Optimal Product Variety in Radio Markets

Steven Berry, Alon Eizenberg and Joel Waldfogel

Summer 2007

Product Variety

In differentiated products markets, competition in the space of horizontal and vertical quality may be as important as competition in price. Understanding competition in quality and in product characteristics is important to

- the welfare analysis of markets,
- anti-trust policy,
- marketing,
- etc.

For example, in anti-trust analysis the U.S. Dept. of Justice focuses on changes in price and not much on changes in product characteristics and quality.

Idea of Paper

In this paper, we present a model of entry into an discrete product space that allows for point estimates of the parameters of variable profits and bounds on fixed costs.

Applying this model to the Radio Industry, we consider optimal product variety in terms of the number of stations in different radio formats ("rock", "country", etc.)

Extensions include: vertical quality, joint ownership, merger analysis.

Background on Radio

- There is a long theoretical literature on the inefficiency of free entry into oligopolistic markets. New firms "steal business" from existing firms: a negative externality. Lower prices for existing consumers and the intro of new varieties create an offsetting positive externality.
- Excessive entry into radio industry has often been suggested.

Berry and Waldfogel, 1999

They use new data and simple methods to estimate the extent of and welfare loss from excess entry in radio broadcasting.

Results from BW '99

- First, look only at market participants: broadcasters advertisers. Welfare loss from free entry, as opposed to the socially optimum N, is 40% of industry revenue. A big number?
- There is still the positive externality to listeners. If listeners value an hour of listening at about 15 cents an hour, then welfare loss to market participants would be just offset by external benefit to listeners.

But they had to assume symmetric stations, no differentiation by format, etc.

Benefits of Variety

- The introduction of new varieties can reverse the finding of excess entry.
- And radio stations offer a variety of "formats".
- Berry and Waldfogel '99 found that as population increases, additional stations are often in existing formats.
- Most likely problem of *insufficient* entry would occur when there are ZERO stations in a given market.

Background on Method

In radio, we observe "price" (advertising price) and "quantity" (of listening) and can estimate a fairly rich model of variable profits.

But to estimate the distribution of fixed costs via MLE, Berry and Waldfogel (1999) had to ensure a unique equilibrium via the very unrealistic assumption of symmetry across products.

In this paper

we use very simple methods to bound fixed costs in each market, within imposing unique equilibrium.

Intuition for "Identification"

Variable profits are estimated from price, quantity in the usual way and then

- "Observed" variable profit in format place an upper bound on fixed cost, because firms have positive total profits.
- Variable profit at $(N_k + 1)$ places a lower bound on fixed cost, because further format entry was not profitable.
- Robust to multiple equilibria, but cannot point-identify distribution of fixed costs, F.
- Can get bounds (sometimes sharp) on Fixed Costs under very weak assumptions on F.

Outline of Model

1. Stations produce listeners, who make a free choice as to listening. Listeners care about format and within format stations are more "similar". Formally, use nested logit.
2. Stations sell listeners to advertisers. Advertisers' demand is downward sloping in the share of the population who listen. Simple constant elasticity functional form.
3. There is free entry into discrete product space (formats) and a static Nash equilibrium. No unique equilibrium: entry problem is no longer a Bresnahn-Reiss style ordered probit.

(1) and (2) give variable profit function, (3) adds fixed costs.

Observed Data and Variable Profits

No Variable Cost (but add endogenous fixed cost of "quality" later).

In market t, format k, We observe:

- ad price p_t,
- format share s_{kt},
- stations numbers N_{kt},
- market demographics x_t,
- population M_t.

At observed vector N_{kt}, observed variable profits are

$$v_{kt} = p_t(s_t) M_t s_{kt}$$

At market outcome, v is just observed revenue, R_{kt}.

Counter-Factual Variable Profits

To create bounds on fixed cost, also need variable profits at $N_{kt} + 1$.

To get this counter-factual, need to

1. Estimate model of listening demand $s_{kt}(x_t, N_{kt}, N_{-k,t}, \theta_d, \xi_{kt})$,
2. Estimate model of Advertising Price $p_t(x_t, s_t, \omega_t)$.

