markets are even higher than my BBL estimates, but he estimates that they are zero (to 2 dp) in smaller markets, even though switching rates are similar across the market-types. This provides some additional evidence that BBL does not provide appealing estimates of these costs, which are not Jeziorski's focus, when applied to radio data.

[45] The legislation did envisage that stations with very low revenues, which would apply to most stations outside urban markets, would be charged flat fees.

[46] Markets are ordered by their 2002 Arbitron market rank (based on population), and I then select every other market. This market selection is done to reduce the computational burden, while still allowing me to make broad statements about what would happen in the industry as a whole.

[47] The most obvious change that has taken place is that broadcast stations increasingly compete with satellite and online radio (where fees are paid). If this competition affects formats differently then this would affect format choices, and possibly the effects of fees. However, nationally the amount of listening to the different formats that I use was very similar in 2010 and 2004, based on the numbers reported by Arbitron in its 2011 *Radio Today* report, although there were some shifts in the popularity of different types of programming within these aggregated categories.

## 6.1 Implementation

As markets are independent and I do not need to pool information from across markets in the way that I did for estimation, I solve the model for each market separately. Full details are in Appendix B, but I provide an overview here. The first step, following the description in Section 3, solves for $\lambda_m^{*,FEE}$, which will depend on the level of the fee. This is done using the states observed in the data for Fall 2004, and 499 duplicates of this market-quarter where station formats and qualities are permuted. With fees, states with more non-music stations are oversampled as markets are expected to evolve in this direction.

The second step is to simulate the model forward. The first step provides firms' equilibrium choice probabilities in Fall 2004, and these probabilities and the estimated transition process for unobserved quality are used to move the model forward one period. However, this takes the market to a format/quality configuration that was not used in the first step, so $\lambda_m^{*,FEE}$ is used to solve for equilibrium choice probabilities in this new market structure, and these are used to simulate the model forward one more period. The process of solving for choice probabilities and forward simulating continues for 40 periods (20 years) after the introduction of the fees. For each market, the forward simulation process is repeated 10 times, and the results below are based on the mean and the standard deviation of these simulations.[48]

**Equilibrium Selection.** Multiple equilibria are a common feature of games, and except in relatively simple games, it is almost impossible to enumerate all of the equilibria. In this paper I do not attempt this type of enumeration and will instead rely on the equilibrium that my solution method, detailed in Appendix B, finds for different levels of fees.[49] Experimentation indicates that when I can find alternative equilibria the implications for how many stations and how many listeners switch to non-music formats are similar, although these equilibria differ in which stations are the first to switch formats.

[48] The computational burden of the two steps is high, so I do not try to account for the estimation error in the parameters.

[49] Experimentation with a subgroup of markets indicates that when I change the level of fees by a small amount the conditional choice probabilities in the equilibrium that I identify change quite smoothly. Intuitively this is a desirable property, but this does not mean that the resulting equilibrium would be played and there may be many equilibria which my method, which is based on iteration of best responses, would never find and they may also share this property. Aguirregabiria and Ho (2011) implement a method proposed in Aguirregabiria (2011) for finding counterfactual equilibria based on Taylor expansions. Implementation of this type of approach in my setting is left for future research.

Table 8: Evolution of the Number of Music Stations Under Different Performance Rights Fees

| | Music Stations | | | Music Listening | | | Non-Music Listening | | |
|---|---|---|---|---|---|---|---|---|---|
| Fee Level | 0% | 10% | 20% | 0% | 10% | 20% | 0% | 10% | 20% |
| Period prior to introduction (Fall 2004) | 713 | 713 | 713 | 0.254 | 0.254 | 0.254 | 0.095 | 0.095 | 0.095 |
| +1 period | 714 | 701 | 693 | 0.254 | 0.250 | 0.246 | 0.094 | 0.097 | 0.101 |
| | (3.7) | (4.6) | (4.6) | (0.002) | (0.003) | (0.003) | (0.002) | (0.003) | (0.004) |
| +5 periods | 715 | 682 | 626 | 0.253 | 0.243 | 0.228 | 0.094 | 0.103 | 0.110 |
| | (6.1) | (7.2) | (7.1) | (0.003) | (0.003) | (0.005) | (0.002) | (0.003) | (0.004) |
| +10 periods (5 years) | 717 | 665 | 595 | 0.253 | 0.238 | 0.220 | 0.095 | 0.106 | 0.116 |
| | (8.0) | (9.8) | (9.8) | (0.004) | (0.004) | (0.006) | (0.003) | (0.003) | (0.005) |
| +20 periods (10 years) | 716 | 651 | 582 | 0.252 | 0.237 | 0.220 | 0.096 | 0.106 | 0.117 |
| | (10.1) | (11.5) | (11.8) | (0.004) | (0.005) | (0.006) | (0.003) | (0.003) | (0.003) |
| +40 periods (20 years) | 720 | 652 | 578 | 0.253 | 0.237 | 0.219 | 0.096 | 0.107 | 0.118 |
| | (10.4) | (10.9) | (12.0) | (0.004) | (0.005) | (0.005) | (0.003) | (0.004) | (0.004) |

Standard deviations across 10 simulations in parentheses. Results based on sample of 51 markets.
Music listening measured as the average combined market share of music stations across markets.

## 6.2 The Effects of Performance Fees on Market Structure

Table 8 shows how the number of contemporary music stations and the average (across markets) combined market share of music and non-music stations (based on the market definition used in estimating the model) is predicted to change for the different levels of fees. In the absence of fees the total number of music stations is predicted to remain approximately the same, although this masks the fact that in certain markets, the model predicts that significant changes in the number of stations in a particular formats that look under- or over-served given the demand/revenue models and current demographics. The model predicts that over 20 years (which I will call the long-run in what follows), 10% and 20% fees would reduce the number of music stations by 9.4% and 20% respectively relative to the no fee case, or by 1.3 and 2.8 stations per market.[50] Based on the preferred estimates, \$93 million (std. deviation \$11 million) more is spent on repositioning costs with 10% fees than with no fees (no discounting). This cost increase is greater, by around 20%, than what one would expect based on the change in the number of music stations alone, because these switches cause some additional churn with some stations moving from non-music to music formats, and others between non-music formats to avoid greater competition. Relative to the long-run format structure with no fees, the AC, Country and Rock formats lose roughly equal numbers of stations, while the Urban

[50] The number of Dark stations per market is predicted to increase by 0.07 and 0.11 respectively under these fees.

format loses the least (3% of its stations with 10% fees). This is consistent with the fact that black listeners are relatively unwilling to substitute to other formats, and in markets where Urban stations are common, blacks make up a large proportion of the population. All three non-music formats gain stations in roughly the same proportion, with the gain in Spanish stations concentrated in markets with large existing Hispanic populations (recall that for the purposes of the counterfactual I assume that demographics remain constant).

The predicted changes in music listenership are roughly 30% smaller than the changes in the number of stations, reflecting the fact that a listener who likes a music station which switches to a non-music format will often switch to one of the remaining music stations.[51] Most of the remaining listeners switch to non-music programming, rather than switching off their radios, reflecting the high value of $\gamma^{\sigma}$, although combined radio listening falls slightly. Unfortunately as listeners are not observed paying prices for listening to the radio, it is not possible to quantify a dollar welfare effect of these changes in listenership.

The long-run adjustment does not take place immediately but, for both 10% and 20% fees, at least 40% of the long-run change in the number of music stations is completed within 5 periods ($2\frac{1}{2}$ years). The adjustment in music audiences is predicted to happen more quickly: for example, with 20% fees, 74% of the change in the amount of music listening takes place in 5 periods, compared with 63% of the change in the number of stations. This reflects the fact that, at least in the equilibria that I find, higher quality stations, with the most listeners, tend to be more likely to switch when fees are initially put in place, whereas with no fees, and in the data, it is smaller stations that are more likely to switch formats. For example, with no fees the average per-period revenues of a station that changes format in the first 5 periods is \$0.9 million, while with 10% fees the average (gross) revenue is \$1.4 m..[52]

**The Effect of Repositioning Costs, Payoff Shocks and Tastes Heterogeneity on the Transition.** I investigate how these features of the model affect the predicted speed of adjustment. Table 9 shows how changes in the number of music stations and music listening when 10% fees are

[51] These results imply that a higher level of fees would raise performers and record companies total revenues from fees. However, assessing whether higher fees would be in interests of these parties also requires knowing how sales of recorded music and concert tickets are affected by airplay. Different studies have reached contrasting conclusions about this elasticity (Dertouzos (2008)).

[52] Quality can also be calculated using the demand model. Measured by the exponent of the sum of $\xi_{st}$ and the fixed quality componens (e.g., signal coverage) the average quality of a switching station with no fees is 1.09, and with 10% fees it is 1.30.

Table 9: Evolution of the Number of Music Stations and Music Audiences under a 10% Fees for Different Assumptions on the Structural Parameters Relative to Evolution with No Fees (Fall 2004 indexed to 1)

| | True Parameters ("Base case") | | Less Taste Heterogeneity | | Higher Repositioning Costs | | Higher Repositioning Costs and High $\sigma^{\varepsilon}$ | |
|---|---|---|---|---|---|---|---|---|
| Column | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Stations | Music Audience | Stations | Music Audiences | Stations | Music Audience | Stations | Music Audience |
| Fall 2004 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| +1 period | 0.981 | 0.984 | 0.941 | 0.950 | 0.986 | 0.992 | 0.986 | 0.993 |
| | (0.006) | (0.005) | (0.010) | (0.012) | (0.002) | (0.003) | (0.004) | (0.002) |
| +5 periods | 0.954 | 0.961 | 0.875 | 0.895 | 0.967 | 0.978 | 0.971 | 0.980 |
| | (0.010) | (0.009) | (0.014) | (0.013) | (0.006) | (0.006) | (0.008) | (0.008) |
| +10 periods | 0.930 | 0.941 | 0.810 | 0.851 | 0.940 | 0.959 | 0.952 | 0.964 |
| (5 years) | (0.015) | (0.010) | (0.017) | (0.015) | (0.010) | (0.009) | (0.014) | (0.012) |
| +40 periods | 0.906 | 0.937 | 0.801 | 0.841 | 0.910 | 0.942 | 0.918 | 0.945 |
| (20 years) | (0.014) | (0.010) | (0.018) | (0.016) | (0.011) | (0.009) | (0.016) | (0.009) |

Results based on sample of 51 markets. Standard deviations across 10 simulations in parentheses.
Music listening measured as the average combined market share of music stations across markets.

introduced are affected by several counterfactuals where, for ease of comparison, values with no fees are indexed to 1.[53] Columns (1) and (2) reflect what happens under 10% fees in Table 8.

*Taste Heterogeneity.* The fact that most listeners, and particularly those in the most desirable demographics, prefer music reduces the incentives of stations to switch to non-music formats to avoid fees. To understand how much more repositioning would take place if there was less taste heterogeneity, I resolve the model (for no fees and a 10% fees) when all of the demographic*format parameters are equal to half of their estimated values ($\gamma^{\sigma}$ takes its estimated value, so commercial radio listening remain almost fixed).[54] Columns (3) and (4) of Table 9 show how the industry adjusts given these hypothetical parameters. With less taste heterogeneity, the short-run and long-run changes in the number of music stations are roughly twice as large as in the base case. Because of reduced format loyalty, this results in an even larger decline in the amount of listening to music

[53] I have also investigated the effects of advertisers heterogenous valuations. If advertisers value all demographics in the same way, there are substantial flows of stations into non-music formats (which attract older and Hispanic listeners) without fees. These flows become larger with fees, but there is no clear effect of the different valuations on the speed of adjustment.

[54] This simulation is performed with values for the linear demand parameters, $\delta_{st}$ and $\xi_{st}$ which would be estimated if the random coefficients had these hypothetical values. The parameters of the revenue function, and the repositioning cost and economies of scope parameters are the same as in the full model.

radio.

*Repositioning Costs.* Columns (5) and (6) report the effects of fees when the cost of repositioning stations is increased by 25%. In the long-run, this change has little effect on the number of music stations, consistent with listener tastes, advertiser valuations, fees and the importance of choice-specific payoff shocks being the long-run determinants of the format structure in each market. However, higher repositioning costs slow the speed of the transition that takes place when fees are introduced. This can be rationalized by the fact that, when repositioning costs are high, there is a greater incentive for a firm to wait to see what its competitors will do before making a desired switch, unless it receives a very favorable draw of the associated payoff shock.

