Response to Peer Review Study 4 Section III "Factors that Affect a Radio Station's Propensity to Adopt a News Format"

Craig Stroup

I thank the peer reviewer, Scott Savage, for his time and comments on my study. He had several comments/questions so I respond to them here. I am gratified that he stated that the methodology and assumptions are reasonable and generally consistent with accepted theory and econometric practices, and that the data are of sufficient quality for the study.

The reviewer asks why all other radio format choices are lumped into a single category and not modeled in a multi-choice setting. Because the study's mandate was focused on news, grouping all other formats together seemed to be the best choice. There are 20 radio station formats, so modeling each one simultaneously wasn't feasible. One could try to reduce the number of format choices by grouping similar formats together, but then news would logically to be combined with talk and public/educational formats, which would have blurred the focus of the study. Thus, by lumping all non-news formats together, I was able to get clear estimates of how various variables affected a station's choice of news format.

The reviewer asks about the choice between logistic versus probit specification, and also about my choice of fixed effects rather than random effects. As is often the case, where the dependent variable is binary, the choice of logistic versus probit is largely up to the author's preference. I chose to use the logistic specification because it allowed me to use odds ratios, which helps convey the relative likelihood that a station will choose a news format over a music format. The results from the specification test suggested by Cameron and Trivedi shows that probit and logit specifications are nearly interchangeable.[1]

[1] In fact, the specification test suggested that logistic specification is slightly superior. See Cameron, A. Colin and Trivedi, Pravin K., *Microeconometrics Methods and Applications* Cambridge 2005, p 472.

As for fixed effects versus random effects, my paper discusses the fact that markets tend to have stable but very different percentages of news stations. Therefore I did not wish to potentially bias my estimated coefficients by using random effects and thereby omitting the fixed effects market dummies.

The reviewer also asks about fixed time effects. My logistic regressions did include fixed time effects, but I did not include their coefficients in the table or discuss them in the paper. This was an oversight. I did not include their estimated coefficients in the table because they were totally insignificant. Because I did not include the coefficient estimates in my table, I simply forgot to mention them in my write-up. I should have mentioned them, and I thank the reviewer for alerting me to my oversight.

The reviewer questions whether I could add some "demand side determinants" for completeness, and also asks why I excluded other traditional demographic information other than income. It is difficult to find any demand side or demographic variables that vary on a small geographic basis (and could therefore be rolled up to the Arbitron market level) and by year. The Census Bureau and the Bureau of Economic Analysis respectively publish population (broken down by race) and income estimates annually at the county level, so I was able to include them in my analysis.[2] These variables were insignificant. It seems likely that if I overlooked any other usable demographic data, that they would also show too little variation over the 3-year period to produce measurable results. Thus, the cross-sectional effects from these variables would be picked up by the fixed effects market dummies, and there should be no omitted variable bias in my coefficient estimates.

The reviewer also asks why I focused on blacks and Hispanics for race variables. Although Census data provide population estimates by race, I chose blacks and Hispanics because Arbitron, which provides ratings for radio stations, provides additional ratings

[2] These estimates are not simple straight line averages of the decennial census. I considered using such estimates for other variables but decided against it. In response to the reviewer's question as to why I included demographic variables, they were included simply to determine if I could identify any relationship between news operations and market characteristics.

data focusing on those groups. If Arbitron's customers believed that other demographic groups such as Asians significantly added predictive power to radio station listenership (and therefore, format choice), then I believe it would develop a program to report statistics for them.

Because the station's format data are self reported, the reviewer asks about its reliability. The stations have the incentive to use care when reporting the data, especially reporting its overall format, because many advertisers use that information to decide on which stations to broadcast their ads. Further, BIA has a data-verification process in place. Perfect data are a rarity, however. To wit, subsequent analysis of the cross-owned-with-TV data uncovered a small number of discrepancies involving university-owned radio stations. These stations appeared to be cross-owned with TV stations in 2002 and 2005, but not in the intervening years. After modifying these data points, the principal regression was rerun. The results were nearly identical to the earlier published results and are printed in Table 1.

In several places, the reviewer comments that a structural model would help him understand why the estimated coefficients had the signs they did. The question of "why" was not part of the mandate for this study.[3] It was sufficient to determine the sign and magnitude of the effects. Where economic theory provided obvious intuition, such as economies of scope for combinations of radio stations with newspapers and TV, it was given. Where there was no obvious economic intuition providing *a priori* expectations, I indicated so.

Also, the reviewer asked why the crowdedness of the news format is negatively related to a station's choice of news format. Holding everything else constant, more news stations in a market (i.e., a more crowded news market) means more competition, and fewer likely listeners for the individual radio stations broadcasting news.

[3] The Public Notice announcing the intent of these studies can be found at http://fjallfoss.fcc.gov/edocs_public/attachmatch/DOC-268606A1.pdf.

The reviewer also asks if it is possible that unobserved changes in preferences through time are correlated with cross-ownership variables and the station's choice of format, and suggests controlling for the number of radio and TV stations (potential and actual). I am not sure if the above discussion about the time dummies moots this question for the reviewer, but regardless, I will address it.

It seems unlikely that preferences would change sufficiently over the three years covered by the data to merit concern, and the insignificance of the year dummies is consistent with that. Also, there were no cross-ownership changes between radio and newspapers during this time, so they could not be correlated with preference changes for these years. Cross-ownership between radio and TV stations did rise steadily, though, from 557 to 643 between 2002 and 2005.

I reran the model adding the number of radio stations in the Arbitron market, and the coefficient estimate is totally insignificant. I do not know the number of potential stations in each Arbitron market, so I could not include the potential number of stations in my analysis.[4] Because Arbitron markets and TV markets do not always neatly overlap, determining the number of potential or actual of TV stations in the Arbitron market would be problematic, so I did not include them in my analysis.

The reviewer also mentioned that I should consider including the penetration of new media such as satellite radio and Internet streaming. The availability of satellite radio would be a nation-wide phenomenon captured by the year dummies, so including it would not change the results. Internet streaming is another matter. It works better on broadband than on dial-up, and consumers' ability and eagerness to adopt broadband across the United States has not been uniform. Thus, a variable designed to capture

[4] The largest markets are generally considered to be fully saturated by radio stations, so for these markets, the number of actual stations equals the number of potential stations. As a check to see if the fact that some markets are saturated and others are not affects the estimated coefficients, I reran the regression analysis using the top third (i.e., most likely to be saturated) and bottom third (least likely to be saturated) of the data, as measured by market population. The estimated coefficients were essentially unchanged, although some variables lost significance, which I attribute to the smaller sample size.

radio-listeners' ability to substitute Internet streaming for radio listening is an intriguing suggestion.

I cannot find annual broadband subscribership information on a county or smaller geographic basis, so I cannot measure broadband penetration on an (Arbitron) market-by-market basis. In an effort to address the reviewer's concern, I used the best proxy I could find. The FCC publishes statistics on the number of broadband connections in service at the state level,[5] so I created the variable "broadband lines per person" for each state. I use this variable as a proxy for broadband subscribership in Arbitron markets.[6] I reran the same specification as in the study, except that I added the broadband lines per person. The results are on the right-hand side of Table 1. At negative 6.2, the new variable's coefficient estimate is sizable (in absolute value), but it was insignificant. Perhaps with more granular broadband subscribership data, more precise and significant results could be obtained. Another interesting result of including broadband subscribership is that the other coefficients changed very little from the original specification. It may be that, if more precisely measured, the analysis would show that broadband penetration is an important variable, but current information suggests that its omission does not seem to affect the other coefficient estimates.

The reviewer also suggests examining the potential endogeneity of cross-ownership variables and suggests single-equation instrumental variables or systems of equation estimation methods as a robustness check. I cannot think of a factor outside my model that would influence both a station's willingness to enter into a cross-owned relationship (with either a TV station or a newspaper) and adopt a news format, so I cannot think of an appropriate instrument. Further, the other two general causes of endogeneity (measurement error and simultaneity) seem to apply here.[7] As for measurement error, the dummy variables are doing a good job of capturing what we wished to be captured, and

[5] See Table 10 of the FCC's High Speed Service for Internet Access: Status as of December 31, 2005, available at http://www.fcc.gov/wcb/iatd/comp.html.
[6] Where an Arbitron market was situated entirely within a state, I used that state's figure; where the market crossed state lines, I created a weighted average figure, based on the populations of the counties in the market. The number of broadband lines in Hawaii is non-public, so that market could not be used.
[7] See Wooldridge, Jefferey, "Econometric Analysis of Cross Section and Panel Data" MIT Press, 2002, 50-51.

are not distant proxies.[8] For instance, a station is either a commercial station or not: there is no in-between. This is unlike the problem of attempting to measure a person's ability by noting whether he/she has an advanced degree. Lastly, simultaneity isn't a problem. All the radio stations that are cross-owned with newspapers in our data set and over 85% of the radio stations that are cross-owned with TV stations were cross-owned in 2002, the first year of our data. Additionally, we observed station formats once per year, generally leaving plenty of time for stations to choose their formats after an ownership change is made.[9]

I again thank Scott for his time and comments.

Table 1
Updated Logistic Fixed Effects Regression Results

	Without Broadband Lines Per Person			With Broadband Lines Per Person		
Variable	Coefficient	P-value	Odds Ratio	Coefficient	P-value	Odds Ratio
News Format Prior Period	6.59 ***	0.00	727.49	6.59 ***	0.000	725.18
AM Station	1.81 ***	0.00	6.10	1.85 ***	0.000	6.35
Crossowned With Newspaper	1.46 **	0.04	4.29	1.43 **	0.041	4.16
Crossowned With TV Station	0.79 ***	0.00	2.20	0.79 ***	0.000	2.21
Commercial Station	-1.45 ***	0.00	0.24	-1.50 ***	0.000	0.22
Local Marketing Agreement	-0.42	0.15	0.66	-0.39	0.182	0.67
Sibling News Station in Market	0.39 ***	0.01	1.47	0.41 ***	0.007	1.51
Owner in Same State	0.24 *	0.06	1.27	0.23 *	0.078	1.26
Parent Stations in Market	0.04	0.16	1.04	0.05	0.138	1.05
Broadband Lines per Person				-6.21	0.360	0.00
Year 2002	-0.13	0.41	0.88	-0.77	0.260	0.46
Year 2003	-0.11	0.49	0.90	-0.57	0.259	0.56
Year 2004	0.09	0.59	1.09	-0.17	0.583	0.84
% News Stats in Mkt Previous Period	-28.40	0.00 [1]	0.00 [1]	-28.99	0.000 [1]	0.00 [1]
Fit Information	Intercept Only	With Covariates		Intercept Only	With Covariates	
Akaike Information Criterion	19,035	3,039		18,874 [2]	2,988	
Schwarz Criterion	19,035	3,148		18,874 [2]	3,105	
-2 Log Liklihood	19,035	3,013		18,874 [2]	2,960	
Number of Observations	32,681			32,161 [2]		

[1] This variable is collinear with the market dummies, so these values are meaningless. All other estimates are unaffected by this multicollinearity. This variable is included in the regression because it is an economically important variable, so dropping it could bias the other variables.

[2] Because broadband subscribership data for Hawaii is unavailable, stations in the Hawaiian market had to be dropped, so the number of observations is smaller and and the "Intercept Only" values in the goodness of fit measures are smaller than their counterparts in the regression without broadband lines per person.

[*] Significant at the 0.1 level.

[**] Significant at the 0.05 level.

[***] Significant at the 0.01 level.

[8] While there did seem to be some "measurement error" of another kind (where some stations were misclassified as not being cross-owned for some years), correcting those observations did not change the results.

[9] A TV station wishing to buy a radio station so that it can achieve economies of scope may wish to buy one with a news format, holding all else equal. Thus, one may wonder if a radio station might become a news station in hopes of becoming cross-owned, but I found no evidence such behavior.