1Department of Commerce
United States Census Bureau
OMB Information Collection Request
2020 Census Post-Enumeration Survey Person Interview and Person Followup Operations
OMB Control Number 0607-XXXX

B. Collections of Information Employing Statistical Methods

1. Universe and Respondent Selection

1The 2020 Post-Enumeration Survey (PES) sample design has been developed to support the various objectives of the program, which includes estimating correct enumerations, erroneous enumerations, and omissions in addition to net coverage for the 2020 Census. The PES is designed to measure the coverage of housing units (HUs) and people, excluding group quarters and people residing in group quarters. The PES will be conducted in the U.S. and in Puerto Rico in selected PES sampled areas. Group quarters (such as college dormitories and correctional facilities) are out of scope because populations can change significantly between census enumeration and PES enumeration operations. Remote areas of Alaska are also out-of-scope for PES because of the seasonal nature of addresses and the population throughout the year make it infeasible to accurately conduct the matching and followup operations necessary for dual-system estimation. For this reason, the Census Bureau's past post-enumeration surveys have never included remote Alaska.

The PES sample consists of two parts. The Population Sample, P sample, and the Enumeration Sample, E sample, have traditionally defined the two systems for dual-system estimation. Both the P sample and the E sample measure the same HU and household population. However, the P sample operations are conducted independent of the census. The E sample consists of census enumerations in the same sample areas as the P sample. For net coverage estimation, after matching with the census lists and reconciliation, the P sample provides information about the population missed in the census whereas the E sample provides information about erroneous census inclusions. The correct enumeration rate and match rate provide an estimate of the true population size using dual-system estimation. Comparing this number to the census counts yields an estimate of net coverage.

The PES is a multiphase operation designed to measure the net coverage and components of coverage for the household population and HUs in the 2020 Census. The PES sample design comprises a number of distinct processes from forming Basic Collection Units (BCUs), creating the sample frame, selecting sample BCUs, to eventually selecting addresses for the P sample and E sample. After the PES BCUs are selected, an address list is created independent of the census for each PES sample BCU. The PES listing workload is 9,832 BCUs for the United States and 397 for Puerto Rico. Overall, approximately 580,000 HUs are listed (560,000 in the nation and 20,000 in Puerto Rico). Finally, after subsampling the PES HUs listed during Independent

Listing, the final expected P-sample size is approximately 171,500 HUs for the nation and 8,000 for Puerto Rico. The national sample is distributed among the 50 states and the District of Columbia roughly proportional to population size, although there are slight increases in the sample for small states and for American Indian Reservations.

Table 1 summarizes the expected 2020 PES listing workloads and P-sample sizes for the United States and Puerto Rico. The E-sample size is expected to be about the same as that for the P sample.

Table 1: 2020 PES Universe and Sample Housing Unit Summary

Geography	Expected Listing Sample Size	Expected P-sample Size
U.S.	560,000	171,500
Puerto Rico	20,000	8,000
Total	580,000	179,500

The PES sample has three phases of sampling. In the first phase of the PES sampling, BCUs in each state are classified into mutually exclusive and relatively homogeneous groups known as sampling strata. These strata are based on the BCU size and whether the BCU is located on an American Indian Reservation. The four major strata are:
  (1) BCUs with 0 to 2 HUs (small stratum).
  (2) BCUs with 3 to 57 HUs (medium stratum).
  (3) BCUs with 58 or more HUs (large stratum).
  (4) BCUs on American Indian Reservations with three or more HUs (American Indian Reservation stratum).
Using 2010 Census data, the medium and large strata are further split into renter and owner BCUs, resulting in up to six sampling strata being formed in each state and Puerto Rico. The definition of the large and medium groups in 2020 PES is different from the 2010 Census Coverage Measurement (CCM). For 2010 CCM, the medium group contained block clusters with 3 to 79 HUs and the large group contained block clusters with 80 or more HUs. Since the size of the BCUs is smaller than the block clusters on average, this change preserves the same proportion of frame HUs in the medium and large strata between the two designs (2010 and 2020). This yields a similar HU Independent Listing (IL) workload.

BCUs in the larger stratum are selected with higher probability of selection than BCUs in the medium stratum in this first phase because HUs in large BCUs are expected to be subsampled in the third phase. This allows more BCUs to be selected into sample. The nonowners stratum is sampled at a higher rate (1.5 times higher) than the BCUs in the owner stratum. This is the same differential sampling factor used in the 2010 CCM. The differential sampling factor of 1.5 provides a balance between improving the reliability of estimates for the smaller nonowner domain and improving the precision of other domain estimates and the total for a fixed sample

size. Within each of the six sampling strata for each state, the BCUs are sorted and a systematic sample is selected with equal probability.

