Appendix O. NASS Comments (with responses)

Review of OMB Documentation for Study of Third-Party Processor Services, Fees, and Business Practices Study

December 7, 2017

Reviewed By: Evan Schulz, NASS Summary, Mathematical Statistician

Part A. Justification

The total burden should also include the time needed to review the 2nd invitation letter (Appendix G) for the portion of the sample that you expect will receive the 2nd letter. Make that clear here.

> The completed SNAP retailer survey response burden is estimated at 20 minutes (.33 hour) per respondent. This burden includes the time needed to review the SNAP retailer survey invitation (Appendix F) and respond to the questionnaire in English (Appendix I) or Spanish (Appendix J). The burden also includes a second survey invitation (Appendix G) for those who do not respond to the first invitation and one or more telephone follow-ups (Appendix H) with non-responders.

Part B. Collections of Information Employing Statistical Methods

There are several open-ended response options in the questionnaire. How will IVR handle open-ended responses?

> The IVR system will record open-ended responses as voice messages. Voices messages will be transcribed to text for analysis.

You're not really resampling to assess the incidence of undercounting chain stores, right? If I'm reading the next paragraph correctly, you're really resampling in order to address possible undercounting of chain stores in your sample. You're drawing an initial sample, then using that sample to assess the incidence of undercounting chain stores and correcting units in the sample that are misclassified, then resampling in order to match population proportions for strata within your sample.

> Resampling here means pulling, with replacement, a random sample multiple times in order to estimate the true proportion of chain stores in the universe of 248,188 qualified stores. The researchers will pull a random sample of stores from the database and

calculate the proportion of chain stores (including chain stores without chain store number) in that sample. The researchers will then “put that sample back in the database”. We will pull another random sample from the database and calculate the proportion of chain stores (including chain stores without chain store number), then put that sample back in the database. We will repeat the process several times. The distribution of the proportion of chain stores, calculated from the different samples, will be used to determine the true proportion of chain stores in the target population with a given confidence level. This will help mitigate the risk of non-representativeness of chain stores in the survey sample, a non-representativeness that can occur because not all chain stores in the target population can be identified as such.

How do you square this with the study objectives in Appendix A that seem to require quantitative analysis (e.g. the prevalence of ISOs providing support to SNAP retailers, comparing costs that ISOs and TPPs charge for similar services)?
Also, if you do end up doing quantitative analysis or making population inferences from data from the ISO interviews, how will you address the inherent biases associated with snowball sampling?

The information gleaned from these interviews will be used to describe the relationships between different types of ISOs and TPPs and how these relationships affect the products and services offered to retailers. This information cannot be obtained from the retailers, as they may not know how the services they procure connect them to EBT. The organizations interfacing with SNAP retailers to provide products and services related to EBT will be identified using a nationally representative survey of SNAP retailers.