*Repositioning Costs and Scale of Payoff Shocks.* As can clearly be seen in all of the different estimates from Section 5, higher estimates of repositioning costs are accompanied by larger estimates of the scale of the payoff shocks, as these changes together allow the predicted rate of switching to be relatively unchanged. It is therefore interesting to ask whether changing both of these parameters affects the predictions. In columns (7) and (8) both repositioning costs and the scale of the $\varepsilon$s are increased by 25%. While these changes result in approximately the same amount of switching as the base parameters with no fees, the predictions with fees are not the same as in the base case. In particular, the model predicts a smaller long-run change in the number of music stations, and music listening, as the choice specific payoff-shocks become more important, and there are more choices associated with music formats than non-music ones. The transition is also slower, as the value of waiting for a favorable $\varepsilon$ draw associated with a desired move increases with the variance of these shocks.

# 7 Conclusion

This article use a dynamic model to predict how the format structure of local radio markets would change if broadcast music radio stations had to pay fees for musical performance rights, as proposed in legislation that received broad political support in 2009. This setting provides a natural one for modeling the effects of a policy that favors a particular type of product, because the sets of available products (stations) and possible product types (formats) are well-defined and, even in the absence of fees, significant product repositioning is observed. My results suggest that fees equal to 10 or 20%

of revenues would have significant, and fairly rapid effects, on the number of music stations, but that the declines would not be as dramatic as some people in the broadcasting industry have suggested for the simple reason that lots of people prefer music programming, including many listeners who are particularly valued by advertisers. Of course, all of the counterfactual results are predictions, based on a particular set of modeling assumptions and also a method for approximating the solution. It will therefore be both interesting and important to test the accuracy of the model's predictions if and when performance fees are eventually introduced.

Estimating and solving a dynamic game that captures the type of rich horizontal and vertical differentiation that are features of the radio industry requires some form of approximation. This article's approach is to approximate value functions using a parametric linear function of variables that reflect the current state of the industry. In the counterfactual, this method produces plausible results. In estimating the main parameters of the game, I combine this type of approximation with different estimation routines suggested in the literature, and I compare the resulting estimates with ones based on methods that approximate the value function by forward simulation. While these approaches do not produce identical estimates, many of them are similar and plausible (e.g., a cost of changing formats equal to somewhere between 30 and 60% of annual station revenues).[55] This is very encouraging because many non-trivial choices are required to implement either approach (e.g., choice of a set of variables for approximating the value function or the choice of alternative policies), and the results should provide confidence that these methods can be used in other settings where an exceptionally large state space is required to capture features that may affect an industry's evolution.

[55] In practice, the computational burden of implementing these alternatives also needs to be considered. An advantage of the parametric approximation methods that I use is that almost the same code can be used to estimate and solve the model, and the estimation procedure yields point estimates. On the other hand, the forward simulation procedures can be implemented quite quickly on a large cluster, but because the most sensible moment inequality (PPHI) estimates are bounds, they may be useful primarily as robustness checks.

# References


[1] Aguirregabiria, Victor (2011), "A Method for Implementing Counterfactual Experiments in Models with Multiple Equilibria", *Economic Letters*, forthcoming

[2] Aguirregabiria, Victor and Chun-Yu Ho (2011), "A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments", *Journal of Econometrics*, forthcoming

[3] Aguirregabiria, Victor and Pedro Mira (2007), "Sequential Estimation of Dynamic Discrete Games", *Econometrica*, 75(1), 1-53

[4] Aguirregabiria, Victor and Pedro Mira (2010), "Stability of Equilibria and Properties of Sequential Estimators: Problems and Solutions", mimeo, University of Toronto

[5] Aguirregabiria, Victor and Aviv Nevo (2012), "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games", mimeo, University of Toronto

[6] Albarran, Alan B. and W. Lawrence Patrick (2005), "Assessing Radio Station Value: A Review of Academic Literature and Analysis of Contemporary Industry Models", *Journal of Radio & Audio Media*, 12(1), 3–13

[7] Alho, Juha M. and Bruce D. Spencer (2005), *Statistical Demography and Forecasting*, New York: Springer

[8] Arbitron Company (2002), *Arbitron Radio Market Report Reference Guide*, New York, NY: Arbitron Company

[9] Arbitron Company, *Radio Today*, various years, New York, NY: Arbitron Company

[10] Arcidiacono, Peter, Patrick J. Bayer, Federico A. Bugni and Jonathan James (2012), "Approximating High-Dimensional Dynamic Models: Sieve Value Function Iteration", mimeo, Duke University

[11] Bajari, Patrick, Lanier Benkard and Jonathan Levin (2007), "Estimating Dynamic Models of Imperfect Competition", *Econometrica*, 75(5), 1331-1370

[12] Barwick, Panle Jia and Parag Pathak (2012), "The Cost of Free Entry: An Empirical Study of Real Estate Agents in Greater Boston", mimeo, MIT

[13] Benitez-Silva, Hugo, George Hall, Gunter Hitsch, Giorgio Pauletto and John Rust (2000): "A Comparison of Discrete and Parametric Approximation Methods for Continuous-State Dynamic Programming Problems," mimeo, Yale University

[14] Berry, Steven T. (1994), "Estimating Discrete-Choice Models of Product Differentiation", *RAND Journal of Economics*, 25(2), 242-262

[15] Berry, Steven T., James Levinsohn and Ariel Pakes (1993), "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change", *American Economic Review*, 83(2), 240-246

[16] Berry, Steven T., James Levinsohn and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, 63(4), 841-890

[17] Berry, Steven T. and Joel Waldfogel (1999), "Free Entry and Social Inefficiency in Radio Broadcasting", *RAND Journal of Economics*, 30(2), 397-420

[18] Berry, Steven T. and Joel Waldfogel (2001), "Do Mergers Increase Product Variety? Evidence from Radio Broadcasting", *Quarterly Journal of Economics*, 116(3), 1009-1025

[19] Bertsekas, Dimitri (2010), "Approximate Policy Iteration: A Survey and Some New Methods", Laboratory for Information and Decision Systems Report No. 2833, MIT

[20] Bertsekas, Dimitri and Huizhen Yu (2007), "Solution of Large Systems of Equations Using Approximate Dynamic Programming Methods", Laboratory for Information and Decision Systems Report No. 2754, MIT

[21] Bertsekas, Dimitri and Sergey Ioffe (1996), "Temporal Differences-Based Policy Iteration and Applications in Neuro-Dynamic ProgrammingSolution of Large Systems of Equations Using Approximate Dynamic Programming Methods", Laboratory for Information and Decision Systems Report No. 2349, MIT

[22] BIA Financial Network, Inc. (2006), Media Access Pro User's Manual, version 3.0, Chantilly, VA: BIA Financial Network, Inc.

[23] Borenstein, Severin and Janet S. Netz (2002), "Maximum or Minimum Differentiation? Location Patterns of Retail Outlets", *Review of Economics and Statistics*, 84(1), 162-175

[24] Datta, Sumon and K. Sudhir (2012), "The Agglomeration-Differentiation Tradeoff in Spatial Location Choice", mimeo, Yale University

[25] Dertouzos, James N. (2008), "Radio Airplay and the Record Industry: An Economic Analysis", paper prepared for the National Association of Broadcasters, Washington DC

[26] Draganaska, Michaela, Michael Mazzeo and Katja Seim (2009), "Beyond Plain Vanilla: Modeling Joint Pricing and Product Assortment Choices", *Quantitative Marketing and Economics*, 7(2), 105-46

[27] Doraszelski, Ulrich and Mark Satterthwaite (2010), "Computable Markov-Perfect Industry Dynamics", *RAND Journal of Economics*, 41(2), 215-43

[28] Dube, Jean-Pierre H. and Jeremy T. Fox and Che-Lin Su (2011), "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients", *Econometrica*, forthcoming

[29] Duncan, James H., Duncan's American Radio, issues from Spring 1996 to Spring 2001

[30] Ericson, Richard and Ariel Pakes (1995), "Markov Perfect Industry Dynamics: A Framework for Empirical Work", *Review of Economic Studies*, 62(1), 53-82

[31] George, Lisa and Joel Waldfogel (2003), "Who Affects Whom in Daily Newspaper Markets?", *Journal of Political Economy*, 111(4), 765-784

[32] Fan, Ying (2012), "Ownership Consolidation and Product Characteristics: A Study of the U.S. Daily Newspaper Market", mimeo, University of Michigan

[33] Farias, Vivek, Denis Saure and Gabriel Y. Weintraub (2012), "An Approximate Dynamic Programming Approach to Solving Dynamic Oligopoly Models", *RAND Journal of Economics*, 43(2), 253-282

[34] Federal Communications Commission (2001), Review of the Radio Industry, 2001, Mass Media Bureau Policy and Rules Division, Washington DC: Federal Communications Commission

[35] Fowlie, Meredith, Mar Reguant and Stephen Ryan (2011), "Market-Based Emissions Regulation and the Evolution of Market Structure", mimeo, MIT

[36] Hendel, Igal and Aviv Nevo (2006), "Measuring the Implications of Sales and Consumer Inventory Behavior", *Econometrica*, 74(6), 1637-1673

[37] Jenkins, Mark, Paul Liu, Rosa L. Matzkin and Daniel L. McFadden (2004), "The Browser War: Econometric Analysis of Markov Perfect Equilibrium in Markets with Network Effects", mimeo, Stanford University

[38] Jeziorski, Przemysław (2012), "Estimation of Cost Efficiencies from Mergers: An Application to US Radio", mimeo, University of California at Berkeley

[39] Judd, Kenneth L. (1998), *Numerical Methods in Economics*, MIT Press: Cambridge, MA

[40] Ma, Jun and Warren B. Powell (2009), "Convergence Proofs of Least Squares Policy Iteration Algorithm for High-Dimensional Infinite Horizon Markov Decision Process Problems", mimeo, Princeton University

[41] Mooney, Catherine Tyler (2010a), "Cost Sharing and Strategic Programming Choices in Broadcast Radio", mimeo, University of Oklahoma

[42] Mooney, Catherine Tyler (2010b), "A Two-Sided Market Analysis of Radio Ownership Caps", mimeo, University of Oklahoma

[43] Nevo, Aviv (2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand", *Journal of Economics and Management Strategy*, 9(4), 513-548

[44] Nevo, Aviv (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry", *Econometrica*, 69(2), 307-342

[45] Nishida, Mitsukuni (2012), "Estimating a Model of Strategic Network Choice: The Convenience-Store Industry in Okinawa", mimeo, Johns Hopkins University

[46] O'Gorman, Catherine and Howard Smith (2008), " Efficiency Gain from Ownership Deregulation: Estimates for the Radio Industry", mimeo, University of Oxford

[47] Pakes, Ariel, Berry, Steven T. and Michael Ostrovsky (2007), "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Examples)", *RAND Journal of Economics*, 38(2), 373-399

[48] Pakes, Ariel, Jack Porter, Katherine Ho and Joy Ishii (2011), "Moment Inequalities and their Application", mimeo, Harvard University

[49] Pesendorfer, Martin and Philipp Schmidt-Dengler (2008), "Asymptotic Least Squares Estimators for Dynamic Games", *Review of Economic Studies*, 75(3), 901-928

[50] Pesendorfer, Martin and Philipp Schmidt-Dengler (2010), "Sequential Estimation of Dynamic Discrete Games: A Comment", *Econometrica*, 78(2), 833–842

[51] Petrin, Amil (2002), "Quantifying the Benefits of New Products: The Case of the Minivan", *Journal of Political Economy*, 110(4), 705-729

[52] Roberts, James W. and Andrew Sweeting (2012), "The Potential Entry Defense in Airline Mergers", mimeo, Duke University

[53] Rust, John (2000), "Parametric Policy Iteration: An Efficient Algorithm for Solving Multidimensional DP Problems?", mimeo, Yale University

[54] Ryan, Stephen P., "The Costs of Environmental Regulation in a Concentrated Industry", *Econometrica*, 80(3), 1019-1061

[55] Seim, Katja (2006), "An Empirical Model of Firm Entry with Endogenous Product-Type Choices", *RAND Journal of Economics*, 37(3), 619-640

[56] Snider, Connan (2009), "Predatory Incentives and Predation Policy: The American Airlines Case", mimeo, University of California at Los Angeles

[57] Speck, Paul S. and Michael T. Elliott (1997), "Predictors of Advertising Avoidance in Print and Broadcast Media Predictors of Advertising Avoidance in Print and Broadcast Media", *Journal of Advertising*, 26(3), 61-76

[58] Su, Che-Lin and Kenneth L. Judd (2011), "Constrained Optimization Approaches to Estimation of Structural Models", *Econometrica*, forthcoming

[59] Sweeting, Andrew T. (2010), "The Effects of Mergers on Product Positioning: Evidence from the Music Radio Industry", *RAND Journal of Economics*, 41(2), 372-97

[60] Sweeting, Andrew T. (2011), "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry", mimeo, Duke University

[61] Watson, Randal (2009), "Product Variety and Competition in the Retail Market for Eyeglasses", *Journal of Industrial Economics*, 57(2), 217-251

[62] United States Government Accountability Office (2010), "The Proposed Performance Rights Act Would Result in Additional Costs for Broadcast Radio Stations and Additional Revenue for Record Companies, Musicians, and Performers", GAO-10-826, August 2010

# Appendices for "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry"

Andrew Sweeting

July 2012

## A State Variables

Table 1 lists all of the state variables in the model. As some station characteristics are assumed to be fixed and ownership is assumed not to change over time, a state is firm-specific, and the value of the state variables will depend on the characteristics of the stations owned by the firm (some of which are fixed), the characteristics of the firm's competitors (the set of competitors is fixed) and market characteristics. The table notes (i) whether a variable is fixed over time, (ii) the information assumption made about the variable, and (iii) whether the value is observed or estimated by the researcher in the first stage.

