

**OMB Information Collection Request
Supporting Statement B
U.S. Department of Commerce
U.S. Census Bureau**

High Frequency Surveys Program

Household Pulse Survey Phase 4.0

OMB Control Number 0607-XXXX

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

As noted in Part A, the Census Bureau developed the Household Pulse Survey as an experimental endeavor in cooperation with five other federal agencies. to produce data in near real-time in the face of a national emergency. The survey is conducted under the auspices of the U.S. Census Bureau's Experimental Data Products Series (see <https://www.census.gov/data/experimental-data-products.html>). The High-Frequency Surveys Program was established as a natural progression from the creation of the Household Pulse Survey. High-frequency surveys are designed to develop and deploy data collection instruments quickly and for data to be released in near real-time.

- 1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.**

The universe for the Household Pulse Survey is the approximately 140,000,000 housing units in the United States. For operational efficiency, only housing units linked to one or more email addresses or cell phone numbers will be eligible for the sample. One respondent from each housing unit is selected to respond for him or herself. The sample for each data collection cycle is approximately 1,040,000 housing units.

In Phase 1, the Census Bureau observed response rates in the range 6-7 percent for households who were invited to participate for the first time and asked to respond in a one-week period. In Phase 2, we extended the response period to 13 days and sent additional reminders to respondents. These efforts had the effect of increasing response rates to a range of 8.1 to 10.3 percent. In Phase 3, we continued with the 13-day response period and observed a response rate of between 5.3 and 7.5 percent. In Phase 3.1, Phase 3.2, and Phase 3.3, we have observed a response rate of between 5.4 and 7.4 percent. Phase 3.4 averaged a response rate of 6.6% and the response rate of Phase 3.5 started at 6.2% and eventually dropped to 4.4%. The response rate over the most recent three waves of data collection is in the range of 3.9 - 4.7 percent. For Phase 3.6, the response rate ranged from 3.9% to 5.6%. There was an increase in response rates for Phase 3.7, with rates ranging from 6.4% to 7%. Phase 3.8 had an average response rate of 5.9% and Phase 3.9 averaged a response rate of 6.3%. The first two data collection periods of Phase 3.10 yielded an average response rate of 6.4%. Based on these recent experiences, we are assuming a 6.3% percent response rate for Phase 4.0 and expect to receive approximately 66,150 responses each data collection cycle for a total of 198,450 for the entire phase. Sample from each data collection cycle will be independent from the prior cycles.

2. Describe the procedures for the collection of information including:

- **Statistical methodology for stratification and sample selection,**
- **Estimation procedure,**
- **Degree of accuracy needed for the purpose described in the justification,**
- **Unusual problems requiring specialized sampling procedures, and**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

Housing units linked to one or more email addresses or cell phone numbers are eligible for the sample. These housing units are stratified by state and the top 15 metropolitan areas. For an estimate of 40 percent, the sample size was determined to produce a margin of error of 2.8 percent for most states and 3.3 percent for the 11 smallest states with eligible sample units (Alaska, Delaware, District of Columbia, Hawaii, Maine, Montana, North Dakota, Rhode Island, South Dakota, Vermont, and Wyoming) with 90 percent confidence. The sample was then allocated to states and metropolitan areas proportionally, using the number of housing units as the measure of size. Samples cases will be matched to the Census Bureau's Contact Frame to assign email addresses and cell phone numbers. See Attachment A for information on the Contact Frame.

A sampled address may have up to five associated email addresses and up to five cell phone numbers. Only the top three of each will be used to contact the sampled units. Email invitations will be sent in the morning and an SMS invitation in the afternoon. Each piece of contact information may receive an initial invitation and up to two reminders. Once a unit has responded, they will no longer receive contacts.

For Phase 4.0, we will incorporate continuous data collection spread throughout the month of data collection. This method will utilize invitation and reminder distributions across 12 sample groups, with each group starting at different points throughout the month. See Attachment C for an illustration of this sample distribution plan.

