Cycling of Non-Self-Representing Primary Sample Units in the National Health Interview Survey

Padraic Murphy[1, 2, 3]
U.S. Census Bureau, Washington, D.C. 20233

Abstract
Recently the National Health Interview Survey tested the cycling of non-self-representing (NSR) Primary Sample Units in six states. A pilot study was conducted during the 2010-based design period to test operational feasibility and to evaluate the reliability of multi-year state estimates where the sample of NSR first stage units is not the same in all years. The pilot study was considered a success, and the National Center for Health Statistics requested that the Census Bureau include the cycling of NSR first stage units as a design feature for all states with NSR units in the upcoming sample redesign based on the 2020 Decennial Census. The primary goal of the cycling is to ensure sufficient degrees of freedom to allow reliable estimates for each state when pooling three years of data. In planning this sample redesign, we initially included a maximal set of counties in the overall cycling sample for each state. However, we found that this approach was neither necessary nor desirable to meet the requirements set forth by the survey sponsor. In this paper we explain how we arrived at our final first stage sample design and describe how the cycling of the NSR units is carried out.

Key Words: Survey Sample Design, Multi-stage Sampling, Primary Sample Units, Cycling, National Health Interview Survey

1. About the National Health Interview Survey

The National Health Interview Survey (NHIS) is a major, continuous health survey conducted by the National Center for Health Statistics (NCHS), part of the Centers for Disease Control and Prevention (CDC). Established in 1957, it collects data on the health status, healthcare access, and health behaviors of the U.S. population through in-person household interviews. The NHIS focuses on chronic conditions, health insurance, physical activity, and more. It is a key resource for monitoring national health trends and informing public health policy decisions. (National Center for Health Statistics, 2023)

2. Basic Sample Design for the NHIS

Since data collection for the NHIS is accomplished via in-person interviewing throughout each calendar year, a simple random sample of housing units and noninstitutional group quarters – with the residence units in sample widely dispersed across the nation – would be prohibitively expensive. Therefore, the NHIS uses geographically clustered sampling techniques. (Moriarty, et al., 2022)

---

[1] Email: Padraic.A.Murphy@census.gov
[2] This article is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not those of the U.S. Census Bureau.
[3] The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data (Project No. P-7533782, Disclosure Review Board (DRB) approval number: CBDRB-FY24-POP001-0085).

The sample design for the NHIS (and other demographic surveys conducted by the U.S. Census Bureau) is revised following each decennial census to account for changes in population and housing. For this reason, we refer to sample designs as "2000-based", "2010-based", and "2020-based", and so on.

For the 2020-based NHIS sample design, planned for data collection taking place during the years 2025 through 2036, the NCHS has chosen to use a two-stage sample design. Specifically, in the first stage of the design large contiguous geographic areas called Primary Sample Units (PSUs) were selected using a stratified design. (Tersine, Jr., 2023) The selection of PSUs was done just one time; however, certain PSUs will only be used for half of the years in the design period, as will be discussed below. In the second stage, housing unit addresses (in counties considered to have good address frame coverage) or measures (in counties with less good coverage) will be selected from within the sample PSUs selected in the first stage. This second-stage sampling will be done on an annual basis, which is a significant change from the current (2010-based) design. However, our focus in this paper is on the selection of PSUs in the first stage.

3. Definition of 2020-based Primary Sample Units, and Contrast with 2010-based Design

As is standard practice for other demographic surveys conducted by the Census Bureau, such as the Current Population Survey, the NHIS 2020-based sample design has defined each PSU to consist of a geographically contiguous set of one or more counties, with a requirement that each PSU have a certain minimum population based on the 2020 census. Where feasible, PSUs were also defined so as not to exceed a certain maximum land area. Certain PSUs – those with 200,000 or more housing units in the 2020 census count, usually corresponding to metropolitan areas – were selected with certainty and are called self-representing (SR) PSUs. In a few cases, smaller PSUs were designated SR in order to facilitate interviewing logistics, or to reduce sampling variance for a state. From the remaining PSUs, designated non-SR (NSR), a sample was selected with probability proportional to size within each state. Note that there are a few states with only SR PSUs, and a few others with only NSR PSUs, but most have a mix of both. The starting point for the 2020-based design PSU definitions was the set of geographic areas used for the 2010-based design, which we discuss next.