## B Solution Method for Counterfactuals

This Appendix details how the model is solved when performing counterfactuals. As markets are independent in my model, I solve the model 'market-by-market', so I do not impose that the approximation to the value function is the same across markets that differ greatly in size. I also assume that market demographics stay fixed at their current values to further simplify the model and to focus more closely on the effect of performance fees.

Table 1: State Variables

| Market Variables | **Evolution** | **Information?** | **Observed?** |
|---|---|---|---|
| Population in $d$=1,..,18 mutually exclusive groups | Ethnic group size evolves with growth rates | Public | Observed |
| Growth rates for Black, Hispanic and White populations | AR(1), iid innovations | Public | Observed, AR(1) process estimated |
| Advertising prices per listener | Fixed | Public | Estimated |
| | | | |
| Station Variables (for firm's own stations) | | | |
| Format | Changes with choice | Public | Observed |
| Changed Format in Previous Period | Changes with choice | Public | Observed |
| Station Quality: | | | |
| observed characteristic component (e.g., based on signal coverage, station location) | Fixed | Public | Characteristics observed, quality coefficients estimated |
| band | Fixed | Public | Observed (interaction with format estimated) |
| $\xi_{st}$ | AR(1), iid innovations | Public | Implied by demand estimates AR(1) process estimated |
| Station Variables (for each station owned by competitors) | | | |
| Owner | Fixed | Public | Observed |
| Format | Changes with choice | Public | Observed |
| Changed Format in Previous Period | Changes with choice | Public | Observed |
| Station Quality: | | | |
| observed characteristic component (e.g., based on signal coverage, station location) | Fixed | Public | Characteristics observed, quality coefficients estimated |
| band | Fixed | Public | Observed (interaction with format estimated) |
| $\xi_{st}$ | AR(1), iid innovations | Public | Estimated |
| | | | |
| Choice Specific Payoff Shocks | | | |
| $\varepsilon$ for each choice for each firm | iid across firms, choice & time | Private | No, scale estimated |

**Selection of states.** It is necessary to solve for values and policies at a fixed subset of $N$ states because some state variables (e.g., unobserved station quality) are continuous and the state space is large. For each market I choose the states that are observed in the data, and then create 499 duplicates of the observed states where the formats of the stations and their unobserved qualities are perturbed. In each duplication, the unobserved quality of each station is chosen as a uniform random draw on $[-2, 2]$, a range which comprises almost all of the values of $\xi_{st}$ in the data. When there are no fees, the probability that a station's format is the same as in the data is 0.3, the probability that the station is Dark is 0.05, and otherwise a new active format is drawn where the probability of each format is the same. With fees, it is likely that markets will evolve to situations with more non-music stations and it is desirable to approximate the value function more accurately in these states. When the format of a station is to be changed, I therefore make the probability of choosing each non-music format twice as large as the probability of choosing each music format.

In the description of the solution procedure, a particular state $j$ is denoted as $\mathcal{M}_{j,o,t}$ where $o$ indicates the firm of interest in the state and $t$ denotes the initial period.

**Variables used in approximating the value function.** I assume that firm's value functions can be approximated by a linear parametric function of functions of the state variables. These functions include measures based on revenues calculated in several different ways, which I now describe.

- 'actual revenues': a station's revenues in a format given demographics, formats and station characteristics including the $\xi_{st}$s
- 'no $\xi$ revenues': a station's revenues in a format given demographics, formats and station characteristics except the time-varying $\xi_{st}$s (i.e., the $\xi$s of all stations are set equal to zero)
- '$\overline{revenue}$': a measure of the station's average revenue potential excluding the $\xi_{st}$s formed by averaging the revenues that it would get across a large set of format configurations for all stations, for a fixed set of demographics.[1]

[1] Note that this measure does not depend on stations' current formats and so, for a particular station, it stays the same across periods.

The approximating function for a state $\mathcal{M}_{j,o,t}$ includes the following variables[2]:

- a constant
- the sum of station actual revenues in music and non-music formats for firm $o$'s stations in period $t$[3]
- a count of how many stations the firm has in formats where it has more than one station;
- the sum of the $\overline{revenue}$ measures for firm $o$'s stations
- the sum of the $\exp(\xi_{st})$ measures for firm $o$'s stations
- the sum of the $\exp(X_{st}\gamma^S)$ measures for firm $o$'s stations (excluding the AM*format component)
- the sum of the AM dummies for firm $o$'s stations
- the sum of the AM dummies interacted with $\exp(X_{st}\gamma^S)$ (excluding the AM*format component) for firm $o$'s stations
- the sum of the AM dummies interacted with $\exp(\xi_{st})$ for firm $o$'s stations
- the sum of the $\exp(X_{st}\gamma^S)$ (excluding the AM*format component) interacted with $\exp(\xi_{st})$ for firm $o$'s stations
- the sum of $\overline{revenue}$ interacted with $\exp(\xi_{st})$ for firm $o$'s stations
- the last 8 variables interacted with a count of the number of stations owned by firm $o$[4]
- the sum of the AM dummies for firm $o$'s stations in the News/Talk format
- the sum of the AM dummies for firm $o$'s stations in the News/Talk format * the total number of AM stations in the market

[2] These variables were chosen based primarily on a set of Monte Carlo experiments, using a simplified model, discussed in the Appendices of Sweeting (2011). These experiments revealed that solutions can become less accurate, due to overfitting, when too many variables are included in the approximation.

[3] When I am assuming that fees are charged these revenues are calculated net of fees.

[4] Note that I do not need to include interactions with market demographics or market prices because the game is solved market-by-market and these variables are assumed to be fixed.

- for the largest competitor faced by the firm (call it firm $x$)[5]:
  - the number of stations owned by $x$
  - the sum of the $\overline{revenue}$ measures for firm $x$'s stations
  - the sum of the $\exp(\xi_{st})$ measures for firm $x$'s stations
  - the sum of interactions between the $\overline{revenue}$ measures and $\exp(\xi_{st})$ for firm $x$'s stations
- for the second largest competitor faced by the firm (call it firm $y$):
  - the number of stations owned by firm $y$
  - the sum of the $\overline{revenue}$ measures for firm $y$'s stations
  - the sum of the $\exp(\xi_{st})$ measures for firm $y$'s stations
  - the sum of interactions between the $\overline{revenue}$ measures and $\exp(\xi_{st})$ for firm $y$'s stations
- for each active format (i.e., the coefficients can vary freely across formats):
  - number of rival stations in the format
  - number of rival stations in the format owned by firms that own more than one station
  - sum of rival stations' $\exp(\xi_{st})$s
  - sum of rival stations' $\exp(X_{st}\gamma^S)$s (excluding the AM*format component)
  - sum of the $\overline{revenue}$ measures for rival stations
  - sum of the $\overline{revenue}$ measures for rival stations that are in the AM band
- a dummy for whether firm $o$ moved a station in the previous period[6]
- the $\overline{revenue}$, the $\exp(\xi_{st})$ and the interaction of $\overline{revenue}$ and $\exp(\xi_{st})$ of a station moved by $o$ in the previous period

[5] The size of competitors is determined by the number of stations owned and, where this is equal, the sum of the $\overline{revenue}$ measures for the different firms.

[6] Recall that I allow for a different cost of moving a recently-moved station.

- a set of measures of the potential gains in revenue that firm $o$ could achieve in the following period if the formats of other stations were held fixed. These are calculated using the 'no $\xi$ revenues' measures (which negates the need to try to integrate over a set of changes in unobserved qualities)
  - the number of moves firm $o$ could make that would raise expected revenues
  - the sum of revenue gains the firm could make from these moves, plus an interaction with the number of stations that the firm owns
  - the maximum revenue gain from moving a station in \$m. and as a proportion of firm $o$'s current revenue
  - the sum of moves that would increase revenues and would involve moving a recently-moved station
  - the sum of the count of how many stations the firm would have in formats where it has more than one station based on moves that would increase the firm's revenues;

In the following description of the solution procedure, $\Phi_{j,k}(\mathcal{M}_{j,o,t})$ is the value of the $k^{th}$ approximating variable in state $j$ and $\Phi$ is the matrix where these variables are stacked for the $N$ states.

**Solution Procedure.** An iterative procedure is used to find the coefficients of the parametric function that is used to approximate the value function. I now detail each of the steps in a particular iteration $i$. In state $\mathcal{M}_{j,o,t}$, $P_o^i(a|\mathcal{M}_{j,o,t})$ is the iteration $i$ guess of the probability that firm $o$ chooses action $a$, $P_o^i(\mathcal{M}_{j,o,t})$ is the collection of these probabilities, and $P_{-o}^i(\mathcal{M}_{j,o,t})$ is the set of choice probabilities of $o$'s competitors in the state.

Step 1. For each of the $N$ states $\widetilde{\pi}(P_o^i(\mathcal{M}_{j,o,t}))$ is calculated as

$$\widetilde{\pi}(P_o^i(\mathcal{M}_{j,o,t})) = \sum_{s \in S^o} R_s(\mathcal{M}_{j,o,t}|\gamma) + \sum_{a \in A_o(\mathcal{M}_{j,o,t})} P_o^i(a|\mathcal{M}_{j,o,t}) \begin{pmatrix} \beta C_o(a)\theta^C - W_o(a)\theta^W \\ +\theta^\varepsilon(\varkappa - \log(P_o^i(a|\mathcal{M}_{j,o,t})) \end{pmatrix} \quad (1)$$

where the $\theta$s are the fixed parameters, $\varkappa$ is Euler's constant, $C_o(a, \mathcal{M}_{j,o,t})$ is the number of stations that the firm will have operating in the same format as one of its other stations in the next period if it chooses action $a$ and $A_o(\mathcal{M}_{j,o,t})$ is $o$'s choice set. $\widetilde{\pi}(P_o^i(\mathcal{M}_{j,o,t}))$ is a function only of $o$'s choice probabilities. $\widetilde{\pi}(P^i)$ is the vector that stacks these values for the $N$ states.

