Coverage Gap: Out-of-State Phone Numbers for State Surveys

Marcus E. Berzofsky[1], Caroline B. Scruggs[1], Howard Speizer[1]
Bo Lu[2], Timothy Sahr[3]
[1]RTI International, 3040 Cornwallis Rd., Research Triangle Park, NC 27709
[2]Division of Biostatistics, The Ohio State University, 1841 Neil Avenue, Columbus, OH 43210
[3]Ohio Colleges of Medicine Government Resource Center, 150 Pressey Hall, 1070 Carmack Road, Columbus, OH 43210

Abstract
As survey designers contemplate using a single frame cellphone design for telephone RDD surveys they consider the potential for coverage bias. From a coverage perspective, there has been discussion about the impact of excluding landline only households. There has not been as much consideration to the impact on coverage bias regarding residents of an area of interest who do not have a telephone number associated with that area. Unlike with landline numbers, a cellphone number is portable and therefore many residents of an area to be studied might not have a phone number associated with the area. Moreover, the latest data indicate that 50.5% of adults can only be reached by cellphone (Blumberg & Luke 2017). As more cellphone only persons move from where they purchased their phone, the undercoverage rate in the area to which they move increases and the potential impact on bias increases. Based on auxiliary data, the undercoverage rate averaged 3% in Ohio and was as high as 40% in some counties.

This phenomenon creates a coverage issue for random cellphone designs. Random cellphone designs utilize the area code associated with an area to determine the universe of phone numbers. In this presentation, using data from the 2017 Ohio Medicaid Assessment Survey, we try to measure the undercoverage rates due to cellphone-only population that don't appear to live within the survey area. We also examine two approaches to reduce coverage bias for persons with out-of-state telephone numbers that live within the area of interest. First, we assess telephone numbers selected from a rate center in a neighboring state to increase coverage in a rural county. Second, we assess coverage gains and the characteristics of persons obtained through MSG's Consumer Cellular Database whose address is known regardless of the area code of the person's cellphone number. Our research found those with out-of-state cellphone numbers are different from those with an in-state phone number and there is likely coverage bias in a traditional RDD frame which excludes person with an out-of-state phone number.

Key Words: coverage bias, cellphone sample, RDD, out-of-state phone numbers, Ohio Medicaid Assessment Survey (OMAS)

1. Introduction

1.1 Understanding the Problem

As a greater portion of the population takes on a cellphone-only or cellphone-mostly status – that is they only have a cellphone in their household or have a landline, but use their cellphone a large majority of the time – persons who move to a state other than where their cellphone was activated are not included in a traditional RDD cellphone frame. In fact, the latest available data indicates 50.5% of adults can only be reached by a cellphone (Blumberg & Luke, 2017). Furthermore, those who are mobile are often younger and more affluent than those who are not. The combination of these two facts creates a potential coverage gap for survey designers.

As depicted in Figure 1, for a state or local-area based survey, the RDD cellphone frame (i.e., the sampling population) does not include persons with an out-of-state phone number but live in the area of interest and, therefore, are a part of the target population. Nationally, according to Census data, between 2014 and 2017 an average of 2.3% of the population moves from one state to another (Census Bureau, 2014 – 2017). This rate varies by age with 5.2% of 18 – 24 year olds and 4.0% of 25 – 34 year olds migrating to a new state each year compared to about 1.3% of persons 55 years old or older. Furthermore, this rate varies by state. For example, Ohio has an average migration into the state rate of 1.7% over the past 4 years (Census Bureau, 2014 – 2017). Still, even in a low migration state like Ohio, over 5 years almost 10% of the population has moved from another state.

