WILEY

Free Entry and Social Inefficiency in Radio Broadcasting
Author(s): Steven T. Berry and Joel Waldfogel
Source: *The RAND Journal of Economics*, Vol. 30, No. 3 (Autumn, 1999), pp. 397-420
Published by: Wiley on behalf of RAND Corporation
Stable URL: https://www.jstor.org/stable/2556055
Accessed: 22-12-2019 15:47 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

RAND Corporation, Wiley are collaborating with JSTOR to digitize, preserve and extend access to *The RAND Journal of Economics*

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

RAND Journal of Economics
Vol. 30, No. 3, Autumn 1999
pp. 397–420

Free entry and social inefficiency in radio broadcasting

Steven T. Berry*

and

Joel Waldfogel**

In theory, free entry can lead to social inefficiency. We study the radio industry in a first attempt to quantify this inefficiency. Using cross-sectional data on advertising prices, the number of stations, and radio listening, we estimate the parameters of listeners' decisions and of firms' profits. Relative to the social optimum, our estimates imply that the welfare loss (to firms and advertisers) of free entry is 45% of revenue. However, the free entry equilibrium would be optimal if the marginal value of programming to listeners were about three times the value of marginal listeners to advertisers.

1. Introduction

■ It is now well known, at least in theory, that free entry can lead to social inefficiency. See Chamberlin (1933), Spence (1976a, 1976b), Dixit and Stiglitz (1977), Mankiw and Whinston (1986), Sutton (1991), and Anderson, DePalma, and Nesterov (1995).

Excessive entry can result when two conditions hold: first, entrants' products are substitutes for existing firms' products, so that entry "steals business" from incumbents; second, average costs are decreasing in output. An extreme example, with perfect substitutes, fixed prices, and exclusively fixed costs, illustrates this clearly. A second entrant garners half of the market and halves the incumbent's output. Consumers derive no additional benefit from the new entrant's product, but resource use on fixed costs is now doubled, reducing social surplus. The logic of free entry dictates that firms enter as long as the private benefit accruing to an entrant exceeds fixed costs. When new products are substitutes for existing products, the business stolen from incumbents

*Yale University and NBER; steven.berry@yale.edu.

**University of Pennsylvania and NBER; waldfogj@wharton.upenn.edu.

We are grateful to Shelly Cagner of Arbitron for helping us get access to data and to James Duncan for helping us understand his data. We thank the Social Science Research Fund at Yale, as well as NSF grant no. SBR-9617887, for financial support. Seminar participants at Northwestern, Princeton, Yale, MIT, Chicago, UCLA, and the NBER helped to improve the article, while Severin Borenstein and Ariel Pakes provided especially helpful comments. We were also greatly aided by referees and by an Editor of this Journal. Andres Lederman and Jakub Rehor provided able research assistance.

Copyright © 1999, RAND

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

places a wedge between private and social benefits of entry. In general, the business-stealing effect can be offset if entry reduces prices or increases available product variety, so that entry can be either excessive or insufficient.[1]

Although these theoretical arguments have been advanced repeatedly in the past decades, we are aware of no empirical studies quantifying the inefficiency associated with free entry. This may be because the data needed to document such a finding are hard to obtain. To calculate the optimal number of firms in an industry, one needs information on revenues and costs. In particular, one needs to know how revenue per firm changes with entry. Recent studies of entry (for example, Bresnahan and Reiss (1987) and Berry (1992) are built around a more easily observed datum, the number of firms in a market.

Here we fill a gap in this literature with a study of entry into the U.S. commercial radio broadcasting industry. The idea that entry into radio markets may be inefficient goes back to Steiner (1952), who constructed examples with wasteful duplication. More recently, Rogers and Woodbury (1996) consider the correlation between diversity in station formats and the size of the radio audience; they find a weak correlation, consistent with our finding that new programming does not greatly increase total listening. Also, with a very different dataset, Borenstein (1986) considers a set of similar questions.

Two features of the radio industry suit it well for our study. First, detailed data on firms' listening shares and advertising revenue allow estimation of firms' revenue functions. Second, the radio industry is characterized by free entry (up to some technological limit) and high fixed costs.

Our model of the radio industry is simple: firm revenue equals a price (the annual advertising revenue per listener) times the number of listeners, while costs are fixed. The share of people listening to radio increases with available variety, which in turn increases with the number of stations available in the market. Thus, cities with large population can support more stations than small cities can. Availability of data on both listening and advertising prices allows us to estimate two functions associated with firm revenue, the listening share function and the inverse demand curve for advertising. The listening share function is derived from a nested logit model of consumers' listening decisions estimated from market-level data, as suggested by Berry (1994). The inverse demand curve for advertising gives advertisers' marginal willingness to pay for listeners as a function of the listening share. These functions allow us to estimate how firm revenue varies with entry. In the same spirit as Bresnahan and Reiss (1990) and Berry (1992), we infer the distribution of fixed costs from entry decisions. (See also Dranove, Shanley, and Simon (1992).) In contrast to earlier work, our fixed-cost estimates employ an explicit revenue function that is identified without the entry model. The entry model is based on the equilibrium condition that we will observe N firms in a market if and only if N firms are profitable while $N + 1$ firms are not.

Like earlier work, we employ the simplifying assumption that firms are symmetric, so that postentry profits depend only on the number of firms. To fully relax this assumption we would need to allow stations to choose their characteristics (e.g., format and quality), which would significantly complicate the entry model. However, we do consider empirically how our estimates of the business-stealing effect are affected when we add station heterogenity to the listening model.

[1] The term "business stealing" is well established in the literature and so we use it; however, many seminar participants have objected to what they see as an implication that any firm "owns" the output. This implication, of course, comes from the term and not from the theory.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

Using the estimated revenue functions and fixed-cost distributions, we can calculate the number of firms under free entry and monopoly. Given a notion of social welfare, we can also calculate the socially optimal number of firms. In most of this study we consider only the welfare of the paying customers, that is, the advertisers. We defer until the end of the article the issue of the externality created by the production of advertising, the value of programming to listeners. Because programming is an unpriced good, we have no data on its value, so we calculate the implied value of programming to listeners that would render observed entry optimal. We calculate social welfare under both free entry, monopoly, and optimality. The difference between welfare at the social optimum and welfare under free entry is a measure of the inefficiency of free entry into radio broadcasting.

The plan of the article is as follows. Section 2 reviews the standard theory of entry into oligopolistic markets. Section 3 describes the data used in the study and documents some relationships in the data. Section 4 adapts the standard theory to radio broadcasting and presents our econometric specification. Section 5 presents estimation results. These results depend on some strong assumptions, so Section 7 presents some robustness analysis. In particular, the "symmetry" assumption of the basic model is relaxed. Section 6 provides estimates of the social inefficiency of free entry into radio broadcasting and also estimates of the value of programming to listeners needed to render free entry optimal. The conclusion provides some directions for future research.

2. Theory of entry

■ We model radio broadcasting as a homogeneous-goods industry, where the product is listeners who are "sold" to advertisers. In this section we review the standard theory of entry into such markets. (The empirical section below specifies particular functional forms that are appropriate for this industry and our data.) We focus on a traditional version of the theory, with symmetric firms. Later we shall discuss the robustness of our results to some more realistic assumptions.

