The Health and Retirement Study: Seeking More Effective Ways to Identify the Next Birth Cohort

Frost Hubbard, Daniel Tomlin, Theresa Camelo
Institute for Social Research, 426 Thompson St., Ann Arbor, MI 48106

Abstract

Interviewer administered household surveys with age specific target populations generally have three options for identifying age eligible households: (1) full enumeration of all household members and their ages; (2) directly asking the household informant if anyone in the household is in the desired age range; and (3) asking a few, brief questions about the age range of household members to isolate if the household is eligible without explicitly identifying the desired age range. Tourangeau et al. (2012) conducted a centralized telephone experiment examining these three methods and found that the full enumeration method produced the highest eligibility rates and lowest completion rates, indicating a tradeoff between coverage and non-completion errors when choosing between the three methods. In an experiment sponsored by the Health and Retirement Study, we attempted to replicate these findings for an in-person screening of persons in a 6 year age cohort. Our results showed trends that the full enumeration of all household members produced higher eligibility and similar completion rates compared to the other two methods. Potential reasons for these findings will be discussed.

Keywords: Coverage, Nonresponse, Household rostering methods

INTRODUCTION

Given the rising costs of survey research activities, primarily driven by decreased completion rates, and specifically the increased costs related to screening activities, the National Institute of Aging (NIA) provided funding to explore more cost efficient ways for the Health and Retirement Study (HRS) to identify age-eligible cohorts. The knowledge gained from these explorations will be used to update the HRS sample design and birth cohort screening protocols. To develop a plan for improving screening efficiency in 2016, the HRS and University of Michigan Survey Research Center (SRC) are identifying areas to improve the HRS screening method and procedures. This paper documents one of these explorations, the HRS Screening Experiment.

The main goal of the HRS Screening Experiment was to determine the household (HH) rostering method that achieves the best balance of high HH age eligibility rates and screening completion rates. In most SRC cross-sectional studies, the full HH rostering method has achieved approximately the same eligibility rates as expected based on the latest population data, yet has resulted in continually decreasing screening completion rates. One potential reason for these decreasing screening cooperation rates is the length of the HH rostering procedure, especially for the HRS.[1] On the other hand, asking the informant directly if there is anyone in the HH in the desired age range would likely lead to higher completion rates, since it is one question, yet potentially generate lower

[1] To determine if a HH is eligible for the study, the HRS screening procedure requires informants to provide the names and ages of all HH members age 18 and above, the dates of birth and relationship status of all HH members between the ages of 45-59, and the race of the selected respondent.

eligibility rates because the informant may be more likely to say that no one in the HH is eligible for the study because they do not want the HH to participate.

As survey completion rates continue to decline, practitioners have been searching for ways to boost or simply maintain completion rates. Tourangeau et al. (2012) recently found that for a telephone study looking for persons within a specific age range, having one to two brief questions that identify whether a household contains an age eligible member, rather than insisting on all selected units completing a full household roster, helped boost screener completion rates. The authors used three versions of screening questions to encompass all the different approaches to screening: a *direct question* to identify if household members fell into a target age group, a *complement question* which asked household members if they were above or below age range limits, and a *full roster* version to collect information on each household member without identifying an age range. As predicted, the full roster version yielded the highest eligibility. Unfortunately, this version simultaneously yielded the lowest screening completion rates. Our experiment used a different age range and a lengthier screening questionnaire than the Tourangeau et al. study. Additionally, all interviews we conducted were done in-person as opposed to the telephone method used above. Given that the authors noticed a decline in completion rates in the group with the highest eligibility, we hoped to find a method that would strike the best balance of high screening completion rates and eligibility rates close to what would be expected from population estimates. Ideally, these factors would play a significant role in controlling screener completion rate decline while simultaneously yielding high eligibility rates.

We can also look at interview avoidance from a social environmental perspective. Groves and Couper (1998) discovered a tendency for those living in high cost housing to refuse survey requests while simultaneously finding a strong tendency of racial and ethnic minority participation. Additionally, they postulated that high crime rates in high density urban areas may lead to a general unwillingness to provide information to strangers. The authors suggested that the findings could ultimately be explained by socioeconomic status between the racial and ethnic groups. One could also theorize that income may, in fact, be a more relevant component. Our sample design, as described in detail in the Methods section, purposefully selected areas with high densities of Black, Hispanic, and White populations which also happened to have a wide range of median household incomes, allowing us to similarly analyze these relationships.

Another major goal of our study was to find a way to improve eligibility while simultaneously reducing costs. An effective method of achieving this goal is the use of commercial lists in developing the sample. These residential address lists are sold by a variety of commercial vendors and are becoming a widely used tool for survey organizations. According to Valliant et al. (2014), sampling households using commercial lists can effectively reduce costs and identify subgroups within a population in spite of the fact that information on these lists can be incomplete. With the intent of more efficiently identifying LBB households, we oversampled addresses which a commercial vendor indicated contained at least one LBB member. We hope to not only corroborate the Valliant et al finding that this commercial data can accurately identify subgroups, but also to expand on those findings to see how the income of a specific area may affect the accuracy of the data provided by commercial vendors. Since the commercial data vendors rely heavily on credit card purchases and home ownership data, we hypothesized that the MSG data would most accurately stratify households into LBB

eligible and not LBB eligible groups in high median income areas compared to low median income areas.