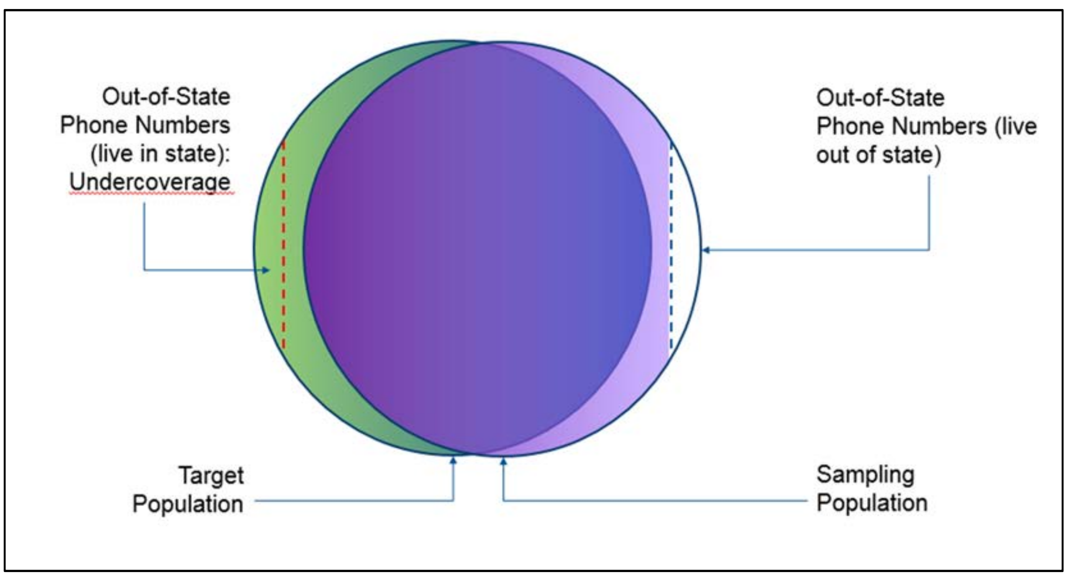


Figure 1: Depiction of potential undercoverage in an RDD cellphone frame due to the exclusion of residents with out-of-state cellphone numbers

These normal migratory shifts in the population coupled with the ability to maintain a single phone number regardless of location of residence lead to an increasing amount of coverage error for surveys which primarily use an RDD cellphone frame. In fact, studies have shown that up to 12% of cellphone only persons reside in a state which differs from the state associated with their phone number (Skalland and Khare, 2013). Considering young adults are more like to move, surveys studying outcomes which are correlated to age such as access to health care and health care status, have the potential for experiencing coverage bias.

In addition to the likely disparities in the age of persons who have out-of-state phone numbers, there are disparities about where in a state persons with out-of-state phone numbers live. Using Marketing System Group's (MSG) Consumer Cellular Database, which links a person's cellphone number to an address, we are able review where persons with out-of-state cellphone numbers live. For example, in Ohio, the Consumer Cellular Database contains 2.5 million households – 48% of households in Ohio – of which 3% have an out-of-state phone number. Figure 2 presents the distribution of persons with an out-of-state by population and percentage of households in a county. Interestingly, these distributions are not the same. When examining by population, the counties containing urban centers such as Cuyahoga county where Cleveland is or Franklin County where Columbus is. Whereas, when examining by percentage of households in the county, counties along the border are more likely to have persons with an out-of-state phone number. For example, Lawrence County – rural county – has 41% of households with a person with an out-of-state cellphone number.

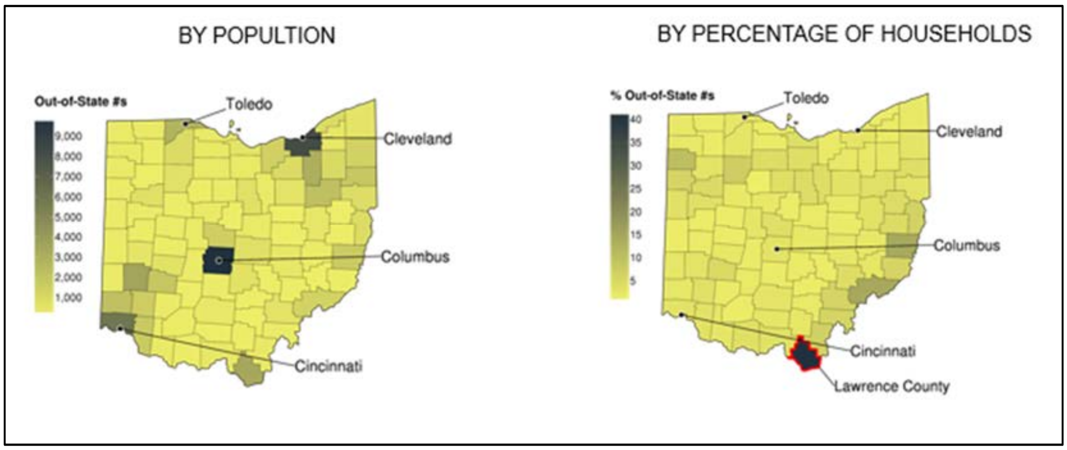


Figure 2: Distribution of persons with an out-of-state phone number in Ohio by population and percentage of households in the county

1.2 Purpose of Paper

In this paper, we aim to determine if there is potential coverage bias in a general population survey, in the state of Ohio, on health care access and health care status and if the potential for bias is different depending on where in the state person lives. Furthermore, if there is bias, determine what characteristics are likely being biased. To that end, we tested two hypotheses:

1. *Hypothesis 1*: Those who move into a state will be younger, more minority and have better health outcomes and excluding those persons bias estimates
2. *Hypothesis 2*: Those who obtained a cellphone number from a neighboring state are not different from those who obtained a cellphone number from Ohio.