The production process in broadcasting is unusual in that the primary inputs, listeners, are not purchased by the firm but rather make a free choice about listening to radio. Listeners' choices result in the share, $s(N)$, of the population listening to a given station as a function of the number of entering stations N. Total listening to radio is then $Ns(N)$, which we also term $S(N)$. The price, p, of advertising (revenue per listener) is assumed to decline in the total listening share:

$$p(N) = p(Ns(N)).$$

Two assumptions are implicit in this formulation. First, our treatment of demand implicitly assumes a variant of the Cournot model: given the number of listeners "produced," price is determined by the market demand curve. Our approach deviates from the usual Cournot model in that output is determined by listener behavior rather than a traditional production function. Second, we model price as a function of listening share, rather than total listeners. Our specification is consistent with an explicit model of advertiser behavior in which the number of advertisers varies proportionately with market size M. We also assume that there is a fixed cost, F, of setting up a radio station and that the costs of a station do not vary with the number of listeners. Given our homogeneous-goods treatment of advertising demand, the entry problem is exactly that of Mankiw and Whinston (1986). In a free entry equilibrium, firms enter until profits are driven to zero, with profits given by

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

$$\pi(N) = Mp(N)s(N) - F, \tag{1}$$

where M denotes market size. Formally, given the integer constraint on N, the number of firms under free entry, N_e, satisfies the condition

$$\pi(N_e) \geq 0 \quad \text{and} \quad \pi(N_e + 1) < 0. \tag{2}$$

The total benefit to market participants is the total benefit to advertisers (which is split between radio-station revenue and advertisers' surplus) minus the costs of operating the stations. Initially, we restrict our examination of social benefit to those benefits captured by market participants, and we ignore the important (unpriced) externality captured by listeners.

Assuming that a social planner cannot set the price of advertising but can only control entry, the planning problem is to choose N to maximize social welfare,

$$M \int_0^{Ns(N)} p(x)\, dx - NF. \tag{3}$$

The first-order condition for this problem is

$$Mp(N)\left[s + N\frac{\partial s}{\partial N}\right] - F = 0, \tag{4}$$

or

$$\pi(N) + MNp(N)\frac{\partial s}{\partial N} = 0. \tag{5}$$

The second term in this expression is negative as long as per-station listening share declines in N. If the free-entry number of firms, N_e, sets profits exactly to zero, then marginal social welfare is negative at N_e. Thus, we expect the free-entry number of firms to exceed the social optimum. Given the integer constraint, Mankiw and Whinston note that the free-entry number of firms must exceed the socially optimum number minus one. The goal of our empirical work is to quantify the extent of this excess entry and its effect on welfare. As an alternative entry model, we might consider the entry decisions of a monopolist who controls all radio entry. The objective function is then to maximize

$$N\pi(N) = R(N) - NF.$$

Note that the monopolist internalizes the business-stealing effect. However, it is easy to see that the monopolist values increases in output less than the social-welfare-maximizing planner, who values the inframarginal benefit to advertisers of the reduction in price caused by new entry. Thus, the monopolist will choose a smaller N than is socially optimal.

3. Data

■ The data in this study are from a cross section of U.S. metropolitan radio markets. The endogenous variables are the number of firms located inside the metropolitan areas,

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

the share of population listening to radio, and the price of advertising. Exogenous variables include demographic characteristics of the metropolitan areas, notably the level of population. In some specifications we also treat the number of firms broadcasting from outside the metropolitan area as exogenous.

□ **Sources.** Data for this study come from two sources, Arbitron (1993) and Duncan (1994). Arbitron provides data on the number of listeners of each commercial radio station in each major U.S. market in spring 1993, as well as whether the station broadcasts from within a metropolitan area. The Arbitron data are generated from listening diaries submitted by compensated survey participants. Arbitron receives listening diaries from roughly 1 in 500 persons in the markets (proportionately fewer in large markets, for example, 1 in 1,800 in New York). The spring Arbitron survey includes data on every commercial radio station with positive reported listening in over 260 metropolitan statistical areas (MSAs). Duncan (1994) reports aggregate 1993 advertising revenue for the top stations in each of 135 MSAs. Most major radio stations share their annual revenue figures with one of two accounting firms that serve the radio industry. Duncan (1994) reports the sum of the participating firms' revenue figures in each MSA and identifies which firms report. This allows us to calculate annual advertising revenue per listener in each MSA based on data for participating firms. Participating firms typically account for over three-quarters of total listening. We use this information to calculate annual advertising revenue per listener, which is our measure of advertising price, p. The data on city characteristics, including total population and other demographics, are from Arbitron (1993), derived from U.S. Census figures.

The Arbitron listening figure we use is the average quarter-hour rating (AQH). A station's AQH shows the number of persons listening to it for at least five minutes during a quarter-hour, averaged over quarter-hours throughout the week (Monday–Sunday, 6:00 AM to midnight). At the city level we calculate the share of population listening to stations broadcasting from inside the metropolitan area S_1, the share listening to stations broadcasting from outside the metro area S_2, the number of commercial stations inside the metro N_1, and the number outside the metro but received inside the metro N_2. The dataset used in the study includes information on 3,285 stations in 135 markets (2,509 inside and 776 outside).[2] Because our entry model assumes postentry symmetry of inside-metro firms, the listening share of each inside firm is simply $s_1 = S_1/N_1$.

Table 1 reports means, standard deviations, minima, and maxima of the market-level variables for the 135 metro areas included in the estimates. During an average 15-minute period, 14.4% of the population listens to at least five minutes of radio. Unlike Arbitron's reported total market AQH figures, which include both public and very small commercial stations, our total AQH listening includes only commercial stations attracting enough listeners to be listed independently in Arbitron.[3] The vast majority of the listening (12.9 percentage points of the 14.4%) is to stations broadcasting from inside the metro. The population of MSAs in the sample ranges from 133,000 to over 14 million. The number of inside-the-metro commercial radio stations varies from 6 to 47, with an average of 18.6. MSAs in the sample have an average annual household income of $36,000 and an average fraction having some college of 47%. Finally, annual advertising revenue per listener averages $277. Average annual

[2] An observation is a city-station, so that a city that is heard in multiple markets can appear in the dataset more than once, although only once as an inside-the-metro station.

[3] Over 95% of listening is to commercial stations. In Berry and Waldfogel (1999) we consider some data on public radio.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

TABLE 1 Description of City-Level Data

Variable	Units	Mean	Standard Deviation	Minimum	Census Population Survey
Share in-metro	%	12.909	12.909	5.172	17.841
Share out-metro	%	1.536	1.536	.000	10.422
N_1 (in-metro)	integer	18.585	18.585	6.000	47.000
N_2 (out-metro)	integer	5.748	5.748	.000	28.000
Population	millions	1.070	1.070	.133	14.034
Ad price	$100	2.766	2.766	1.466	6.213
Income	$1,000	35.531	35.531	21.860	51.936
College	%	46.969	46.969	28.300	65.100

To scale coefficients, the income and college variables are divided by 10 in the empirical work and Ad Price is per AQH listener-year.

revenue per market is $37.8 million, which sums to $5.1 billion for all 135 markets in the sample.

While our entry model will assume symmetry, our model of listening behavior can incorporate station heterogeneity. Table 2 includes a list of station-level variables and their means. The variables include measures of the signal quality of the station as well as a description of the station programming (i.e., the "format" of the station). Signal quality is a function of wattage and tower height and tower location. We have tower height data only for FM stations, while our only measure of tower location is a dummy variable for outside-the-metro stations. About two-thirds of stations broadcast on the FM band and the average station broadcasts at about 38,000 watts.

Duncan (1994) classifies stations according to 15 industry-standard formats. The last part of Table 2 gives the means of dummy variables that are equal to one if a station is classified in the listed format. The formats are a bit odd in places; for example, Spanish describes the language of the station, not the content, and Black presumably also characterizes the listeners, not the content. Across all stations, the most common of the 15 commercial formats is country music, followed by album-oriented rock, news/talk, and adult contemporary. The omitted category, classical, is the least popular. Other uncommon formats are jazz and "classic album-oriented rock."

□ **Relationships in the data.** The basic question we seek to answer with listening data is whether the share of population listening to a format grows as stations enter the market. Do stations simply split a pie (business stealing), or do they add listeners (market expansion)? If stations are identical, then listeners will be indifferent between stations, and entry will not expand the market. Instead, a new entrant simply steals business from the incumbent(s). On the other hand, if stations are differentiated, then an entrant will draw listeners from both market expansion and business stealing.