METHODS

EXPERIMENTAL DESIGN

To determine if a HH is eligible for the HRS and select a financial unit (i.e., unmarried resident or married/partnered couple) within an eligible HH, an interviewer must obtain the following pieces of information in sequential order about the household residents:

1. First name, last name[2] and age of all HH residents age 18 and above

2. Year of birth of anyone in HH age 40-59[3]

3. Married/partnered status and to whom for all in eligible age range based on year of birth

4. Race/ethnicity of the selected respondent as a proxy for the race/ethnicity of the HH

The experimental design contained three groups, as shown in Figure 1. In the first group, or Path 1 in Figure 1, interviewers asked informants the following three questions:

How many people in the household are between the ages of:

a) 18-39?

b) 40-59?

c) 60 and over?

If the informant said that no persons were between the ages of 40-59, the interviewer obtained the informant's name, verified the address for quality control purposes and stopped the screening, without obtaining a full household roster. However, if the informant answered that one or more household residents were between the ages of 40-59, the interviewer proceeded to obtain a full household roster.

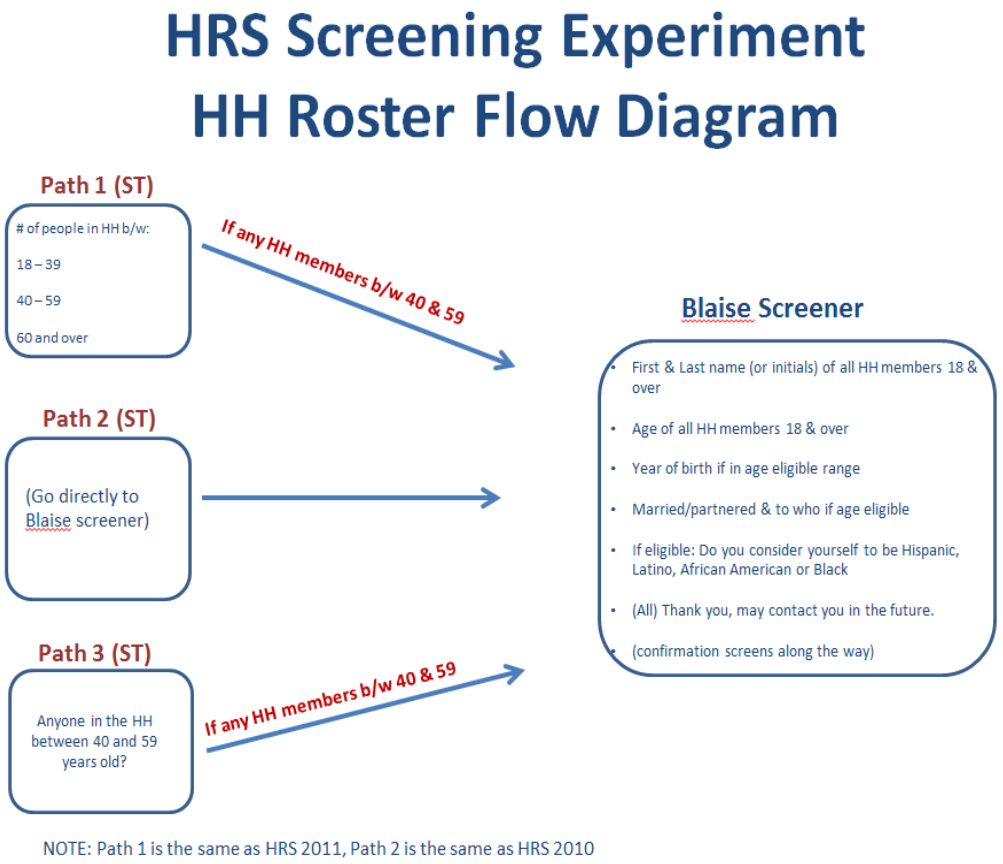
[2] HRS allows interviewer to collect initials rather than full first and last names if the informant was wary of providing full names, but discourages the collection of initials during interviewer training.

[3] We chose a wider age range than the LBB birth cohort (born between the years of 1960-65) to avoid potential measurement error based on discrepancies between the reported age and actual year of birth. These discrepancies had the potential to reduce our LBB eligibility rate. For example, imagine the informant lives with her brother and cannot remember his exact age at the moment when asked and reports his age to be approximately 45. When next asked his year of birth, the informant can recall that easily and answers 1965 (age 48).

In the second group (Path 3 in Figure 1), the interviewer asked the informant directly, "Is anyone in the household between 40 and 59 years old? If the informant answered no, the interviewer confirmed the informant's name and address for quality control purposes and ended the screener. If the informant answered yes, the interviewer obtained a full household roster.

In the final group (Path 2), the interviewer went straight to the process of obtaining a full household roster. Cases assigned to this experimental group were not asked any age eligibility questions.