2. Methods

2.1 Ohio Medicaid Assessment Survey

The Ohio Medicaid Assessment Survey (OMAS) is a periodic general population survey of residents in Ohio. The survey has been conducted approximately every two years since 2004. The survey collects information on and produces estimates related to access to health care and health care status. The survey is a dual frame RDD survey. In 2017, OMAS collected 36,000 interviews – 70% of which were collected through the RDD cellphone frame.

In addition to using traditional RDD cellphone and landline frames, which only include phone numbers with an Ohio area code, the 2017 survey included telephone numbers from two additional sources. First, MSG's Consumer Cellular Database was used to create a stratum of phone numbers with out-of-state area codes. The Database contained 110,700 numbers with an out-of-state area code, but were associated with an address in Ohio. A random sample of 108,600 numbers from the database was selected. Second, a rate center in West Virginia was identified, through the use of the Consumer Cellular Database, was identified as having a large number of Ohio residents. This rate center contained 236,000 possible cellphone numbers and is located near Lawrence County, OH (see Figure 3). For our study, we drew a sample of 6,472 cellphone numbers. Cell-Wins was used on both samples to identify inactive phone numbers. Approximately 35% of each sample was flagged as being inactive. Cases flagged as inactive were not fielded.

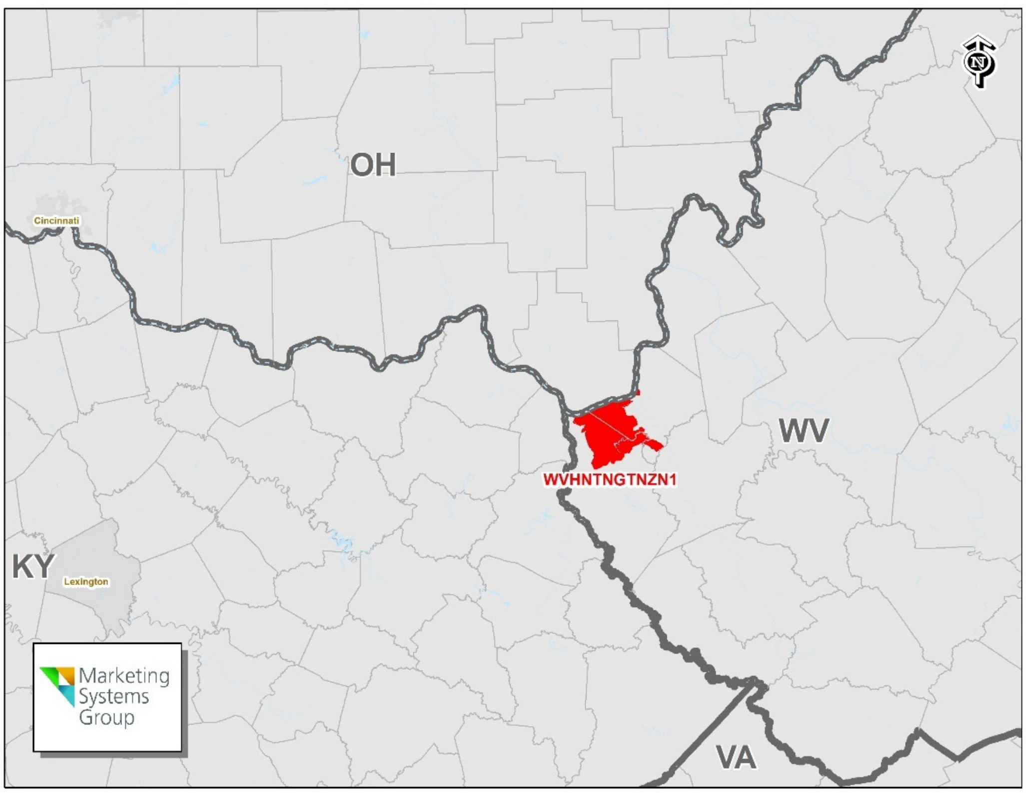


Figure 3: Location of West Virginia rate center selected for inclusion in the 2017 OMAS

The 2017 OMAS collected a total of 31,314 cellphone interviews. Of those collected through the traditional RDD frame, 6,987 were flagged as also being in the Consumer Cellular Database. Among the cases from the Consumer Cellular Database out-of-state phone number stratum, 1,370 interviews were collected. Additionally, the West Virginia rate center stratum yielded 45 interviews. Our analysis utilized the set of completed interviews.