How to Model the Product Space

- **"Ex-Ante" vs. "Ex-Post"** differentiation: with ex-ante, have to specify number and characteristics of *potential* competitors. For airlines (Berry '91) and Chain Stores (Jia, '06) this might might sense, but other times is quite arbitrary.
- **Continuous v.s Discrete** product space. Easier to specify "counterfactual profits" (profits of the "next entrant") with discrete space.

Here

we use ex-ante identical entrants into a discrete space of product "segments".

The Model of Listening.

- Within format, stations are symmetric post-entry, but each new station brings some unique benefit.
- Motivate functional form for listening equation via nested logit utility function for listeners.
- Simplest Nested Logit nests only on formats. Also look at two level nests: listen/don't listen and then format.
- Natural extension is to BLP-style demand with random coefficients logit (see Sweeting, 2007).

Simplest Nested Logit

Utility to listener i tuned to station j in format k in market t is

$$u_{ijt} = \delta_{kt} + \nu_{ikt}(\sigma) + (1-\sigma)\epsilon_{ijt},$$

with

$$\delta_{kt} = x_t\beta_k + \xi_{kt}$$

where

- δ_k is the mean taste for format k,
- ν_{ikt} is a random variable that introduces correlated tastes within format, parameterized by σ,
- ϵ is an station/listener i.i.d. match component,
- x_t are market attributes (demographics),
- ξ_{kt} is an unobserved (to us) taste for the format in this market,
- β_k is a format specific parameter.

Format Nests

For the one-level-nest models, the estimation equations are derived as

$$ln(s_{kt}) - ln(s_{0t}) = x_{kt}\beta_k + (1-\sigma)ln(N_{kt}) + \xi_{kt}.$$

Complications:

- Note the endogeneity of RHS N_{kt}.
- We let the mean utility levels (and the ξ's) vary by "in" and "out" metro stations.
- The two-level nests (e.g. "formats" and "in/out" of listening) add an additional parameter ρ that captures the correlation within the upper level nest.

Endogeneity

.

For the endogenous "within nest share", three main instruments are used:

- population (exogenous, and correlated with N)
- number of out-metro stations, N^2, (assumed exogenous)
- number of out-metro stations in the same format

Demand from Advertisers

We treat stations as "producing" listeners and then selling them to advertisers. For now, a very simple inverse ad-demand function.

The demand from advertisers for listeners in market t is modeled by a downward-sloping, constant-elasticity specification:

$$ln(p_t) = x_t\alpha - \eta ln(s_t) + \omega_t \quad (1)$$

Popl. and out-metro stations are instruments for endogenous share. Might be able to have this vary by format / demographic, but data is pretty bad for this.

Equilibrium in Product Segments

Once we have listening demand and the (inverse) advertising demand equation, we have estimated variable profits.

Segment Fixed Costs

To recover fixed-costs (constant across products within segments) need to have a model of equilibrium market structure.

Static Complete Info Nash

A good assumption for work that relies on the cross-sectional nature distribution of market structure. With no explicit dynamics, we would like firms to choose the best-response to to rival's actions – otherwise why don't they move? Justification for cross-sectional study is [i] population and demographics are strong instruments and [ii] firms are in "long-run" equilibrium.

In a dynamic model, some private info makes more sense – firms might be surprised to find themselves in a bad location and then move away.

Problems with Non-unique Equilibria

- With multiple equilibria, knowing the primitives of the model (including unobservables) is not enough to determine outcome
- Can't run MLE, in discrete models can't run GMM.
- Equilibrium selection rule is a further unobservable
- Incomplete Model, as in Manski.
- In simple settings, can model and/or estimate the equilibrium selection rule, but with very many outcomes across many diverse markets, this gets difficult.