Figure 1. HRS Screening Experiment HH Roster Flow Diagram



As discussed in the introduction, the key statistics of interest were the screening completion rate and eligibility rate. Our overall hypothesis was that Path 1, the "Three Age Range" experimental group, would exhibit the best balance of high screener completion rates and high age eligibility rates among the three experimental groups. Specifically, we predicted that the three age range (path 1) and one age range (Path 3) groups would have higher screening completion rates than the "Straight to Full HH Roster screener" group because the lure of answering just a few short questions rather than a full HRS screening would lead to higher completion rates. In addition, we thought that the "One Age Range" group, would have a lower eligibility rate than the other two groups because informants in eligible households would be more likely to lie and report no age eligible household members, resulting in the HH not being asked to participate in a future study.

SAMPLE DESIGN

FIRST STAGE OF SELECTION (PSUS)

The three PSUs of LA County, Denver County and Dallas County were hand-picked because the major cities in each of these counties had a significant percentage of African-Americans and Hispanics. In addition, the SRC had at least 5 on-staff interviewers in each county with interviewing experience. This made hiring and training much easier and eliminated recruiting costs.

SECOND STAGE OF SELECTION (SSUS OR SEGMENTS)

Like the PSUs, we purposively chose the three segments in each PSU. In other words, the segments did not have a random chance of selection. From the latest American Community Survey (ACS) data, we obtained the population counts of all the tracts within each of the three PSUs by race/ethnicity (White, Black, and Hispanic) and LBB age eligibility (born between 1960-1965). Within each PSU, we then purposively selected the three tracts within the county that had the highest percent of LBB-Hispanics, LBB- Non-Hispanic Blacks, and LBB-Non-Hispanic Whites. We did this to ensure that large enough numbers of each group were asked to complete the screener in order to compare the three race/ethnicity groups on the three separate HH screening techniques.

We conducted screeners in both English and Spanish to follow the expected screening procedure for screening in the LBB birth cohort in 2016. As such, the questionnaire was designed to enable screening in whatever language is most convenient for the selected HH. We required interviewers assigned to the "High Hispanic" segments to be bilingual. Though we selected a "Denver, High Hispanic" segment and addresses within that segment, we never released those addresses for data collection because we did not have a Spanish speaking interviewer in Denver. Additionally, to reduce the potential for interviewer effects, we aimed for an interpenetrated design with two interviewers assigned to each segment at the beginning of data collection. While we achieved this in Los Angeles (3 segments, 7 interviewers) and Denver (2 segments, 4 interviewers), we fell short of the goal in Dallas (3 segments, 5 interviewers).

While we initially chose tracts based on the density of persons of a specific race/ethnicity within the tract, we realized during data collection that this method also created three distinct groups based on the affluence of the residents within the tract (as shown in Table 1). To measure affluence, we used the 2008-2012 American Community Survey tract level measure for median HH income. The median HH income in the tracts we chose in Dallas, Denver and Los Angeles with the highest density of LBB Non-Hispanic Whites ranges from $106,000 to $223,000. We classified these three tracts as the "High Income" group. The three

tracts we identified as the "Medium Income" group were: (1) the Los Angeles, high density LBB Non-Hispanic Black, (2) Los Angeles, high density LBB Hispanic, and (3) the Denver high density LBB Non-Hispanic Black. Each of these tracts had median HH incomes ranging from $34,000 to $43,000. Finally, we defined the "Low Income" group of as the two remaining Dallas tracts (high density LBB Non-Hispanic Black and high density LBB Hispanic) whose median incomes ranged from $22,000 to $26,000. Since we present results by these three groups in the Results section, we provide the definition of the three income groups in Table 1 below as a quick reference for the reader.

Table 1: Income Group Classification

Income Group	Range of ACS 2008-2012 Median HH Income	Tracts in Group
High	$106,000 to $223,000	1. Dallas high density LBB Non-Hispanic White 2. Denver high density LBB Non-Hispanic White 3. Los Angeles high density LBB Non-Hispanic White
Medium	$34,000 to $43,000	1. Los Angeles high density LBB Non-Hispanic Black 2. Los Angeles high density LBB Hispanic 3. Denver high density LBB Non-Hispanic Black
Low	$22,000 to $26,000	1. Dallas high density LBB Non-Hispanic Black 2. Dallas high density LBB Hispanic

THIRD STAGE OF SELECTION (ADDRESSES WITHIN SEGMENT STRATA)

As a test of Marketing Systems Group's (MSG) ability to identify HHs likely to be eligible for the Late Baby Boomer (LBB) birth cohort, we oversampled addresses based on the expected age eligibility of household members. To do so, we created 2 strata – (1) expected age eligible and (2) expected age ineligible/eligibility unknown - from the MSG age data we received when we purchased USPS addresses in the nine specific segments.