2.2 Hypothesis #1

For our first hypotheses we utilized the 1,370 completed interviews from the Consumer Cellular Database out-of-state phone number stratum. Prior to conducting our main analyses, we compared the cases out-of-state phone number cases with in-state phone numbers flagged as being in the Consumer Cellular Database. This was done to determine if simply being in the database predisposes a person to be different in some manner than others in the state. Second, we compared the demographic characteristics of persons with an out-state cellphone number to those with an in-state phone numbers (using the entire RDD cellphone sample). The demographic characteristics considered were

- Age group (< 45, >=45)
- Gender
- Race/ethnicity (White, non-White)
- Marital status (married, non-married)
- Income (Below 138% of FPL, above 138% of FPL)
- Employer offers health insurance (yes, no)
- County type (metro, suburban, rural non-Appalachian, rural Appalachian)

Next, we conducted bivariate analyses comparing persons with out-of-state area code cellphone numbers to those with in-state cellphone numbers for the following outcomes:

- Percentage uninsured
- Percentage having problems getting healthcare
- Percentage with high blood pressure or hypertension
- Percentage who have ever had a heart attack
- Percentage who have ever had coronary heart disease
- Percentage who have ever head congestive heart failure
- Percentage who have ever had diabetes

Finally, to determine if any difference in the demographics where causing bias, for each outcome, we fit a logistic model containing the demographic variables previously considered.

2.3 Hypothesis #2

For the second hypothesis, we focused our analyses on cellphone respondents in Lawrence County, Ohio. In total, Lawrence County had 284 cellphone respondents (81% of all respondents in Lawrence County). We focused this analysis on Lawrence County because the rate center in West Virginia we selected borders Lawrence County and, as such, was selected to help supplement the in-state Lawrence County sample. Therefore, we first assessed how many of the interviews obtained through the West Virginia rate center resided in Lawrence County compared to a different Ohio county and determined how many of the completed interviews came from the Consumer Cellular Database out-of-state stratum.

Next, we compared the demographics of the respondents across the three strata type: in-state cellphone, WV cellphone, and out-of-state cellphone. Due to sample size concerns, the demographics considered were age (percentage less than 45), gender (percentage male), race/ethnicity (percentage non-White non-Hispanic), marital status (percentage married), and income level (percentage below 138% of the federal poverty level).

Finally, bivariate analyses were conducted across each of the three strata the key outcomes

- Percentage uninsured
- Percentage having problems getting healthcare
- Percentage with high blood pressure or hypertension
- Percentage who have ever had a heart attack
- Percentage who have ever had coronary heart disease
- Percentage who have ever head congestive heart failure
- Percentage who have ever had diabetes

3. Results

3.1 Hypothesis #1

3.1.1 Comparison of Consumer Cellular Database

Figure 4 presents a comparison of respondents who were linked to the Consumer Cellular Database but selected through the traditional RDD cellphone frame and those selected through the Consumer Cellular Database out-of-state stratum. Except for gender, all demographic characteristics are statistically different. Specifically, those with an out-of-state cellphone number are more likely to be under 44; less-White; married; have income over 138% of the FPL; and have employer-based health insurance.