Similar Models

- Bresnahan and Reiss looked at symmetric entry, ex-post differentiation,
- Reiss and Spiller, Berry and Waldfogel estimated variable profits outside the entry model,
- Mazzeo considered discrete product segments ("quality") and ex-post differentiation, needs strong assumptions on order to get unique equil.
- Seim uses private info
- Manski – use incomplete models, maybe get bounds.
- Iishi, Iishi-Ho-Pakes-Porter – similar ordered models plus bounds estimation.

Bounding the Distribution of Fixed Costs

Complete Info Static Nash Equilibrium

- No variable costs. F has to be less than observed revenue.
- Also, F has to be greater than counterfactual revenue at $(N_{kt} + 1)$.
- Construct counterfactual revenue from listening demand and ad-price equation (including values of unobservables.)
- Can't do this for markets with $N_t = 0$; selection discussed below.

Upper Bound on F

We know that

$$R_{kt} > F_{kt}$$

This provides an upper found for F, making only the assumption that R and F are constant within segment.

Further, the empirical CDF of R_{kt} is a lower bound to the empirical CDF of F across sample markets.

If we further assume that F is i.i.d., then the true CDF of R_{kt} is a lower bound to the true CDF $\Phi(F)$.

Lower Bound on F

In equilibrium,

$$V_k(N_{kt}+1, y_t, x_t, \theta_0) < F_{kt}.$$

This provides an upper found for F_k, again making only the assumption that R and F are constant within segment k.

Further, the empirical CDF of $V_k(N_{kt}+1, y_t, x_t, \theta)$ is an upper bound to the empirical CDF of F across sample markets.

Again, if we further assume that F is i.i.d., then the true CDF of $V_k(N_{kt}+1, y_t, x_t, \theta)$ is an upper bound to the true CDF $\Phi(F)$.

Sampling Error of the Bounds on F

If we want to do within market prediction, holding all market characteristics fixed, then the upper bound R_t has no sampling error (except from the Arbitron survey) and the estimated lower bound $V(N_t + 1, y_t, x_t, \hat{\theta})$ has sampling error only from $\hat{\theta}$.

If we think of the estimate of $\Phi(F)$, then sampling variance comes both directly from the sampling error in estimating the empirical CDFs of R and $V(N_k + 1)$, but also again from $\hat{\theta}$.

Selection Problem

Big problem: sometimes $N_{kt} = 0$, so don't see S_{kt}, p_t, etc.

Problem for

- Estimating Listening equation,
- Calculating Upper and Lower bounds on CDF

Selection in Listening Equation

Difficult problem: multivariate selection on ξ's & F's of all formats, without any known selection rule (possible multiple equilibria).

Solution: estimate only on markets where probability $N_k > 0$ is one. Here assuming (reasonably) a bound to the support of F. There is zero probability that a market the size of New York will have no rock station. Market with large enough Hispanic population will certainly have a "Hispanic format" station.

This solution is worse the finer is the definition of format. Intermediate solution: formats that vary in observables, but share an unobservable ξ.

Selection and Bounds on Φ

For the lower bound on Φ

$$\tilde{R}_t = \begin{cases} R_t & \text{if } N_t > 0 \\ \bar{F} & \text{if } N_t = 0 \end{cases}$$

Where $\bar{F}$ is largest F in the "large" markets not subject to selection. The distribution of $\tilde{R}_t$ is a lower bound on Φ.

Same idea when can't compute $V(N_t + 1)$ – replace with $\underline{F}_t$.

Data Sources

A cross-section of metropolitan radio markets. The market definitions are those of Arbitron (close to MSA definitions) Data from

- *American Radio, By Duncan's American Radio, Spring 2001* – Arbitron's listening figures for its 286 metro markets. Use Average Quarter Hour listeners
- Duncan's Radio Market Guide, 2001-02 Editions. – market-level revenue estimates. There are some problems with these. Also – market demographics (% black, ave. income, college, etc.)
- For now, 163 markets. Can probably expand to 200 or even 286.