For each of the $N$ states, the choice probabilities of all firms are used to calculate $E_{P^i}\Phi^i$, a vector which contains the expected value of each of the approximating variables for the following period given strategies. For a particular $k$,

$$E_{P^i}\phi_{j,k}^i = \int \phi_{h,k}(\mathcal{M}_{h,o,t+1})g(\mathcal{M}_{h,o,t+1}|P_o^i(\mathcal{M}_{j,o,t}), P_{-o}^i(\mathcal{M}_{j,o,t}), \mathcal{M}_{j,o,t})d\mathcal{M}_{h,o,t+1} \tag{2}$$

where $g$ is the transition density. This integral is approximated by reweighting variables for a pre-specified sample of $t+1$ states, as calculating $\phi_{h,k}(\mathcal{M}_{h,o,t+1})$ requires solving a random coefficients demand model. Specifically for a given state $\mathcal{M}_{j,o,t}$, I consider a set of $H$ states $\mathcal{M}_{h,o,t+1}$ which is equal to the set of states that can be reached by any move by $o$, a set of $S^{\xi}$ draws for innovations in $\xi$ and $S^{-o,m}$ moves by other firms in the same local market. During the solution procedure the integral is approximated by

$$E_{P^i}\phi_{j,k}^i \approx \sum_{h=1}^{H} \phi_{h,k}(\mathcal{M}_{h,o,t+1}) \frac{g(\mathcal{M}_{h,o,t+1}|P_o^i(\mathcal{M}_{j,o,t}), P_{-o}^i(\mathcal{M}_{j,o,t}), \mathcal{M}_{j,o,t})}{\sum_{h'=1}^{H} g(\mathcal{M}_{h',o,t+1}|P_o^i(\mathcal{M}_{j,o,t}), P_{-o}^i(\mathcal{M}_{j,o,t}), \mathcal{M}_{j,o,t})} \tag{3}$$

$S^{\xi} = 10$ and $S^{-o,m} = 500$. To be accurate, the $S^{-o,m}$ moves should include those that are most likely to be made. With no fees, I choose the ones that are most likely to be made based on the first-stage estimates of the conditional choice probabilities. With fees, non-music formats will be more likely to be chosen, all else equal, and this needs to be accounted for. The first step is to estimate a simpler multinomial logit model using the observed data where the covariates are the elements of $W_o(a_o)$ and $C_o(a_o)$ and the revenues that the firm would earn in the next period given the particular choice if no other firms' made a format switch and unobserved station qualities and demographics remained unchanged. The estimated coefficients in this model are consistent with stations moving towards increased revenues. I then recompute the revenue variable taking into account the effects of the tax, giving a new set of choice probabilities. I then use these choice probabilities to choose the $S^{-o,m}$

moves that are most likely to be made.

Step 2. Create matrices $(\Phi - \beta E_{P^i}\Phi)$ $(\beta = 0.95)$. As the parameters $\lambda$ are overidentified $(N > K)$, use an OLS regression to calculate the coefficients $\widehat{\lambda^{P^i}}$

$$\widehat{\lambda^{P^i}} = ((\Phi - \beta E_{P^i}\Phi)'(\Phi - \beta E_{P^i}\Phi))^{-1}(\Phi - \beta E_{P^i}\Phi)'\widetilde{\pi}(P^i) \tag{4}$$

Step 3. New choice probabilities for each state are calculated using the fixed parameters $\theta$ and the multinomial logit choice formula

$$P_o'(a|\mathcal{M}_{j,o,t}) = \frac{\exp\left(\frac{FV(a,\mathcal{M}_{j,o,t},P^i_{-o}(\mathcal{M}_{j,o,t})) - W_o(a)\theta^W + \beta C_o(a)\theta^C}{\theta^\varepsilon}\right)}{\sum_{a' \in A_o(\mathcal{M}_{j,o,t})} \exp\left(\frac{FV(a',\mathcal{M}_{j,o,t},P^i_{-o}(\mathcal{M}_{j,o,t})) - W_o(a')\theta^W + \beta C_o(a')\theta^C}{\theta^\varepsilon}\right)} \tag{5}$$

where

$$FV(a,\mathcal{M}_{j,o,t}, P^i_{-o}(\mathcal{M}_{j,o,t})) = \sum_{h=1}^{H}\sum_{k=1}^{K} \phi_{h,k}(\mathcal{M}_{h,o,t+1}) \left\{ \frac{g(\mathcal{M}_{h,o,t+1}|a, P^i_{-o}(\mathcal{M}_{j,o,t}), \mathcal{M}_{j,o,t})}{\sum_{h'=1}^{H} g(\mathcal{M}_{h',o,t+1}|a, P^i_{-o}(\mathcal{M}_{j,o,t}), \mathcal{M}_{j,o,t})} \right\} \widehat{\lambda^{P^i}_k} \tag{6}$$

i.e., it reflects the states that can be reached given that action $a$ is chosen. The same formulae are used to calculate updated choice probabilities of competitors.

Step 4. If the maximum absolute difference between $P'$ and $P^i$ is less than 1e-5, the procedure stops, and the values of $\widehat{\lambda^{P^i}}$ are saved as $\lambda^*$. Otherwise, $P^i$ (i.e., both $P_o$ and $P_{-o}$) is updated as a weighted combination of $P^i$ and $P'$

$$P^{i+1} = \psi P' + (1-\psi)P^i \tag{7}$$

where $\psi = 0.1$, and step 1 is repeated for iteration $i+1$.

To start the procedure it is necessary to have an initial set of guesses $P^1$. I use the first-stage multinomial logit approximation to the CCPs.

**Forward Simulation.** The solution procedure gives conditional choice probabilities for each firm in the initial (observed) state of the market (period $t$). These choice probabilities and the AR(1) process for $\xi$ are used to simulate the model forward one period to $t+1$. In this new configuration, it is necessary to solve for a new set of choice probabilities. This involves a further iterative procedure. Before this procedure begins, a set of $H$ states for $t+2$ is drawn, and the approximating variables are calculated for these states. The $H$ states are chosen as described in Step 1 above. A set of initial choice probabilities to start the iterative procedure is also required, and, once again, I use the choice probabilities implied by the first-stage multinomial logit approximation to the CCPs. At iteration $i$, with choice probabilities $P^i$, the following scheme is followed:

Step 1. $P^i_{-o}(\mathcal{M}_{j,o,t+1})$ is used to calculate the $FV$ value for each possible action by each firm

$$FV(a,\mathcal{M}_{j,o,t+1}, P^i_{-o}(\mathcal{M}_{j,o,t+1})) = \sum_{h=1}^{H}\sum_{k=1}^{K}\phi_{h,k}(\mathcal{M}_{h,o,t+2})\left\{\frac{g(\mathcal{M}_{h,o,t+2}|a, P^i_{-o}(\mathcal{M}_{j,o,t+1}), \mathcal{M}_{j,o,t+1})}{\sum_{h'=1}^{H} g(\mathcal{M}_{h',o,t+2}|a, P^i_{-o}(\mathcal{M}_{j,o,t+1}), \mathcal{M}_{j,o,t+1})}\right\}\lambda^* \tag{8}$$

and the multinomial logit formula is used to calculate new choice probabilities

$$P'_o(a|\mathcal{M}_{j,o,t+1}) = \frac{\exp\left(\frac{FV(a,\mathcal{M}_{j,o,t+1},P^i_{-o}(\mathcal{M}_{j,o,t+1}))-W_o(a)\theta^W+\beta C_o(a)\theta^C}{\theta^\varepsilon}\right)}{\sum_{a'\in A_o(\mathcal{M}_{j,o,t+1})}\exp\left(\frac{FV(a',\mathcal{M}_{j,o,t+1},P^i_{-o}(\mathcal{M}_{j,o,t+1}))-W_o(a')\theta^W+\beta C_o(a')\theta^C}{\theta^\varepsilon}\right)} \tag{9}$$

Step 2. If the maximum absolute difference between $P'$ and $P^i$ is less than 1e-5, the procedure stops. Otherwise, $P^i$ (i.e., both $P_o$ and $P_{-o}$) is updated as a weighted combination of $P^i$ and $P'$

$$P^{i+1} = \psi P' + (1-\psi)P^i \tag{10}$$

where $\psi = 0.1$, and step 1 is repeated for iteration $i+1$.

The converged choice probabilities are used to simulate the model forward to the next period and the procedure is repeated, until the market is advanced for 40 periods.

# C Estimation

As explained in the text, estimation is separated into two stages. The parameters of the listener demand model, the revenue function and the process governing demographic growth rates are estimated in the first stage, along with a set of initial estimates of firms' conditional choice probabilities (CCPs) that are based on a multinomial logit choice model. In the second stage, these estimates are used to estimate the remaining parameters (repositioning costs, economies of scope and the scale of the payoff shocks associated with each format choice) using the dynamic model. This Appendix provides full details of these procedures.

## C.1 First Stage: Estimation of the Listener Demand Model and the Evolution of Unobserved Station Quality ($\xi$)

The listener demand model is a random coefficients demand model. There is no price variable, but there is a potential endogeneity problem as unobserved station quality may affect firms' format choices. I avoid this problem by forming quasi-differenced moments based on innovations in station quality ($\upsilon_{st}^{\xi}$), that are assumed to be unknown to firms when period $t$ format choices are made in period $t-1$. The model has 37 non-linear parameters ($\rho^{\xi}$, $\gamma^{\sigma}$ and 35 demographic taste parameters), collectively labelled $\gamma^{NL}$, and a set of linear parameters ($\gamma^{L}$) that capture format tastes, time effects and observable differences in station quality. Estimation involves minimizing a GMM objective function based on three sets of moments.

### C.1.1 Quasi–Differenced Moments

The quasi-differenced moments are formed from the mean utility equations for listener mean utilities and the AR(1) process that determines the evolution of the component of station quality that is not associated with observed station characteristics, $\xi$. For stations that do not change format

$$\nu_{st}^{\xi} = \delta_{st}(q, \gamma^{NL}) - \rho^{\xi}\delta_{st-1}(q, \gamma^{NL}) - (1-\rho^{\xi})X_{st}\gamma^{L} - (1-\rho^{\xi})F_{st}\overline{\gamma^{F}} \tag{11}$$

where the mean utility $\delta_{st}$ is uniquely defined by observed market shares $q$ and the non-linear taste parameters (Berry (1994), Berry et al. (1995)). The $X$ variables include station characteristics and format*AM interactions and, per the discussion in Section 3, time dummies. The assumption that the quality innovations $\nu_{st}^{\xi}$ are unknown when format choices are made implies that $\nu_{st}^{\xi}$ will be uncorrelated with $X$. The moments are formed as

$$E(Z'\upsilon^{\xi}(\rho^{\xi}, \gamma^{NL}, \gamma^{L})) = 0 \qquad ? \tag{12}$$

where the instruments $Z$ include $X_{st}$, $F_{st}$ and the log of the station's market share in the initial period of the data, which should be correlated with $\delta_{st-1}$. As in Nevo (2000, 2001), given estimates of the non-linear parameters, the linear parameters $\gamma^{L}$ can be estimated by linear regression where the dependent variable is $\delta_{st}(q, \gamma^{NL}) - \rho^{\xi}\delta_{st}(q, \gamma^{NL})$. $\sigma_{\upsilon^{\xi}}^{2}$ is estimated using the residuals from this regression.

I assume that the AR(1) process that governs the evolution of quality is the same for stations that change format, apart from a fixed quality change $\gamma^{\xi}$. This is potentially controversial, so I choose to estimate the model using only stations that stay in the same format, and then estimate $\gamma^{\xi}$ using the residuals implied by the estimated coefficients and the mean utilities of stations that do switch formats. I then examine how well the model does at matching the distribution of share changes for switching stations in the data. As shown in the text, the model does very well in this dimension, providing support for my assumption.

### C.1.2 Demographic Moments

Petrin (2002) illustrates how the accurate estimation of coefficients for demographic tastes using aggregate market share data can be aided by using demographic-specific moments. I form this type of moment based on the average demographic composition of the audience of different formats reported in Arbitron's annual *Radio Today* reports. Specifically these reports list the average proportion of a format's listeners who are in particular age (12-24, 25-49, 50 plus), gender and ethnic/racial (white, black or Hispanic) categories based on a particular set of markets. I specify 35 moments (which match the 35 demographic taste parameters) based on the difference between these reported averages and the averages predicted by my model for the quarters used by Arbitron and the set of markets that are common to my sample and Arbitron's calculations[7]

$$E(prop_{ftd}^{ARB} - prop_{ftd}(\delta(q, \gamma^{NL}), \gamma^{NL})) = 0 \tag{13}$$

where *prop* is the proportion of a format's listeners who are in a particular demographic group.

### C.1.3 One Additional Moment

The quasi-differenced moment with the lagged share instrument, and the demographic moments provide only 36 moments for identifying 37 non-linear parameters. Intuitively, the parameter which lacks an obvious identifying moment is $\gamma^\sigma$ which determines how much substitution takes place between radio listening and the outside good when the number or quality of stations changes. For example, a high value of $\gamma^\sigma$ implies that, all else equal, listening will increase slowly as the number of stations increases. To provide an additional moment I assume that the expected value of $\xi_{st}$, which could also affect how audiences increase with the number of stations in a market, is independent of market size, measured by log population.[8] This is similar in spirit to Berry and Waldfogel's (1999) use of population as an instrument to identify the nesting parameter in a nested logit model of station listenership.[9]

### C.1.4 Estimation Algorithm

Berry et al. (1995) and Nevo (2000) outline a simulation-based estimator for random coefficient demand models. I follow their algorithm, adding the additional moments outlined above. The algorithm involves solving for values of $\delta$ for each guess of the non-linear parameters, using analytic formula for the gradients of the objective function. The tolerance on this contraction mapping is

[7] Arbitron uses different markets for its age/gender and ethnic calculations. There are some markets included in Arbitron's calculations which are not in my sample. I have verified that the demographic taste coefficients remain similar if I include all of the markets used by Arbitron in the demand estimation. Creating the moments requires aggregating some of the formats used in Arbitron's reports, which is done by weighting these formats by average listenership.