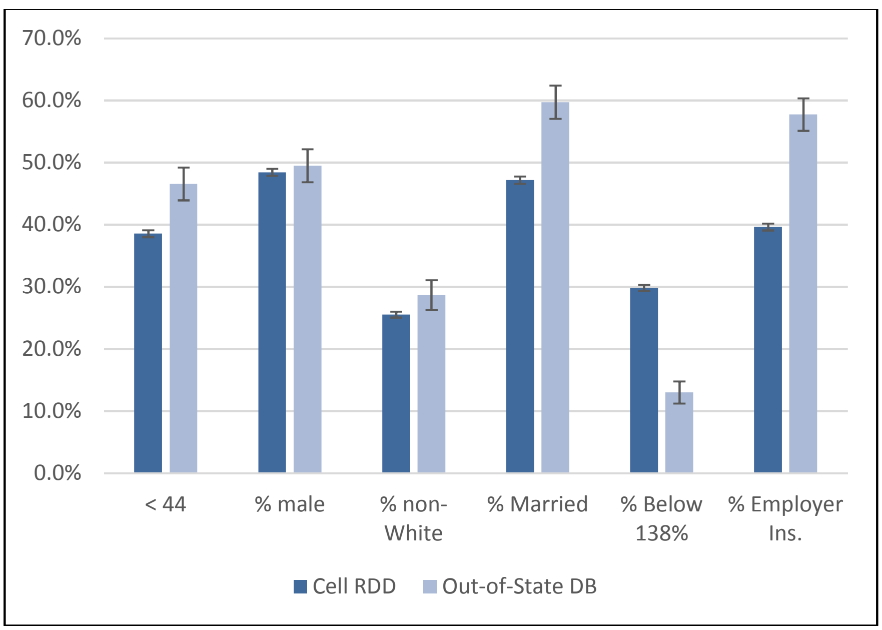


Figure 4: Percentage of respondents contained in the Consumer Cellular Database by in-state phone number status and demographic characteristics.

Figure 5 presents the comparison of Consumer Cellular Database respondents by their in-state phone number status by county type. Persons in the database were more statistically likely to have an out-of-state phone number in suburban counties and significantly less likely to in rural non-Appalachian counties.

The differences in the respondent characteristics indicate that any comparisons between those with an out-of-state phone number and all cellphone respondents with an in-state cellphone number.

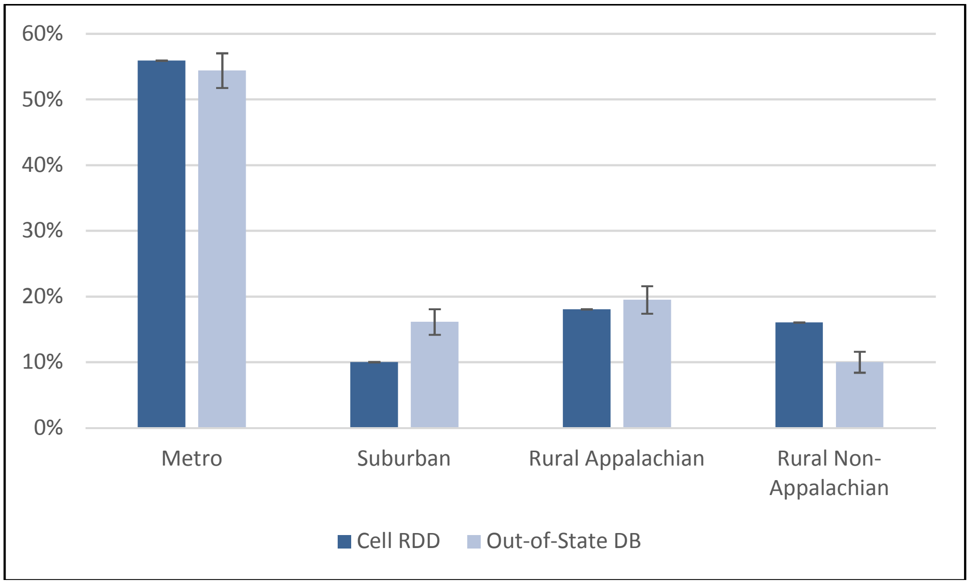


Figure 5: Percentage of respondents contained in the Consumer Cellular Database by in-state phone number status and county type.

3.1.2 Bivariate Comparison

Figure 6 presents the bivariate comparison of outcomes between respondents in the Consumer Cellular Database out-of-state stratum and those in the traditional RDD cellular frame strata. For all outcomes, those with an out-of-state phone number were statistically less likely to be uninsured or have a health problem.