How does the share of population listening to radio vary across cities with the number of stations available in those cities? Figure 1 shows that although the listening share increases with the number of stations, it increases rather slowly. The flat relationship between listening and stations depicted in Figure 1 offers suggestive evidence that the effect of a marginal entrant on listening is small, at least in markets with many incumbents. However, we cannot infer the causal link between the number of stations

© RAND 1999.

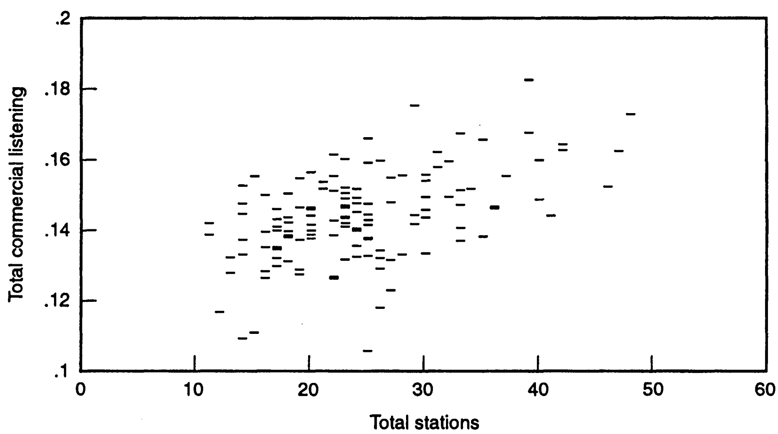
This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

TABLE 2 Station-Level Variables

Variable	Mean
FM	.667
Mega watt	.038
Tower height*FM	.608
Out metro	.221
Country	.146
Album rock (AOR)	.107
News/Talk	.096
Adult contemp	.091
Top 40	.083
Soft adult contemp	.073
Other	.070
Black	.065
Religious	.064
Big band	.048
Full service	.046
Spanish	.042
Classic AOR	.028
Unknown	.021
Jazz	.010

FIGURE 1

STATIONS AND LISTENING BY MARKET



© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

and the share of listeners from such a figure. The number of stations in a market is endogenously determined by the willingness of people in the market to listen to such programming. Hence, we need stronger techniques to allow us to infer the effect of entry on listening.

Figures 2 and 3 show the relationships among some other important variables relevant to our modelling exercise. Figure 2, which shows positive relationships between population and stations (both total and in-metro), demonstrates that markets with more people can support more stations. Figure 3 documents a negative relationship between the in-metro listening share and the advertising price. This is consistent with advertisers having a downward-sloping demand curve for listeners. The relationships in Figures 1–3 help motivate our modelling strategy, but an explicit model is needed to guide our interpretation of the relationships among variables.

4. Econometric specification

■ To estimate the model, we specify a process that generates the data, including functional forms for the listening share function, the advertising demand function, and the distribution of fixed costs. Our discussion will specify an error structure for each equation as well as a set of exogenous data.

□ **The listening share function.** We use a simple discrete-choice formulation for listeners' choices. Each person in the market chooses among a set of choices that includes each station in the market and also includes the choice of not listening. Each additional station brings some unique benefit to consumers, but it may also steal listening from existing stations. We use a nested logit utility function to parameterize the degree to which stations offer unique, as opposed to redundant, programming.

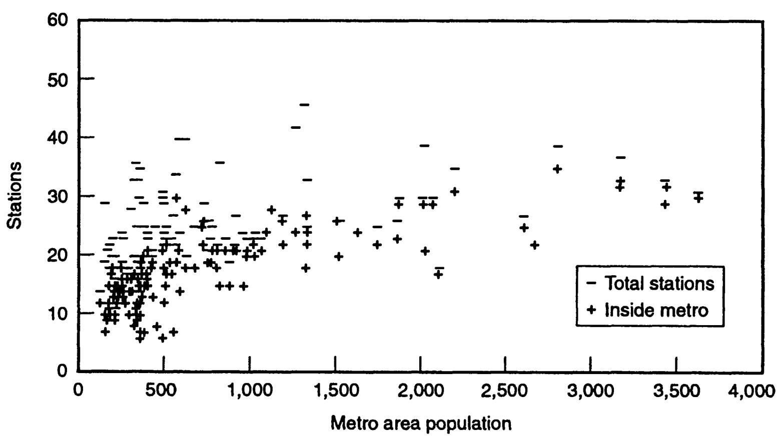
The utility of potential listener i for station j is chosen to give a "nested logit" functional form for utility:

$$u_{ij} = \delta_j + v_i(\sigma) + (1 - \sigma)\epsilon_{ij}, \tag{6}$$

where δ_j is the mean utility of listening to station j and ϵ_{ij}, an identical and independently distributed extreme value deviate, is the idiosyncratic benefit of this station for

FIGURE 2

POPULATION AND STATIONS BY MARKET

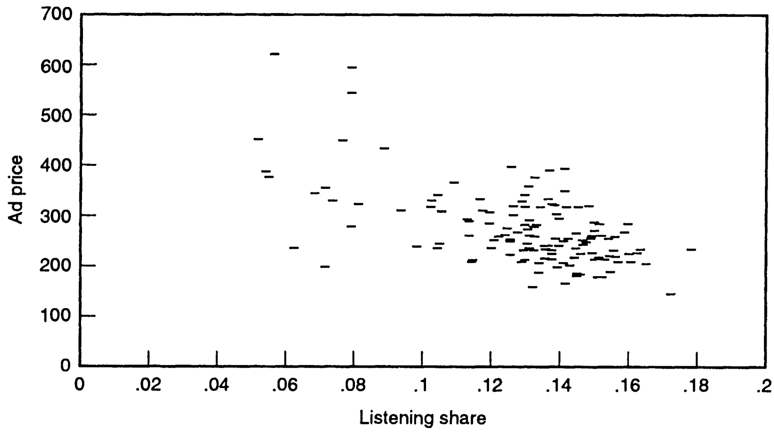


© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

FIGURE 3

IN-METRO LISTENING SHARE AND AD PRICE



this person. As the parameter σ goes to one, stations become identical and the only random term in the utility function is v, which is constant across all stations. The common term v has a distribution, derived in Cardell (1997), that goes to zero as σ goes to zero. Therefore, when σ is zero, we return to the logit model and stations give completely idiosyncratic benefits. To complete the specification, the utility of not listening is random and given by the extreme value deviate ϵ_{i0}.

In the nested logit specification, the term δ_j then captures the average benefit of listening, as opposed to not listening, while σ parameterizes the business-stealing effect. As σ goes to one, the business-stealing effect is complete and an additional station does not increase total listening. For smaller values of σ, total listenership increases in the number of stations, with a maximum rate of increase when σ is zero. This can be seen from the nested logit choice probabilities. For station j, the share of the market listening is

$$s_j(\delta, \sigma) = \frac{e^{\delta_j/(1-\sigma)}}{D} \frac{D^{1-\sigma}}{(1 + D^{1-\sigma})}, \tag{7}$$

where

$$D \equiv \sum_j e^{\delta_j/(1-\sigma)}.$$

The first term in the expression for $s_j(\delta, \sigma)$ is the share of station j as a fraction of total radio listening, while the second expression is the total radio listening share.

To maintain the tractability of the entry model, our simplest assumption is that all stations in a given market have identical mean utility levels and therefore identical postentry market shares. While this assumption is necessary for the tractability of the entry model, it is not necessary to the estimation of the listening model. Therefore, we will also present results of the listening model under the more realistic assumption that stations are differentiated by observed and unobserved characteristics.

