2023 NSDUH Small Data Collection Effort in Puerto Rico Addendum

1. Introduction

This document summaries the 2023 NSDUH small data collection effort in Puerto Rico for the possible expansion of full-scale data collection starting in a future NSDUH survey. Respondent data gathered during the pilot were used for operational purposes only and are not included in the main study dataset. The document is organized by the following tasks:

- Sampling
- Instrumentation
- Staff recruitment
- Staff training
- Address verification, mailing, and shipping
- Data collection
- Feedback from field interviewers
- Summary of findings and recommendations

2. Sampling

Sample frame construction was facilitated using electronic dwelling unit (DU) enumeration (i.e., eListing) in 10 segments. The eListing effort was more challenging in Puerto Rico than in the mainland United States due to non-standardized, wide variation in Puerto Rico address formats. Additional information, such as an urbanization code,[1] subdivision, condominium number or name, or other geographic indicators, is often needed to locate residential DUs. Therefore, when eListing the segments for the small data collection effort, descriptions were frequently used, which led to additional effort by the sampling, data editing, and programming teams to process and manage these data. Several changes were made to the main study eListing application to account for descriptions and preserve consistency across training programs.

For the small data collection effort, an equal number of sample dwelling units (SDUs) (i.e., 136) was selected in each segment except in one segment where a smaller number of DUs (i.e., 130) was selected. The remainder were reserved for potential full-scale data collection in 2024. A total of 1,354 DUs were selected for the small data collection effort. The sample was designed to yield 100 completed interviews (10 completed interviews per segment).

3. Instrumentation

Prior to fielding the small data collection effort, a Spanish language methodologist reviewed the screening and interview specifications and respondent materials to assess the need for any changes warranted because of Puerto Rican dialects. The language methodologist then

---

[1] An urbanization code denotes an area, sector, or residential development within a geographic area.

reviewed the specifications and materials with the field supervisor (FS) for Puerto Rico, who is bilingual and lives in Puerto Rico. These reviews uncovered no warranted changes in the screening instrument and only minimal changes in the interview instrument. The interview instrument changes for the small data collection effort were as follows:

- "Puerto Rico" was added to FIPE4 as a possible location for a respondent to live.

- "Puerto Rico" was added to all STATE FILL locations throughout the instrument.

These changes were incorporated into the existing NSDUH instrument using wording fills triggered by interview location (i.e., "Puerto Rico"). Furthermore, respondent materials were tailored for use in Puerto Rico to include the correct suicide hotline number for Puerto Rico.

4. Staff Recruitment

An employment website was used to source applicants. Job descriptions included two unique requirements: (1) physical location in Puerto Rico and (2) proficiency in English and Spanish. Except for these requirements, all recruitment procedures followed the same as those on the mainland.

5. Staff Training

The eListing and FI training programs mirrored those used in the main NSDUH and included some customization to Puerto Rico. Equipment and hard-copy training materials were shipped to listers and FIs two days earlier than typically scheduled. This extra time accounted for anecdotal information that packages shipped from the mainland take slightly longer to be delivered to home addresses in Puerto Rico.

Puerto Rican FIs attended a regular new FI training session. Changes were made to the training program for Puerto Rican FIs to provide them more opportunities to learn and practice in Spanish. The key changes included additional group and individual exercises in Spanish and extended bilingual training. In addition, Puerto Rican FIs attended a team meeting led by the Puerto Rican FS and the language methodologist to review logistical and data collection considerations unique to Puerto Rico.

6. Address Verification, Mailing, and Shipping

Address verification testing was performed using one USPS (i.e., Coding Accuracy Support System) and one UPS (i.e., UPS Address Verification API) service to identify addresses that may not be deliverable. Both platforms identified addresses that were not recognized by the U.S. Postal Service (USPS). One platform was used to convert unrecognized addresses into valid, deliverable addresses. Ultimately, the decision was made not to utilize either validation platform for two reasons: (1) concerns that doing so could result in removing addresses that could be located by a postal carrier even if not validated by USPS, and (2) that automatically modifying addresses could result in sending lead letters to houses outside of the segment area instead of intended residences within the sample area. Instead, mapping staff reviewed the

addresses and reformatted them for consistency based on the USPS standards for Puerto Rico (USPS, 2023).

In preparation for data collection, the project's case management system was updated to include processes to print labels for Puerto Rico addresses. Test lead letters were mailed to the FS and FIs to evaluate how long it took for mail to be received from the mainland. On average, it took 6 days for mail to arrive compared to about 3 days on the mainland.

After validating addresses, 807 out of 1,354 addresses (59.6 percent) were sufficiently accurate to have lead letters sent. All other addresses were only descriptions and did not have mailable addresses. Out of the 807 lead letters that were mailed, 309 (38.3 percent) were returned to the project's Field Distribution Center.

During the screening, address edits were made to 157 DUs. Multiple edits were made for some DUs if a respondent indicated a different preference for the address during the screening process.

7. Data Collection

Screening and interviewing for the small data collection effort lasted approximately seven weeks (April 20, 2023 – June 12, 2023). Of the 1,354 DUs selected, 1,280 were eligible (94.5 percent). Among these DUs, 357 screenings were completed, for an unweighted screening response rate of 27.9 percent. From the completed screenings, 193 respondents were selected for the NSDUH interview. Of those people selected, 100 completed the interview, for an unweighted interview response rate of 51.8 percent. Initial contact with respondents in Puerto Rico followed the same procedures that were followed on the mainland; however, given the short data collection period, there were no follow-up communications (e.g., web follow-up letters, unable-to-contact letters, and refusal letters). This should be considered when comparing the small data collection response rates to the main study response rates.

Tables 7.1 and 7.2 provide summaries of screenings and interviews completed by mode (in-person or web) and language (English or Spanish), and Table 7.3 shows respondent age distributions.

