

Cell Phone Only Research at Arbitron: Statistical Analyses

Richard Griffiths¹

¹Arbitron, Inc., 9705 Patuxent Woods Drive, Columbia, MD 21046

Abstract

As the size of the U.S. population that lives in households with no landlines has burgeoned, Arbitron has undertaken a plan of study aimed at finding methods to mitigate potential bias due to frame noncoverage of this population in their radio audience measurement surveys. This plan of study encompasses a number of different facets, including an examination of the cost-effectiveness of sampling the cell phone-only (CPO) population. This paper reports the results of statistical analyses designed to study the characteristics of the CPO population and how the information gleaned from these analyses might be used to improve the quality of Arbitron's audience measurement estimates in the face of the CPO issue.

Key Words: Cell Phone Only, Noncoverage, Survey, Sampling, Weighting

1. Introduction

To produce estimates of radio listening audiences in the United States, Arbitron divides the country into about 300 geographical areas called markets. Arbitron then conducts surveys of a sample of households in each market. In several of these markets, Arbitron has begun including cell phone-only (CPO) households as part of the sample.

As the size of the U.S. population that lives in households with no landlines has burgeoned, Arbitron has undertaken a plan of study aimed at finding methods to mitigate potential bias due to frame noncoverage of this population in their radio audience measurement surveys. This plan of study encompasses a number of different facets, including an examination of the cost-effectiveness of sampling the CPO population. This poster-presentation paper reports the results of statistical analyses designed to study the characteristics of the CPO population and how the information gleaned from these analyses might be used to improve the quality of Arbitron's audience measurement estimates in the face of the CPO issue.

Analyses pertaining to the following questions are included in this poster-presentation paper:

1. What do the Arbitron cell phone-only and landline samples look like?
2. What types of radio station formats do cell phone-only respondents listen to and how do they differ from those of landline respondents?
3. Is there geographic information that can be used to make sampling for cell phone-only households more efficient?
4. Is there variation in radio listening that is explained by cell phone/landline status and is not already explained by Arbitron's weighting variables?
5. What is the effect on weighting efficiency of undersampling cell phone-only households?

For the most part, the graphs in the succeeding sections tell the stories of how these questions were answered. The accompanying text provides a brief, running commentary on those stories.

2. What Do the Arbitron Cell Phone-only and Landline Samples Look Like?

Figure 1 below shows the age distributions of the CPO and landline respondents during June 2007. From the graph, we see that CPO respondents were more likely to be young than their landline counterparts.

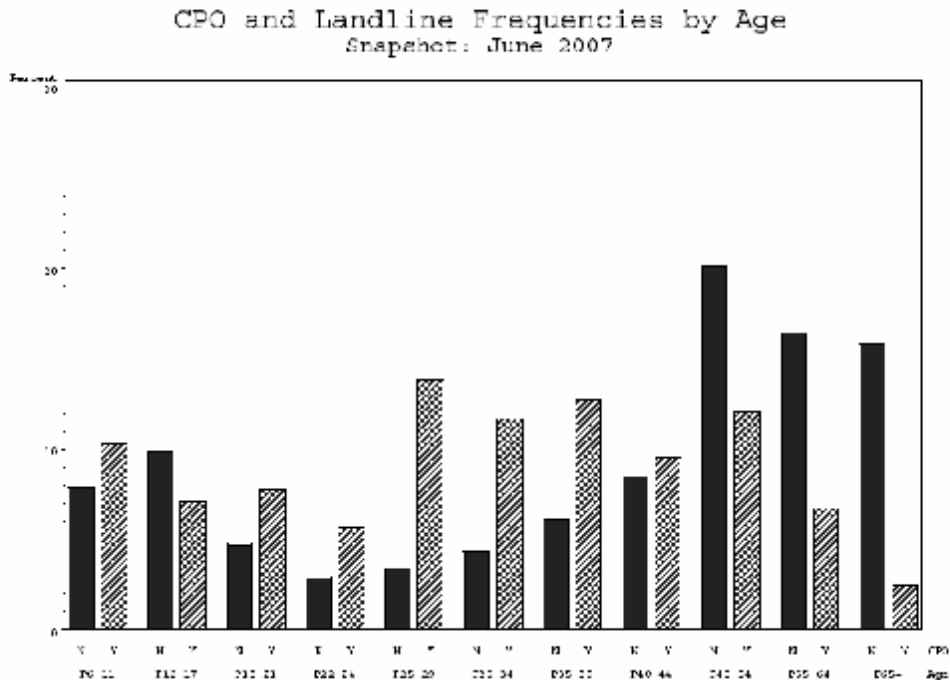


Figure 1: Age Distributions of the CPO and Landline respondents.

Figure 2 shows the race/ethnicity (Black, Hispanic, Other) distributions of the CPO and landline respondents during June 2007. From the graph, we see that CPO respondents were more likely to be black or Hispanic than their landline counterparts.

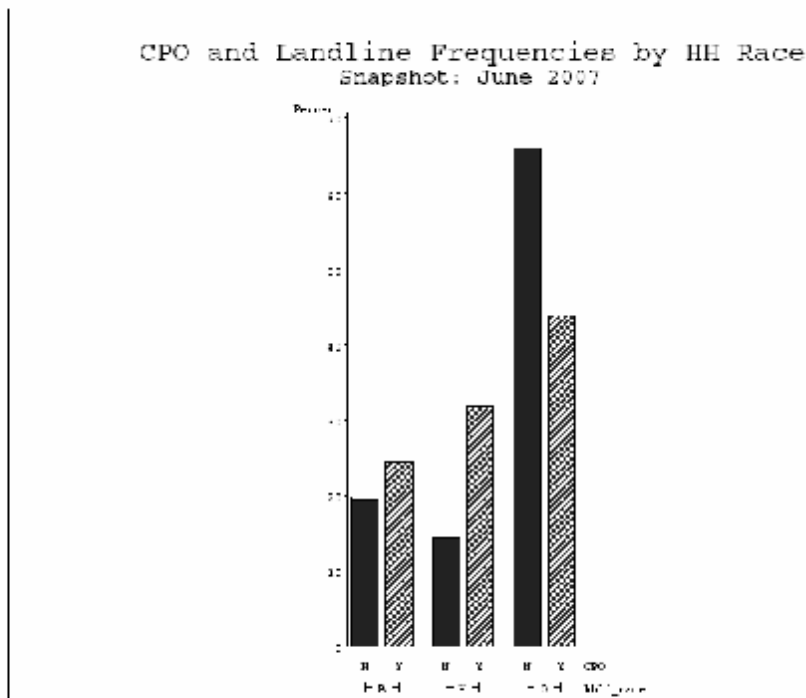


Figure 2: Race/Ethnicity distributions of the CPO and Landline respondents.

3. What Types of Radio Formats Do Cell Phone-only Respondents Listen to and How Do They Differ from Those of Landline Respondents?

The following figure displays the radio station formats that landline and CPO respondents listened to during the April to June 2007 survey period. It shows that CPO respondents were quite a bit more likely than landline respondents to listen to Contemporary Hit Radio, Spanish, and Urban format stations and less likely to listen to News/Talk/Sports stations. This is reflective of the differences in the age and race/ethnic distributions between the two groups.

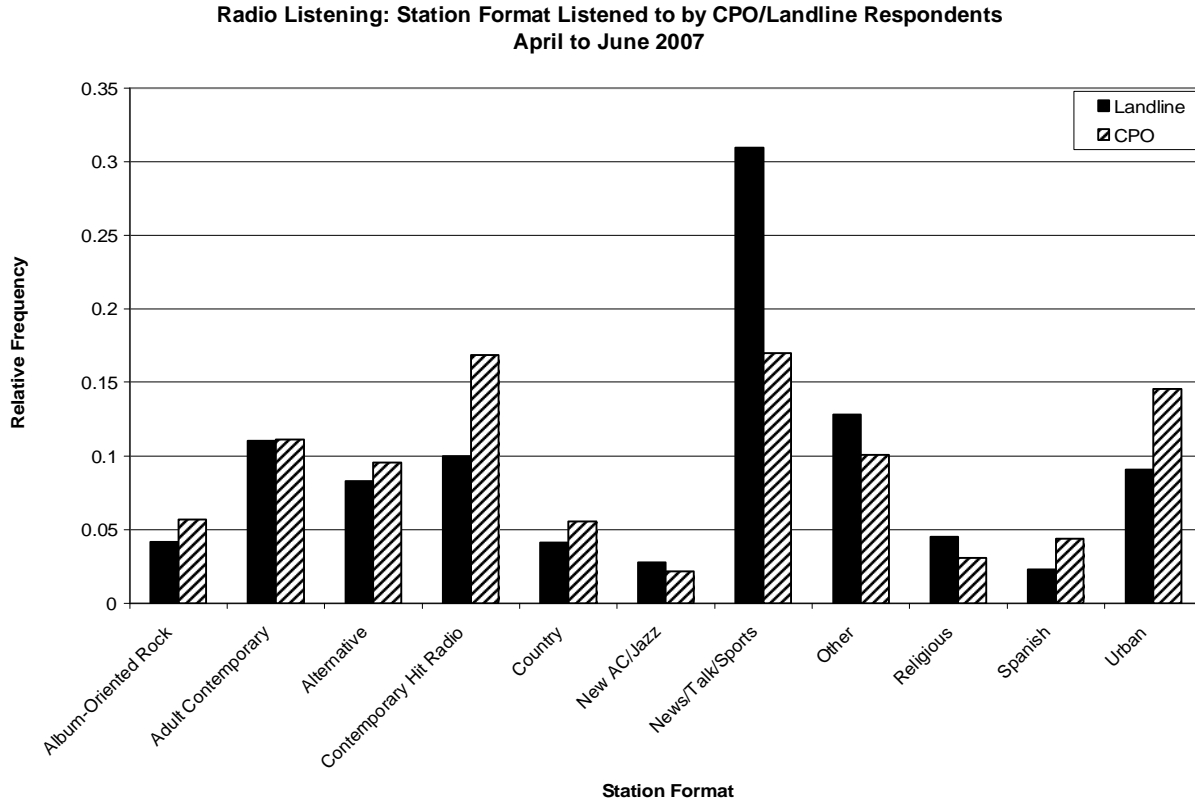


Figure 3: Listening by Radio Station Format.

4. What is the Effect on Weighting Efficiency of Undersampling Cell Phone-only Households?

To control the costs associated with sampling the CPO population, Arbitron looked at undersampling CPO households relative to landline households. As part of that investigation, we examined the effect that undersampling would have on the variability of the weights as calculated by the sample weighting efficiency (or, statistical efficiency). The sample weighting efficiency was calculated as given in the following formula:

$$Stat\ Eff = 100 \cdot \left[\frac{n \cdot \sum_{i=1}^n w_i^2}{\left(\sum_{i=1}^n w_i \right)^2} \right]^{-1},$$

where n is the sample size and w_i is the weight for the i^{th} respondent.

The following figure shows the results of a simulation study that examined this effect. In this simulation, we examined the effect of undersampling the CPO population at different levels (sampling rates of two-thirds, one-half, and one-third the population percent) at different population CPO penetration levels (5, 10, 15, 20, and 25 percent).