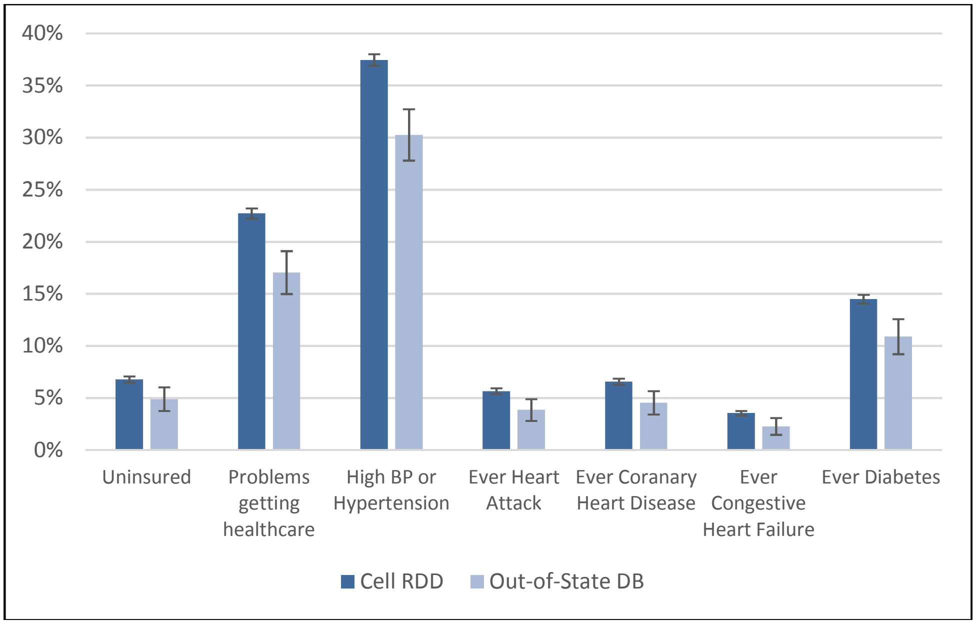


Figure 6: Percentage of respondents in the cellular RDD frame and Consumer Cellular Database strata by health insurance status and health outcome

3.1.3 Model-Based Comparison

Figure 7 presents the marginal probability of a respondent's health insurance status and health outcome status controlling for respondent characteristics. For all outcomes, once differences in the respondent characteristics where controlled for, none of the differences between respondents in the RDD cellphone frame and Consumer Cellular Database out-of-state frame were statistically different.

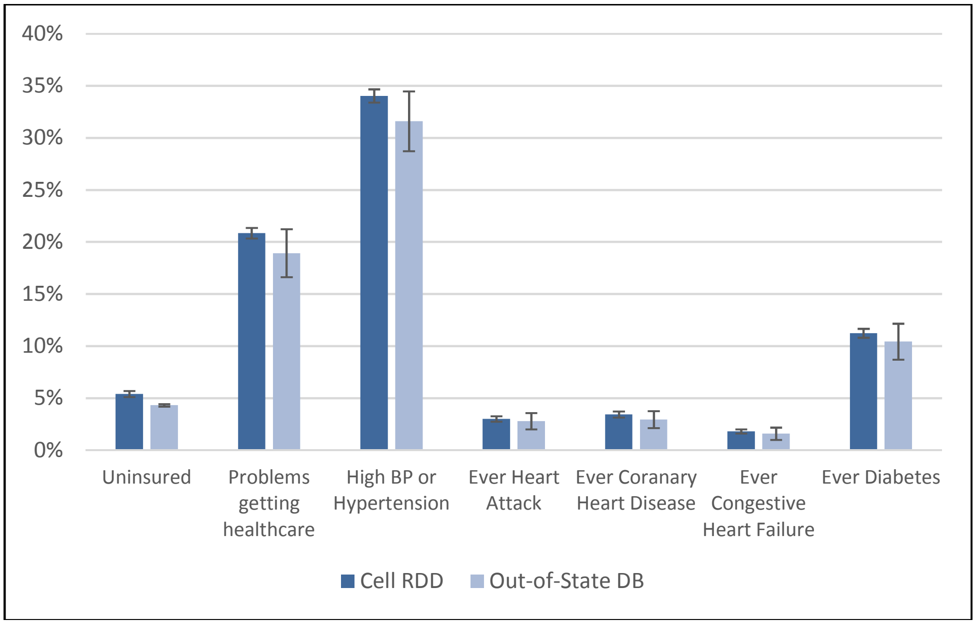


Figure 7: Marginal probabilities of respondents in the cellular RDD frame and Consumer Cellular Database strata controlling for respondent characteristics by health insurance status and health outcome

3.2 Hypothesis #2

3.2.1 Respondent Characteristics

The sample from the West Virginia rate center obtained 45 interviews from person residing in Ohio. Of those 32 (71.1) indicated they lived in Lawrence County. Furthermore, of 284 cellphone respondents in Lawrence County 11.3% came from the West Virginia rate center stratum compared to 6.7% of cellphone respondents coming from the Consumer Cellular Database out-of-state stratum.