In the 2010-based design, covering data collection from 2016 through 2025, while large county-based geographic areas were defined and classified within states as "Type A" or "Type B", there was no explicit selection of these geographic areas in a separate stage. Rather, the areas were used to define strata within each state, each area was assigned a measure of size based on 2010 census data, and a sample of conceptual clusters was selected within each stratum. These conceptual clusters, or "measures" were later "filled in" with actual addresses either from commercial vendor files or via listing operations carried out by the Census Bureau, depending on the coverage quality of the commercial vendor files. (Moriarity, et al., 2022)

Whether a given area was "in sample" for the 2010-based design was determined by the stratum sampling interval, the measure of size of each area, and the random start value used for each stratum. If the stratum sampling interval was larger than an area's measure of size, it was possible for the area not to be in sample, and no interviewer would be hired or assigned for that area. Otherwise, if any measures were selected within an area, that area was in sample. In this sense, certain areas with large measures of size were in sample with certainty and are analogous to PSUs that are designated as SR in a true two-stage design. (These correspond to Type A geographic areas in the 2010-based design.) An area with measure of size smaller than the sampling interval of the stratum containing it is analogous to an NSR PSU in a true two-stage design. (These correspond to Type B geographic areas in the 2010-based design.)

In adjusting the 2010-based geographic areas to create the 2020-based PSU definitions for NHIS, in addition to re-grouping counties to account for changes in the delineation of metropolitan areas, an attempt was also made to have NSR PSUs within a state be as close to equally sized as possible, in terms of the 2020 census housing unit count, in order to reduce sampling variance.

4. Stratification of Primary Sample Units in the 2020-based Sample Design

The selection of PSUs in the NHIS 2020-based design was done independently by state. Within each state, each SR PSU (if any) was a single-PSU stratum. (Equivalently, one could say that all the SR PSUs in a state were in a single stratum from which all elements were selected. Either way, the probability of selection of an SR PSU was 1.0.) All NSR PSUs in a state (if any) were placed in a single stratum. We will discuss the determination of the annual number of NSR PSUs in sample for a state below; but for now, assume we know that this number is an integer k. Suppose the total number of NSR PSUs defined in a state is j. Then we must have $j > k$, since if $j = k$ we would have to select all the NSR PSUs, which would make them SR. In fact, for all states with NSR PSUs it turned out that $j \geq 2k$, which is key for the scheme of cycling NSR PSUs that we will discuss below. This is because we selected all the NSR PSUs to be used during the 12-year design period (2025 through 2036) at one time; and the total number of NSR PSUs selected for the design in each state was $n = 2k$.

5. Cycling of Non-self-representing Primary Sample Units

Several years after starting the collection of data under the 2010-based design, the U.S. Census Bureau, who works with NCHS on planning and implementing the NHIS, proposed a plan to cycle Type B geographic areas to improve state-level estimates when pooling data from three years. A pilot study was proposed in six states to evaluate the feasibility and effectiveness of this cycling strategy. (Tersine, Jr., 2021)

This pilot study was carried out and based on the results the NCHS requested that the Census Bureau include cycling of NSR PSUs in all states (where applicable) in the 2020-based design.

Features of the cycling of NSR PSUs for NHIS in the 2020-based design:

- Once interviewing begins in an NSR PSU, it will continue for at least two years. This facilitates the hiring and retention of interviewers within each state.
- Once the cycling pattern is established in a state, about half of the NSR PSUs in sample in year Y will be cycled out and replaced in year Y+1.
  - In the six states in the pilot program, 2020-based PSUs corresponding to 2010-based geographic areas serve to establish the pattern in those states, so that half of the Type B geographic areas in sample in 2024 (the final year of operations under the 2010-based design) will be replaced by new NSR PSUs in 2025, with these new NSR PSUs remaining in sample through the end of 2026. The PSUs corresponding to Type B areas in sample in 2024 that were not replaced in 2025 will continue interviewing through the end of 2025, and will be replaced at the beginning of 2026, with the replacement PSUs having interviews in 2026 and 2027.
  - For states not in the pilot program, the status of 2010-based Type B areas in 2024 is ignored. From the set of defined 2020-based NSR PSUs in such a state, k of the $2k$ PSUs selected for the full design period (2025 through 2036) will begin interviewing in 2025. Of these, about half will be replaced by some of the remaining k after two years, so interviewing will begin in these replacement PSUs in 2027. The remaining PSUs in the original set of k will continue interviewing for three years, through the end of 2027, and will be replaced at the beginning of 2028 by the remaining half of the other k. At this point the pattern is established, and every year after that, about half of the NSR PSUs in sample the previous year will be replaced.
- The $2k$ NSR PSUs selected in a state are stratified by size (in terms of HU count) into k strata, with two PSUs in each stratum. Every year, there is one PSU from each of these strata in sample. In about half of the strata each year (once the pattern is established) a PSU