In seven of the nine segments, we selected *all* of the cases that were identified as LBB Eligible from MSG commercial data to be a part of the study. Previous analyses based on the HRS 2010-11 data collection results found that addresses to which MSG could not match any demographics were much more likely to be unoccupied housing units than addresses where at least some demographic data could be matched (Valliant et al., 2014). Thus, we could have created three strata

within each segment design (expected LBB eligible, not expected LBB eligible, and no commercial data match). In fact, the third stage of selection in the HRS 2016 sample design will have more than just these two strata, including a 'no data match' stratum. However, with such a small sample size for *this* experiment and under the idea that identifying and screening out unoccupied housing units would not cost too much, we decided not to create a separate stratum for 'no data match'. Instead, we put all of the 'no data match' cases into the expected age ineligible/eligibility unknown stratum.

FOURTH STAGE OF SELECTION: (SUBSAMPLING NONFINAL ADDRESSES AFTER 8 WEEKS OF DATA COLLECTION)

To help increase screening completion rates while minimizing an increase in cost, we implemented a 2nd phase sample design after eight weeks of data collection. Specifically, we subsampled 157 of the 313 non-final cases (i.e. roughly 50%) for which we would continue to attempt to complete screeners for an additional two weeks.[4] We implicitly stratified by segment to ensure that the 157 cases *selected* for the second phase had the same segment distribution as the 313 non-final cases at the end of eight weeks. Thus, the first eight weeks of data collection were the first phase and weeks 9 and 10 were the second phase.

As outlined by the Groves and Heeringa (2006) responsive design paper, to optimize effectiveness of a 2nd phase design, practitioners need to change the essential survey conditions. Outlined below are three ways we changed the essential survey conditions for the 2nd phase:

1. We consolidated our interviewing staff to our 5 most productive interviewers (2 in Dallas, two in LA and one in Denver)

2. Prior to the start of Phase 2, we sent a priority mail envelope to all 157 second phase cases which contained a $5 bill and a letter explaining the importance of the study and our continued efforts.

3. We relaxed our screening protocol to allow interviewers to finalize a case as not being age eligible based on information from someone other than a household resident (e.g., housekeeper, nanny, neighbor).

[4]By the middle of the eighth week of data collection both the LA High Hispanic and Denver High African-American segments had screening completion rates of 82% or higher. To concentrate our second phase efforts in areas where we believed it would produce the greatest increases in screening completion rates, we excluded these two segments from the second phase. Specifically, at the end of the eighth week, we assigned final non-interview or refusal outcome codes to all non-final cases in the LA High Hispanic and Denver High African-American segments. This prohibited any cases from these two segments from being selected for the second phase.

RESULTS

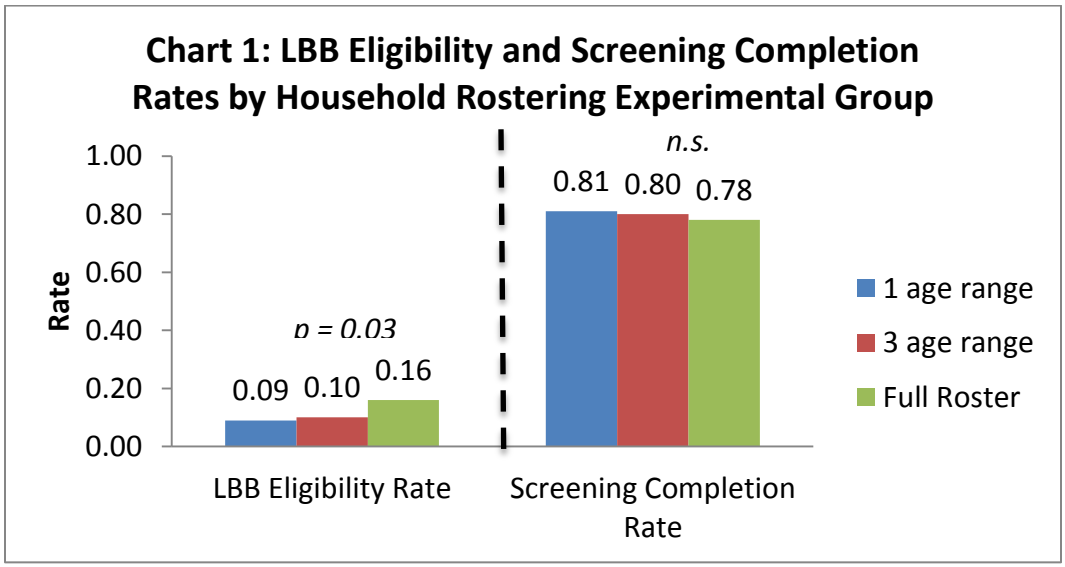
The HRS Screening Experiment data collection period lasted just over ten full weeks, from August 22 through November 3, 2013. We finished with 905 completed screeners and a fully weighted screening completion rate of 80%.

A more detailed breakdown of released cases, completed screeners and interviewer hours per completed screener by PSU can be found in Table 2.