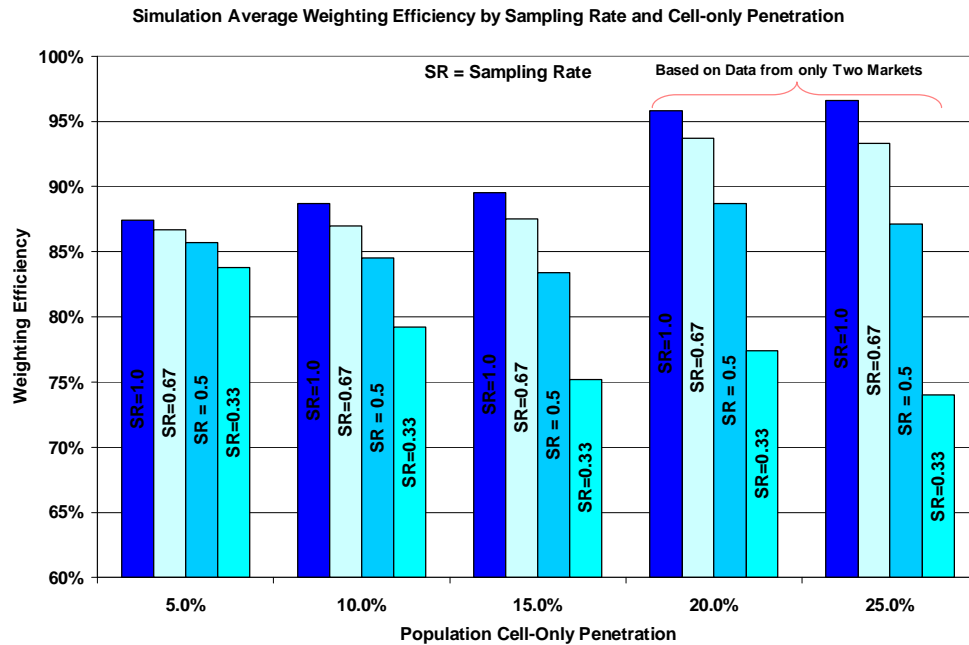


Figure 4: Effect on Weighting Efficiency of Undersampling CPO Households.

The graph shows that the effect on weighting efficiency increases as the size of the population penetration rate increases. At low penetration rates, even sampling at one-third of the population penetration percent had only a mild effect on weighting efficiency. At the largest CPO population penetration rates, even sampling at one-half the penetration rate resulted in notable losses in weighting efficiency.

5. Is There Variation in Radio Listening that is Explained by Cell Phone/Landline Status and is not Already Explained by Arbitron’s Weighting Variables?

An important question to answer in Arbitron’s investigation of CPO sampling was whether or not we could adjust the sample through weighting to account for an undercoverage (or noncoverage) of CPO households. To that end, we performed some regression analyses. The goal of these analyses was to determine how much variation in respondent listening was explained by CPO/Landline status of the respondents but not explained by the typical Arbitron weighting variables.

The following figures show the cumulative R² values from these regression analyses as we add in dependent variables. The leftmost bar in each graph shows the R² value when only market is used to explain the variation in listening levels. The next bar shows the amount that both market and gender explain. This proceeds through a succession of bars, one for each weighting variable, and ends with a bar for CPO/Landline status. By examining the height of this last bar relative to the one to its left, we can see how much variation is being explained by CPO/Landline status beyond that already explained by the other weighting variables.

Each graph shows the results for a different radio format.

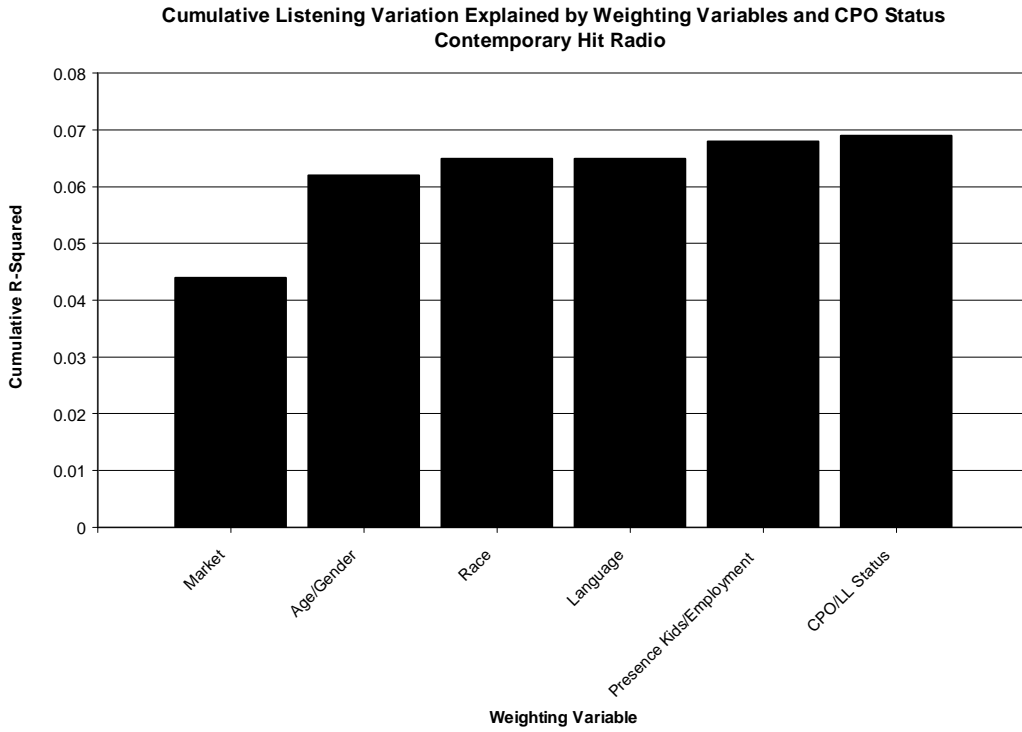


Figure 5: Listening Variation: Contemporary Hit Radio.