[8] Specifically, I assume that the vector of $\xi_{st}$ for stations that are based inside the market should be independent of market size. The assumption would likely not hold for stations located outside of the market, as their signals are likely to cover less of the market in larger markets. The signal coverage of these stations is not observed in the data, so this difference would not be controlled for by the included $X_{st}$ variables.

[9] One might object to this moment on the basis that in larger markets, where fixed costs can be spread across more listeners, investment in quality is likely to be larger. However, if this objection was correct, audiences would increase with market size (correlated with the number of stations), and I would likely underestimate $\gamma^\sigma$. In practice, the estimated value of $\gamma^\sigma$ is very high implying that there is little substitution with the outside good.

set equal to 1e-12. Predicted shares for given non-linear parameters are calculated using 25 Halton draws of $v^R$ for each of the 18 demographic groups. The shares for each of the 450 simulated individuals are then weighted using the frequency of each demographic group in the population to calculate the predicted market share (results using more draws are almost identical).

This algorithm has been criticized based on examples where it fails to find the minimum of the objective function (Dube et al. (2011)). However, a feature of my model is that it is exactly identified, so I know that the minimized value of the objective function should be equal to zero (up to numerical tolerance). At the parameter estimates, the value of the objective function is 2.90e-12.

# The moments # parameters

## C.2 First Stage: Estimation of the Revenue Model

The revenue model assumes that station $s$'s revenues for a listener with demographics $D_d$ are

$$r_{st}(Y_s, D_d, \gamma) = \gamma_{my}(1 + Y_{st}\gamma^Y)(1 + D_d\gamma^D) \tag{14}$$

where $\gamma_{my}$ are market-year effects. However, only annual station revenues are reported in the data, so, for estimation, I assume that the mean annual revenues per listener, derived from BIAfn's revenue and share estimates, is

$$\overline{r_{sy}^{BIA}} = \frac{\sum_{t \in y} \sum_{\forall d} r_{st}(Y_s, D_d, \gamma) \widehat{l_{sdt}}(\delta, \widehat{\gamma^{NL}}, \widehat{\gamma^L})}{\sum_{t \in y} \sum_{\forall d} \widehat{l_{sdt}}(\delta, \widehat{\gamma^{NL}}, \widehat{\gamma^L})} + \varepsilon_{sy}^R \tag{15}$$

where $\widehat{l_{sdt}}(\delta, \widehat{\gamma^{NL}}, \widehat{\gamma^L})$ is the estimated listener demand model's prediction of $s$'s audience in demographic group $d$ in period $t$. The residual $\varepsilon_{sy}^R$ is assumed to be uncorrelated with station characteristics, local tastes or format choices, as if, for example, it is random measurement error in BIAfn's revenue formula. The model is estimated using Non-Linear Least Squares, and the standard errors are corrected, by expressing the first-order conditions as moments, for uncertainty in the estimated demand parameters.

## C.3 First Stage: Estimation of Demographic Transition Process

The population of ethnic group $e$ in market $m$ is assumed to evolve according to the following process

$$\log(pop_{met}) - \log(pop_{met-1}) = \tau_0 + \tau_1 (\log(pop_{met-1}) - \log(pop_{met-2})) + u_{met} \tag{16}$$

where $pop$ is the level of population. This model cannot be estimated directly as the *County Population Estimates* are annual (July each year), so they are only observed every other period. However, adding the equations for $t$ and $t-1$ and substituting for $(\log(pop_{met-1}) - \log(pop_{met-3}))$ in the resulting equation yields

$$\log(pop_{met}) - \log(pop_{met-2}) = 2\tau_0(1+\tau_1) + \tau_1^2 (\log(pop_{met-2}) - \log(pop_{met-4})) + \widetilde{u}_{met} \tag{17}$$

where $\widetilde{u}_{met} = u_{met} + (1+\tau_1)u_{met-1} + \tau_1 u_{met-2}$. The population numbers in this equation are observed, but $(\log(pop_{met-2}) - \log(pop_{met-4}))$ will be correlated with $\widetilde{u}_{met}$. I estimate (17) by 2SLS using $(\log(pop_{met-4}) - \log(pop_{met-6}))$ as an instrument for $(\log(pop_{met-2}) - \log(pop_{met-4}))$.[10] I estimate this equation using data on the black, white and Hispanic populations in all radio markets (not just the 102 markets in the sample) from 1996 to 2006, where the particular ethnic /racial group makes up at least 10% of the market population.[11] The estimates are $\widehat{\tau_0} = 0.00014$ (0.000039) and $\widehat{\tau_1} = 0.96968$ (0.00335). The standard deviation of the innovations $u_{met}$ is 0.0027.

## C.4 First Stage: Estimation of Firm CCPs

I calculate initial estimates of firms' choice probabilities by estimating a multinomial logit model, where, as in the true model, the choices for each firm are to keep its stations in the same format or to move one of them to a new format. In an ideal world, these CCPs would be estimated non-parametrically, but this is not possible given the size of the state space, the large number of choices that each firm has and the size of the observed sample. The small number of observations where a firm moves more than one station are not included when calculating the likelihood. The following explanatory variables are included in the logit model for each choice:

[10] The instrument will be correlated with the endogenous variable if $\tau_1 \neq 0$ (serial correlation in population growth rates) and it should be uncorrelated with $\widetilde{u}_{met}$ if the innovations in growth rates are independent.

[11] Including observations on smaller population groups leads to more volatile growth rates, which can create some implausible population changes when applied to larger populations.

- a dummy for whether the choice involves a station moving to an active format, and interactions with a measure of market revenues per share point (to capture market size effects), and for station being moved: the current period revenue, $\exp(\xi_{st})$, the exponent of the fixed quality component (e.g., based on signal coverage), the interaction of these two exponentiated qualities and an interaction of $\exp(\xi_{st})$ with a dummy for whether the station is an AM station
  - interactions of these variables with the total number of stations that the firm owns and the current period revenues of the firm
- a dummy for whether the choice involves a station moving from an active format to Dark, and interactions with the prior revenue of the station being moved and the measure of market revenues per share point
  - interactions of these variables with the total number of stations that the firm owns and the current period revenues of the firm
- a dummy for whether the moving station has made a format switch in the previous period and an interaction with market revenues per share point
  - interactions of these variables with the total number of stations that the firm owns and the current period revenues of the firm
- a dummy for whether the moving station is coming from the inactive dark format, and an interaction with market revenues per share point
  - interactions of these variables with the total number of stations that the firm owns and the current period revenues of the firm
- a count of how many of the owner's stations will be located in a format with another station with the same owner and the interaction of this variable with market revenues per point;
  - interactions of these variables with the total number of stations that the firm owns and the current period revenues of the firm
- for the active format that the station is being moved from:

  - a dummy (e.g., a dummy for a move from Country)
  - this dummy interacted with a dummy for whether the moving station is an AM station
  - this dummy interacted with the following demographic measures: proportion black, proportion Hispanic, proportion aged 50+, proportion aged 12-24, the ratio of the growth rate of the black population and the white population, the ratio of the growth rate of the Hispanic population and the white population

- five measures of the intensity of competition that the firm faces in the format that it is moving a station from:
  - the number of stations owned by other firms
  - the number of stations owned by other firms that own multiple stations in the format
  - the sum of $\exp(\xi)$s for stations owned by other firms
  - the sum of exponents of the fixed quality component (e.g., based on signal coverage) for stations owned by other firms
  - the sum of the "revenue potential" measures of stations owned by other firms (see the discussion of this variable in Appendix B)

- for the active format that the station is being moved to:
  - a dummy (e.g., a dummy for a move to Rock)
  - this dummy interacted with a dummy for whether the moving station is an AM station
  - this dummy interacted with the following demographic measures: proportion black, proportion Hispanic, proportion 50+, proportion 12-24, the ratio of the growth rate of the black population and the white population, the ratio of the growth rate of the Hispanic population and the white population

- five measures of the intensity of competition that the firm faces in the format that the firm is moving a station to:
  - the number of stations owned by other firms;

- the number of stations owned by other firms that own multiple stations in the format
- the sum of $\exp(\xi)$s for stations owned by other firms
- the sum of exponents of the fixed quality component (e.g., based on signal coverage) for stations owned by other firms
- the sum of the revenue potential measures of stations owned by other firms (see the discussion of the revenue potential variable in Appendix B)

Excluding certain format dummies to avoid perfect multi-collinearity, this specification has 160 explanatory variables so it is quite flexible but it is, of course, quite a coarse representation given the richness of the state space. Table 2 gives the coefficient estimates from this model. Relative to a baseline model where the only dummies are for a switch to active format, a switch to Dark and a switch from Dark, the pseudo-$R^2$ of the estimated model is 0.142, suggesting that while the included variables have explanatory power, many observed switches look quite random.

## C.5 Second Stage: Estimation of the Dynamic Model Using Value Function Approximation

As explained in the text, I consider a number of different estimators of the dynamic model. In this Appendix I detail the estimators that use parametric approximations to the value function. Appendix D details the estimators that use forward simulation to approximate the value function. I begin the discussion by specifying features that are common to all of the parametric approximation estimators.

**Selection of states.** While I can only estimate the model using firms' observed choices, I am not limited to using only observed states when I approximate the value function. I therefore use the 6,061 observed states from the 612 observed market-quarters where I observe firms' choices for the next period, and then create 9 duplicates of each of these market-quarters ($N = 60,610$ states in total) where station formats, unobserved qualities and market demographics can take on different values from those that are observed. Ownership and observed station characteristics are held fixed as this is assumed in the model. In each duplication, the unobserved quality of each station is chosen as a uniform random draw on $[-2, 2]$, a range which comprises almost all of the values implied by the estimated demand model. The probability that a station's format is the same as in the data is 0.3, and with probability 0.7 it receives a new format draw. If this happens, the probability of Dark is set equal to 0.05 and the probability of each of the other formats are set to be equal to each other. Demographics are altered by varying the size of the white, Hispanic and black populations by iid draws uniformly drawn from $[-20\%, 20\%]$ (the size of each age-gender group within the ethnic/racial group changes by the same percentage). In the description of the estimation procedure, a particular state $j$ is denoted as $\mathcal{M}_{j,o,t}$ where $o$ indicates the firm of interest in the state and $t$ denotes the initial period.

**Variables used in approximating the value function.** I assume that firm's value functions can be approximated by a linear parametric function of functions of the state variables. I assume the same function holds across markets, but a number of interactions with market characteristics provide flexibility.[12] The functions include measures based on revenues calculated in several different ways, which I now describe.

- 'actual revenues': a station's revenues in a format given demographics, formats and station characteristics including the $\xi_{st}$s
- 'no $\xi$ revenues': a station's revenues in a format given demographics, formats and station characteristics except the time-varying $\xi_{st}$ (i.e., the $\xi$s of all stations are set equal to zero)
- '$\overline{revenue}$': a measure of the station's average revenue potential excluding the $\xi_{st}$s formed by averaging the revenues that it would get across a large set of format configurations for all stations, for a fixed set of demographics.[13]

The approximating function for a state $\mathcal{M}_{j,o,t}$ includes the following variables:

[12] Implementation of methods that use initial estimates of the conditional choice probabilities usually assumes that these CCPs can also be approximated by the same parametric function across markets.

[13] Note that this measure does not depend on stations' current formats and so, for a particular station, it stays the same across periods.