in sample the previous year will be replaced by the other PSU in its stratum. This ensures that every year's sample of NSR PSUs is representative with respect to PSU size.

- There is a caveat if $k = 1$. In that case, only two PSUs are involved in cycling, and so each one remains in sample for two years at a time, with the two PSUs alternating. Thus, over the 12-year design period each of the two is in sample for three two-year periods.
- If $k > 1$ is even, exactly half of the strata will swap out PSUs each year. If k is odd, the number of strata swapping out PSUs alternates from year to year between the ceiling and floor of $k/2$.

- Each selected NSR PSU involved in cycling will be out of sample for two years at a time but will come back in sample after each two-year hiatus. This is also intended to facilitate retention (or possibly re-hiring) of experienced interviewers.
- Beginning after NHIS data collection is completed in Y=2028, a three-year period ending in year Y (e.g., 2026 through 2028; or if Y=2031, 2029 through 2031) will include data collected from all $2k$ selected NSR PSUs. Note that given the annual size of k and the requirement that each PSU be in sample at least two years, $2k$ is the maximum number of sample NSR PSUs possible over a three-year period.

Figure 1 illustrates the cycling of NSR PSUs in a hypothetical state where $k = 3$, and the six NSR PSUs selected for the design period are labeled A-F. The three size strata consist of (1){A,D}, (2){B,E}, and (3){C,F}. Note there is always exactly one PSU from each of the size strata in sample each year. Also, except for PSUs B and C being in sample for three years at the beginning of the design to establish the cycling pattern, each PSU alternates being in and out of sample for two years at a time. Finally, as illustrated by the shaded section, note that three-year periods include all six PSUs. (This does not include 2025 through 2027, when the cycling pattern is being phased in.)

Figure 1: Illustration of NSR PSU Cycling for a Hypothetical State Where $k = 3$

		Year											
Size stratum	***PSU***	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036
(1)	A	x	x			x	x			x	x		
(2)	B	x	x	x			x	x			x	x	
(3)	C	x	x	x			x	x			x	x	
(1)	D			x	x			x	x			x	x
(2)	E				x	x			x	x			x
(3)	F				x	x			x	x			x

6. Determination of the Annual Number of Non-self-representing Primary Sample Units in Sample

In this section, we describe the process for determining the annual first-stage NSR sample size k for each state.

The starting point is a national requirement that NHIS have 30,000 completed adult interviews nationally each year.[4]

Note that when a household (occupied housing unit) is selected for sample, assuming the interviewer has successfully contacted an adult in the household and assembled a roster of the household

---

[4] After the final determination of annual first-stage NSR sample sizes, the annual completed adult interview requirement was changed to a lower number, but no effort was made to revise the first-stage sample sizes, which would not have changed very much, if at all.

members, at most two interviews are conducted, for one adult and one child, if the household includes any children. It sometimes happens that a completed interview is obtained for a child, but no adult interview is completed. When discussing response rates in the following, only adult interviews (attempted or completed) are considered. Since the assumption is that every household includes at least one adult, all sample households are included in the denominator of a response rate calculation. However, a household with a completed child interview but no completed adult interview would NOT be included in the numerator.

Using the 2020 census housing unit count as a measure of size, the national completed adult interview requirement was allocated proportionally among the 50 states and DC. Based on NCHS data presentation standards for proportions (Parker, Talih, Malec, & et al., 2017), as applied to a generic ten percent characteristic for a state level estimate using three years of data, and assuming a design effect of 3.0, it was determined that the minimum number of completed adult interviews for a state each year is approximately 84. There were a few states for which the proportional allocation was below this minimum. Therefore, we re-allocated from states with a surplus over this minimum to get the states with a shortfall up to the minimum, while keeping the national total constant at 30,000. The method of re-allocation was to calculate the total shortfall across states, divide by the total surplus from states without a shortfall, and reduce the surplus in in those states by that fraction. Each state with a shortfall had its allocation set to the minimum 84.