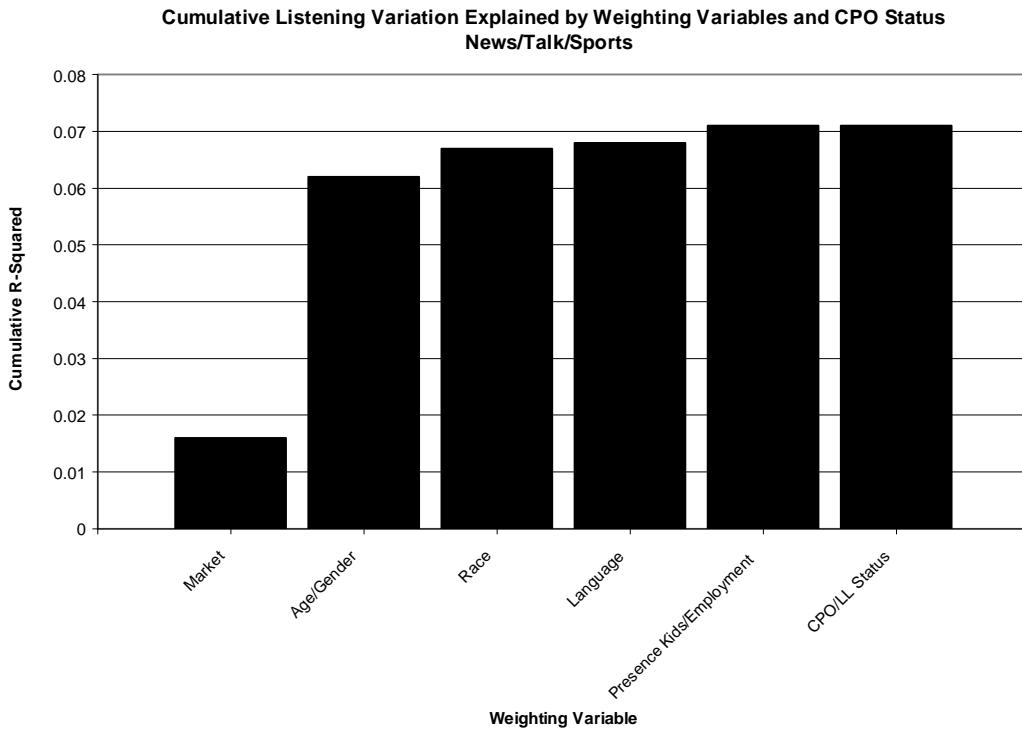


Figure 6: Listening Variation: News/Talk/Sports.

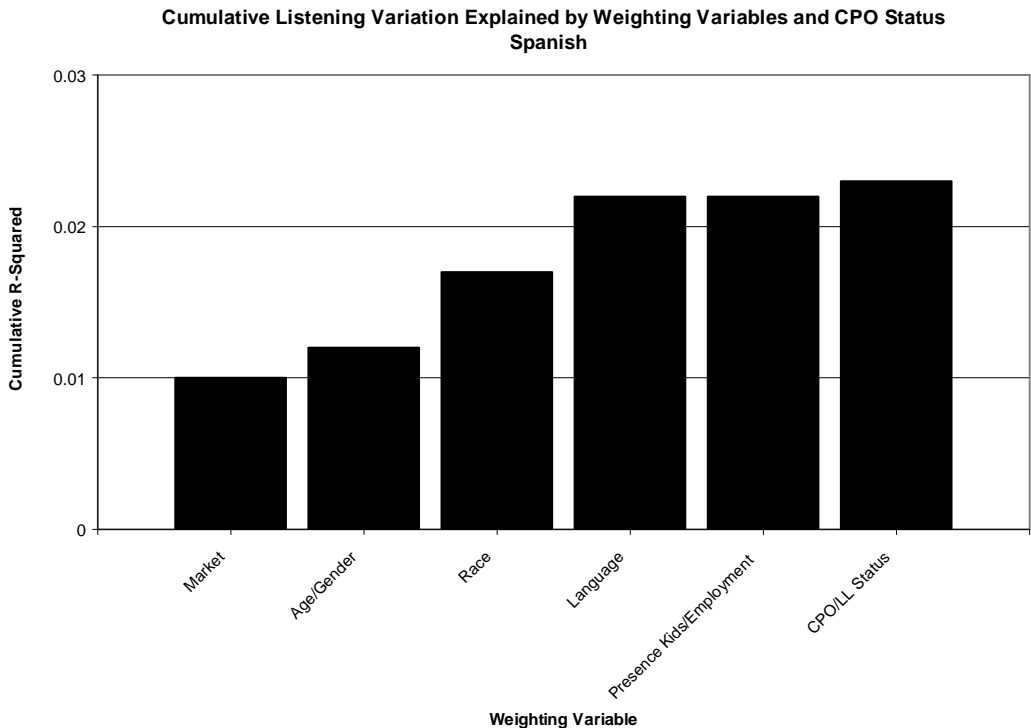


Figure 7: Listening Variation: Spanish Format.