To estimate the model, we parameterize the mean utility of station j in market k as

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

$$\delta_{jk} = x_{jk}\beta + \xi_{jk}. \tag{8}$$

The term x_{jk} is a vector of observed market (and, possibly, station) characteristics, β is a vector to be estimated, and ξ_{jk} is an unobservable assumed to be mean independent of the exogenous data. To estimate the parameters, β and σ, of listener utility, we follow the method in Berry (1994) exactly. Given the observed station market share, $s_{jk}(N)$, we can invert the nested logit market-share function to solve for each δ_{jk} as a function of the parameter σ. The unobservable ξ_{jk} is then defined as

$$\xi_{jk}(\delta, \sigma) = \delta_{jk}(\sigma) - x_{jk}\beta.$$

Given a vector of exogenous data in each market, z_k, a method of moments estimator can then be formed from the moment conditions:

$$E[\xi_{jk}(\delta, \sigma) \,|\, z_k] = 0. \tag{9}$$

The exogenous data consist of the x vector, plus population, and the number of stations broadcasting from outside the metro. Berry (1994) shows that δ is linear in s, so two-stage least squares, for example, can be used as the method of moments estimator.

In the symmetric case that we carry over to the entry model, the j subscript is not necessary as all stations in a market are identical. In this case (with $\delta_j \equiv \delta$), the listening share of a station is just a function of σ, δ, and the number of entering stations, N:

$$s_j(N, \delta, \sigma) = \frac{1}{N}\frac{N^{1-\sigma}}{e^{-\delta} + N^{1-\sigma}}. \tag{10}$$

The same method from Berry (1994) still works to estimate the parameters β and σ, but now we solve for only one δ in each market, as a function of S and N.[4]

□ **The demand for listeners.** Given a number of listeners, a station then produces revenue by "selling" these listeners to advertisers. We assume that there is a fixed number of advertising minutes sold per hour and that the price of a single advertisement sold by a station is proportional to its number of listeners. Consequently, total revenue of a station is the market ad price per listener times the average number of listeners.

There are interesting problems about the endogeneity of the number of advertisements per hour and about different advertising rates for different demographic groups that, for lack of data, we do not address here. These problems could change the welfare results, and we consider some possible implications at the end of Section 7.

We allow advertisers' marginal willingness-to-pay for listeners to decline in the share of population listening to radio. In the empirical work, we adopt a simple constant elasticity specification for the inverse advertising demand curve:

$$p = \alpha(S(N))^{-\eta}, \tag{11}$$

where η is the inverse elasticity of demand, α is a parameter that shifts demand, and $S(N)$ is the share of population listening to radio.

[4] In fact, in the empirical work we do allow for some station heterogeneity even in the simplest cases, so we never use equation (10) directly in the estimation.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

We assume that the demand parameter α is a function of observed demand shifters x_k and an unobserved error ω_k. Assuming that

$$\ln(\alpha) = x_k\gamma + \omega_k, \tag{12}$$

the inverse demand curve for advertising is defined by

$$\ln(p_k) = x_k\gamma - \eta \ln(S_k) + \omega_k, \tag{13}$$

where $S_k = N_k s_k(N_k)$ is the observed total listening share of radio in market k, and γ and η are parameters to be estimated. The term ω_k is an unobservable shock to advertising prices that is assumed to be mean independent of the exogenous data:

$$E[\omega_k(\gamma, \eta) | z_k] = 0. \tag{14}$$

The endogenous data are p and S_k, while the exogenous data z_k include instruments for the endogenous variable S_k. Again, given the linearity of our functional form, this equation may be estimated via instrumental-variables methods, such as two-stage least squares. The instruments are the same as in the estimation of listening demand.

□ **The distribution of fixed costs.** The last primitive of the model is the distribution of fixed costs, which helps to determine the observed number of firms. We estimate the distribution of fixed costs from the entry behavior of firms.[5]

Because entry is discrete, the estimation problem is more difficult and we have to impose stronger assumptions. In particular, when estimating fixed costs we use a fully symmetric version of the model: listeners are divided equally among stations, and fixed costs are equal across stations. Modelling station heterogeneity in fixed costs would require a major methodological advance that is beyond the scope of this article. In Section 7 we discuss possible biases from the symmetry assumption.

For the empirical specification, we assume that fixed costs are

$$\ln(F_k) = x_k\mu + \lambda v_k, \tag{15}$$

where v_k is distributed standard normal, while μ and λ are parameters to be estimated. That is, fixed costs are assumed to have a log-normal distribution. As in Bresnahan and Reiss (1990), the assumption that fixed costs are equal for all firms leads naturally to the "ordered probit" likelihood function. Unlike Bresnahan and Reiss (1990), however, we have data on postentry outcomes and need only estimate the parameters of fixed costs from the ordered probit (as opposed to estimating all the parameters of the profit function from the ordered probit).

A number of firms equal to N_k is observed in equilibrium if and only if fixed costs are such that N_k stations make a profit but $N_k + 1$ stations would not. That is, F_k must fall between the value of variable profits given N_k stations,

$$v_k(N) = Mp(N)s(N),$$

and the value of variable profits given N_{k+1} stations. The term $v_k(N_k)$ is just the product

[5] The notion of estimating cost parameters from behavior, as opposed to cost data, goes back at least to the earliest supply and demand analyses, which infer marginal-cost parameters from prices. For another early example, see Rosse (1970). The practice is now commonplace.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

of observed population, advertising price, and per-station market share. The term $v_k(N_{k+1})$ is easy to calculate given our functional form assumptions.

The likelihood function is then

$$L(\theta) = \Phi\left(\frac{\ln(M_k p_k(N_k) s_k(N_k)) - x_k \mu}{\lambda}\right) - \Phi\left(\frac{\ln(M_k p_k(N_k + 1) s_k(N_k + 1)) - x_k \mu}{\lambda}\right), \quad (16)$$

where Φ is the standard normal cumulative density function. Our estimates of the parameters of fixed cost maximize the log-likelihood function of the data, conditional on our earlier estimates of the parameters of variable profit. The estimation error in the listening and advertising price parameters will cause the usual maximum-likelihood standard errors to be incorrect. For this reason, we also estimate all the parameters of the problem jointly.

□ **Joint estimation.** Each of the estimation problems above can be thought of as method of moments. The moments associated with the listening and ad price equations are formed from the interactions of the equation "errors" and the exogenous data. The maximum-likelihood estimation first-order conditions (with respect to the parameters of the distribution of fixed costs) are also properly thought of as moment conditions.[6] To estimate the parameters of all three equations jointly, we simply stack the three sets of moment conditions and estimate by generalized method of moments. The vector of sample moment conditions, as a function of all the parameters, is

$$g(\theta) = \sum_k \begin{pmatrix} \xi_k(\beta, \sigma) z_k \\ \omega_k(\gamma, \eta) z_k \\ \partial \ln(L_k(\theta)) / \partial(\mu, \lambda) \end{pmatrix}. \quad (17)$$

As we will see in the empirical section, equation-by-equation and joint estimation give nearly identical results. However, the standard errors of the joint estimates are correct.

5. Results from the basic model

■ In this section we present parameter estimates of the simplest model and discuss the fit of those estimates to the data. The following subsections will discuss the implications of these parameters and the robustness of those implications to some extensions of the model.

□ **Parameter estimates.** Table 3 reports estimated parameters. Columns 1 and 2 report listening share parameters from single-equation specifications of the listening share function. In addition to OLS estimates (column 1), we report two-stage least squares (TSLS) results using metro population and stations outside the metro (N_2) as instruments for N_1 as well (column 2). We obtain very similar results using only population as an instrument. All of the parameter estimates—including the estimates of the important parameter σ—are quite similar across specifications. The per-station listening share is higher in each of the included regions (northeast, north central, and

[6] It has often been noted that MLE is equivalent to zeroing the expected value of the derivatives of the likelihood function with respect to the parameters. These expected values can then be thought of as the moments that generate a method of moments estimator.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