Figure 8 presents the respondent characteristics for respondents in Lawrence County by the cellphone frame – traditional RDD frame, Consumer Cellular Database out-of-state, or West Virginia rate center. Due to sample size constraints, the statistical power of the comparisons was low. Given this deficiency only those in the Consumer Cellular Database tested differently than the traditional RDD frame. The respondent characteristics of those in the West Virginia rate center did not test different compared to the traditional RDD frame.

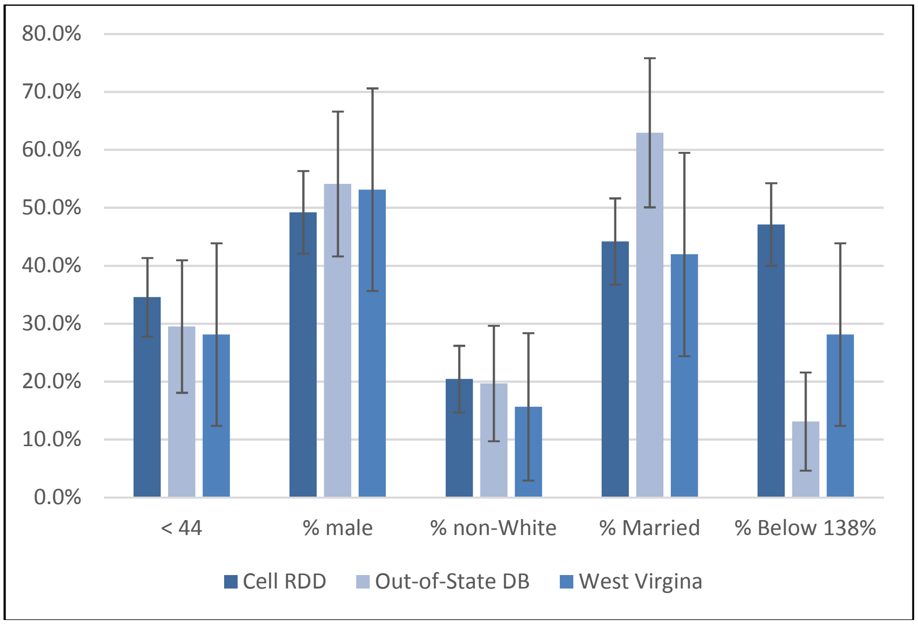
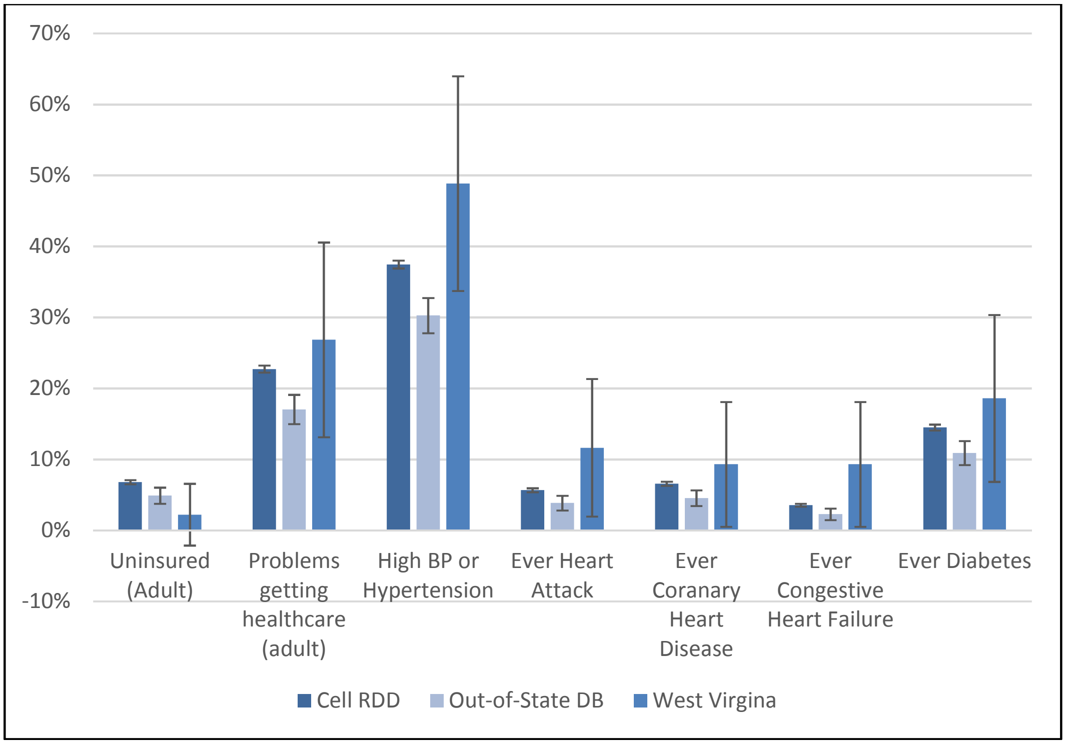