- a market-quarter fixed effect (note that this takes the same value for each of the perturbed versions of the same market-quarter);
- the following measures of market demographics:
  - the proportion of blacks in the population
  - the proportion of Hispanics in the population
  - these variables interacted with market population multiplied by the market-year price ($\gamma_{my}$) effect from the estimated revenue equation, and interacted with the number of stations owned by firm $o$
- the sum of station actual revenues for firm $o$'s stations in period $t$
- a count of how many station the firm has in formats where it has more than one station;
- the sum of the $\overline{revenue}$ measures for firm $o$'s stations
- the sum of the $\exp(\xi_{st})$ measures for firm $o$'s stations
- the sum of the $\exp(X_{st}\gamma^S)$ measures for firm $o$'s stations (excluding the AM*format component)
- the sum of the AM dummy for firm $o$'s stations
- the sum of the AM dummy interacted with $\exp(X_{st}\gamma^S)$ (excluding the AM*format component) for firm $o$'s stations
- the sum of the AM dummy interacted with $\exp(\xi_{st})$ for firm $o$'s stations
- the sum of the $\exp(X_{st}\gamma^S)$ (excluding the AM*format component) interacted with $\exp(\xi_{st})$ for firm $o$'s stations
- the sum of $\overline{revenue}$ interacted with $\exp(\xi_{st})$ for firm $o$'s stations
- the last 8 variables interacted with a count of the number of stations owned by firm $o$, and interactions with market population multiplied by the market-year price ($\gamma_{my}$) effect from the estimated revenue equation

- the sum of the AM dummy for firm $o$'s stations in the News/Talk format
- the sum of the AM dummy for firm $o$'s stations in the News/Talk format * the total number of AM stations in the market
- the last two variables interacted with the proportion of blacks and Hispanics in the population, and these four variables interacted with market population multiplied by the market-year price ($\gamma_{my}$) effect from the estimated revenue equation
- for the largest competitor faced by the firm (call it firm $x$)[14]:
  - the number of stations owned by firm $x$
  - the sum of the $\overline{revenue}$ measures for firm $x$'s stations
  - the sum of the $\exp(\xi_{st})$ measures for firm $x$'s stations
  - the sum of interactions between the $\overline{revenue}$ measures and $\exp(\xi_{st})$ for firm $x$'s stations
  - the interactions of these four variables with market population multiplied by the market-year price ($\gamma_{my}$) effect from the estimated revenue equation
- for the second largest competitor faced by the firm (call it firm $y$):
  - the number of stations owned by firm $y$
  - the sum of the $\overline{revenue}$ measures for firm $y$'s stations
  - the sum of the $\exp(\xi_{st})$ measures for firm $y$'s stations
  - the sum of interactions between the $\overline{revenue}$ measures and $\exp(\xi_{st})$ for firm $y$'s stations
  - the interactions of these four variables with market population multiplied by the market-year price ($\gamma_{my}$) effect from the estimated revenue equation
- for each active format (i.e., the coefficients can vary freely across formats):
  - number of rival stations in the format
  - number of rival stations in the format owned by firms that own more than one station

[14] The size of competitors is determined by the number of stations owned and, where this is equal, the sum of the $\overline{revenue}$ measures for the different firms.

- sum of rival stations $\exp(\xi_{st})$s
- sum of rival stations $\exp(X_{st}\gamma^S)$ (excluding the AM*format component)
- sum of the $\overline{revenue}$ measures for rival stations
- sum of the $\overline{revenue}$ measures for rival stations that are in the AM band
- the interaction of these 6 variables with (i) the proportion black in the population, (ii) the proportion Hispanic in the population, (iii) the growth rate of the black population, (iv) the growth rate of the Hispanic population and (iii) the interaction of these 30 variables with market population multiplied by the market-year price ($\gamma_{my}$) effect from the estimated revenue equation

- a dummy for whether firm $o$ moved a station in the previous period[15]
- the $\overline{revenue}$, the $\exp(\xi_{st})$ and the interaction of $\overline{revenue}$ and $\exp(\xi_{st})$ of a station moved by $o$ in the previous period
- a set of measures of the potential gains in revenue that firm $o$ could achieve in the following period if the formats of other stations were held fixed. These are calculated using the 'no $\xi$ revenues' (which negates the need to try to integrate over a set of changes in unobserved qualities)
  - the number of moves firm $o$'s could make that would raise expected revenues
  - the sum of revenue gains the firm could make from these moves, plus an interaction with the number of stations that the firm owns
  - the maximum revenue gain from moving a station in \$m. and as a proportion of firm $o$'s current revenue
  - the sum of moves that would increase revenues that would involve moving a recently-moved station
  - the sum of the count of how many stations the firm would have in formats where it has more than one station based on formats that would increase the firm's revenues;

[15] Recall that I allow for a different cost of moving a recently-moved station.

In the following description of the estimation procedure, $\Phi_{j,k}(\mathcal{M}_{j,o,t})$ is the value of the $k^{th}$ approximating variable in state $j$ and $\Phi$ is the matrix where these variables are stacked for the $N$ states.

**Initial choice probabilities.** Initial values of the CCPs are required to start the estimation procedure, and in the case of the modified procedure the choice probabilities of *other* firms are held constant at these initial values during estimation. In both cases I use the first stage estimates from the multinomial logit choice model. This involves extrapolation using the parametric form of this model for the states that are not observed in the data.

### C.5.1 Modified Pseudo-Likelihood and Moment-Based Procedures

I begin by describing the procedures used to produce the estimates in the first two columns of Table 6 in the text. In these estimators the choice probabilities of other firms ($P_{-o}$) are held fixed at their initial (first stage multinomial logit) estimates. The logic of these procedures follows Aguirregabiria and Mira (2010) (discussed in Aguirregabiria and Nevo (2012)), although the moment-based version is also inspired by the discussion and Monte Carlo results in Pakes et al. (2007, POB).

Estimation is based on an iterative procedure with the following steps in iteration $i$. In state $\mathcal{M}_{j,o,t}$, $P_o^i(a|\mathcal{M}_{j,o,t})$ is the iteration $i$ guess of the probability that firm $o$ chooses action $a$, $P_{-o}(\mathcal{M}_{j,o,t})$ are the (fixed) choice probabilities of $o$'s competitors in that state.

Step 1. For each of the $N$ states, $\widetilde{\pi}(P_o^i(\mathcal{M}_{j,o,t}), \theta^i)$ is calculated as

$$\widetilde{\pi}(P_o^i(\mathcal{M}_{j,o,t}), \theta^i) = \sum_{s \in S^o} R_s(\mathcal{M}_{j,o,t}|\gamma) + \sum_{a \in A_o(\mathcal{M}_{j,o,t})} P_o^i(a|\mathcal{M}_{j,o,t}) \left( \begin{array}{c} \beta C_o(a)\theta^{C,i} - W_o(a)\theta^{W,i} \\ +\theta^{\varepsilon,i}(\varkappa - \log(P_o^i(a|\mathcal{M}_{j,o,t})) \end{array} \right) \quad (18)$$

where $\theta^i$ denotes the current guess of the parameters, and $\varkappa$ is Euler's constant, $C_o(a)$ is the number of stations that the firm will have operating in the same format as one of its other stations in the next period if it chooses action $a$ and $A_o(\mathcal{M}_{j,o,t})$ is $o$'s choice set. $\widetilde{\pi}(P_o^i(\mathcal{M}_{j,o,t}),\theta^i)$ is a function only of $o$'s choice probabilities. $\widetilde{\pi}(P^i,\theta^i)$ is the vector that stacks these values for the $N$ states.[16]

For each of the $N$ states, the choice probabilities of all firms are used to calculate $E_{P_o^i}\Phi$, a vector which contains the expected value of each of the approximating variables given strategies. For variable $k$,

$$E_{P_o^i}\phi_{j,k} = \int \phi_{h,k}(\mathcal{M}_{h,o,t+1})g(\mathcal{M}_{h,o,t+1}|P_o^i(\mathcal{M}_{j,o,t}),P_{-o}(\mathcal{M}_{j,o,t}),\mathcal{M}_{j,o,t})d\mathcal{M}_{h,o,t+1} \tag{19}$$

where $g$ is the transition density. This integral is approximated by reweighting variables for a pre-specified sample of $\mathcal{M}_{h,o,t+1}$ states, as calculating $\phi_{h,k}(\mathcal{M}_{h,o,t+1})$ requires solving a random coefficients demand model. Specifically for a given state $\mathcal{M}_{j,o,t}$, I consider a set of $H$ states $\mathcal{M}_{h,o,t+1}$ which is equal to the set of states that can be reached by any move by $o$, a set of $S^{\xi}$ draws for innovations in $\xi$ and $S^{-o,m}$ moves by other firms in the same local market. The integral is approximated by

$$E_{P_o^i}\phi_{j,k}^i = \sum_{h=1}^{H}\phi_{h,k}(\mathcal{M}_{h,o,t+1})\frac{g(\mathcal{M}_{h,o,t+1}|P_o^i(\mathcal{M}_{j,o,t}),P_{-o}(\mathcal{M}_{j,o,t}),\mathcal{M}_{j,o,t})}{\sum_{h'=1}^{H}g(\mathcal{M}_{h',o,t+1}|P_o^i(\mathcal{M}_{j,o,t}),P_{-o}(\mathcal{M}_{j,o,t}),\mathcal{M}_{j,o,t})} \tag{20}$$

$S^{\xi}=10$ and $S^{-o,m}=500$. To be accurate the integration procedure requires the $S^{-o,m}$ moves to include those that are most likely to be made. I choose the ones that are most likely to be made based on the first-stage estimates of the conditional choice probabilities.

Step 2. Create matrices $(\Phi-\beta E\Phi^i)$ and, as the parameters $\lambda$ are overidentified when $N>K$, use an OLS regression to calculate the coefficients $\widehat{\lambda^i}$

$$\lambda^i\widehat{(\theta^i,P_o^i)} = ((\Phi-\beta E_{P_o^i}\Phi^i)'(\Phi-\beta E_{P_o^i}\Phi^i))^{-1}(\Phi-\beta E_{P_o^i}\Phi^i)'\widetilde{\pi}(P_o^i,\theta^i) \tag{21}$$

Step 3. Use $\widehat{\lambda^i}$ to calculate the future value of each firm when it makes choice $a$ (note this is not quite the same as the choice-specific value function as defined in the text as that also includes

[16] An initial guess of the structural parameters is required. I assume a common switching cost of \$2m for all switches between active formats, a cost of \$1m for a switch from Dark, a cost of \$4m for a switch to Dark, $\theta^C=0.1$ and $\theta^{\varepsilon}=0.5$.

current revenues and repositioning costs associated with $a$)

$$FV(a,\mathcal{M}_{j,o,t},P_{-o}(\mathcal{M}_{j,o,t}))=\sum_{h=1}^{H}\sum_{k=1}^{K}\phi_{h,k}(\mathcal{M}_{h,o,t+1})\left\{\frac{g(\mathcal{M}_{h,o,t+1}|a,P_{-o}(\mathcal{M}_{j,o,t}),\mathcal{M}_{j,o,t})}{\sum_{h'=1}^{H}g(\mathcal{M}_{h',o,t+1}|a,P_{-o}(\mathcal{M}_{j,o,t}),\mathcal{M}_{j,o,t})}\right\}\lambda_k^i\widehat{(\theta^i,P_o^i)} \tag{22}$$

Step 4. Estimate the structural parameters $\theta'$ using firms' observed choices. The probability that $a$ is chosen is

$$P_o(a|\mathcal{M}_{j,o,t})=\frac{\exp\left(\frac{FV(a,\mathcal{M}_{j,o,t},P_{-o}(\mathcal{M}_{j,o,t}))-W_o(a)\theta^W+\beta C_o(a)\theta^C}{\theta^\varepsilon}\right)}{\sum_{a'\in A_o(\mathcal{M}_{j,o,t})}\exp\left(\frac{FV(a',\mathcal{M}_{j,o,t},P_{-o}(\mathcal{M}_{j,o,t}))-W_o(a')\theta^W+\beta C_o(a')\theta^C}{\theta^\varepsilon}\right)} \tag{23}$$

Current revenues drop out because they are common across choices. The pseudo-likelihood and moment-based estimators differ in how these probabilities are used.

For the *pseudo-likelihood estimator*, the probabilities are used in what is similar to a standard multinomial logit estimation, except that the scale parameter differs across markets. Observations for firms moving more than one station are excluded from the calculation of the pseudo-likelihood. One advantage of this estimator is that the log-likelihood objective function is well-behaved and easy to maximize.

For the *moment-based estimator*, the probabilities are used to match a number of informative switching rates in the data, where the rates are formed by averaging across states. Specifically for the three market-size groups (population less than 0.25 m., 0.25m.-1m., 1m.+) the estimator matches: (i) the (average-across-states) probability that a station is moved from one active format to another active format, (ii) the probability that a station is switched to Dark, (iii) the probability that a Dark station is moved to an active format; (iv) the probability that a station that switched formats in the previous period makes a further switch, the probability that a non-Urban station switches to Urban in markets with (v) a below median proportion of blacks and (vi) an above median proportion of blacks, the probability that a non-Spanish station switches to Spanish in markets with (vii) a below median proportion of Hispanics and (viii) an above median proportion of Hispanics, (ix) the probability that a non-News AM station is moved to News and (x) the probability that a multi-station firm chooses a move that increases the number of stations that it operates in the same format. One more moment is provided for each market group by matching the average revenue of a switching station. Observations for firms moving more than one station are excluded when calculating both the data and predicted moments. The identity matrix is used to weight the moments. One disadvantage of this estimator is that the objective function can have multiple local minima. The estimation routine therefore uses a both Nelder-Mead and derivative-based optimization routines from different starting points to search for the global minimum.

Step 5. Use $\theta'$ to compute

$$P'_o(a|\mathcal{M}_{j,o,t}) = \frac{\exp\left(\frac{FV(a,\mathcal{M}_{j,o,t},P_{-o}(\mathcal{M}_{j,o,t}))-W_o(a)\theta'^W+\beta C_o(a)\theta'^C}{\theta'^\varepsilon}\right)}{\sum_{a'\in A_o(\mathcal{M}_{j,o,t})}\exp\left(\frac{FV(a',\mathcal{M}_{j,o,t},P_{-o}(\mathcal{M}_{j,o,t}))-W_o(a')\theta'^W+\beta C_o(a')\theta'^C}{\theta'^\varepsilon}\right)} \tag{24}$$

Step 6. If the maximum absolute difference between $P'_o$ and $P^i_o$ is less than 1e-4 and the maximum absolute difference between $\theta^i$ and $\theta'$ is also less than 1e-4, the procedure stops. Otherwise, $P_o$ is updated as a weighted combination of $P^i_o$ and $P'$

$$P^{i+1}_o = \psi P'_o + (1-\psi)P^i_o$$

and $\theta$ is updated as

$$\theta^{i+1}_o = \psi\theta' + (1-\psi)\theta^i$$

where $\psi = 0.1$, and the procedure returns to step 1 for iteration $i+1$.

### C.5.2 Iterated Pseudo-Likelihood Procedure

The iterated pseudo-likelihood procedure follows Aguirregabiria and Mira (2007) in that the choice probabilities of other firms are also updated during estimation. Specifically, in the description set out above $P_{-o}$ should be replaced by $P^i_{-o}$ and in step 6 the choice probabilities of all players are updated.

# D Estimation of the Dynamic Model Using Forward Simulation

An alternative approach to estimating dynamic games, and approximating value functions, involves the use of forward simulation, an approach most closely associated with Bajari et al. (2007, BBL). I implement two estimators that use forward simulation: one based on the objective function proposed by BBL, and one that is based on moment inequalities following Pakes et al. (2011, PPHI) which involves averaging across states.

As explained in the text, it can be difficult to estimate a large number of parameters using these methods, so I only consider a model with 3 parameters (a cost of switching to an active format $\theta^W$, the economy of scope from operating multiple stations $\theta^C$ in the same format and the scale parameter of the iid payoff shocks to each format choice, $\theta^\varepsilon$) and estimate the model separately for each of the three market-size groups.[17] I begin by describing the BBL estimator, and then explain the changes made to implement the PPHI estimator.

*BBL.* The BBL estimator uses the equilibrium assumption that, given the strategies of other firms, a firm's observed policy should give it higher expected payoffs than any alternative policy. Given the linear form of the payoff function, a firm's value when it uses strategy $\Gamma_o$ and other firms use strategies $\Gamma^*_{-o}$ can be expressed as

$$V(\mathcal{M}_{j,o,t}|\Gamma_o,\Gamma^*_{-o},\theta) = \mathbf{V}_{o,\Gamma_o,\Gamma^*_{-o}}\theta = \mathbf{R}_{o,\Gamma_o,\Gamma^*_{-o}} - \theta^W \mathbf{W}_{o,\Gamma_o,\Gamma^*_{-o}} + \theta^C \mathbf{C}_{o,\Gamma_o,\Gamma^*_{-o}} + \theta^\varepsilon \boldsymbol{\varepsilon}^F_{o,\Gamma_o,\Gamma^*_{-o}} \tag{25}$$

$$\text{where } \mathbf{R}_{o,\Gamma_o,\Gamma^*_{-o}} = E_{o,\Gamma_o,\Gamma^*_{-o}} \sum_{t'=0}^{\infty} \beta^{t'} \sum_{s\in S^o} R_s(\mathcal{M}_{o,t+t'}|\gamma),$$

$$\mathbf{W}_{o,\Gamma_o,\Gamma^*_{-o}} = E_{o,\Gamma_o,\Gamma^*_{-o}} \sum_{t'=0}^{\infty} \beta^{t'} \sum_{s\in S^o} I(f_{st+t'} \neq f_{st+t'+1}, f_{st+t'+1} \neq \text{DARK}) \tag{26}$$

$$\mathbf{C}_{o,\Gamma_o,\Gamma^*_{-o}} = E_{o,\Gamma_o,\Gamma^*_{-o}} \sum_{t'=1}^{\infty} \beta^{t'} C_o(\mathcal{M}_{o,t+t'}),\ \boldsymbol{\varepsilon}^F_{o,\Gamma_o,\Gamma^*_{-o}} = E_{o,\Gamma_o,\Gamma^*_{-o}} \sum_{t'=0}^{\infty} \beta^{t'} \varepsilon^F_{ot+t'}(a_{ot})$$

where I use $\mathcal{M}_{o,t+t'}$ to denote whatever state firm $o$ is in period $t+t'$. The necessary equilibrium

[17] In simplifying the model I assume that the cost of moving from Dark to an active format is the same as moving between a pair of active formats, and that there is no cost to moving to Dark. I have estimated specifications with separate coefficients for these costs, but without the imposition of arbitrary constraints, found that the estimates produced were often completely implausible (e.g., a cost of \$100 million for switching to Dark).

condition is that

$$V(\mathcal{M}_{j,o,t}|\Gamma_o^*, \Gamma_{-o}^*, \theta) - V(\mathcal{M}_{j,o,t}|\Gamma_o^a, \Gamma_{-o}^*, \theta) \geq 0 \ \forall \Gamma_o^a, \mathcal{M}_{j,o,t} \tag{27}$$

where $\Gamma_o^*$ is firm $o$'s observed equilibrium strategy and $\Gamma_o^a$ are alternative strategies.

$\mathbf{R}_{o,\Gamma_o^*,\Gamma_{-o}^*}$, $\mathbf{W}_{o,\Gamma_o^*,\Gamma_{-o}^*}$, $\mathbf{C}_{o,\Gamma_o^*,\Gamma_{-o}^*}$ and $\boldsymbol{\varepsilon}_{o,\Gamma_o^*,\Gamma_{-o}^*}$ can be approximated using forward simulation of the model and initial estimates of firms' conditional choice probabilities. $\mathbf{R}_{o,\Gamma_o^a,\Gamma_{-o}^*}$ (and the equivalent for the other terms) can be approximated by using a different set of choice probabilities. BBL propose finding the parameters that make these inequalities hold in the data for a finite set of alternatives using an objective function

$$\widehat{\theta^{BBL}} = \arg\min_{\theta} \sum_{o} \sum_{\forall a} \max\{(\mathbf{V}_{o,\Gamma_o^a,\Gamma_{-o}^*} - \mathbf{V}_{o,\Gamma_o^*,\Gamma_{-o}^*})\theta, 0\}^2$$

where the estimates will be a set if there are parameters that satisfy all of the inequalities. The estimator has a manageable computational burden because it is not necessary to recalculate $\mathbf{R}, \mathbf{W}, \mathbf{C}$ and $\boldsymbol{\varepsilon}$ as the parameters change. It is straightforward to add the additional parameter restriction that $\theta^{\varepsilon} \geq 0$ (scale of the payoff shocks must be non-negative).

The iterative forward simulation procedure is straightforward. Suppose that we want to simulate the values of $\mathbf{R}, \mathbf{W}, \mathbf{C}$ and $\boldsymbol{\varepsilon}$ for a particular firm $o$ using observed policies $\Gamma_o^*$. For a given simulation $sim$, we start from an initial state, setting $\mathbf{R}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$, $\mathbf{W}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$, $\mathbf{C}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$ and $\boldsymbol{\varepsilon}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$ to zero, and then iterate the following steps:

Step 1. Given the state, calculate $o$'s revenues by solving the random coefficients model of listener demand and then using the estimated revenue model to calculate the total revenues each station receives. For $o$, increase the value of $\mathbf{R}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$ and $\mathbf{C}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$ based on its revenues and the format configuration of its stations.

Step 2. Given the state, form a matrix that contains the same explanatory variables used for the first-stage multinomial logit estimates of conditional choice probabilities. Then use the estimated coefficients of this model to calculate the CCPs for all firms in the market.

Step 3. Using these CCPs, choose an action for each firm and update all station formats. For $o$, update $\mathbf{W}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$ if it changes the format of one of its stations. Update $\boldsymbol{\varepsilon}_{sim,o,\Gamma_o^*,\Gamma_{-o}^*}$ by adding $(\varkappa - \log(P_o^*(a)))$ where $\varkappa$ is Euler's constant and $a$ is the action that firm $o$ chooses.

Step 4. Use the estimated transition processes for $\xi$ and market demographics to update these variables.

Step 5. Repeat steps 1-5 for the next period, and continue until the model has been simulated forward 50 periods.

Given that a market can evolve in many ways, it is necessary to average across many simulations in order to reduce simulation error, although this increases the computational burden, especially as it is necessary to solve a random coefficient demand model in each period for each simulation and policy. I use 500 simulations and construct inequalities based on all of the observed states in the data. I experimented using 2,000 simulations for small markets, where the BBL estimates of $\theta^W$ and $\theta^\varepsilon$ are larger than the parametric approximation estimates and lie outside the PPHI bounds that I describe below. The BBL estimates were slightly further away from the other estimates and the PPHI bounds in this case, suggesting that the number of simulations does not explain the results. However, one advantage of the PPHI formulation discussed below is that it may reduce the effects of simulation error in the estimates of the components of the value function.

Any deviation from $\Gamma_o^*$ provides a possible alternative policy that can be used for estimation. My experience from estimating this model and other models with large state spaces is that the choice of alternative policies can significantly affect the results, especially with the BBL objective function.[18] Out of the alternatives I tried, the ones described below provided the estimates that were most similar to those implied by my other estimators, and they also appeared to be among the most robust to varying the set of states used in calculating the objective function. There is also an intuitive reason

[18] Srisuma (2010) discusses an example where a commonly used alternative, which involves adding noise to the choice probabilities, cannot identify the parameters.

why each of the alternatives used should help to identify the parameters. In forming them, I use information on what the model predicts about a firm's revenues in the following period for each of its possible actions, assuming that the formats of other stations are held fixed and the unobserved qualities of all stations are set equal to 0.[19] It is these revenues that I am referring to when I talk about 'next period revenues' in the following descriptions.

Alternative Policy 1: More format switching: if a format choice, involving a switch, gives a higher next period revenue than maintaining the same format configuration but under the estimated actual policies (i.e., estimated CCPs) it would be chosen with lower probability[20], make the probabilities of making this switch and maintaining the same configuration equal to each other. Intuitively this alternative policy should tend to increase a firm's expected future revenues, but also increase the amount of switching that it does, and the fact that this policy is not optimal should identify a lower bound on repositioning costs.

Alternative Policy 2: Less format switching: reduce the probability that a firm makes each choice involving moving a station to another active format by 90%, increasing the probability that the firm chooses to maintain its existing formats. Assuming that moves that increase revenues are more likely to be chosen, this change will reduce switching and expected future revenues, and the fact that this policy is not optimal should identify an upper bound on repositioning costs.

Alternative Policy 3: Higher probability of making format choices that increase clustering of stations: identify format choices that would increase the number of stations in the same active format relative to the choice of no move (call these the 'increase options'), and those that would reduce it (the 'reduce options'). Reduce the probability of choosing each of the reduce options by two-thirds, and proportionally increase the probability of choosing each of the increase options. As clustering of stations will result in cannibalization it will tend to reduce expected future revenues.

[19] Results are similar if instead the $\xi$s are assumed to stay fixed at their current values, but experimentation indicated that using the $\xi = 0$ revenues gives a slightly better prediction about which switches will increase a firm's revenues in the long-run, consistent with the fact that the $\xi$s are transitory.

[20] If multiple moves would produce higher expected revenues, I use the one that has the highest expected revenue.

Intuitively the fact that this policy is not optimal should identify an upper bound on the value of economies of scope.