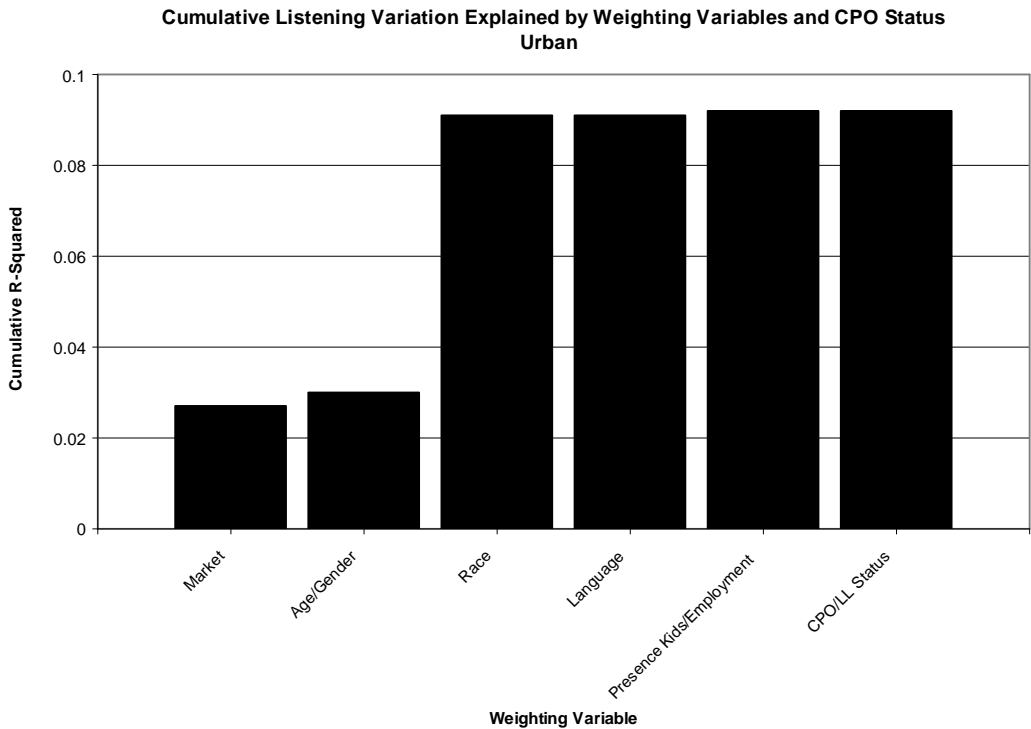


Figure 8: Listening Variation: Urban.

The results indicate that very little variation in radio listening is explained by CPO/Landline status after accounting for the other weighting variables.

6. Is There Geographic Information that Can be Used to Make Sampling for Cell Phone-only Households More Efficient?

The following figures offer images of the distribution of the Arbitron CPO household sample relative to Census 2000 household counts on several variables for which we expected the distributions to be different. To the extent that the distributions are different, the variables offer a possible means of stratification for identifying large concentrations of CPO households. This information could be used in a sampling plan designed to efficiently sample the CPO population.

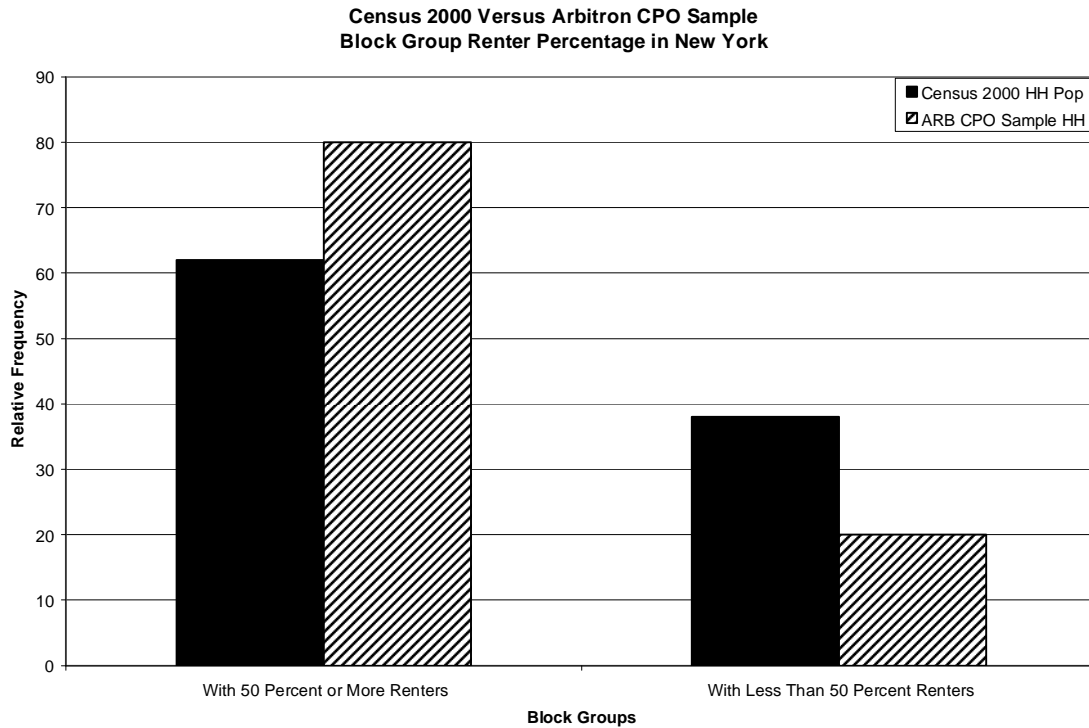


Figure 9: Census 2000 V. Arbitron CPO Respondents: Block Group Renter, New York.