The second phase of the 2020 PES selects a subsample of BCUs from the first-phase small sampling stratum using a similar method as in the 2010 CCM design. We use a double-sampling technique by selecting a slightly larger sample of small BCUs in the first phase then selecting a subsample of small BCUs for Person Interview (PI) using an updated measure of size. This is done to reduce a BCU's influence on the estimates when more HUs are found than expected. Additionally, small BCU subsampling reduces costs, as conducting interview and Followup operations in small BCUs is more expensive per HU than in medium or large BCUs.
Using HU counts from both the Independent Listing and the updated census address list, the small BCUs selected in the first phase are restratified by these counts within each state. A systematic sample of BCUs is selected within each stratum with equal probability. All BCUs from the small sampling stratum with ten or more HUs, based on the updated counts, are retained in sample. All BCUs from the small sampling stratum that are in American Indian Country are also retained in sample. (American Indian Country includes American Indian Reservations and associated trust lands, as well as the American Indian statistical areas.)

The first and second phases of the 2020 PES select the BCU sample. In the third phase of PES sampling, we select a subsample of HUs within large BCUs. For a BCU with 57 or fewer HUs observed, all the HUs are included in the sample. For a BCU with 58 or more HUs observed, a subsample of segments of contiguous HUs is selected to facilitate data collection in the field and to reduce the impact of intraclass correlation on the variance. This phase of sampling results in more similar overall selection probabilities for HUs because the large BCUs will have a higher probability of selection during the first phase.

The sampling frame for the P-sample HUs is the result of the PES initial HU matching and follow-up operations. The intent of these HU operations is to identify matches between the independent HU list and an early census HU list. In addition to sending the P sample to the PI operation, a sample of census units that were missed during the Independent Listing operation will be sent to the Person Interview. After person data collected from the PI operation are matched to data collected by the census enumeration in the PES sample areas, certain cases with discrepancies between the PES PI and census or incomplete codes in the PES will be sent for another PES interview called the Person Followup (PFU). The P-sample people result from the person interviewing in the P-sample HUs.

The sampling frame for the E-sample HUs consists of the HUs in PES sample areas from the list of 2020 Census enumerations that is available after the P sample is selected. While these two samples are selected at different points in time, we attempt to geographically overlap them to the extent possible. The E-sample people are the census enumerations in the E-sample HUs with enough information collected.

2. Procedures for Collecting Information

During PI, interviewers use a computer-assisted data collection instrument to obtain information about the current residents of the sample housing unit and certain people who moved out of the sample housing unit between Census Day and the time of the PES interview. The instrument collects names, addresses, where they lived on Census Day and where they lived at the time of the interview. Demographic information (as in the census) is also collected. The expected workload for this operation is about 190,000 housing units, including approximately 181,000 cases for the nation and approximately 9,000 cases in Puerto Rico.

For the Person Interview Reinterview operation, cases will be selected from the Person Interview work as it is returned. All cases are eligible for reinterview except cases where the outcome of the PI was a refusal, language problem, no knowledgeable respondent, or if more than one interviewer completed the original PI case.

There are four sampling components to the PI Reinterview operation: random, analytic, supplemental, and rework. The random method selects a random sample of the eligible cases completed by every interviewer for reinterview. The analytic method is based on statistical calculations and paradata collected in the PI instrument. The supplemental selections allow staff to select additional cases for reinterview for any interviewer at any time during the PI operation if they suspect an interviewer may not be following procedures. Rework occurs when the outcome for an interviewer is a hard fail, regardless of how the case was selected, and all of that interviewer's eligible cases are reworked. The PI Reinterview operation includes a reinterview caseload of approximately 15 percent of the total PI workload.

After PI Reinterview cases are selected, the cases are made available to the regional census center (RCC). Staff at the RCC will conduct the reinterview over the phone if a phone number is available for the case. If a telephone number is not available or a case is not able to be completed by telephone, the case will be made available for assignment to a reinterviewer for a personal visit. The reinterviewer will complete the PI Reinterview case with the original PI household respondent, the original proxy respondent, or a new respondent, depending on the unit status, respondent type, and the availability of contact information from the PI case. The reinterview instrument will collect information about the unit status on the PI date, whether the original respondent was contacted, and household roster for cases determined to be occupied on the PI date.