Summary Stats

Table A1: Description of Market-Level Data

Variable	Units	Mean	Std. Deviation
Share in-metro	%	0.111	0.026
Share Out-metro	%	0.015	0.023
N1 (in-metro)	integer	19.577	7.557
N2 (out-metro)	integer	7.184	8.299
Population	millions	1.016	1.687
Ad Price	$	570.480	237.653
Income	10,000$	4.584	0.860
College	%	21.200	5.370

Statistics computed over the 163 markets for which we have full data

Table 1: 10-format configuration

Format Group	Formats Included				
"Mainstream"	Adult Cont. Classic Hits	Hot AC 80s Hits	Modern AC	Soft AC	Adult Altern.
CHR	CHR				
Country	Country	Classic Cntry.	Trad. Country		
Rock	Rock	Active Rock	Modern Rock	Classic Rock	
Oldies	Oldies				
Religious	Religious	Cont. Christ.	Black Gospel	Gospel	S. Gospel
Urban	Urban	Urban AC	Urban Oldies	Rhythmic Old	
Spanish	Spanish Span.-Cl. Hits Span.-Talk Ranchero	Span.-Oldies Span.-EZ Tejano Romantica	Span.-Adult Alt Span.-Hits Tropical	Span.-C. Christ Span.-NT Reg'l Mex.	Span.-CHR Span.-Relig. Span.-Stand.
News/Talk	News/Talk Sports	News Farm	Talk	Hot Talk	Bus. News
Other	Variety Pre-teen A30 Dance	Bluegrass Ethnic N Classical	Blues Silent N A Adult Stand.	cp-new A22 Jazz Easy List.	Americana A26 Smooth Jazz

Format Presence

Table 2: Format Numbers

Format Group	Frequency	Mean N	Max N	Mean format share
Mainstream	100.00%	4.48	11	2.31%
CHR	93.25%	1.66	6	1.16%
Country	99.39%	2.99	9	1.85%
Rock	100.00%	3.42	9	1.88%
Oldies	98.16%	1.48	5	0.79%
Religious	79.75%	1.88	6	0.37%
Urban	73.62%	2.10	6	1.24%
Spanish	40.49%	1.63	15	0.40%
News/Talk	100.00%	4.31	13	1.55%
Other	94.48%	2.80	9	1.09%

Table 3: Comparing listening models

	In/Out	Formats	2-level
in-market	0.1333**	0.6388**	0.1325**
	[0.0248]	[0.0829]	[0.0253]
hispXspan	0.0095	0.3519**	0.0192
	[0.0075]	[0.0358]	[0.0135]
blackXurban	0.0238*	0.5057**	0.0378*
	[0.0100]	[0.0506]	[0.0191]
southXreligious	0.0555**	0.8091**	0.0768*
	[0.0206]	[0.0953]	[0.0323]
southXcountry	0.0087	0.3164**	0.0181
	[0.0159]	[0.0721]	[0.0194]
σ	0.9043**	0.5192**	
	[0.0188]	[0.0630]	
Upper level corr			0.886
			[0.028]**
Lower level corr			0.167
			["tba"]
Observations	1919	1919	1919
Adjusted R-squared	0.9851	0.7199	0.9846

Uninteracted demographics, format dummies and region dummies not shown

Selection Solved via sue of Large Markets?

We are in the process of looking at robustness to various means of solving for selection – so far choice of method does not greatly change result.

Following graphs (and probits not presented here) demonstrate that large markets almost certainly have $N_k > 0$. (Although "religious format" may still be a problem.)

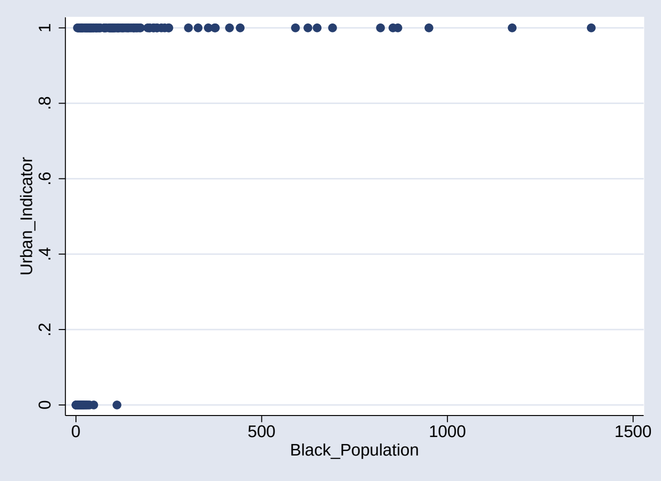

Figure: Presence of Urban Station Plotted against Black Metro Population, in 1000s (NYC Excluded)