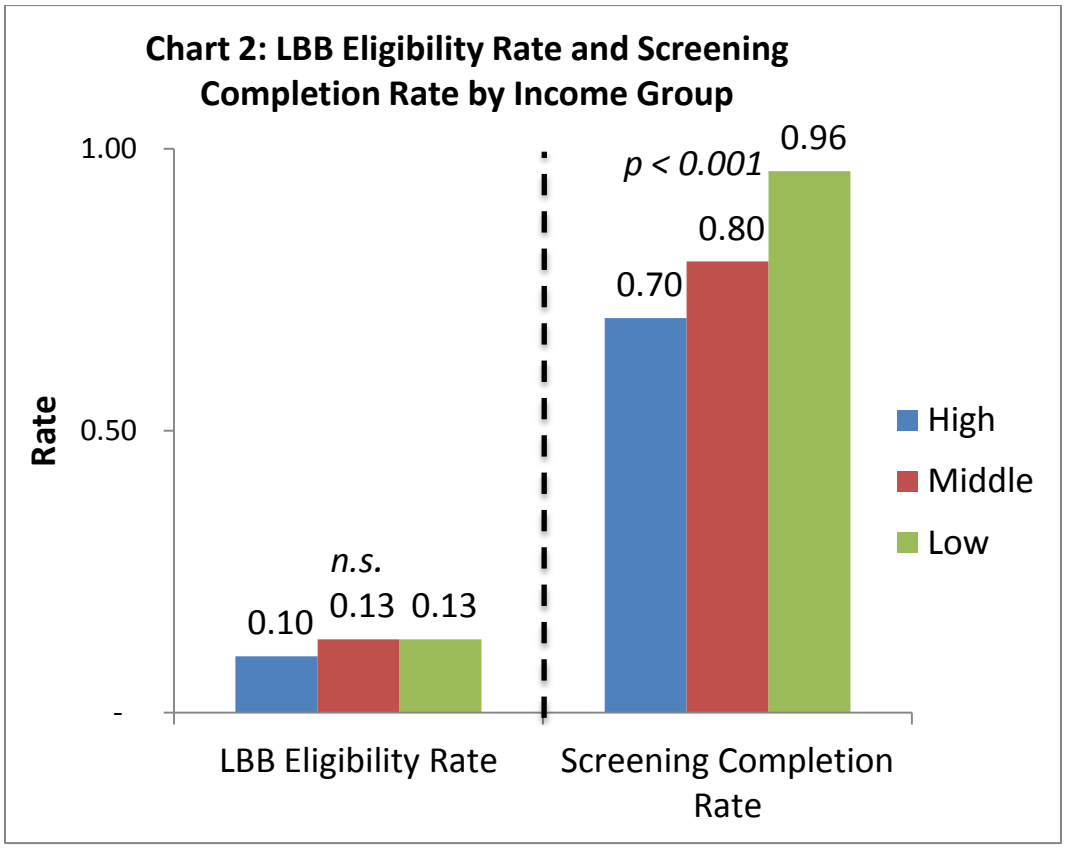
Table 2: Key Production Statistics by PSU

PSU	Released Cases	Completed Screeners (Unweighted)	Hours Per Screener
Dallas County	486	335	1.99
Denver County	326	242	2.46
LA County	488	328	2.84
Overall	1300	905	2.42

Chart 1 displays the two key statistics for the experiment – LBB age eligibility rates and screening completion rates - by household rostering procedure. The full roster procedure (0.16) produced significantly higher LBB age eligibility rates than the 1 age range (0.09) and 3 age range (0.10) procedures, $\chi^2 = 3.505$, $p = 0.03$. In terms of the screening completion rates by household rostering procedure, the small differences in rates were not significant. By these two metrics, the full roster technique, the technique used by the HRS since the beginning of the study, outperformed the one age and three age range groups. While this result refutes our initial hypothesis that among the three techniques, the 3 age range procedure would provide the best balance of high screening completion rates and LBB eligibility rates similar to the full roster procedure, we are pleasantly surprised. We would prefer to implement the full HH roster procedure, since it best ensures accurate and higher eligibility rates.