Historical NHIS rates of completed adult interviews were used to project state level rates for 2025. We divided the re-allocated number assigned to each state as described in the previous paragraph by the projected rate of completed adult interviews to inflate to an initial designated sample size for each state.

Within a state define a domain as either a single SR PSU, or the set of all NSR PSUs in the state. (So for example, if a state has three SR PSUs and nine NSR PSUs, there are four domains in the state.) The initial designated sample size for each state was sub-allocated proportionally among the domains in each state. Denote the initial designated sample size for the NSR domain in a state as A. It is assumed that each NSR PSU in sample in a year will have a single workload, and that the nominal workload size is 120. Therefore, the annual first-stage NSR sample size for a state is, initially, $k = \frac{A}{120}$, rounded to the nearest integer.

Assumptions regarding the cluster sizes used for variance estimation in each SR PSU (which, in this context, is treated as a first stage stratum from which a sample of clusters is selected) were used to estimate the number of clusters and hence the degrees of freedom contributed to a generic proportion estimate by each year of data collected in each SR PSU, assuming three years of data. Each NSR PSU is considered a single variance estimation cluster, and the data collected within an NSR PSU in different years of a three-year period are considered all part of that cluster. These assumptions are used to estimate the degrees of freedom contributed by the NSR stratum in each state. The degrees of freedom contributed by a domain is the number of clusters minus one. Since each three-year period will include all $2k$ NSR PSUs selected for a state, the degrees of freedom estimate for the NSR domain of a state is $2k - 1$. Using the proportions for sub-allocation together with the SR and NSR estimated degrees of freedom, the Satterthwaite approximation formula allows an estimate of total degrees of freedom (using three years of data). (Cochran, 1977, pp. 95-96) A simplified version of the Satterthwaite formula, relevant for estimating a proportion in this context, is:

$$d_{state} = \frac{1}{\frac{P_{SR}^2}{d_{SR}} + \frac{P_{NSR}^2}{d_{NSR}}}$$

Here, d_{state} is the estimated degrees of freedom for the state; P_{SR} is the proportion of the state measure of size in SR PSUs, P_{NSR} is the proportion in NRS PSUs, d_{SR} is the total SR degrees of freedom, and d_{NSR} is the NSR degrees of freedom.

Another NCHS requirement for the presentation of proportion estimates is that an estimate must have a minimum of eight degrees of freedom, that is: $d_{state} \geq 8$. (Parker, Talih, Malec, & et al., 2017) Also, we know that $d_{NSR} = 2k - 1$. Putting this together with the Satterthwaite approximation, and doing some algebra, one arrives at the following minimum constraint for k:

$$k \geq \frac{1}{2}\left(1 + \frac{8d_{SR}P_{NSR}^2}{d_{SR} - 8P_{SR}^2}\right)$$

There were some states where the initial designated sample size resulted in values that did not satisfy this constraint, so the state total degrees of freedom estimate was less than eight. To fix this, for each state failing the degrees of freedom requirement, we set k to the smallest integer satisfying the above inequality, assuming the variables on the right side are fixed. Recalling the nominal NSR workload size of 120, this allows a calculation of how much each state's designated sample size must be increased. Suppose the adjusted value of k for a state is k_B, corresponding with a new NSR designated size of $B = k_B \times 120$. Then the deficit that must be made up for the state is given by $B - A$. We re-allocate the initial designated sample sizes so that each state with a degrees of freedom shortfall gains enough sample to satisfy the degrees of freedom requirement, maintaining the total designated sample size at a fixed amount. Those states that had shortfalls when allocating completed adult interviews are not reduced in this degrees-of-freedom re-allocation. All other states that have a surplus in both the first re-allocation and this one have their initial designated sample size reduced proportionally in this round. Now every state has a designated sample size that (a) is expected to result in a number of completed adult interviews over three years that will satisfy the basic precision requirement, and (b) has at least the minimum number of NSR PSUs in sample annually to ensure the estimated total degrees of freedom for state-level proportion estimates is no less than eight.