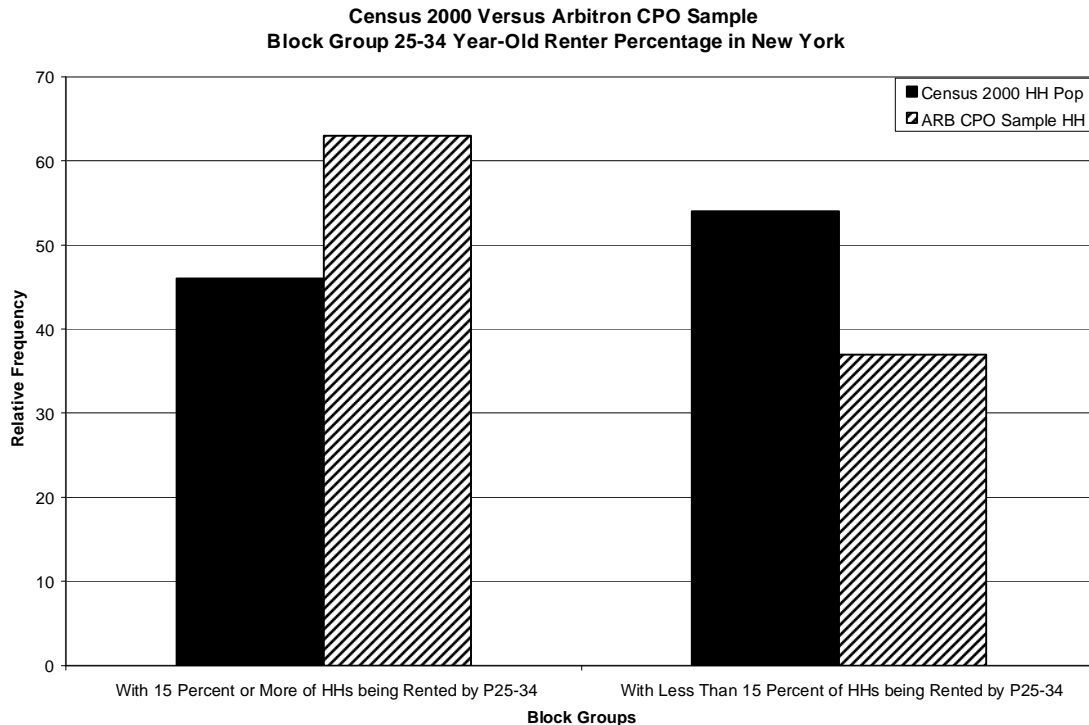


Figure 10: Census 2000 V. Arbitron CPO Respondents: Block Group 25-34 Year-old Renters, New York.

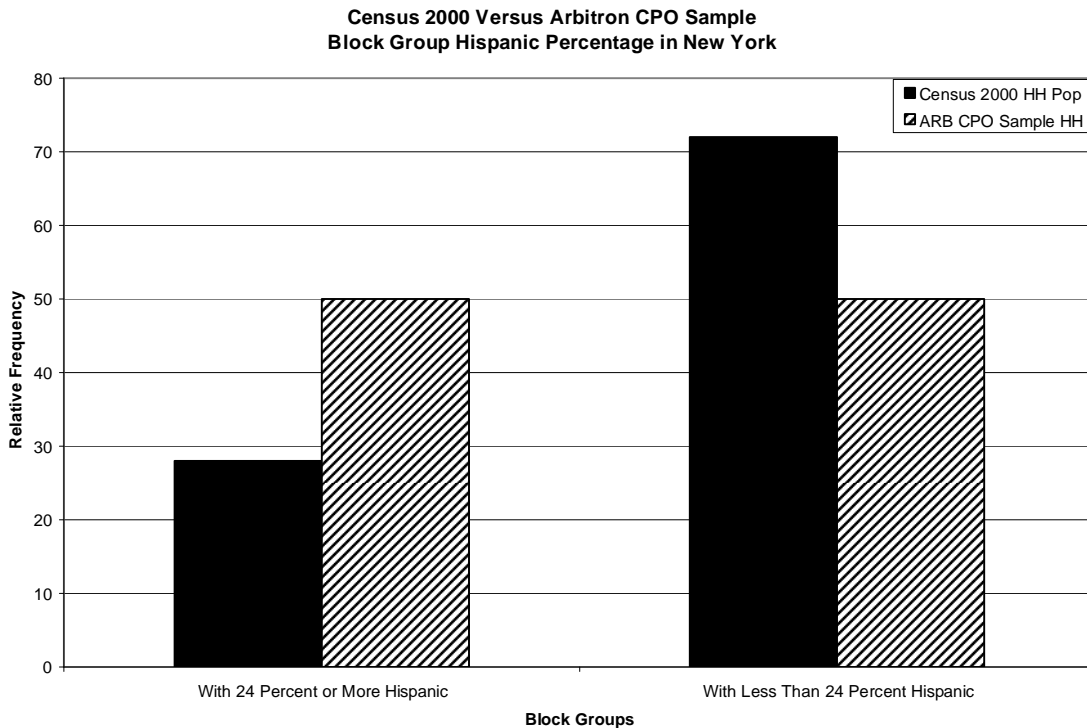


Figure 11: Census 2000 V. Arbitron CPO Respondents: Block Group Hispanic Percentage, New York.