The Census Bureau will conduct this information collection online using Qualtrics as the data collection platform. Qualtrics is currently used at the Census Bureau for research and development and production surveys and provides the necessary agility to deploy the Household Pulse Survey quickly and securely. It operates in the Gov Cloud, is FedRAMP authorized at the moderate level, and has an Authority to Operate from the Census Bureau to collect personally identifiable and Title 13-protected data.

Attachment B sets forth details on the Census Bureau's weighting plan, which will use a multi-step process to help reduce nonresponse bias in the estimates.

- 3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

The survey is designed to meet the goal of rapid-response, accurate, and timely estimates. It will be conducted by an internet questionnaire, with links sent by email and SMS text message. Up to three email addresses and cell phone numbers will be used for each unit to increase response for nonrespondents. Two reminders will be sent to each piece of contact information if a response is not received. These modes are expected to yield response rates much lower than traditional surveys. The benefits are implementation efficiency, cost, and timeliness of responses.

Standard errors will be large but will be available for data users to understand the quality of the estimates. Nonresponse bias is likely to be an issue with this design, but measures such as the demographic distribution of the survey respondents compared to benchmarks will be produced for data users to consider in their analyses. Limitations due to collection and contact modes may exacerbate issues associated with non-response by limiting options for reducing bias. Weighted distributions of key demographic covariates are included in the detailed tables and are accompanied by companion standard error tables. These distributions of these characteristics can be compared with other federal and non-federal data sources. The microdata allows for the comparison of weighted and unweighted estimates and evaluations of gaps in response. Source and Accuracy documents include coverage rates both before and after weighting raking procedures and are released with each data cycle release. There are no other known data sources that can provide comprehensive, relevant, and timely information. At the very least, this survey will provide non-anecdotal evidence about the effects of

the pandemic on the population even if the quality is lower than would be traditionally acceptable.

Source and Accuracy statements and other technical documentation is available at <https://www.census.gov/programs-surveys/household-pulse-survey/technical-documentation.html>.

- 4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.**

Prior to Phase 1, the Center for Behavioral Science Methods completed a thorough expert review that consisted of nine independent experts reviewing the questionnaire independently, followed by a consensus meeting to discuss comments and come up with recommendations. The recommendations were then passed to the Demographic Directorate, who accepted most of the expert recommendations and responded with subject-matter justifications for those recommendations that were not accepted. In the end, the questionnaire comprised a consensus of the expert survey methodologist recommendations and the subject matter needs. A similar process was followed prior to subsequent phases as the Census Bureau considered questionnaire revisions to keep up with evolving data needs. As was necessary, we implemented questionnaire experiments during past phases to refine the questionnaire. Results from these experiments were used to improve the data collection instrument. Results from these tests will be published.

We will take a similar course for Phase 4.0: feedback from Federal agency partners, data users and other stakeholders on diminishing, changing or emerging data needs will inform changes to the questionnaire.

Phase 4.0 will include methodological research to understand different methods of weighting and producing estimates, impacts of including a longitudinal design to estimates (see OMB Control 0607-0125) and alternate means of handling item missingness. Ongoing research investigates the use of supplemental administrative records on these estimates. The Census Bureau will publish results of these investigations for the public.

The Census Bureau will release these data under the auspices of its Experimental Statistical Product Series. Information on the Series is available at <https://www.census.gov/data/experimental-data-products.html#>.

5. **Contacts for Statistical Aspects and Data Collection**

Statistical Design:

Anthony Tersine
Demographic Statistical Methods Division
Demographic Programs Directorate
anthony.g.tersine.jr@census.gov

Data Collection/Survey Design:

Jason Fields
Social Economic and Housing Statistics Division
Demographic Programs Directorate
jason.m.fields@census.gov

Jennifer Hunter Childs
Demographic Programs Directorate

jennifer.hunter.childs@census.gov

Statistical Analysis:

David Waddington
Social Economic and Housing Statistics Division
Demographic Programs Directorate
david.g.waddington@census.gov

Attachments

- A. Census Bureau Contact Frame
- B. Household Pulse Survey: Weighting Plan
- C. Sample Distribution Plan