Figure 8: Percentage of respondents in Lawrence County, Ohio by respondent characteristics and cellphone frame

3.2.2 Bivariate Comparison

Figure 9 presents the results of the bivariate comparison of health insurance status and health status outcomes. Compared to respondents in the traditional RDD cellphone frame the respondents from the Consumer Cellular Database out-of-state stratum were statistically significant. In all cases, the Consumer Cellular Database out-of-state stratum respondents were more likely to have health insurance and less likely to have a poor health status. However, the West Virginia rate center respondents were only significantly different in their health insurance status. For all health status outcomes the statistical power was poor leading to very wide confidence intervals.



4. Discussion

4.1 Hypothesis #1

Our findings confirm hypothesis 1 is correct. We found respondents with an out-of-state phone number have statistically significantly better health outcomes from those with an in-state phone number; however, once the respondent characteristic differences between the two groups are taken into account the differences are not longer statistically different. This finding suggests that there is potential coverage bias in samples which rely only on a traditional RDD cellphone frame (or a dual frame where the second frame is an RDD landline frame). Given the direction of the differences the coverage bias will understate how healthy the population is – in other words, the true population is healthier than the survey estimate would suggest.

The major limitation of this study is the only persons with out-of-state phone numbers included in our sampling population are those in the Consumer Cellular Database. While the database does contain 48% of households in Ohio, it is not clear how well those with out-of-state cellphone numbers represent all persons in Ohio with an out-of-state phone number. To our knowledge there is no data source which provides information on those in a state with a non-in-state phone number. While there is information from the Census Bureau regarding those who migrated from another state, it only covers those who moved in the past year and does not indicate if their cellphone number (if they have a cellphone) was changed or not.

4.2 Hypothesis #2

Our findings, while confirming of hypothesis 2, are inconclusive due to the fact that the statistical power to detect differences was low. While it was anticipated the vast majority of phone numbers selected from another state are linked to persons in that other state. That said, our research does provide some indication that people in Ohio who activate their phone in West Virginia are simply doing so because the larger town with a cellphone store is in the neighboring state. In other words, it is simply easier to buy a phone in West Virginia then Ohio, but, that, in and of itself, does not indicate a difference in the persons characteristics or outcomes related to health status.

Not surprisingly, the major limitation with evaluating this hypothesis was obtaining residents of Ohio through the West Virginia rate center stratum. The low eligibility rate requires a very large starting sample size which makes this approach to minimizing coverage a costly option.

5. Conclusions

Based on our research, there does appear to be coverage bias in a traditional RDD cellphone frame. Given the increasing proportion of residents in an area a cellphone only or mostly person and have an out-of-area phone number, addressing this coverage bias is critical for surveys who use RDD frames. Those with an out-of-state phone number are younger, less likely to be White and more stable (e.g., more likely to be married, have higher income, and have employer-based health insurance). Persons with these differences are more likely to have better health status measures. However, when these respondent characteristic differences are controlled for the differences in health outcomes disappear. Therefore, including those with out-of-state phone numbers in the sample would reduce the coverage bias due to their exclusion from a traditional RDD cellphone frame.

Databases which link cellphone numbers to addresses can be used to identify persons in the area of interest who have an out-of-area phone number. However, these databases need to be evaluated to determine how well they represent the population. Likely differences in the database need to be taken into account. Selecting samples from rate centers bordering a state are not an effective method because they are costly and only help cover those in border counties. Those living in central counties with an out-of-state phone number will not be covered through this approach.

References

Blumberg, S.J. and Luke, J.V. (2017). Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January – June 2017.

U.S. Census Bureau (2017). American Community Survey, Table S701. Downloaded on September 20, 2018 from https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml