

EFFECTS OF MONETARY PREMIUM VARIATIONS ON RESPONSE/NONRESPONSE BIAS:  
REPRESENTATION OF BLACK AND NONBLACK RESPONDENTS IN SURVEYS OF RADIO LISTENING

Roberta M. McConochie and Claude A. Rankin, The Arbitron Ratings Company

BACKGROUND

The Arbitron Ratings Company has a long standing commitment to the broadcasting and advertising industries to accurately measure and report the television viewing and radio listening behavior of respondents in our Syndicated ratings surveys. In addition to the metered electronic measurement of television in the top eleven TV markets, Arbitron utilizes a seven-day mail diary to measure every television and radio market in the country on a regular basis. Historically this task has been complicated by the underrepresentation of Black persons in our survey results, i.e. nonresponse bias.

Beginning in the late 1960's, Arbitron instituted special procedures to identify, reach, and measure Black respondents. Our early procedures included the creation of special geographic sampling units, race weighting, and daily telephone retrieval of survey information in lieu of our traditional mail procedures. More recently, Arbitron, in 1982, revamped its minority measurement procedures. Daily telephone retrieval was replaced by a program of "Differential Survey Treatments," or DST, for Black respondents in Metros where their population exceeded 10% of total. DST is a mail diary survey program with augmented premiums and extra follow-up telephone contacts designed to enhance returns from targeted groups. The highest augmented per-person premium, \$5, is sent to persons in Black households with a Male 18-34. Prior Arbitron studies have supported the findings reported in the reviews of premium research that the use of monetary incentives improves response with minimal respondent bias effects (1,2,3).

Despite empirical support to the contrary, some quarters of the broadcasting industry feared that augmented premiums might bias survey behavior, encouraging higher Black returns and the over-reporting of actual Black listening. The term "bought listening" was coined to describe this hypothesized response bias.

The present study was conducted in order to better understand how Black and non-Black diary return rates and reported listening vary with per-person premium amounts. To support other Arbitron research interests, the test samples were restricted to households with a Male 18-34 -- a particularly difficult demographic group to represent in survey research.

EXPECTATIONS

The "bought listening" hypothesis holds that increased premiums tend to influence or "buy" respondents' reported listening behavior, causing them to over-report actual listening. While this hypothesis is difficult to prove in the field setting, since only reported listening can be measured and even that will be contaminated by varying levels of nonresponse, survey results can be examined for support or the lack of support for this contention. Manifestations of "bought listening" would include significant,

substantial, and progressive gains in audience estimates as premium amounts increased.

Previous Arbitron research into the use of augmented premiums to enhance minority representation had not shown any support for the "bought listening" hypothesis. What our research had shown is that higher premiums tend to improve representation of these groups vis-a-vis their true proportion in the general population. While we had every reason to believe that augmented premiums would improve Black return rates, we did not expect to see significant, substantial, and progressive increases in reported listening levels.

METHOD

The test was fielded from April 14 to June 8, 1983, during Arbitron's Spring quarterly measurement of radio listening. The test sample, both persons and markets, was substantially larger than in any single previous Arbitron DST study. Five large and medium Metros were selected from the four main Census regions. All five had Black populations in excess of 10%, with the average percent of Black persons 12+ being 16.2%. The initial sampling frame in the five test Metros was identical with respect to geography and listed/unlisted telephone characteristics to our Syndicated survey sample. A combination random telephone book and random digit dialing sample was used. The sample was evenly distributed across the weeks of the survey period.

Under Arbitron's Syndicated survey methods, seven-day mailed personal diaries and per-person diary premiums are sent to households where a respondent has agreed to cooperate in a "placement" telephone call made four weeks earlier. At the time of placement, households in areas with greater than a 10% Black population are asked a race/nationality question. Respondents identifying themselves as Black are additionally asked whether a Male 18-34 resides in the household. This information is used to target the appropriate augmented diary premium treatment to persons in each residence.

With two exceptions, all procedures and materials for both Black and non-Black target samples were identical to those used in Arbitron's Syndicated surveys. The two exceptions were:

- o The asking of a Male 18-34 question of all test households, both Black and non-Black, during placement calling; and
- o Per-person diary premium variations.

Once identified as having a Male 18-34, households in each test group were randomly divided into three diary premium treatment groups:

- o A baseline 50¢/\$1 group (with the larger amount going to the historically poorer performing markets per Arbitron's standard methodology);

- o A \$2 per-person group; and
- o A \$5 per-person group.

Persons sent a diary, both Black and non-Black, received controlled follow-up treatments to encourage diary completion and return according to Arbitron's standard survey methodology.

Following the one-week survey period, returned diaries were edited for usability according to Arbitron's established procedures.

#### DATA ANALYSIS

Two industry-accepted measures were utilized to assess premium effects on sample performance and audience estimates. These are presented below along with a brief definition of how they are used at Arbitron.

- o Return Rate: The number of complete/on time diaries received divided by the number of usable agreeing persons in the sample, expressed as a percent.
- o Average-Quarter Hour Persons Using Radio (PUR): The percent of persons listening to radio for at least five minutes during an average quarter-hour in a particular time period. In our Syndicated Market Reports, Arbitron performs sample balancing to compensate for disproportional nonresponse bias among the various demographic groups. The combined Metro average PUR measures shown in this report were sample balanced to equate in-tab age/sex distributions with respect to population distributions.

A total of over 5,000 usable diaries were analyzed in this study. Results were examined for significance using a test of differences between proportions with  $p \leq .05$ . Because each respondent reported their listening over seven days and over all the hours during each day, each of these diaries was actually worth somewhat more than a single unit of analysis for the measurement of listening. Arbitron takes this replication effect into account in testing the significance of the listening results (4).

#### RESULTS -- SAMPLE PERFORMANCE AND REPRESENTATION

Return Rate: Based on previous premium research we expected increases in diary premiums to result in generally progressive return rate gains. Test results were consistent with our expectations (Tables 1 and 2). Premium increases over the 50¢/\$1 baseline gave substantially higher return rates in contrast to the baseline performance results for both the Black and non-Black samples. For both samples there were progressive return gains for each premium increase over the baseline amount. However, the extent of the return rate gain from premium increases appeared to vary with the baseline premium level: the 50¢ baseline groups showed greater increases for the \$2 and \$5 premiums than did the \$1 baseline.

The striking impact of premium variations implies that listening results could be determined, at least in part, by return rate effects, that is, the varying levels of nonresponse across the test

samples. Thus, test results can neither prove nor disprove the "bought listening" hypothesis. However, in two samples, Black Pacific and non-Black East South Central #1, premium variations had a non-significant impact. In these cases, audience results are less contaminated by return rate effects than in the other samples.

The sample performance results imply that progressive premium increases do increase return rates, both from the Black and non-Black samples. While consistent with the results of most premium research, the replication of results between ethnic groups expands the field's knowledge of respondent-premium behavior.

#### RESULTS -- AUDIENCE ESTIMATES

The "bought listening" hypothesis predicts an overall progressive increase in reported listening with progressive increases in premiums. However, on the basis of previous Arbitron findings, we expected to see no differences in reported listening for premium variations. Performance results reported above showed significant and substantial increases in return rates with premium increases. Because of these return rate differences among the premium groups, we would not be able to conclusively disprove the "bought listening" hypothesis, even if we consistently found the expected level of differences in reported listening. The presence of varying levels of nonresponse bias, an element beyond our control, prevents a definitive attribution of results to premiums per se.

Average Quarter-Hour Persons Using Radio (PUR): Premium variations showed no significant differences in PURs between the baseline 50¢/\$1 group and the higher premium groups in either Black or non-Black samples (Tables 3 and 4). Furthermore, there were very few instances of progressive PUR increases for the higher premium groups. In the two samples with non-significant return rate effects, Black Pacific and non-Black East South Central #1, premium increases also showed no significant impact. Black Pacific did show a slight progressive PUR increase across treatment groups, albeit the rise was neither significant or substantial. Premium increases in the non-Black East South Central #1 Metro produced mixed non-significant results, with the \$5 premium actually depressing PURs compared to the baseline.

The PUR results provide virtually no support for the "bought listening" hypothesis. There was no evidence that premium variations significantly or substantially affected reported quarter-hour listening. These results do not disprove "bought listening" because they are contaminated by considerable variations in nonresponse. However, in the two samples where nonresponse was not significant, we saw neither significant or consistently progressive premium effects, again suggesting the lack of support for "bought listening."

#### IMPLICATIONS

Even though we were not able to conclusively disprove the "bought listening" hypothesis, we can present four reasons why there is a lack of evidence for this contention.

- o First, increases in premium amounts did not produce consistently significant, substantial, or progressive increases in audience estimates for the Black sample in this study.
- o Second, the absence of significant, substantial, or progressive increases across the premium treatment groups was replicated in the non-Black sample.
- o Third, in the two test Metros where nonresponse was relatively constant across groups, the audience estimates examined showed no significant effects attributable to premium increases.
- o Fourth, the results obtained in this analysis were in accordance with the findings of two previous Arbitron studies concerning the effects of premium variation on audience estimates.

While no single one of these reasons disproves the response bias "bought listening" hypothesis, they each weaken any argument made on its behalf. When combined, these four reasons indicate a strong lack of support for the contention that

increased premiums change respondents' reported radio listening.

Thus, it appears that augmented monetary premiums can be utilized to reduce nonresponse bias from Black populations with minimal risk of increased response bias. This conclusion is supportive of the main body of premium research results which have found no systematic bias with the use of monetary premiums.

#### REFERENCES

1. Armstrong, J. Scott (1975). "Monetary Incentives in Mail Surveys," Public Opinion Quarterly, 39, 111-6.
2. Kanuk, Leslie, and Conrad Berenson (1975), "Mail Surveys and Response Rates: A Literature Review," Journal of Marketing Research, 12, 440-53.
3. Linsky, Arnold S. (1975), "Stimulating Response to Mailed Questionnaires: A Review," Public Opinion Quarterly, 39, 82-101.
4. Occhiogrosso, Michael G. and Martin R. Frankel (1982), Arbitron Replication II: A Study of the Reliability of Radio Ratings. New York: Arbitron Ratings Company.

**TABLE 1**  
Return Rate  
Persons 12+ in Black Households with a Male 18-34  
Individual and Combined Metros

	.50/\$1	\$2		\$5	
		RETURN RATE	DIFF VS .50/\$1	RETURN RATE	DIFF VS .50/\$1
<b>.50 Premium Metros†</b>	%	%		%	
East South Central No. 1	39.3	54.8	15.5*	63.2	23.9*
East South Central No. 2	15.9	47.7	31.9*	55.0	39.1*
South Atlantic	30.8	43.9	13.1*	42.3	11.5*
<b>\$1 Premium Metros‡</b>					
East North Central	40.1	37.3	- 2.8*	54.6	14.5*
Pacific	36.0	42.2	6.2	29.7	- 6.3
<b>Combined Metros</b>	33.6	43.4	9.8*	49.4	15.8*
(Consenting Persons)	(934)	(1000)		(1139)	

**FOOTNOTES**

- † — Metros where listed, non-ethnic sample receive a .50 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- ‡ — Metros where listed, non-ethnic sample receive a \$1 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- \* — Statistically significant at the 95.0% level of confidence.

**TABLE 3**  
Average Quarter-Hour Persons Using Radio Ratings  
Persons 12+ in Black Households with a Male 18-34  
Individual and Combined Metros

	.50/\$1	\$2		\$5	
		RATING	DIFF VS .50/\$1	RATING	DIFF VS .50/\$1
<b>.50 Premium Metros†</b>					
East South Central No. 1	17.2	23.5	6.3	19.6	2.4
East South Central No. 2	— ‡	19.8	—	24.6	—
South Atlantic	20.7	23.5	2.8	26.4	5.7
<b>\$1 Premium Metros‡</b>					
East North Central	22.6	22.4	- 0.2	20.9	- 1.7
Pacific	23.5	25.6	2.1	26.9	3.4
<b>Combined Metros</b>	22.2	22.8	0.6	22.9	0.7
(In-tab)	(312)	(434)		(557)	

**FOOTNOTES**

- † — Metros where listed, non-ethnic sample receive a .50 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- ‡ — Rating not shown; base less than 30.
- § — Metros where listed, non-ethnic sample receive a \$1 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- \* — Weighted to equate in-tab age/sex distributions.
- NB — None of the differences is significant at the 95.0% level of confidence.

**TABLE 2**  
Return Rate  
Persons 12+ in Non-Black Households with a Male 18-34  
Individual and Combined Metros

	.50/\$1	\$2		\$5	
		RETURN RATE	DIFF VS .50/\$1	RETURN RATE	DIFF VS .50/\$1
<b>.50 Premium Metros†</b>	%	%		%	
East South Central No. 1	43.3	48.5	5.2	51.4	8.1
East South Central No. 2	39.8	55.8	16.0*	69.3	29.5*
South Atlantic	50.4	64.8	14.4*	64.7	14.3*
<b>\$1 Premium Metros‡</b>					
East North Central	53.8	58.8	5.0	65.5	11.7*
Pacific	42.2	47.8	5.6*	57.1	14.9*
<b>Combined Metros</b>	46.3	54.1	7.8*	61.1	14.8*
(Consenting Persons)	(4315)	(1874)		(1795)	

**FOOTNOTES**

- † — Metros where listed, non-ethnic sample receive a .50 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- ‡ — Metros where listed, non-ethnic sample receive a \$1 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- \* — Statistically significant at the 95.0% level of confidence.

**TABLE 4**  
Average Quarter-Hour Persons Using Radio Ratings  
Persons 12+ in Non-Black Households with a Male 18-34  
Individual and Combined Metros

	.50/\$1	\$2		\$5	
		RATING	DIFF VS .50/\$1	RATING	DIFF VS .50/\$1
<b>.50 Premium Metros†</b>					
East South Central No. 1	18.9	19.8	0.9	17.2	- 1.7
East South Central No. 2	15.5	22.1	6.6	18.5	3.0
South Atlantic	19.3	18.4	- 0.9	18.3	- 1.0
<b>\$1 Premium Metros‡</b>					
East North Central	18.3	19.6	1.3	19.0	0.7
Pacific	16.8	17.3	0.5	17.4	0.6
<b>Combined Metros</b>	17.8	18.7	0.9	18.0	0.2
(In-tab)	(1984)	(1012)		(1096)	

**FOOTNOTES**

- † — Metros where listed, non-ethnic sample receive a .50 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- ‡ — Metros where listed, non-ethnic sample receive a \$1 diary premium. All unlisted telephone sample receive a \$1 diary premium.
- \* — Weighted to equate in-tab age/sex distributions.
- NB — None of the differences is significant at the 95.0% level of confidence.