Table 7.1 Screenings Completed, by Mode and Language

Mode	English	Spanish	Total
In Person	1	332	333 (93%)
Web	8	16	24 (7%)
Total	9 (2%)	348 (98%)	357

Table 7.2 Interviews Completed; by Mode and Language

Mode	English	Spanish	Total
In Person	10	84	94 (94%)
Web	2	4	6 (6%)
Total	12 (12%)	88 (88%)	100

Table 7.3 Main Study and Puerto Rico Small Data Collection Effort Respondent Age Distribution

Result	2022 Main Study Percentage	Puerto Rico Small Data Collection Effort Percentage
Rostered Individuals, by Age Group		
12-17	9.16%	7.76%
18-25	11.42%	9.12%
26-34	13.65%	9.93%
35-49	22.13%	21.22%
50+	43.64%	51.97%
Selected Interview Respondents, by Age Group		
12-17	23.30%	26.42%
18-25	24.92%	18.65%
26-34	15.21%	14.51%
35-49	20.53%	19.17%
50+	16.04%	21.24%
Interviews Completed, by Age Group		
12-17	20.76%	25.00%
18-25	24.18%	19.00%
26-34	15.88%	8.00%
35-49	21.26%	21.00%
50+	17.93%	27.00%

8. Feedback from Field Interviewers

FI feedback was collected via two methods: an online debriefing survey followed by an online focus group. Overall, FIs indicated data collection went well. FIs reported that the data collection equipment and software applications were easy to use. They also reported the DU locations displayed on the electronic maps were largely accurate, but the addresses listed in the screening program were only "somewhat accurate." This was primarily due to lack of standardization and the variation of address formats in Puerto Rico.

One major outcome of the address inaccuracies was respondents not receiving a lead letter prior to an FI's visit. FIs reported that respondents had "seldom" or "never" received a letter, which made respondents more reluctant to participate. FIs suggested that more public outreach or "marketing" about the survey could be helpful to make respondents more aware of the survey and feel reassured about its legitimacy.

9. Summary of Findings and Recommendations

Following is a summary of findings and recommendations from the small data collection effort in Puerto Rico.

- Sampling

4

- Satellite image quality is poor in many areas (e.g., heavily wooded, rural areas), and there is often no street view. This makes it more difficult to provide in-house sampling support to field staff and contributes to higher costs and frame error.
- Establishing procedures for collecting addresses using the eListing application and training of listers was difficult because Puerto Rico does not have standard address conventions like in the mainland United States.
- Because of heavy use of address descriptions, extra effort to compile address data in data management systems is required.

- Instrumentation
  - Minor changes to the interview instruments (in-person and web) are warranted to include U.S. territories, such as Puerto Rico, as response options for where a respondent was born.
  - FIs should receive additional training on how to address respondent questions and concerns about the sex assigned at birth and gender identity questions.
  - In full-scale data collection, the frequency of responses for questions tailored for Puerto Rico should be compared with data from mainland respondents. Any potential concerns with respondent comprehension should be assessed, followed by adjustments to questions or FI instructions as warranted.

- Staff recruitment
  - Puerto Rico should have an FS located in the territory. This facilitates better coaching of FIs due to familiarity with the culture and geography, creates a stronger personal connection with FIs, and eases travel to help resolve barriers to access, meet with community leaders, and mentor FIs.
  - We expect the vast majority of interviews to be completed in Spanish. This requires all FIs to be proficient in Spanish and English.

- Staff training
  - Additional training on locating DUs is warranted due to differences in Puerto Rico address standards and poor satellite imagery in many areas.
  - Puerto Rican FIs would benefit from additional practice addressing respondents' concerns in Spanish.
  - Additional instructions for onboarding are warranted due to different legal employment requirements in Puerto Rico.

- Data collection
  - Logistics
    - All data collection materials should be translated into Spanish.
    - Because of mailouts requiring more time in Puerto Rico than on the mainland, lead letters should be sent at least one week prior to the commencement of data collection.
    - Data collection equipment and bulk supplies should be shipped to FIs in Puerto Rico several days earlier than to FIs on the mainland.
  - Production and response rates

- Some respondents reported wanting to participate to help the United States but not Puerto Rico because of concerns about corruption.
- There was very limited web screening and interviewing production in Puerto Rico because some households did not receive lead letters, and some areas had low Internet connectivity.
- Adverse weather is an unpredictable variable. We cannot extrapolate from the small data collection effort given that it was done in a very short period with no adverse weather events encountered.
- Physical barriers to DU access are very common. During the small data collection effort, roughly a quarter of SDUs were not accessible because they were behind gates. This should be considered when predicting the sample size needed to meet interview targets.
- Given the short data collection period, there were no follow-up communications (e.g., web follow-up letters, unable-to-contact letters, refusal letters). For this reason, we cannot reasonably predict response rates for Puerto Rico.
  - Data quality
- Efforts to verify the quality of FIs' work are feasible to implement for full-scale data collection.
- The quality of the data collected cannot be assessed without a much larger set of completed cases.
  - Cost
- Although Puerto Rico is a small area, a significant amount of overnight travel is needed due to geographical obstacles and location of FIs relative to sample location.
- The higher prevalence of physical barriers to DU access requires more FI time to identify and contact gatekeepers and residents.
- A very large proportion of interviews were completed in Spanish. This requires considerable FI time to contact DUs and complete interviews because the time to conduct Spanish interviews is generally longer than in English.
- As a result of the aforementioned factors, the average cost per case is higher in Puerto Rico than in the mainland United States.

Based on these findings, full-scale data collection in Puerto Rico is feasible with sufficient planning time, assuming Office of Management and Budget and Institutional Review Board approval.