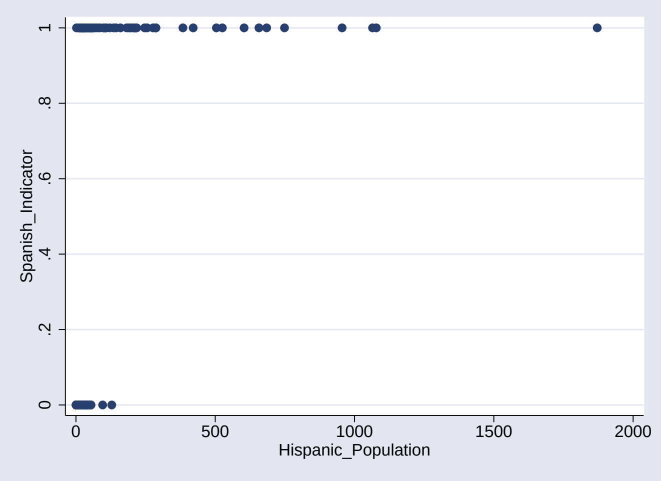

Figure: Presence of Spanish Station Plotted against Hispanic Metro Population, in 1000s (NYC, LA Excluded)

Table 7: Ad Price Equation

	IV Coeff	SE
northeast	-0.0739	[0.0645]
midwest	0.0799	[0.0609]
south	0.0132	[0.0602]
income	0.0606	[0.0302]
college	0.1639	[0.0434]
black	-0.0242	[0.0208]
hisp	-0.0124	[0.0138]
η	0.5101	[0.0737]
Constant	4.5537	[0.1885]
Observations	163	
Adjusted R-squared	0.4929	

Instruments are Population, N_2

Estimated Bounds

We graph these by format.
Preliminary, and we ought to

- Provide Confidence Regions
- Show robustness to treatment of formats, etc.
- Allow distributions to vary with x_t
- Add quality choice (perhaps reduce within segment spread of F)

Could also consider a parametric, multivariate dist. of F.
(Technique more difficult.)

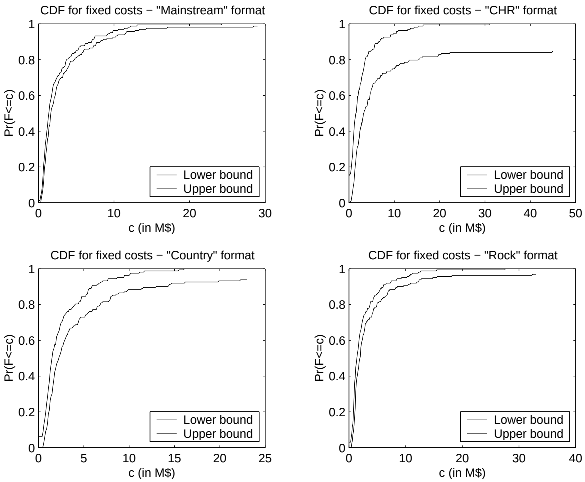

Figure: Estimated bounds on the CDF of fixed costs

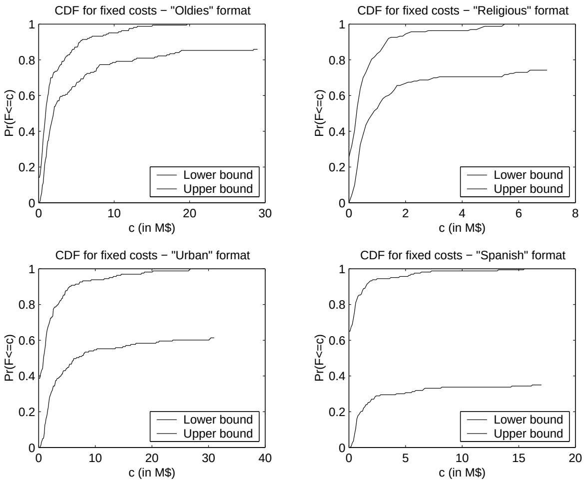

Figure: Estimated bounds on the CDF of fixed costs

Optimal N

Once we have θ and bounds on F, we can place bounds on the optimal number of stations.
Caveats:

- As in B-W '99, we can only look at the welfare of market participants – producers (stations) and consumers (advertisers). Listeners are an unpriced input, who do receive social value.
- Easiest is to hold F at mid-point of the bounds for the market, and then get point estimate of optimal N vector.
- But can also use bounds on F to create bounds on N vector.

Table 9: Comparison of Observed and Optimal Mean Number of In-metro Stations

Format	Observed	Optimal	Percentage Difference
Mainstream	3.35	1.38	0.59
CHR	1.06	0.85	0.20
Country	2.10	1.05	0.50
Rock	2.33	1.09	0.53
Oldies	1.02	0.88	0.14
Religious	1.66	0.81	0.51
Urban	1.50	0.72	0.52
Spanish	1.34	0.60	0.56
News/Talk	3.08	1.35	0.56
Other	2.12	1.07	0.50
Sum	**19.58**	**9.79**	**0.50**

Bounds on Optimal N

The last table used the mid-point of the market-specific bounds on F. It is better to use the bounds themselves.

We can get a upper bound on optimal N_k by setting F_k to its market-specific lower bound and for all other formats $(r \neq k)$ setting N_r equal to its upper bound.

The per-market bounds on F are tight enough that it doesn't matter that much. (See the following table).

Observed vs. Optimal Mean Number of In-metro Stations

Format	Observed	Optimal (low)	Optimal (upp)	Optimal ("mid interval")
Mainstream	3.35	1.29	1.60	1.38
CHR	1.06	0.85	0.86	0.85
Country	2.10	0.99	1.10	1.05
Rock	2.33	1.01	1.21	1.09
Oldies	1.02	0.85	0.88	0.88
Religious	1.66	0.75	0.90	0.81
Urban	1.50	0.68	0.77	0.72
Spanish	1.34	0.54	0.67	0.60
News/Talk	3.08	1.22	1.56	1.35
Other	2.12	1.01	1.19	1.07
Sum	**19.58**	**9.20**	**10.75**	**9.79**

Comparison

These are about 50% reductions in stations, compared to 75% in Berry-Waldfogel '99. Adding quality may change this further (which direction?)

And, need to add benefit to listeners

Extensions

- Consider product segments as quality (high, low) plus format. Have to observe "quality" – based on observed station quality (e.g. wattage) and/or mean utility, δ, in station-specific listening equation.
- Consider multi-product firms – now counter-factual profit of one more or one fewer station has to consider effect on jointly owned stations.
  - Let distribution of fixed costs include an economy of joint-ownership,
  - Consider mergers and anti-trust policies toward mergers,
  - Because of bounds on F – will only get bounds on optimal policies.

Extension to Quality segments

How to meaure quality. In terms of x – power in watts? Or in terms of discretized estimated "quality", δ, from demand?

Extension to Multi-Product Firms

Do Multi-product firms have lower costs?

Bounds now need counter-factual effect on other stations in the market.

Also, need to actually estimate (bounds on) distribution of fixed costs.

Bounds on Multiproduct Firms

Condition for "entry" to be profitable is now:

$$R_{kt} > F_{kt} + \left(\sum_{k' \in Z_k} \left[V_{k't}(N_{k't}, N_{kt} - 1) - R_{k't} \right] \right)$$

Conclusion

- Dealing with interesting horizontal and vertical variety is now feasible
- In radio, adding horizontal variety in formats means that "optimal" reduction in the number of stations goes from 75% to 50%
- Extensions include radio "quality" and multi-product firms