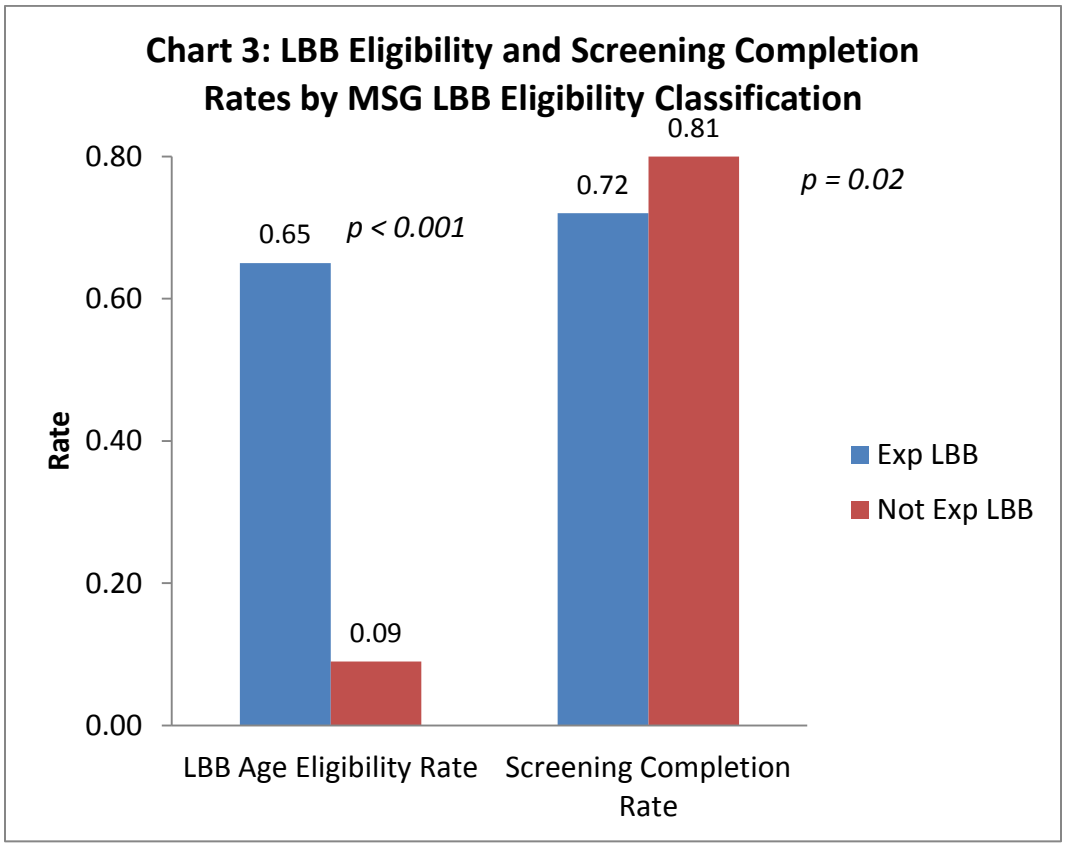
While Chart 1 shows the key results of this experiment that dispelled our initial hypotheses, there were other significant results which were also informative. One interesting outcome was the comparison of screening completion rates and eligibility rates by the median income of the area, as Chart 2 illustrates. Across the three PSUs, segments with the lowest median income ($22,000 to $26,000) had significantly higher screening completion rates (0.96), than areas with a medium median income (0.80) and areas with a high median income (0.70), $\chi^2 = 20.154$, $p < 0.0001$. There were no significant differences in LBB eligibility rates across the three median income groups. This result supports Groves and Couper's (1998) finding that residents in high income housing are more likely to refuse requests to participate.



Additionally, we observed significant differences in LBB eligibility rates and screening completion rates based on whether MSG's data classified an address as containing an LBB eligible member (as seen in Chart 3). Valliant et al. (2014) showed that MSG age data, though imperfect and incomplete, does a reasonable job of accurately identifying whether the household contains a member of a specific age range (especially ages 40 and above), when people are more financially stable and likely to be on such commercial lists. The addresses which MSG's commercial data identified as being LBB eligible had a significantly higher LBB age eligibility rate (0.65) than address not identified as being LBB eligible (0.09), $\chi^2 = 250.006$, $p < 0.0001$. This finding corroborates and builds upon Valliant et al.'s results, showing that MSG's commercial data is quite good at correctly

classifying whether households contain a member within narrow age ranges, when looking for persons age 45 and older.

Chart 3 reveals that the addresses not expected to be LBB eligible had a higher screening completion rate (0.81) than addresses expected to be LBB eligible (0.72), $\chi^2 = 5.357$, $p = 0.021$. Thus, it appears that at least some likely LBB age eligible HHs were able to discern that they would be eligible for the study and thus refused to participate, despite specifically training interviewers to use language and provide materials that suggested we were looking to see if anyone in the HH was eligible for one of a number of SRC studies.



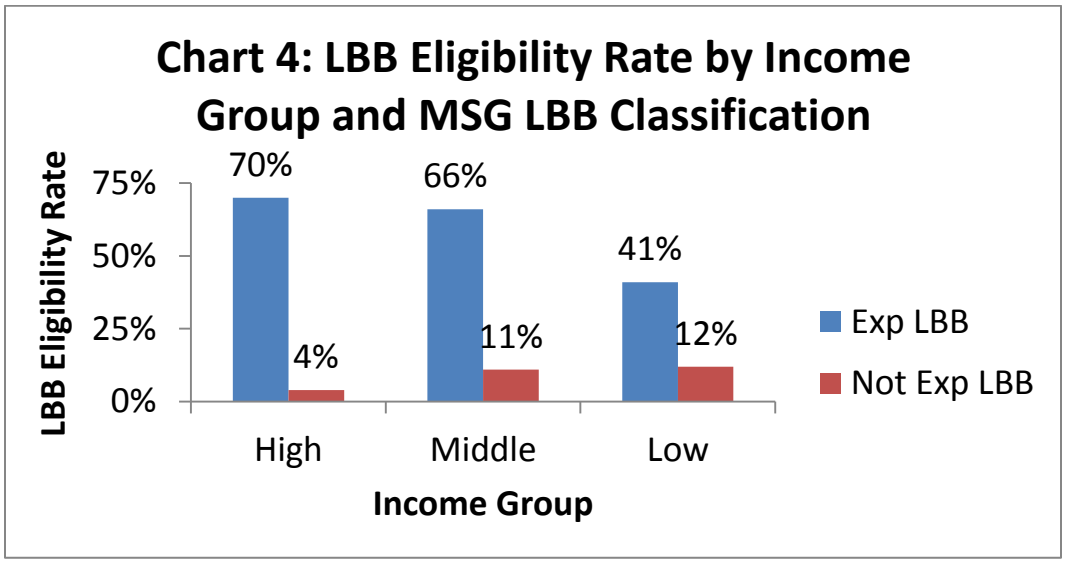
Presumably, residents in areas with higher household incomes are less likely to move frequently and more likely to have mortgages and own credit cards than those in lower household income areas. MSG's household level database is a compilation of various consumer data sources, which all rely heavily on credit card purchases and home ownership data. We suspected that the MSG data would most accurately stratify households into LBB eligible and not LBB eligible groups in high median income areas compared to low median income areas. In other words, we believed that when looking at the achieved LBB eligibility rates, there might be an interaction between median household income (High, Medium and Low) and MSG's LBB eligibility classification groups (Expected LBB, Not Expected LBB). To test for this interaction while controlling

