Survey Research Methods (2020)
Vol. 14, No. 3, pp. 289-300
doi:10.18148/srm/2020.v14i3.7465


ISSN 1864-3361
http://www.surveymethods.org

# Response patterns in a multi-day diary survey: implications for adaptive survey design

Mengyao Hu
Institute for Social Research
University of Michigan
Ann Arbor, USA

Edmundo Roberto Melipillán
Facultad de Psicología
Universidad del Desarrollo
Concepción, Chile

Brady T. West
Institute for Social Research
University of Michigan
Ann Arbor, USA

John A. Kirlin
Kirlin Analytic Services
Fredericksburg, USA

Ilse Paniagua
Institute for Social Research
University of Michigan
Ann Arbor, USA

In multi-day diary surveys, respondents make participation decisions every day. Some respondents remain committed throughout, whereas others drop out after the first few days or in the later days of the survey, leading to item nonresponse. Such item nonresponse at the day level can introduce nonresponse and underreporting error, reduce statistical power and bias survey estimates. Despite its critical influence on survey data quality, the important issue of day-level item nonresponse in diary surveys has received surprisingly little attention. This study evaluates different response patterns in a seven-day diary survey and considers how these patterns might inform adaptive designs for future diary surveys. We analyzed data from the U.S. National Household Food Acquisition and Purchase Survey (FoodAPS), a nationally representative survey designed to collect comprehensive data on household food purchases and acquisitions during one-week time periods. In total, there were 4,826 households with 14,317 individuals that responded to the survey. To evaluate how response patterns differed across respondents and across the diary period, we employed a latent class analysis (LCA), which enables the identification of different groups of respondents based on their reporting patterns. Our analysis identified six classes of respondents, ranging from "stayers" (individuals with a high probability of participation for all seven days) to those exhibiting minimal effort. To inform adaptive designs in future diary data collections, we evaluated respondent profiles for each derived class based on variables that were known before the diary portion of the survey. Supplemental Nutrition Assistance Program (SNAP) status and any evidence of refusing the screening interview were linked with classes having a higher probability of dropping out. Our findings have implications for future designs of multi-day diary surveys.

*Keywords:* Multi-day diary surveys; Latent class analysis; Response patterns; Adaptive survey designs

## 1 Introduction

Multi-day diary surveys collect data on daily events in many fields, including time use, consumer expenditure, food consumption, nutrition and health. They often involve a diary book that respondents use to record events when or shortly after they occur. In an ideal situation, respondents will closely follow the survey instructions and truthfully report the events as soon as they occur. However, not all respondents are as diligent as survey researchers would like them to be. Respondents may drop out over the course of a diary survey, or may simply report that no events occurred for a day to reduce the time burden of survey participation. Prior research has found that rates of reporting events in diary

Mengyao Hu, Institute for Social Research, 426 Thompson St., Ann Arbor, MI, USA (E-mail: maggiehu@umich.edu)

surveys will likely go down over time (Hu, Gremel, Kirlin, & West, 2017; Stone, Kessler, & Haythomthwatte, 1991), the dropout rate is high in long-term diary surveys (Carson et al., 2006), and respondents may report on some but not all days of the diary period, leading to item nonresponse (Hu et al., 2017). Such item nonresponse at the day level in diary surveys can introduce nonresponse and underreporting error (Hu et al., 2019), reduce statistical power and even bias survey estimates. Despite its critical influence on survey data quality, the important issue of item nonresponse at the day level in diary surveys has received surprisingly little attention.

In past survey research, respondents who fail to complete all days in a diary period, regardless of their response patterns, have often been perceived and treated similarly. For example, all of those respondents, no matter on which day or for how many days they drop out, can be classified as one group and compared to those who stay for the entire diary period (Wiseman, Conteh, & Matovu, 2005). The drawback of this approach is that it ignores the possibility that those who fail to report for all diary days may have different response patterns. For example, some respondents may respond only for the first few days and then drop out, some may drop out at a much later stage, and others may respond infrequently with an unclear pattern. Different patterns may have different underlying causes and may require different interventions, imputation strategies, and analysis methods. To better understand and address these dropout issues, it is important to understand the types and implications of different response patterns in diary surveys. The objectives of this current study are two-fold: first, to evaluate different response patterns in a 7-day diary survey, and second, to examine how these patterns can inform future diary survey design and analysis in an adaptive survey design framework.

Based on survey satisficing theory (Krosnick, 1991), item nonresponse at the day level in diary surveys can be related to the following factors: task difficulty, respondents' motivation, and respondents' fatigue. Greater motivation is likely associated with less item nonresponse. The greater the task difficulty, the more likely item nonresponse is to occur. The same is true for respondents' fatigue. Two other factors discussed in the panel survey literature may also be related to day-level item nonresponse in diary surveys: habits and sudden shocks (e.g., sudden dropout due to an unexpected event) (Lugtig, 2014; Stone et al., 1991). With better habits of completing diaries and fewer shocks during the diary survey process, item nonresponse is less likely to occur. We now review how each of these factors might impact diary surveys in turn.

**Task difficulty** A diary survey is different from other types of surveys—it requires respondents to fill out diaries for a specific period of time (e.g., usually a week for TV diaries and food expenditure diaries; Dillman, 1991) and ideally to record events as soon as they occur instead of recalling events. Most diary surveys use self-administered modes, such as paper-and-pencil and the web. As discussed in previous literature, a long and complicated diary instrument will likely result in a high dropout rate (Stone et al., 1991). To make sure that respondents fill out the diaries in a timely manner, survey researchers: 1) often conduct face-to-face or telephone interviews prior to the diary data collection period to go through the instructions with respondents (Silberstein & Scott, 1991); 2) make several contacts (e.g., by telephone and/or email) during the diary period to remind respondents about filling out their diaries as soon as the events occur and/or to confirm reported activities for that day or the past several days; and 3) establish a final contact with respondents to discuss returning the diary book or getting their feedback on the survey after the diary data collection (Silberstein & Scott, 1991). All these procedures are important steps to ensure data quality, but at the same time they may impose more burden on respondents, let alone the tedious nature of recording frequent events repetitively for themselves or for the whole household for a lengthy period of time.

**Respondent motivation** Not all respondents are motivated to complete the diary survey following the instructions in a timely fashion, especially those who are not interested in the survey topic and participate only for the incentives (Stone et al., 1991), or those who never wanted to participate but were convinced in the initial in-person or screener interview (Lugtig, 2014). These low-motivation respondents may be more likely to drop out in the diary survey. Respondent motivation can be influenced by many sources. As discussed in Krosnick (1991), respondent motivation is greater among those who are interested in the survey topic and those who understand the importance of the survey. Jäckle, Burton, Couper, and Lessof (2019) found that people who keep track of their expenditures and who use store loyalty cards were more likely to participate in a daily spending diary, suggesting that people who are more diligent in tracking their daily activities and used to repetitive tasks would have higher motivation to participate in this type of study. Those with greater motivation and higher commitment to the survey are likely to participate until the end of the diary survey period.

**Respondent fatigue** The tedious nature of diary survey data collection can pose a high burden on respondents. As the diary survey progresses, respondents are likely to become increasingly fatigued and disinterested (Schmid, Balac, & Axhausen, 2019). Their likelihood of dropping out may also increase across days. Prior studies suggest that nonresponse at the day level can increase over the diary period, suggesting increased respondent fatigue (Hu et al., 2017; Hu et al., 2019). The point where the survey becomes too burdensome to continue participation may be different for each respondent (Lemay, 2009; Lugtig, 2014).

**Habits** This cause for dropping out of diary surveys has mainly been discussed in the panel attrition literature (e.g., Lugtig, 2014). In diary surveys, respondents who fill out the diary daily may become habituated to the process. Those with the reporting habit established are more likely to fill out the diary every day. Similarly, those who missed reporting for the first several days may develop the habit of non-reporting and produce item nonresponse for all seven days. For those who developed a habit of reporting, once the habit is broken, respondents will be at higher risk of not reporting for the following days or drop out suddenly in diary surveys. This may explain possible dropout at later stages of the diary period.

**Shocks** Another cause of drop out that has been discussed in the panel attrition literature is "shock," referring to sudden dropout due to 1) life-changing events, like moving, illness and death, or 2) an unpleasant experience with the survey (Lugtig, 2014). Given the relatively shorter time span of a diary survey compared to panel surveys, shocks in diary surveys are relatively rare and can take different forms when compared to longer-term panel studies—e.g., a short business trip or a sudden visitor at the household. Shock due to an unpleasant experience, e.g., unpleasant interactions with telephone interviewers during the diary process, can be prevented by optimizing the survey protocol—e.g., improving the training of interviewers for all possible contacts with respondents.

To the best of our knowledge, no prior research has evaluated response patterns across respondents and across fielding days in multi-day diary surveys. This may be related to the challenge of obtaining day-level response status information (e.g., reported or refused to report) in diary surveys. Given the nature of diary surveys, no reported events on a particular day is often viewed as an actual survey response—zero events occurred on this day (e.g., Bee, Meyer, & Sullivan, 2012; Hanley & Lippman-Hand, 1983). However, without confirmation with respondents, it is sometimes unknown whether this type of event is an actual zero event response or respondents simply failed to report for that day. If the latter is true, treating it as a zero-event response will lead to biased survey estimates. In this study, we evaluate the response patterns in a 7-day diary survey—the U.S. National Household Food Acquisition and Purchase Survey (FoodAPS, 2016). FoodAPS is unique in having a rich set of paradata, including information about whether non-reporting on a given day is a nonresponse or a non-event. We take advantage of these unique paradata in this study to evaluate different response patterns across respondents and across time.