There are a few additional steps needed on an annual basis to determine the second stage sample size for each PSU in sample for a year, taking into account constraints on interviewer workloads. However, the focus in this paper is on the one-time selection of the first stage sample; and the information needed for PSU selection is determined once we know the annual NSR first stage sample size k for each state.

7. Identifying Quasi-self-representing Primary Sample Units ("QSR PSUs")

In states with NSR PSUs, as mentioned earlier, all NSR PSUs were placed in a single stratum, and a sample of $2k$ PSUs selected. The method of selection was nominally systematic with probability proportional to size (PPS). The measure of size used was the 2020 census housing unit count. The first step in this procedure for each state is to determine the size of the NSR domain (or stratum) and the size of each NSR PSU. Initially, the size of the first stage NSR sampling frame, denoted M^*, was the total measure of size of all NSR PSUs in the state; and the number of PSUs to select, denoted n^*, was $n^* = 2k$. The initial skip interval for PPS systematic sampling, then, was $S^* = M^*/n^*$. Although we defined NSR PSUs to be as close to equally sized as possible within each state, we were constrained by the requirement to group whole counties; therefore, the NSR PSUs varied in size. Within each state, some larger PSUs had measures of size that exceeded the initial value of S^*. We designated such large PSUs "quasi-SR" or "QSR" because while not in sample every year of the design, they would be in the set of NSR PSUs selected for the twelve-year design period with certainty. Assuming the number of QSR PSUs was less than the initial value of n^*, we reduced n^* by the number of QSR PSUs, also reducing M^* by the cumulative size of the QSR PSUs identified. We then recalculated $S^* = M^*/n^*$, using these updated values. As in the first round, any PSUs with measure of size greater than S^* were designated QSR. These steps are repeated until all remaining NSR PSUs not designated QSR have measure of size less than S^*. The final number of QSR PSUs in each state was less than or equal to the value of $2k$. If $n^* > 0$ after the final step, a sample of n^* non-QSR NSR PSUs was selected, as described below.

8. Selecting the Sample of Non-self-representing Primary Sample Units in Each State

After the process described in the previous section, each state with NSR PSUs had $2k - n^*$ QSR PSUs, where n^* could be any integer between 0 and $2k$, inclusive. If $n^* > 0$, we selected a sample of this many non-QSR NSR PSUs, using a systematic PPS procedure.

In order to have the sample of NSR PSUs be representative of the entire set of NSR PSUs, with respect to the measure of size (2020 census housing unit count), we sorted the set of non-QSR NSR PSUs by descending measure of size. The skip interval used was the final value of S^* from the previous section. A random start R was generated in the usual manner, multiplying S^* by a $U(0,1)$ random number, so that $0 < R < S^*$. A series of n^* hit values, $h_1, \ldots, h_{n^*}$ was calculated, with $h_1 = R$, and, in general,

$$h_i = R + (i-1)S^* \ , \ \ i = 1 \ to \ n^*$$

For each value of h_i, the first PSU for which the cumulative measure of size (in the sort order specified) was greater than or equal to h_i was selected. Note that since the measure of size of all non-QSR NSR PSUs is less than S^*, exactly one PSU will satisfy this criterion for each hit value.

Combining any QSR PSUs identified as in the previous section with any non-QSR PSUs selected as described here, we had the final sample of $2k$ NSR PSUs for each state.

In the six states used for the pilot study, those PSUs corresponding to the Type B areas that began interviewing in 2024 and continued through 2025 may or may not be part of the sample of $2k$ actually selected in the first-stage sampling process described in this and the previous section. If such a PSU is *not* selected, it must be paired with a PSU that is selected (not another PSU that continued from the pilot) which will take its place in the rotation beginning in 2029.

9. Assigning Years in Sample to Selected Non-self-representing Primary Sample Units

As described in section 5 above, in each state with an NSR domain, the sample of $2k$ NSR PSUs selected will cycle in and out of sample annually, with k in sample each year, during the period 2025 through 2036. Since, with sufficient sample, state-level estimates *may* be produced using single years of data, we wanted the sample of k NSR PSUs in sample each year to be representative, with respect to size, of the total sample of $2k$, which is in turn representative of the whole NSR domain. Therefore, after selection, we formed each state sample of NSR PSUs into k size strata of two PSUs each and assigned the sample NSR PSUs to years so that each year will have one PSU from each stratum. The year assignment also satisfied the other criteria listed in section 5.