for other factors, we constructed a HH level logistic model with LBB age eligibility as the dependent variable and the following independent variables:

- MSG LBB Eligibility Classification (Exp LBB, Not Exp LBB)
- Median HH Income of Area (High, Medium, Low)
- Experimental Group (1 age range, 3 age range, Full roster method)
- Interaction of MSG LBB eligibility classification and Race/Ethnicity density of the area

As we suspected, we found a significant interaction between MSG LBB eligibility classification and median HH income of the area for LBB age eligibility rate. As seen in Chart 4, the difference between the LBB eligibility rates of expected LBB households and not expected LBB households in high income areas (0.70 – 0.04 = 0.66) was greater than this same disparity in the middle income (0.66 -0.11 – 0.55), $\beta = 3.134$, $p = 0.009$, and low income (0.41 – 0.12 = 0.29) areas, $\beta = 3.649$, $p = 0.004$.[5]

Chart 4 shows a trend that MSG data is better able to accurately classify HH's into LBB eligible and non-LBB eligible in the Middle median Income group versus the Low median income group, this interaction is not significant. The reason this trend does not produce a statistically significant difference is that 97% of all completed screeners in the "Middle" median income group were in the "Not Expected LBB" group while at the same time 98% of all completed screeners in the "Low" median income group were also in the "Not Expected LBB" group. Thus, any significant differences between the three income groups are driven by the LBB Eligibility Rate in the "Not Expected LBB" households.



[5] See Table 3 in the Appendix for the full model results.

DISCUSSION

Our experiment produced four findings in regards to our key measures of screening completion rates and eligibility rates for an in-person household screening procedure. There was an essential finding central to the experimental design itself, and three additional findings not associated with the HH rostering experiment. These additional findings related instead to the different levels of affluence in areas where data was collected, information about the ages of HH residents provided by a commercial vendor, and the interaction between the affluence of a neighborhood and the age information from a commercial vendor.

Concerning the household rostering method experiment, our crucial finding was that while there were no significant differences in screening completion rates between the three experimental groups, the full enumeration group achieved a significantly higher eligibility rate than both the one age range and three age range methods. In other words, full enumeration of all household members produced higher eligibility and similar completion rates compared to the other two methods.

We modeled much of our experimental design on Tourangeau et al.'s (2012) innovative work comparing the effectiveness of different HH screening methods, but implemented our experiment in a much different survey design with the idea that a different environment might produce significantly different results. Tourangeau and colleagues work was a centralized, telephone study screening for households with persons aged 35 to 55. Our study was a decentralized, in person study screening for households with persons born between 1960 and1965 and not married or partnered with someone born before 1960. The difference in eligibility rates we found match the results from Tourangeau et al.'s (2012) work on comparing HH screening methods, however the lack of a difference in completion rates does not.

One potential reason for this lack of difference in completion rates is the extensive amount of information an interviewer needs to collect to determine if a HH is eligibility for the HRS. In the Tourangeau study, informants only needed to provide age information. In our study, informants in age eligible HHs needed to provide age information as well as years of birth and relationship status to other HH members for age eligible HH members. During our interviewer debriefing after data collection, many interviewers indicated that for the 1 age range and 3 age range methods, they did not feel comfortable emphasizing to potential HH informants that they "only needed to answer a few short questions". If the informant indicated that there was at least one HH member age 40-59, the informant would then need to complete the full HH rostering procedure and answer these questions for each eligible HH member. The interviewers felt that using this point of emphasis, when it might not necessarily be true, could undermine the rapport they had established in convincing the informant to participate.

A second significant result showed a difference in screener completion rates by median HH income of the area. In highly affluent areas, we obtained significantly lower screening completion rates than middle or low income areas. These results strengthen Groves and Couper's (1998) hypothesis that people with a lower socioeconomic status would be more likely to respond to a household survey. In this study, it was difficult to tease out if the difference in completion rates was due to SES or racial/ethnicity

compositions since the three most affluent areas were also the three high density white areas.