TABLE 3 Results from the Structural Model

Parameters	Share OLS	Share TSLS	Ad OLS	Ad TSLS	Entry	Full
Constant	−2.224	−2.335				−2.328
	(.104)	(.138)				(.077)
Northeast	.068	.080				.066
	(.032)	(.033)				(.023)
North central	.054	.070				.068
	(.028)	(.031)				(.023)
South	.032	.040				.034
	(.025)	(.026)				(.022)
Income	.023	.012				.008
	(.022)	(.024)				(.020)
College	−.029	−.025				−.030
	(.014)	(.015)				(.012)
σ	.847	.804				.794
	(.029)	(.045)				(.028)
Constant			3.667	3.728		3.781
			(.204)	(.226)		(.256)
Northeast			.043	.044		.047
			(.063)	(.063)		(.076)
North central			.184	.183		.181
			(.054)	(.054)		(.060)
South			.109	.107		.104
			(.051)	(.051)		(.048)
Income			.052	.055		.070
			(.043)	(.043)		(.048)
College			.095	.093		.087
			(.029)	(.029)		(.032)
η			.579	.550		.514
			(.067)	(.081)		(.115)
Constant					−.844	−.808
					(.199)	(.038)
Northeast					.245	.254
					(.067)	(.004)
North central					.405	.411
					(.069)	(.042)
South					.270	.270
					(.065)	(.022)
Income					.195	.189
					(.051)	(.011)
College					.083	.080
					(.037)	(.004)
Population					.630	.652
					(.021)	(.043)
λ					.224	.224
					(.015)	(.009)
R^2	.90	.90	.47	.47		

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

south) than in the excluded western region. There is weak evidence that per-station listening is higher in cities with higher per-capita income and lower in cities with a large fraction of college-educated residents. Estimates of σ are also robust to exclusion of market characteristics variables (region dummies, income, and percent college educated). The estimates of σ are all roughly .8, indicating that commercial stations are strongly substitutable for one another. The similarity of OLS and TSLS estimates of σ indicates that N_1 is determined largely by our instruments (population and outside-the-metro competition) and not by city-specific variation in tastes for radio that we do not observe.

To illustrate the meaning of our parameter estimate, consider a hypothetical market with one station attracting 10% of the population as listeners. If σ is one, the second station has no effect on overall listening. If σ is zero (the logit case), a second station increases total listening by over 80% (to 18.2% of the population). With $\sigma = .8$, the second station increases total listening only slightly, from 10% to 11.3%.

The estimate of σ will drive our results, so we are concerned with the robustness of this estimate to richer specifications. This is one topic of Section 7, which considers the effect of adding observed and unobserved station heterogeneity to the listening model. For now, we note that the estimate of σ appears to be quite robust to various changes in specification, including a much more careful model of station heterogeneity.[7]

Columns 3 and 4 of Table 3 report results of single-equation estimation of advertising demand. The price of advertising per listener is highest in the north central region, followed by the south, the northeast, and the west. The advertising price is higher in high-income cities and in highly educated cities. In both specifications the demand for advertising is elastic, and the elasticity is roughly 1.82 (1/.550). Once again, instrumenting has little effect on the results.

Column 5 reports estimates of the fixed-cost parameters from the entry model estimated via the ordered probit holding the listening and advertising coefficients fixed at their estimated (TSLS) values. The ordered probit coefficients describe the mean of the distribution of log fixed costs. The mean of station fixed costs has the same ordinal pattern as the advertising price. It is highest in the north central region, followed by the south and northeast, then the west. Fixed costs are higher in high-income and high-education cities. We include population in the specification for fixed costs to reflect the possibility that inputs may be more costly in large cities, and fixed costs rise in population. Station fixed costs may also rise in population because of license rents. We explore this possibility below by reestimating the model excluding the largest 25 markets. According to industry analyst James Duncan, license scarcity would be likely to arise only in those markets. The parameter λ, the estimated standard error of log fixed costs, is .22. This turns out to imply that the variance in fixed costs conditional on city characteristics is about one-quarter of mean fixed costs. The last column of Table 3 reports results from joint GMM estimation of the entire model. These results are similar to the single-equation results, but the standard errors of the fixed-cost coefficients are corrected for the presence of other estimated coefficients. We concentrate on these results below.

□ **Fit.** To gauge how well our model fits the data, we compare the model's implied number of stations inside each market with the actual number. We calculate the model's implied N_1 by taking a large number of draws from the fixed-cost distribution for each market. For each draw we solve for the highest integer number of stations yielding positive profit. The model will predict the observed number of stations (as a free-entry

[7] See the results in Table 6, discussed below.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

equilibrium) only if fixed costs are drawn from the region of the fixed-cost distribution consistent with observed N_1. We repeat this exercise 1,000 times. The correlation of model and actual N_1 is .780, implying an R^2 of .608. By contrast, OLS regression (which by design maximizes R^2) of N_1 on variables used in the study gives an R^2 of .735. The OLS regression parameters have no direct economic interpretation and cannot be used to calculate social welfare.

6. Policy simulations

■ We first discuss some implications for the welfare of the market participants only; that is, we consider only the value captured by the stations and the advertisers. We then consider what, if anything, can be said about the externality to listeners.

□ **The market participants.** Table 4 reports simulated values of the number of in-metro stations, costs, revenue, welfare, listening, and the advertising price for our sample. For all simulations, we draw from the part of the distribution of fixed costs consistent with the observed number of stations (recall that N stations in a market implies specific bounds on fixed costs). This ensures that the model's free-entry number of inside stations equals the actual number in every simulation. We choose this approach because our concern here is not with measuring fit, but rather with the contrast among free entry, monopoly, and social optimality.

Table 4 clearly indicates that free entry into radio is excessive when only the welfare of the market participants is considered. While there are 2,509 commercial stations in the 135 markets under free entry, the "socially optimal" number is 649 (with a standard error of 46). Compared with the current average of 18.6 inside stations per market, the social optimum has 4.8 (.31) inside stations per market. This is a reduction of 74% in the number of stations. Ignoring the value of programming to listeners—as commercial radio broadcasters naturally do—social welfare with the current (free entry) configuration of stations is $5.33 billion per year ($3.06 billion). With the optimal configuration of stations—again, ignoring listener welfare—social welfare is $7.64 billion per year ($3.04 billion), indicating that the deadweight loss of free

TABLE 4 Comparison of Free Entry, Optimality, and Monopoly

	Free Entry	Optimal	Monopoly
In-metro entry	2,509	649	341
		(46)	(55)
Aggregate costs ($ millions)	5,007	1,144	602
	(3)	(92)	(101)
Aggregate revenue ($ millions)	5,100	4,334	3,959
		(204)	(173)
Welfare ($ millions)	5,331	7,640	7,422
	(3,064)	(3,037)	(2,878)
Ad price	277	326	375
		(11)	(48)
Listening share (%)	12.91	9.28	7.53
		(.19)	(.50)

The free-entry numbers without standard errors are calculated directly from data. The difference between free entry and optimal welfare has a standard error of 167.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

entry into radio broadcasting in these markets is \$2.3 billion per year (\$167 million). This is about 45% of current free-entry revenue.

The main source of welfare improvement from moving to the social optimum is the reduction in station operating costs, from \$5.01 billion (\$3 million) per year under free entry to \$1.14 billion (\$92 million) in the social optimum. The direct benefit of this \$3.86 billion annual operating cost reduction is considerably offset, however, by the increase in prices paid by advertisers, from \$277 to \$326 (\$11) per listener annually. In principle, regulators could target the number of firms at the social optimum by two means. First, regulators could directly limit entry. Second, regulators could levy a tax on entry. Stations facing an entry tax equal to per-station profit under the social optimum (averaging about \$5 million per station annually across markets) would freely enter up to the optimal number of stations. Such a tax would collect about \$3 billion in annual revenue.

Table 4 also shows the consequences of monopoly. Interestingly, monopoly would generate outcomes closer to the social optimum than free entry does. A monopolist (in each market) would operate fewer stations than is socially optimal (341 as opposed to 649.) Social welfare and revenue under monopoly are rather close to those in the social optimum. Welfare under monopoly is only 2.9% below the welfare maximum.

To give some specifics on variation across markets, Table 5 reports simulation results for markets at the 10th, 25th, 50th, 75th, and 90th population percentiles. The table illustrates that the optimal number of inside stations increases little with population, while the free-entry number of inside stations is closer to being proportional to population. Consequently, the deadweight loss is larger in more-populous markets.