We first placed the size strata in random order, with the two PSUs within each stratum also in random order. Then the first PSU in each stratum is assigned to be in sample for 2025 and 2026. At the beginning of 2027, the first half of the strata will swap out PSUs. In 2028, the remaining strata will swap out their PSUs. Refer to Figure 1 for an example. If k is odd, the swapping in 2027 will involve a number of strata equal to the *floor* of half the annual sample size, $\left\lfloor \frac{k}{2} \right\rfloor$. Then the swap in 2028 will involve the remaining $\left\lceil \frac{k}{2} \right\rceil$, where this denotes the *ceiling* of half the annual sample size.

In the six states in the pilot study, the procedure just described was amended to account for some PSUs starting their rotations in the previous design period, in 2024. As mentioned earlier, such a PSU may or may not be part of the sample of $2k$, and if not must be paired with a sample PSU that did not start its rotation in 2024. In this case, the sample PSU will take the place of the pilot PSU in the cycling pattern, entering in 2029 (or possibly 2030 in one case). In all six pilot states, the number

PSUs that started in 2024 was no greater than the ceiling of $\frac{k}{2}$. In one pilot state it was decided to retain a PSU that started its rotation in 2023 for a third year, through 2025. This was because the annual number of sample NSR PSUs (k) in this state was larger than the number of PSUs associated with Type B areas that started their rotations in 2024. Also, the continuing pilot PSUs must be placed in size strata, but since all pilot PSUs must be in sample in 2025 and only one PSU per stratum may be in sample, no two pilot PSUs may be in the same size stratum; and the random ordering within a stratum must be ignored because the pilot PSU in a stratum must be first. The stratum ordering was also affected, since we wanted the strata with pilot PSUs to be in the first half. There was one pilot state where the number of pilot PSUs was equal to $\left\lceil \frac{k}{2} \right\rceil$, which means that one of them will continue interviewing through 2026, not being swapped out until the beginning of 2027. Otherwise, all pilot PSUs will be swapped out at the beginning of 2026. If a pilot PSU was in the sample of $2k$, it will cycle back into sample two years after being swapped out. Otherwise, it will be replaced by the sample PSU with which it was paired. These adjustments allow for state-level estimates using three years of data to begin in 2027 for the pilot states, rather than 2028. (Tersine, Jr., 2023)

10. Primary Sample Unit Weights

The NHIS uses baseweights in estimation. The baseweight for a sample household has two components, from the first and second stages of sample selection. We describe in this section the calculation of the first stage component for each PSU that will be applied to all second stage sample units selected within the PSU. We refer to this component as the PSU weight.

The PSU weight is simply the inverse of the probability that a PSU will be in sample for a given year. For NSR PSUs, this seems to imply the weight might change from year to year, depending on the other PSUs in sample that year. However, it turns out the PSU weight is constant across years, as will become evident.

For an SR PSU, since it is in sample with certainty for all years, its weight is simply 1.0.

For an NSR PSU, the unconditional probability it is in sample for a given year is the product of the conditional probability it will be assigned to a given year, given its selection in the twelve-year sample, times the probability of that selection. Since there are two PSUs in each stratum, and the determination of which years each is in sample depend on the random ordering within the stratum – essentially determined by an unbiased coin flip – the conditional probability given selection is simply 0.5, regardless of whether the PSU is QSR or not.

For a QSR PSU, selection in the twelve-year sample is certain, so the unconditional probability of being in sample for a given year is 0.5, making the PSU weight 2.0.

For a non-QSR PSU with index $\boldsymbol{i}$, let the PSU measure of size (2020 HU count) be denoted $\boldsymbol{m_i}$. Let the total number of NSR PSUs in a state be $\boldsymbol{n_{NSR}}$, and let the number of QSR PSUs be $\boldsymbol{n_{QSR}}$, where $\boldsymbol{n_{QSR} \leq n_{NSR}}$. Also, let the corresponding cumulative measures of size for these sets be $\boldsymbol{M_{NSR}}$ and $\boldsymbol{M_{QSR}}$, respectively. If the probability of selection in the twelve-year sample is denoted $\boldsymbol{p_i}$, then

$$\boldsymbol{p_i = \frac{(n_{NSR} - n_{QSR})m_i}{M_{NSR} - M_{QSR}}}$$

Thus, if we denote the PSU weight of a non-QSR NSR PSU with index $\boldsymbol{i}$ as $\boldsymbol{w_i}$, then

$$\boldsymbol{w_i = \frac{2}{p_i}}$$

Note that

$$\sum_{i \notin QSR} p_i = n_{NSR} - n_{QSR}$$

That is, the sum of non-QSR NSR twelve-year selection probabilities is the number of non-QSR PSUs to be selected. Since the twelve-year selection probability of an SR or QSR PSU is 1.0, this generalizes to the sum of twelve-year selection probabilities of all PSUs being equal to the twelve-year number of sample PSUs in a state. Further, taking into account the conditional probability, 0.5, of an NSR PSU being assigned to a given year, the sum of unconditional probabilities of PSUs being in sample in a given year is equal to the annual number of PSUs in sample for a state. This is a familiar result from sampling theory, which is reassuring.

11. Final Note on the Twelve-year First Stage NSR State Sample Sizes

A motivation for cycling NSR PSUs in the 2020-based NHIS sample design is to improve the geographic coverage within each state. This raises the question of why we limited the sample to just twice the number of annual NSR PSUs in sample. If there are enough NSR PSUs in total, one could theoretically cycle through as many as six times the annual sample size over the course of twelve years. In fact, we took this approach early in the process of developing a scheme to cycle NSR PSUs, selecting a "maximal" sample in each state. However, three considerations led to the final decision to select just twice the annual NSR sample size.

First, NCHS will never use more than three years of data to create state-level period estimates. Given the constraints of keeping each PSU in sample for at least two years and replacing only half of the NSR sample each year, having a maximal sample would not actually improve the reliability of any given estimate, even if PSUs were perfectly uniform in size but geographically diverse with respect to characteristics of interest.

Second, the presence of more QSR PSUs, and more generally the variability of PSU size, resulted in empirically greater variance in simple estimates of housing unit totals observed in sampling simulations using maximal samples, as opposed to using the sample size of $\mathbf{2k}$.

Finally, additional costs associated with interviewing in new areas – either hiring and training new interviewers locally or paying for experienced interviewers to travel – would likely be substantial with a maximal sample, with no real benefit in terms of reliability.

Acknowledgements

The author is grateful to his colleagues at the US Census Bureau and staff at the National Center for Health Statistics who provided both an historical foundation and feedback during the process of developing the first stage (PSU) component of the 2020-based NHIS sample design. In particular, he would like to thank John Chesnut for presenting this work at the 2024 JSM, and Michael White for the many brainstorming sessions, the invaluable oversight and double-checking, and for suggesting the presentation of this topic to a wider audience.

References

Cochran, W. G. (1977). *Sampling Techniques (3rd edition).* New York: John Wiley & Sons.

Moriarity, C., Parsons, V. L., Jonas, K., Schar, B. G., Bose, J., & Bramlett, M. D. (2022). Sample Design and Estimation Structures for the National Health Interview Survey, 2016-2025. *Vital Health Statistics, 2*(191). doi:https://dx.doi.org/10.15620/cdc:115394

National Center for Health Statistics. (2023, November 22). *About NHIS*. Retrieved September 24, 2024, from CDC: https://www.cdc.gov/nchs/nhis/about_nhis.htm

Parker, J. D., Talih, M., Malec, D. J., & et al. (2017). National Center for Health Statistics Data Presentation Standards for Proportions. *Vital and Health Statistics, 2*(175).

Tersine, Jr., A. G. (2021). *Memorandum for Stephen J. Blumberg, Director, Division of Health Interview Statistics, National Center for Health Statistics: "Benefits of Cycling Geographic Areas in the National Health Interview Survey".* U.S. Census Bureau, Demographic Statistical Methods Division. Washington, DC: U.S. Census Bureau.

Tersine, Jr., A. G. (2023). *Internal Census Bureau Memo for Jeffrey D. Sisson, Chief, Demographic Systems Division: "2020 Sample Redesign Requirements for the National Health Interview Survey Selection of Primary Sampling Units (WG4_DOC004_PSUSelection_NHIS, Version 1.1)".* Washington, DC: US Census Bureau.