Data collected from the original interview and the reinterview are evaluated to determine if the original interviewer falsified data. The Census Bureau uses a three-stage approach to match the original interview and reinterview to evaluate the cases. The first stage uses an automated system to determine if the original interview and reinterview information matches. If it does not match, the case is forwarded to the National Processing Center (NPC). In this second stage, clerks in NPC review the case information and try to determine if the discrepancies can be explained (such as discrepancies created by the transposition of letters or misspellings). If an

NPC clerk determines there is evidence of deliberate falsification or intentional violation of procedures in the case, as well as a pattern of such activity from related cases, they may suggest coding the case for a Hard Fail Recommendation. If this is approved by an NPC supervisor, it will be forwarded to the RCCs. The RCC staff only review hard fail recommendations or interviewers with a pattern of fails. In the third stage, managers in the RCCs will code these cases accordingly. All codes assigned during this stage are considered final.

During PFU, interviewers use the PFU paper questionnaire, Form D-1301 for stateside or Form D-1301(PR) for Puerto Rico, to obtain information about the people selected for follow-up. The expected workload for this operation is approximately 65,000 cases, including approximately 62,000 stateside cases and 3,000 Puerto Rico cases.

For the PFU Reinterview operation, cases will be selected from the PFU cases as they are returned. Cases are eligible for reinterview if the PFU outcome was complete or partial, if there was only one respondent for the case (either a household member or a proxy), and if the case was not observed by a field supervisor. There are only two sampling components of the PFU Reinterview operation: Random and Supplemental, which are described above. About 15 percent of all PFU cases will be selected for PFU Reinterview.

After PFU Reinterview cases are selected, staff at the RCC will conduct the reinterview over the phone if a phone number is available for the case. If a telephone number is not available or a case is not able to be completed by telephone, the case will be made available for assignment to a reinterviewer for a personal visit. The reinterviewer will use the reinterview paper questionnaire, Form D-1301.2 for stateside or Form D-1301.2(PR) for Puerto Rico, to complete with the original PFU respondent or a new respondent. The reinterview questionnaire will collect information about whether the original respondent was contacted. If the original respondent was not contacted, the reinterviewer will conduct a new PFU interview for that case.

3. Methods to Maximize Response

The Person Interview and Person Followup questionnaires contain the minimum number of questions necessary to obtain the data required for the 2020 PES. The interviewer will make up to six attempts for each operation to obtain an interview. The interviewer will explain the reason the Census Bureau is conducting the operation, and respondents will be informed of their legal responsibility to answer the questions. In addition, respondents will be assured that their answers are confidential.

4. Testing of Procedures

Nothing will be tested during the Person Interview and Person Followup operations.

5. Contacts for Statistical Aspects and Data Collection

Timothy Kennel
Statistical Methods
Decennial Statistical Studies Division
U.S. Census Bureau
301-763-6795

Definition of Terms

Components of Census Coverage – The components of census coverage include correct enumerations, erroneous enumerations, whole-person imputations, and omissions. Correct enumerations are people or HUs that were correctly enumerated in the census. Erroneous enumerations are people or HUs that were enumerated in the census but should not have been. Examples of erroneous enumerations are duplicates, nonexistent HUs or people, and people or HUs that were enumerated in the wrong place. Omissions are people and HUs that were not correctly enumerated in the census but should have been. Lastly, whole-person imputations are census records for which all of the demographic characteristics were imputed. Many of these imputations represent people in HUs where we knew the household count but did not obtain information about the people residing at the HU.

Net Coverage – Reflects the difference between the true population (which is estimated by the Post-Enumeration Survey) and the census count. If the census count was less than the actual number of people or HUs in the population, then we say there was an undercount. If the census count was more than the actual number of people or HUs in the population, then we say there was an overcount.

List of Attachments

1. D-1409(E/S), Person Interview Introductory Letter (Privacy Act Notice)
2. D-1709(E/S), Person Followup Introductory Letter (Privacy Act Notice)
3. D-1301 and D-1301 (PR), Person Followup Questionnaire
4. D-1301.2 and D-1301.2(PR), Person Followup Reinterview Questionnaire
5. D-1400, D-1400(S), Person Interview Information Sheet
6. D-26, Notice of Visit
7. D-ID, Language ID card
8. PES Person Interview 2020 Instrument Spec
9. PES Person Interview Reinterview 2020 Instrument Spec