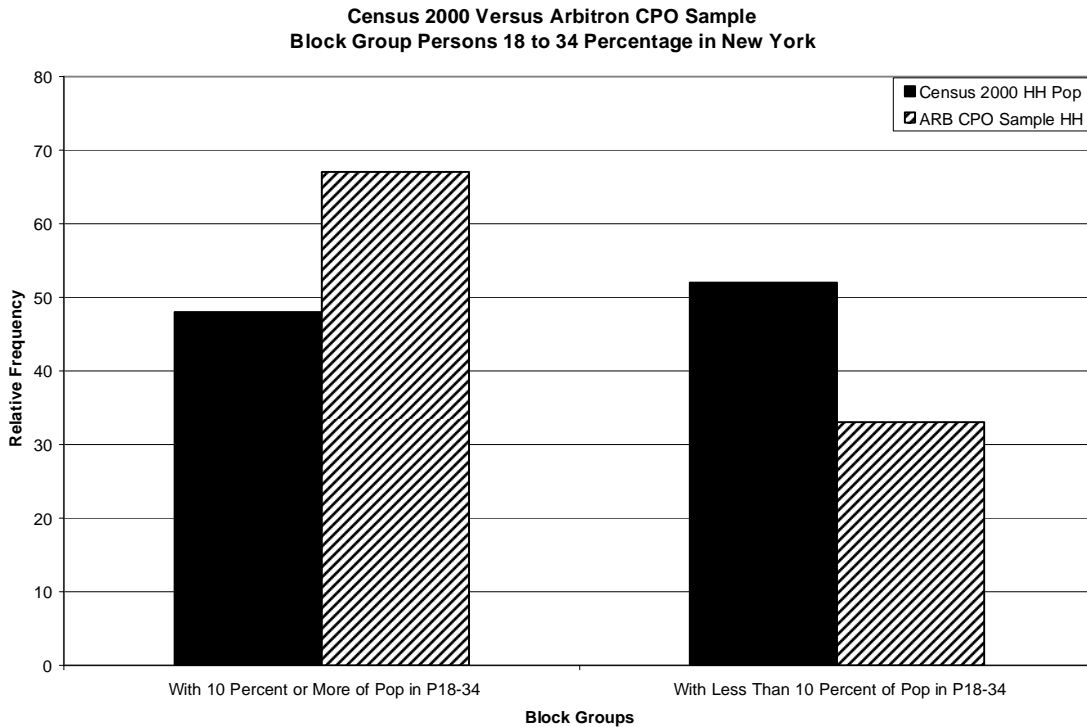


Figure 12: Census 2000 V. Arbitron CPO Respondents: Block Group 18-34 Percentage, New York.

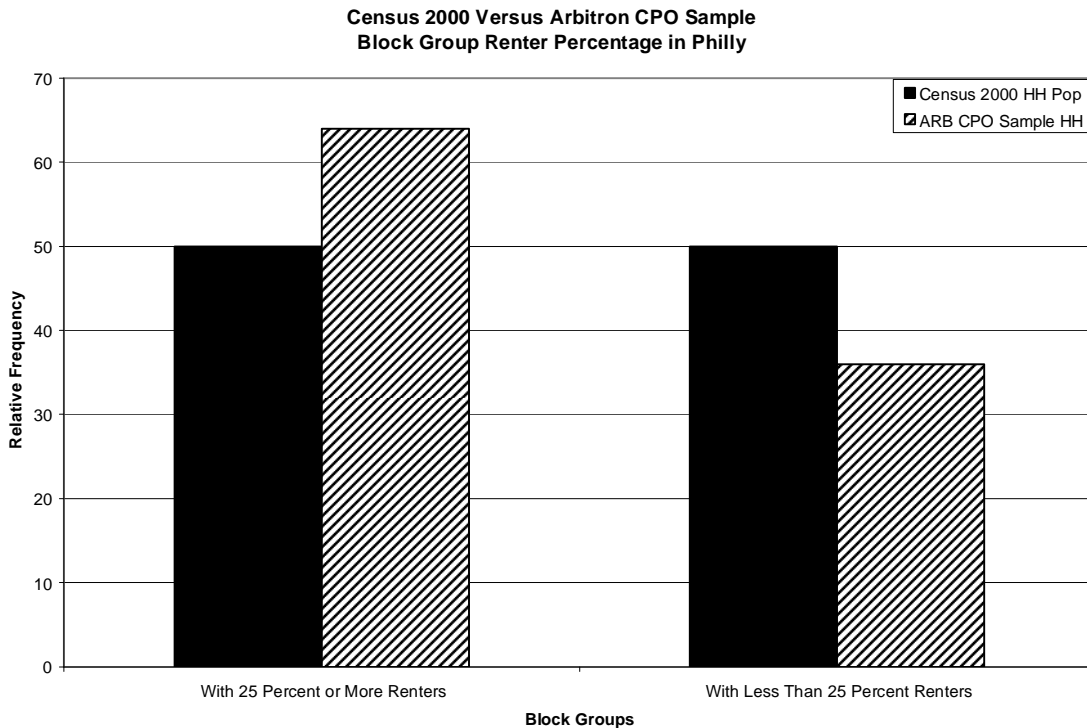


Figure 13: Census 2000 V. Arbitron CPO Respondents: Block Group Renter Percentage, Philadelphia.