□ **The value of programming to listeners implied by free entry.** Up to this point we have ignored the external benefit of programming to radio listeners. The social optimum calculated above ignores listeners' valuation of radio programming. Here we ask: What listener valuation of an additional station renders free entry optimal? To answer this question we augment the welfare maximand of Section 2 to include a component reflecting listeners' valuation of radio programming. We denote by $L(N)$ the per-capita benefit to listeners as a function of the number of stations, giving total welfare of

$$M \int_0^{Ns(N)} p(x)\, dx - NF + ML(N) = W(N) + ML(N). \tag{18}$$

We solve the associated first-order condition, evaluated at the free entry N, for the value of the marginal station to listeners that implicitly renders free entry optimal:[8]

$$M\frac{\partial L}{\partial N} = -\frac{\partial W}{\partial N}. \tag{19}$$

We can then calculate the implied value of a marginal station in each market required to render free entry optimal.

Note that advertisers value increases in listening share, S. Listening share is a monotonic function of N, so we can also think of the value to listeners as a function of S,

[8] Waldfogel (1993) performs a similar exercise, inferring welfare weights underlying policy, in the context of criminal sentencing.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

TABLE 5 **Simulation Results for Selected Markets**

	Rockford	Jackson	Toledo	Charlotte	San Diego
Description of city					
Population (millions)	.2	.3	.5	1.0	2.2
Population percentile	10	25	50	75	90
Outside stations	11	0	8	4	4
Number of in-metro stations					
Free entry	9	17	15	20	31
Optimal	4	3	5	5	9
	(.3)	(.5)	(.5)	(.4)	(.5)
% In-metro listening					
Free entry	11.9	13.0	12.5	12.7	13.1
Optimal	8.7	9.5	8.6	8.9	9.5
	(.3)	(.3)	(.3)	(.2)	(.2)
Revenue ($ millions)					
Free entry	7.5	12.2	16.2	39.8	85.1
Optimal	6.4	10.4	13.4	33.4	72.6
	(.4)	(.5)	(.7)	(1.6)	(3.5)
Costs ($ millions)					
Free entry	7.2	11.9	15.7	38.9	83.9
	(.1)	(.0)	(.0)	(.0)	(.0)
Optimal	3.2	2.1	5.2	9.7	24.4
	(.2)	(.3)	(.5)	(.9)	(1.4)
Welfare ($ millions)					
Free entry	8.1	12.8	17.0	41.7	88.6
	(4.5)	(7.3)	(9.7)	(23.8)	(50.8)
Optimal	9.9	19.1	21.9	58.0	122.8
	(4.5)	(7.2)	(9.5)	(23.5)	(50.3)
Ad price ($/listener-year)					
Free entry	286.9	282.8	252.1	303.6	292.9
Optimal	336.1	331.6	305.8	363.3	344.8
	(11.9)	(13.9)	(10.3)	(13.1)	(13.1)
Implied $ value per listener-year	559.5	1,281.4	629.7	1,088.1	1,028.4
	(49.4)	(248.3)	(68.5)	(166.9)	(151.3)

Note: Standard errors are in parentheses. Numbers without standard errors are calculated directly from data.

$$L(N) = L(S(N)).$$

The marginal value to listeners of an increase in S is then

$$\frac{\partial L}{\partial S} = \frac{\partial L/\partial N}{\partial S/\partial N}. \tag{20}$$

This number is easy to calculate from the value of $\partial L/\partial N$ in (19), and we find an average value for $\partial L/\partial S$ of $893 ($146) per year of listening. In our data, a year of listening is eighteen hours per day for each day of the year, so $893 per listener-year is about 13.5 cents per hour of listening. Because marginal entry is more wasteful in larger cities, the implied value of programming to listeners that renders free entry optimal grows with population (see Table 5 for examples). We can compare our implied value of increases in listening to the average observed advertising price of $277 per

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

listener-year or about 4.2 cents per listener hour. Thus, to justify free entry, the external benefit of programming to listeners must be over three times the market value. Note that the external benefit of listening share has two components. First, an increase in N attracts new listeners to radio. Second, the new station allows existing listeners to switch to its possibly higher-valued programming. Our implied value of dL/dN, and thus our value of $\partial L/\partial S$, includes both of these benefits.

We have no direct evidence on the value that listeners in our sample attach to radio programming, so we leave it to other analysts to determine whether the actual value of programming to listeners renders free entry optimal in this context. We have some indirect evidence on (minimum) valuations of radio programming from European license fees. As of 1988, only Belgium and Switzerland offered separate radio licenses but did not offer combined radio and television licenses. The license fees for Belgian and Swiss radio use were \$23.2 and \$59.4 per household per year, respectively, or \$8.5 and \$22.8 per capita annually in the two countries. Almost half of Belgian and Swiss households purchased radio licenses in 1987 (European Broadasting Union, 1988).

7. Robustness: heterogeneity and endogenous fixed costs

■ The model estimated above corresponds closely to simple theoretical models in the literature, but it misses much of the richness of real-world radio markets. An important question is how our estimates would change in the presence of more realistic assumptions. Here, we examine two of our strongest assumptions: that stations are symmetric and fixed costs are exogenous. We provide some empirical discussion of the first assumption and some theoretical discussion of the second. We also briefly consider some other problems with the specification.

□ **Differentiated products.** Several readers of early versions of this article suggested that the shape of Figure 1 might be determined by station-level heterogeneity rather than by the business-stealing effect. One version of this argument is that stations might enter in approximate order of their popularity. Thus, as N increases, total market listening share increases little because the marginal station is not popular, rather than because of business stealing. Stations could differ because of format, with more popular formats (such as country music) entering earlier and less popular formats (such as classical music) entering only as market size increases. Alternatively, new stations might enter into existing formats, but at a lower quality level. For example, they might be low-wattage stations with very little geographic range and therefore low listenership.

In the absence of any empirical evidence to the contrary, this argument seems plausible. Adding plausibility is the fact that within-market station listening shares are quite variable. Therefore, as an empirical test, we return to the listening model with differentiated stations. Now, the mean consumer utility for a station, δ_{jk}, depends on observed characteristics of the station, x_{jk}, and on unobserved quality, ξ_{jk}. The x vector includes measures of station formats and broadcast power. The unobserved ξ vector presumably includes the quality of talent employed by the station. Our observed station-level x measures are taken from Duncan, while the station-level listening shares are from Arbitron.

Aside from introducing the station-level data on characteristics and listening shares, we retain the simple nested logit framework, with the parameter σ measuring the degree of business stealing. Now, however, the shape of Figure 1 can be explained by the varying x's and ξ's, not just by σ. To begin, we leave the utility "nests" as in/out of radio listening. We also look at some results in which the nests are industry-defined station formats (like "Top 40," "News/Talk," and "Jazz").

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

Table 6 presents results of the listening demand equation with station heterogeneity. The parameters are the β's from equation (8) plus the parameter σ. The table lists the variables in x and the mean values of those variables. The first set of variables again describes the city markets. We have added a new variable here, which is the percentage of the population that drives to work (% Commute). The next set of variables comprises the "physical" characteristics of the individual station. These are FM broadcast, power of the station (in mega watts), the height of the tower, and whether the station is outside of the metro. Next comes the list of dummy variables for station formats, in order of the number of stations offering each format. We also include some interactions between the different types of x variables (these interactions seemed to tell us more than the simple format dummies.) For example, we interact % Hispanic (in the city) with the dummy for whether the station is a Spanish-language station.

Once again, the estimation method follows Berry (1994) exactly. Columns 3–4 of Table 6 give the parameter estimates (and standard errors) of OLS estimation, followed in columns 5–6 by the two-stage least-squares results. The instruments are the same as in the earlier demand results. We see that many of the station-level variables are important and make sense. For example, high-wattage stations with tall broadcast antennas gain more listeners. In the south, religious stations are more popular than elsewhere. These results formally reject the pure symmetric model (which, of course, was always a stylized assumption). The pure station formats do not seem to describe tastes very precisely, but in some cases the interactions terms are significantly different from zero.

Columns 7–8 give the results from a similar estimation strategy, with the sole exception that the grouping for the nested logit is assumed to be given by the formats, rather than by in/out of radio. Now, the utility specification in (6) becomes (for consumer i considering station j, which is in format f)

$$u_{ij} = \delta_j + \nu_{if}\sigma + (1 - \sigma)\epsilon_{ij}. \tag{21}$$

Here the random taste term ν_{if} is constant within format but varies across formats, and the variation of within-format taste is measured by σ.[9]

In this case there are 17 nests instead of 2. This allows for tastes that are correlated within format, so that a listener to one country music station is likely to stay within the country format as the number of stations increases. Contrast this to our earlier specification, which allows for correlation within radio listening but not for format-specific tastes. For simplicity, we impose that the σ parameter measuring the correlation of within-format tastes is equal across all groups.[10]

In this special case, many of the coefficients on the format dummy variables are significantly different from zero, and they have a very intuitive pattern. The coefficients typically decline as we move down the table (as expected), except for the natural result that, for example, Spanish stations are popular only in cities with a large Hispanic population. These results are consistent with the notion that station heterogeneity is important for understanding radio listening.

Importantly, however, we see that station heterogeneity does *not* account for the shape of Figure 1. We still get high estimates of σ, now in the range of .9, which indicates a very large amount of business stealing, either within format or within radio

[9] Given more data, we could try to let the parameter σ vary across formats.

[10] We experimented with nested logits that have multiple levels of nests—e.g., an upper level including all radio stations that is then subdivided into lower-level nests defined by formats. However, the data did not seem to be able to distinguish between the two levels of nests.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

TABLE 6 Listening Equations with Station Heterogeneity

	Groups are In/Out				Groups are Formats	
	OLS		2SLS		2SLS	
Variable	Parameter	SE	Parameter	SE	Parameter	SE
City Variables						
Constant	−2.006	.069	−2.310	.115	−5.217	.174
North central	.008	.013	.013	.013	−.002	.032
South	−.033	.014	−.032	.014	−.193	.036
West	−.029	.016	−.032	.016	−.013	.039
Black	.307	.060	.214	.067	−.353	.151
Hispanic	.053	.045	.023	.047	−.548	.112
% Commute	.141	.122	.181	.126	−.261	.306
% College	−.275	.064	−.303	.066	−.177	.161
Station Variables						
FM	.048	.014	.091	.020	.044	.049
Mega watt	.612	.120	.956	.161	1.428	.374
Tower height	.022	.008	.030	.009	.041	.021
Out metro	−.129	.011	−.194	.023	−.216	.054
Station Formats						
Country	−.007	.044	.044	.048	1.787	.123
Album rock (AOR)	.006	.043	.047	.046	1.488	.118
News/Talk	−.089	.064	−.069	.066	1.207	.169
Adult contemp	−.035	.044	−.001	.046	1.302	.117
Top 40	−.000	.044	.052	.048	1.439	.115
Soft adult contemp	−.019	.044	.023	.047	1.126	.113
Other	−.015	.044	.014	.046	.804	.111
Black	−.039	.055	−.015	.057	.221	.135
Religious	−.052	.047	−.071	.049	−.423	.117
Big band	.004	.047	.069	.052	.492	.118
Full service	.052	.047	.135	.055	1.055	.117
Spanish	−.025	.056	−.013	.058	.025	.138
Classic AOR	−.000	.048	.054	.052	.615	.118
Unknown	−.071	.050	−.094	.052	−.674	.125
Jazz	−.009	.057	.013	.059	.148	.141
Interactions						
% Hisp*Spanish	.062	.099	.120	.104	4.210	.309
% Black*Black	.322	.174	.487	.186	7.319	.502
AM*News/Talk	.141	.053	.201	.057	.159	.138
South*Religion	.061	.034	.099	.036	.983	.086
South*Country	.008	.024	.002	.024	.569	.069
Substitution						
σ	.928	.004	.864	.020	.923	.052

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

listening more generally. Therefore, while the symmetric model is not literally true, neither does station heterogeneity explain away the large degree of business stealing that drives our results.

It would obviously be preferable to empirically follow through on the implications of station heterogeneity for the entry model, but that requires a model of entry into characteristics space that is well beyond the scope of this article (and indeed is beyond the current empirical literature.) The large values of σ shown in Table 6 are suggestive, however, that such a model would still demonstrate excess entry (from the point of view of market participants). However, more complicated patterns could arise in the multiformat model; for example, there could be too many stations overall and yet some formats could be underserved.[11]

□ **Endogenous fixed costs.** The last subsection looked at how product differentiation affects radio listening. An even more ambitious criticism of our article is that the degree of product differentiation is endogenous. Surely, each firm decides into which format to enter and also decides, at some cost, what quality of station to operate. That is, both x_{jk} and ξ_{kj} are endogenous. Further, the level of fixed costs is endogenous; for example, a station can spend money to purchase a popular syndicated morning radio personality and thereby increase demand. The cost of programming is fixed with respect to listening (in the sense that when an additional listener tunes in, costs do not increase), but listening and costs are endogenously related.

Some readers have gone so far as to suggest that the endogeneity of fixed costs removes the business-stealing problem altogether. They suggest that "buying" listeners by increasing product quality is just like producing more product at a given marginal cost. Therefore, the argument continues, costs aren't really fixed and so there is no problem of excess entry. This argument has been made to us repeatedly and is therefore worth addressing. The argument is wrong because increases in product quality also result in "stolen" business. The quality of one product enters the demand for another, so increases in product quality create an unpriced, negative effect on rivals' profits. The correct model here is not a traditional model of variable production costs, but a model of endogenous fixed costs, on the lines of Sutton (1991). The nature of the entry problem in such models is, in general, a very difficult theoretical proposition, but we can outline some of the incentives in choosing a quality level conditional on entry.

Formally, consider our nested logit model but let the firms *choose* a quality level δ_j at some cost $F(\delta_j)$. The firm's profit depends on the entire vector of quality levels, δ:

$$\pi_j(\delta) = M_p(S(\delta))s_j(\delta) - F(\delta_j). \tag{22}$$

Let us begin with the simple case where advertisers' demand is perfectly elastic (e.g., ad prices are set in some larger market), so that $p'(S) = 0$. This lets us focus on business stealing as opposed to other oligopoly distortions. The first-order condition is then

$$Mp\frac{\partial \pi_j}{\partial \delta_j} = Mp\frac{\partial s_j}{\partial \delta_j} - \frac{\partial F}{\partial \delta_j} = 0. \tag{23}$$

This just says that the marginal revenue product of quality should be set equal to the

[11] See Berry and Waldfogel (1999) for a discussion of possible underprovision of Jazz, News, and Classical formats.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

marginal cost of quality, which does sound like a condition for social optimality. However, the welfare function of the market participants is

$$M \int_0^{S(\delta)} p(x)\, dx - F(\delta), \tag{24}$$

which has the first-order condition

$$\left[Mp \sum_k \frac{\partial s_k}{\partial \delta_j}\right] - \frac{\partial F}{\partial \delta_j} = 0. \tag{25}$$

This differs from the private first-order condition because of the business-stealing effect:

$$Mp \sum_{k \neq j} \frac{\partial s_k}{\partial \delta_j}. \tag{26}$$

That is, the firm does not privately take into account that increases in own-quality reduce the listening to other stations, but the social planner does.

The last paragraph assumes perfectly elastic demand. Once we assume that demand slopes down, the private first-order condition gains a term that reflects the traditional oligopoly incentive to reduce output (in this case, by reducing the use of the input called quality.) The additional term is

$$Mp'(S) \sum_k \frac{\partial s_k}{\partial \delta_j}. \tag{27}$$

This term is negative and so could reduce the level of quality choosen by the firm. The sum $\Sigma_k\ \partial s_k/\partial \delta_j$ is positive but goes to zero as σ goes to one (i.e., with perfect business stealing). Therefore, the absolute size of this oligopoly incentive increases with the slope of inverse demand and decreases in σ.