Alternative Policy 4: Lower probability of making format choices that will increase station clustering: this is simply the reverse of Alternative Policy 3 (i.e., the probability of the reduce options is increased, and the probability of increase options is reduced). As spreading out stations will tend to increase expected future revenues because it reduces cannibalization, intuitively the fact that this policy is not optimal should identify a lower bound on the value of economies of scope.

Alternative Policy 5: More random switching: identify all of the format choices that will raise a firm's next period revenues relative to keeping its current format configuration, and set the probability of all of these choices equal to each other. Conditional on one of these choices being made, setting the probabilities to be equal to each other maximizes the expected value of the payoff shock associated with the choice. Intuitively, this alternative policy should also reduce expected future revenues (because the probability of the best of these options will have fallen) or reduce economies of scope, so the fact that this policy is not optimal should identify an upper bound on $\theta^{\varepsilon}$ (the scale of the payoff shocks).

Alternative Policy 6: Less random switching: identify all of the format choices that will raise the firm's next period revenues relative to keeping its current format configuration and set the probability of all of these choices except the one that maximizes next period revenues equal to zero, attributing these probabilities to the choice that does maximize next period revenues. Intuitively, this alteration should reduce the expected value of $\varepsilon$s while increasing expected future revenues, and the fact that this policy is not used should identify a lower bound on $\theta^{\varepsilon}$.

For each of these alternative policies the forward simulation calculations are repeated. Other firms continue to use the conditional choice probabilities implied by $\Gamma^*_{-o}$, and the draws of demographics and innovations in $\xi$ are the same as in the simulations for $\Gamma^*_o$. Applied to my data, the BBL estimator always produces point estimates because there are no parameters which satisfy the inequalities for all firms.

*Moment Inequalities (PPHI)***.** The BBL estimator uses the fact that the inequality (27) should hold for any state and any alternative policy, but in practice the inequalities may not hold at the estimated parameters for a significant number of states and alternative policies, and the estimates may not be consistent when the estimates of the first-stage conditional choice probabilities are inaccurate or there is significant simulation error in the approximation of $\mathbf{R}, \mathbf{W}, \mathbf{C}$ and $\varepsilon$. An alternative estimation approach makes the weaker assumption that the same set of inequalities should hold when an average is taken across $O$ states (in practice, all states in the observed data for a given market-size group), producing an estimating moment inequality of the form

$$\frac{1}{O}\sum(\mathbf{V}_{o,\Gamma_o^*,\Gamma_{-o}^*} - \mathbf{V}_{o,\Gamma_o^a,\Gamma_{-o}^*})\theta \geq 0 \text{ for an alternative policy } \Gamma^a$$

Arguments for why averaging might be helpful in the presence of first-stage bias or simulation error are analogous to the arguments presented in Pakes et al. (2007) for why a moment-based estimator using switching rates may be more reliable. I construct one of these linear moment inequalities for each of the alternative policies considered above (using exactly the same simulations used for the BBL estimator), and find the set of parameters which satisfy all of these inequalities (in my data this approach always generates a set). The lower and upper bounds of this set for each of the parameters are reported in Table 7 of the text, and the 'interval inference method' described in Section 3 of Pakes et al. (2011) is used to calculate the 95% confidence interval for each of these bounds.

Of course, the disadvantage of averaging is that information about changes in payoffs in individual states when alternative policies are used is lost. If the only effect of averaging across states was to lose information, then we would expect the BBL estimates to be within the (possibly wide) sets generated by the PPHI estimator, but for at least two of the three market-size groups this is not the case, suggesting that averaging across states is also reducing problems created by using inaccurate estimates of the choice probabilities or forward simulation error.[21]

[21] As noted above, this result does not seem to be affected by increasing the number of simulations. Therefore the primary problem is likely to be the effects of bias in the first-stage estimates so that the CCPs may be inaccurate in some states. It is also noticeable that for medium-sized markets, where the initial estimates may be more accurate for the simple reason that these markets lie in some sense around the mean of the data, the BBL estimates are much more similar to those produced by both PPHI and the other estimators.

# References


[1] Berry, Steven T. (1994), "Estimating Discrete-Choice Models of Product Differentiation", *RAND Journal of Economics*, 25(2), 242-262

[2] Berry, Steven T., James Levinsohn and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, 63(4), 841-890

[3] Berry, Steven T. and Joel Waldfogel (1999), "Free Entry and Social Inefficiency in Radio Broadcasting", *RAND Journal of Economics*, 30(2), 397-420

[4] Dube, Jean-Pierre H. and Jeremy T. Fox and Che-Lin Su (2011), "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients", *Econometrica*, forthcoming

[5] Nevo, Aviv (2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand", *Journal of Economics and Management Strategy*, 9(4), 513-548

[6] Nevo, Aviv (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry", *Econometrica*, 69(2), 307-342

[7] Petrin, Amil (2002), "Quantifying the Benefits of New Products: The Case of the Minivan", *Journal of Political Economy*, 110(4), 705-729

[8] Srisuma, Sorawoot (2010), "Estimation of Structural Optimization Models: A Note on Identification", mimeo, London School of Economics

[9] Sweeting, Andrew T. (2011), "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry", mimeo, Duke University

**Table 2: First Stage Multinomial Logit Model of Format Choice**
(standard errors in parentheses)

| (a) Coefficients on Switch Characteristics and Interactions | | | |
|---|---|---|---|
| | | *** Firm's Revenue** | *** Number of Stations Owned By Firm** |
| Switch to An Active Format | -2.000 | 0.021 | 0.132 |
| | (4.788) | (0.034) | (0.043) |
| * Revenue of Moving Station | 0.134 | 0.012 | -0.105 |
| | (0.157) | (0.009) | (0.043) |
| * Market Revenue Per Point | -0.643 | -0.170 | 1.162 |
| | (2.409) | (0.143) | (0.802) |
| Moving Station Also Moved Last Period | -0.134 | 0.384 | -0.414 |
| | (0.708) | (0.216) | (0.274) |
| * Market Revenue Per Point | 9.061 | -0.879 | -3.446 |
| | (5.479) | (0.812) | (3.959) |
| Switch From Dark Format | -4.628 | -1.518 | 1.634 |
| | (8.436) | (1.077) | (0.730) |
| * Market Revenue Per Point | 6.256 | 8.740 | -8.090 |
| | (9.821) | (6.073) | (5.173) |
| Switch To Dark Format | -0.051 | -0.633 | -0.240 |
| | (8.366) | (1.359) | (0.707) |
| * Revenue of Moving Station | -4.921 | -0.244 | 0.771 |
| | (4.468) | (3.054) | (2.168) |
| * Market Revenue Per Point | -9.608 | -2.421 | 8.865 |
| | (10.131) | (10.427) | (6.901) |
| Count of Stations in the Same Format | 0.471 | -0.062 | -0.117 |
| After Switch | (0.366) | (0.037) | (0.073) |
| * Market Revenue Per Point | 0.083 | -0.044 | 1.353 |
| | (2.938) | (0.111) | (0.883) |

**Table 2 cont.: First Stage Multinomial Logit Model of Format Choice**
**(b) Format Coefficients and Demograpghic Interactions: For Format Station Would Move From**

| | Dummy | * AM Station | * Prop Black | * Prop Hispanic | * Prop 12-24 | * Prop 50+ | * Black Growth/ White Growth | * Hispanic Growth/ White Growth |
|---|---|---|---|---|---|---|---|---|
| AC/CHR | - | -0.340 | 3.812 | 2.665 | -8.782 | -10.899 | -0.031 | 0.010 |
| | | (1.160) | (3.019) | (2.788) | (18.396) | (12.949) | (0.014) | (0.005) |
| Country | -4.798 | 1.744 | 5.926 | 2.779 | 9.073 | -10.711 | -0.027 | 0.010 |
| | (5.014) | (0.647) | (3.253) | (2.794) | (18.456) | (13.584) | (0.015) | (0.005) |
| Rock | -1.200 | 4.294 | 5.809 | 3.252 | -2.766 | -13.019 | -0.023 | 0.008 |
| | (4.929) | (1.106) | (3.168) | (2.900) | (19.412) | (13.440) | (0.014) | (0.005) |
| Urban | -7.568 | 0.361 | -3.877 | -0.316 | 11.583 | 1.666 | 0.038 | -0.013 |
| | (5.841) | (0.632) | (3.761) | (3.382) | (19.817) | (14.028) | (0.038) | (0.013) |
| News | -5.397 | -0.412 | 0.972 | 3.136 | 6.151 | -3.611 | -0.029 | 0.010 |
| | (5.367) | (0.692) | (3.728) | (3.060) | (19.239) | (13.718) | (0.015) | (0.005) |
| Other Programming | -1.527 | 1.021 | 2.798 | 0.673 | 5.376 | -14.540 | -0.027 | 0.009 |
| | (4.086) | (0.582) | (2.921) | (2.640) | (17.344) | (12.615) | (0.014) | (0.005) |
| Spanish | -0.621 | 1.489 | -1.431 | -6.034 | -4.229 | -9.759 | -0.052 | 0.000 |
| | (7.748) | (0.831) | (5.477) | (3.775) | (24.437) | (15.085) | (0.021) | (0.009) |

**Table 2 cont.: First Stage Multinomial Logit Model of Format Choice**
**(c) Format Coefficients and Demograpghic Interactions: For Format Station Would Move To**

| | Dummy | * AM Station | * Prop Black | * Prop Hispanic | * Prop 12-24 | * Prop 50+ | * Black Growth/ White Growth | * Hispanic Growth/ White Growth |
|---|---|---|---|---|---|---|---|---|
| AC/CHR | - | -4.359 | -5.344 | -0.477 | 0.661 | 10.540 | 0.022 | -0.007 |
| | | (1.135) | (3.166) | (2.916) | (18.023) | (13.108) | (0.016) | (0.006) |
| Country | -3.543 | -1.192 | -0.826 | 0.089 | 0.982 | 15.452 | 0.028 | -0.009 |
| | (5.432) | (0.655) | (3.294) | (3.049) | (19.666) | (13.584) | (0.015) | (0.005) |
| Rock | 5.746 | -1.357 | 9.824 | 2.659 | -27.486 | 2.117 | 0.026 | -0.008 |
| | (5.433) | (0.666) | (3.457) | (3.169) | (20.575) | (13.499) | (0.019) | (0.007) |
| Urban | 1.019 | 1.696 | -2.092 | -1.223 | -4.628 | 5.141 | 0.011 | -0.001 |
| | (4.466) | (0.574) | (3.031) | (2.705) | (18.023) | (12.714) | (0.015) | (0.005) |
| News | -2.131 | -0.363 | -3.853 | -4.094 | 6.715 | 10.327 | 0.030 | -0.009 |
| | (4.750) | (0.603) | (3.167) | (3.142) | (18.262) | (13.201) | (0.014) | (0.005) |
| Other Programming | 0.519 | -1.600 | -4.248 | 8.348 | -3.824 | 4.639 | 0.030 | -0.009 |
| | (5.073) | (0.668) | (3.751) | (2.935) | (19.043) | (13.558) | (0.014) | (0.005) |
| Spanish | 1.886 | -3.487 | -8.583 | -4.041 | 0.369 | 6.097 | 0.024 | -0.008 |
| | (4.382) | (0.886) | (3.192) | (3.072) | (17.905) | (13.038) | (0.015) | (0.005) |

**Table 2 cont.: First Stage Multinomial Logit Model of Format Choice**
**(d) Characteristics of Moving Station and Measures of Competition**

| | | | Format Station | |
|---|---|---|---|---|
| **Characteristics of Station Being Moved** | | **Measures of Competition** | **Moving From** | **Moving To** |
| exp(xi) | -0.187 | Number of Other Stations in Format | -0.098 | 0.028 |
| | (0.100) | | (0.069) | (0.066) |
| exp(fixed quality component) | -0.090 | Number of Other Stations in Format that | 0.097 | -0.064 |
| | (0.035) | Have the Same Owner | (0.054) | (0.048) |
| exp(xi)*exp(fixed quality component) | -0.051 | Sum of exp(xi) of Other Stations | 0.048 | -0.006 |
| | (0.034) | | (0.032) | (0.031) |
| exp(xi)*AM dummy | -0.074 | Sum of exp(fixed quality component) | 0.026 | -0.007 |
| | (0.072) | | (0.009) | (0.009) |
| | | Sum of Mean Revenue Measure of | -0.049 | -0.020 |
| | | Other Stations | (0.028) | (0.027) |

**Log-Likelihood: -2170.5, Observations 6,025**