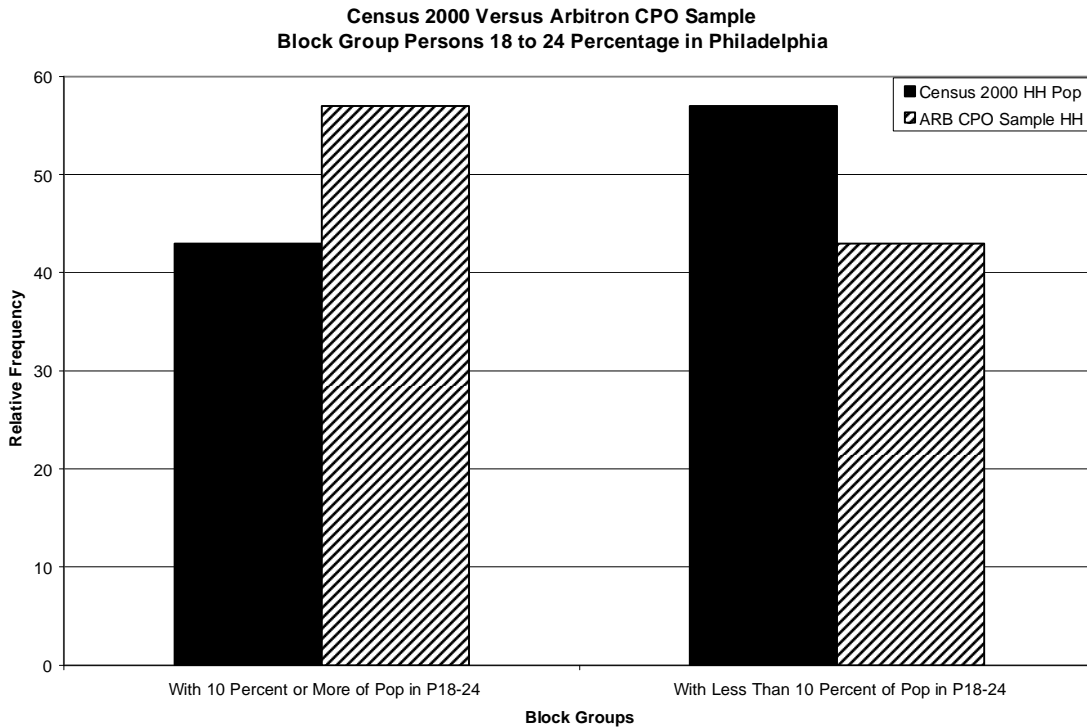


Figure 14: Census 2000 V. Arbitron CPO Respondents: Block Group 18-24 Percentage, Philadelphia.

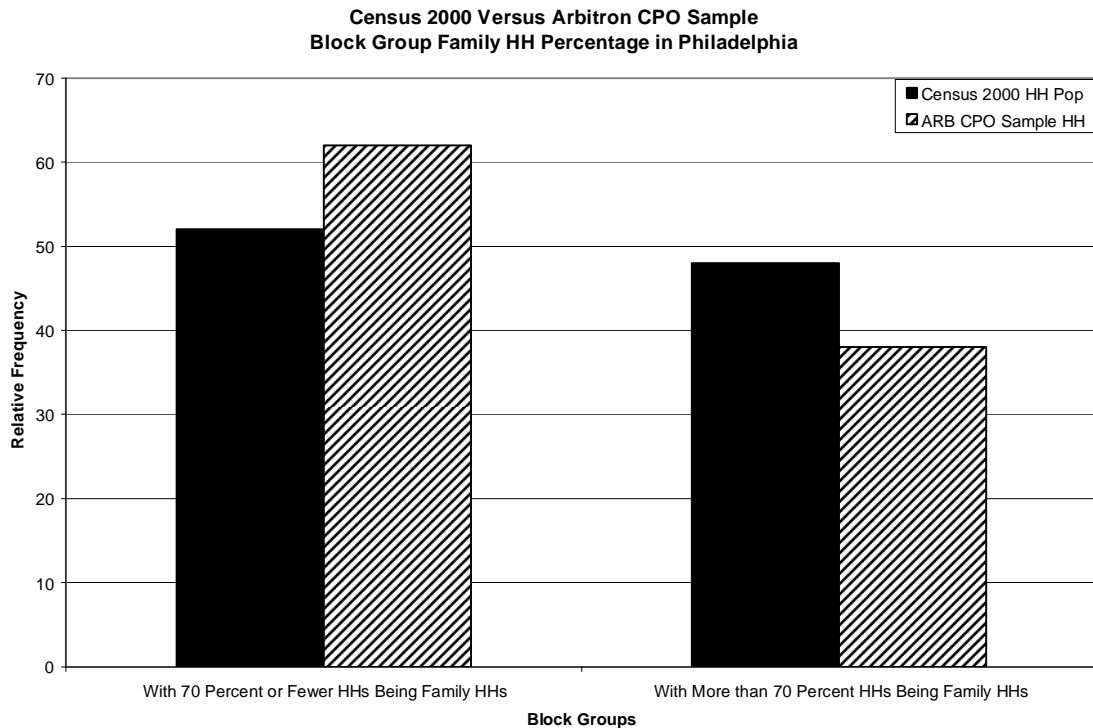


Figure 15: Census 2000 V. Arbitron CPO Respondents: Block Group Family Household Percentage, Philadelphia.

7. Discussion of Results

The following are some of the key points we have taken from the above analyses:

- In terms of estimation of radio ratings, while there are differences in radio listening behavior between cell phone-only and landline respondents, much of those differences are related to factors that Arbitron already uses in weighting its samples.
 - For instance, relative to their landline counterparts, cell phone-only respondents tend to listen to more contemporary hit radio than they do to News/Talk/Sports radio. This reflects the age difference between cell phone-only and landline respondents. Age is one of Arbitron's weighting variables.
- In terms of sampling, we see that it is feasible to sample cell phone-only households at a lower rate than landline households without impacting estimator precision greatly. This represents a potential cost-savings since cell phone-only sample is generally more expensive.
- However, we also see that some cost-savings may be possible if we can effectively use geographic information to help us locate cell phone-only respondents more efficiently.
 - This information cannot be readily accessed with RDD-based frames for cell phone-only sampling. It is more useful for address-based frames.