It would be worthwhile to explore the empirical implications of this model further, but once again the associated entry model requires a major methodological extension of the literature and is beyond the scope of this article.[12] In any case, the theory indicates that we have no particular reason to believe that the bias from assuming exogenous fixed costs runs in the direction of less business stealing. The endogenous fixed-cost model might, however, shift the locus of the problem from entry onto the marginal choice of quality. For example, one could find that stations in large markets spend too much money, relative to the welfare of the market participants, on popular radio personalities. Also, again similar to our results, one might find that this expenditure creates a large external benefit to listeners.

□ **Robustness to other problems.** Our results may be sensitive to other assumptions implicit in our framework and sample. Our estimates of fixed costs may be affected by government restrictions on broadcasting that create license rents in large markets. The positive relationship between population and fixed costs might reflect this rather than other input prices that are higher in large cities. To test this we reestimated the

[12] One feasible possibility would be to *condition* on the number of firms and treat the choice of δ in the same way that the choice of price is treated in the empirical discrete-choice literature.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

model excluding the top 25 markets, and we obtained very similar results. The ratio of welfare loss to revenue was actually somewhat higher than our estimate using the full sample.

Another possible problem with our approach is our treatment of stations located outside the metropolitan areas. Our model—and our simulation—treats the number of outside stations (N_2) as fixed even as we optimally reduce N_1. Many stations that are outside of one market are inside of some other market. Thus N_2 should probably decrease in the social optimum as other markets move to the optimal number of inside stations. We therefore redid our policy simulations using an ad hoc rule that reduced N_2 proportional to our reductions in N_1. These simulations result in a larger deadweight loss from free entry. This is because advertisers actually value the listening stolen from outside-the-metro stations. With fewer outside stations, in-metro entry is more wasteful because business is stolen, to a greater extent, from other inside stations.

We are also concerned about the possible endogeneity of outside-the-metro entry. We could test whether our treatment of N_2 as exogenous affects our results by reestimating the model including only "isolated" markets (with relatively few outside stations). As a rough approximation, we remove from the sample the 28 cities in the densely populated northeast. These cities have an average ratio of outside-the-metro listening to inside-the-metro listening of 22%, while the remaining cities' ratio is less than 13%. The results are very similar. The estimate of σ remains about .8, the estimate of η falls a bit to about .3 (indicating more elastic advertising demand), and the welfare loss from free entry is about 45% of revenue.

Given better data, we could address several concerns about the advertising model. Data on the number of advertising minutes per hour would allow us to model a more interesting oligopoly interaction in the advertising market. Data on individual station advertising rates would allow us to estimate how valuable different demographic groups are to advertisers. One possibility, suggested to us by a referee, is that radio advertisers care about the match between their products and the demographics of the listening audience. In this case, station diversity might increase advertiser welfare by improving the "targeting" of ads even when total listening is not increasing. We note, however, that our measures of advertising revenue per listener *decline* in the number of stations, indicating that any increase in advertising demand from station diversity is offset by some competitive effect. Unfortunately, we do not have the data on advertising prices by station that would allow us to fully address this issue.

8. Conclusion

■ In this study we have filled a gap in the empirical literature on the efficiency of free entry. Because we have data on firms' revenue we are able to estimate how revenue varies with entry. Consequently, our entry model generates direct estimates of the distribution of fixed costs. Given an explicit measure of welfare we determine the optimal number of stations. We then compare the number of stations and social welfare under free entry and optimality. Ignoring the value of programming to listeners we find that free entry into U.S. radio broadcasting causes a welfare loss of over 40% of the size of current industry revenue. We can rationalize the number of stations under free entry as optimal if the value of programming is about 15 cents per hour of listening. Our numerical estimates of welfare effects depend on many assumptions and are best viewed as rough-cut estimates. However, when we examine the likely biases of our assumptions, we find little reason to believe that the basic result—a large business-stealing effect—is wrong.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms

The cost structure of radio broadcasting is an extreme case, with large fixed costs and zero marginal costs. This is the sort of situation in which free entry is most likely to be inefficient. Yet we suggest that radio broadcasting is not unique. Many activities related to the production and distribution of information have similar characteristics, including the computer software industry, television broadcasting, and print media. R&D-intensive industries also share this characteristic. Further study is needed to quantify the degree of inefficiency in such contexts.

References

ANDERSON, S.P., DEPALMA, A., AND NESTEROV, Y. "Oligopolistic Competition and the Optimal Provision of Products." *Econometrica*, Vol. 63 (1995), pp. 1281–1302.

ARBITRON CO. *Radio USA*. New York: Arbitron, 1993.

BERRY, S.T. "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, Vol. 60 (1992), pp. 889–917.

———. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, Vol. 23 (1994), pp. 242–262.

——— AND WALDFOGEL, J. "Public Radio in the United States: Does It Correct Market Failure or Cannibalize Commercial Stations?" *Journal of Public Economics*, Vol. 71 (1999), pp. 189–211.

BORENSTEIN, S. "Too Much Rock and Roll?" Unpublished paper presented at the 1986 Econometric Society meetings.

BRESNAHAN, T.F. AND REISS, P.C. "Do Entry Conditions Vary Across Markets?" *Brookings Papers in Economic Activities*, Vol. 18 (1987), pp. 833–882.

——— AND ———. "Entry in Monopoly Markets." *Review of Economic Studies*, Vol. 57 (1990), pp. 531–544.

CARDELL, N.S. "Variance Components Structures for the Extreme Value and Logistic Distributions with Application to Models of Heterogeneity." *Econometric Theory*, Vol. 13 (1997), pp. 185–213.

CHAMBERLIN, E. *The Theory of Monopolistic Competition*. Cambridge, Mass.: Harvard University Press, 1993.

DIXIT, A.K. AND STIGLITZ, J.E. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, Vol. 67 (1977), pp. 297–308.

DRANOVE, D., SHANLEY, M., AND SIMON, C. "Is Hospital Competition Wasteful?" *RAND Journal of Economics*, Vol. 23 (1992), pp. 247–262.

DUNCAN, J.H., JR. *Duncan's Radio Market Guide*. Indianapolis: Duncan's American Radio Inc., 1994.

EUROPEAN BROADCASTING UNION. *European Broadcasting Review: Programs, Administration and Laws, Volume 39*. Geneva: European Broadcasting Union, 1988.

MANKIW, N.G. AND WHINSTON, M.D. "Free Entry and Social Inefficiency." *RAND Journal of Economics*, Vol. 17 (1986), pp. 48–58.

ROGERS, R.P. AND WOODBURY, J.R. "Market Structure, Program Diversity, and Radio Audience Size." *Contemporary Economic Policy*, Vol. 14 (1996), pp. 81–91.

ROSSE, J.N. "Estimating Cost Function Parameters Without Using Cost Data: Illustrated Methodology." *Econometrica*, Vol. 38 (1970), pp. 256–275.

SPENCE, M. "Product Differentiation and Welfare." *American Economic Review*, Vol. 66 (1976a), pp. 407–414.

———. "Product Selection, Fixed Costs and Monopolistic Competition." *Review of Economic Studies*, Vol. 43 (1976b), pp. 217–236.

STEINER, P.O. "Program Patterns and Preferences, and the Workability of Competition in Radio Broadcasting." *Quarterly Journal of Economics*, Vol. 66 (1952), pp. 194–223.

SUTTON, J. *Sunk Costs and Market Structure*. Cambridge, Mass.: MIT Press, 1991.

WALDFOGEL, J. "Criminal Sentences as Endogenous Taxes: Are They 'Just' or 'Efficient'?" *Journal of Law and Economics*, Vol. 36 (1993), pp. 139–151.

© RAND 1999.

This content downloaded from 193.49.169.59 on Sun, 22 Dec 2019 15:47:15 UTC
All use subject to https://about.jstor.org/terms