Third, we found that LBB eligibility rates were significantly higher in areas where the MSG data predicted high eligibility rates. This finding supports Valliant and colleagues (2014) conclusion that although commercial data can be somewhat incomplete in terms of age and demographics, it can still be useful in stratifying a sample to target specific characteristics.

Lastly, we found a significant interaction between the median HH income of an area and the MSG LBB eligibility classification on eligibility rates. Specifically, our commercial data vendor, MSG, most accurately classified addresses into LBB eligible or not LBB eligible in high income areas, followed by medium income areas, and finally low income areas. As mentioned previously, MSG aggregates data from companies such as Neustar and Experian, who compile credit card purchase and mortgage record data, and sell this consumer information to businesses and individuals. Since HHs with higher incomes are more likely to own a credit card and own their homes, it makes sense that the demographic information about these households from consumer databases are more complete and accurate than HHs with lower incomes.

These final two results provide an interesting story about high income areas when trying to identify and convince a demographic subgroup of the population to participate in a study. Households in high income neighborhoods were, to a remarkable degree, accurately stratified into LBB eligible and non-eligible groups, helping to make oversampling designs in these areas much more efficient and effective. However, these gains in sampling efficiency in high income areas compared to lower income areas wound up being neutralized by lower completion rates.

One area for future research would be to attempt to replicate the screening completion rate and eligibility rate results from the Tourangeau et al. study between the three HH rostering techniques in an in-person, decentralized study. We postulated earlier in this section that the reason we did not replicate the Tourangeau et al screener completion rate findings across our three rostering groups was due to our more detailed, narrow eligibility criteria. Ultimately this led to a longer screener questionnaire, however, we are not able to say decisively whether that was the reason for the difference or whether it was due to differences in mode of data collection.

Another area we plan to examine in more detail is to isolate differences in screening completion rates by combinations of wealth, racial/ethnicity densities and crime rates using paradata from previous large scale SRC screening efforts as well as the HRS 2016 LBB screening. We believe that high income, high density minority areas are likely to have similarly low screener completion rates to what we observed in the high income, high density white areas. In other words, we believe that the lower screening completion rates are driven by affluence rather than race/ethnicity. Furthermore, we feel that in lower income areas, screening completion rates may differ by levels of crime. Thus, we can theorize that less affluent areas with low crime rates would produce higher in-person screening completion rates than those with high crime rates. In the areas with higher crime rates, interviewers will understandably be concerned about their personal safety and households may be less willing to open their doors to strangers.

REFERENCES

Groves, R. M., and Couper, M. (1998). *Noncompletion in household interview surveys*. New York: Wiley.

Groves, R., and Heeringa, S. (2006), "Responsive design for household surveys: tools for actively controlling survey errors and costs." *Journal of the Royal Statistical Society Series A: Statistics in Society*, 169(Part 3): 439-457.

Tourangeau, R., Kreuter, F., and Eckman, S. (2012). "Motivated Underreporting in Screening Interviews." *Public Opinion Quarterly* 76(3):453–469.

Valliant, R., Hubbard, F., Lee, S., &and Chang, C. (2014),). "Efficient Use of Commercial Lists in U.S. Household Sampling." *Journal of Survey Statistics and Methodology Methods* 2 (2): 182-209.

APPENDIX

Table 3: Maximum Likelihood Estimates from Logistic Model Predicting LBB Eligibility of Selected Households

Parameter	Beta	SE	*p*
Intercept	-3.99	1.02	***
HH Roster Grp (ref = 1 age range)			
3 age range	0.01	1.44	
Full HH Roster	1.90	1.15	
MSG LBB Eligibility Classification (ref = Not Exp LBB)			
Exp LBB	5.34	1.09	***
HH Median Income (ref = High Income)			
Low Income	1.69	1.10	
Middle Income	1.84	1.08	
HH Roster Grp * MSG LBB Eligibility Classification (ref =1 age range * Not Exp LBB)			
3 age range * Exp LBB	-0.96	1.53	
Full HH Roster * Exp LBB	-2.28	1.25	
HH Roster Grp * HH Median Income (ref = 1 age range * High Income)			
3 age range * Low Income	0.53	1.55	
3 age range * Middle Income	-0.37	1.54	
Full HH Roster * Low Income	-1.45	1.28	
Full HH Roster * Middle Income	-1.51	1.25	
MSG LBB Eligibility Classification * HH Median Income (ref = Not Exp LBB * High Income)			
Exp LBB * Low Income	-3.65	1.27	***
Exp LBB * Middle Income	-3.13	1.20	***
HH Roster Grp * MSG LBB Eligibility Classification * HH Median Income (ref = Full HH Roster * Not Exp LBB * High Income)			
3 age range * Exp LBB * Low Income	0.66	1.79	
3 age range * Exp LBB * Middle Income	2.21	1.71	
Full HH Roster * Exp LBB * Low Income	2.25	1.57	
Full HH Roster * Exp LBB * Middle Income	2.97	1.45	*

* $p < 0.05$

** $p < 0.01$

*** $p < 0.001$