What else do we know from panel surveys? Due to the lack of prior research on response patterns across respondents and across fielding days in diary surveys, we broaden our literature review to also include relevant articles on response patterns in panel surveys, which share many similarities with multi-day diary surveys. Diary surveys can be viewed as a special type of longitudinal panel survey, with each fielding day corresponding to each wave in a panel survey. For each day in the diary period, as for each wave in panel surveys (Lugtig, 2014), respondents have a propensity to report for that day, ranging from 0 to 1, where 0 means respondents will definitely not report for this day and 1 means respondents will definitely report.

Several studies have evaluated response and dropout patterns in panel surveys. Most of this research focuses on studying dropout rates either across time or across respondents. Very few studies have evaluated response patterns across respondents and across time simultaneously. For example, Zhao (2002) evaluated the change in dropout rates in time sequence for each wave in a panel survey but did not evaluate response propensity differences between different groups of respondents. Some other studies have simply classified respondents into two groups—those who stay and those who ever drop out (Nicoletti & Peracchi, 2005; Watson & Wooden, 2009; Wiseman et al., 2005) , or three groups—stayers, attriters who do not come back, and returners (i.e., ever drop out) (Fitzgerald, Gottschalk, & Moffitt, 1998), and these studies have not evaluated possible changes in response propensities across time. It is likely that respondents have different reporting patterns in panel surveys—some respondents may drop out as early as the first few waves, while others may stay many more waves and then drop out. These prior approaches cannot discern respondents with different response patterns across time (Lugtig, 2014).

We are aware of only one study that allowed the response propensities to vary across time and across respondents in panel surveys. Lugtig (2014) used a latent class modeling approach to separate different groups of respondents based on their distinct response patterns, using the Longitudinal Internet Studies for the Social Sciences (LISS) panel data. Lugtig (2014) found three response patterns among those who ever drop out: early attriters, late attriters, and those with infrequent answers ("lurkers"). The advantage of this approach is that it allows researchers to evaluate different types of respondents defined by when they drop out in studies involving repeated measurement (Lugtig, 2014).

In this paper, we use a similar latent class modeling approach to identify and profile different response patterns across time and respondents in a multi-day diary survey. Despite their similarities with panel surveys, diary surveys differ in many ways. For example, panel respondents receive invitations for each wave of the panel survey, while diary-survey respondents generally receive only one invitation prior to the diary data collection period. The cognitive burden of the two types of surveys will also likely be different, resulting in different patterns of nonresponse. In addition, compared to an internet panel survey like the LISS panel, the process of a multi-day diary survey often involves more interviewer-based administration, including an initial interview prior to the diary data collection period and several contacts during the diary survey period. Given these differences between diary surveys and panel surveys, it is likely that the response patterns in a diary survey will differ from the patterns found in a panel survey. We focus on using the identified response patterns to suggest adaptive design approaches for future diary surveys.

## 2 Methods

### 2.1 FoodAPS Overview

The FoodAPS was a nationally representative survey designed to collect comprehensive data on household food purchases and acquisitions during one-week time periods. In total, there were 4,826 households with 14,317 individuals that agreed to participate in FoodAPS. FoodAPS employed a multistage sample design. Detailed descriptions of the FoodAPS design can be found in Hu et al. (2017), Ong, Hu, West, and Kirlin (2018), and at the FoodAPS website[1].

Figure 1 presents the overall data collection process of FoodAPS. Before the fielding period (marked as fielding day 1-7), the primary food shopper in each household was assigned as the primary respondent (PR), who needed to participate in two in-person interviews (one initial interview before the fielding period and one final interview after) and 3 telephone interviews. The PRs also reported food acquisitions for children under 11 years old in the household. Household members aged 11 years and above were provided with a diary book to record their food acquisitions. In FoodAPS, all food acquisitions collected were classified into two types: food-at-home (FAH), including all "food and drinks brought into the home" and food-away-from-home (FAFH), including "meals, snacks, and drinks got outside the home" (FoodAPS, 2016). On the evenings of fielding days 2, 5 and 7, PRs were supposed to call the telephone interview center to confirm each reported food acquisition. If a PR did not call in on specified scheduled day, the center would call the PR the following morning.

### 2.2 Variables

One important measure in FoodAPS is daily response status, which is obtained through the three planned telephone interviews where interviewers ask the PR about each household member's food acquisitions (including the PR) for each day since the prior contact. Respondents' daily reporting status is recorded and classified as confirmed, refused and unconfirmed. The confirmed status includes two situations: 1) the respondent reported one or more acquisitions in the diary book ("confirmed yes") for a given fielding day, or 2) the respondent reported no acquisition in the diary book, and the PR confirmed that there were no food acquisitions for that day ("confirmed no"). The refused status represents respondents who refused to provide information about possible acquisitions on a given day to the PR for reporting to the telephone center. The unconfirmed status means respondents reported no food acquisitions that fielding day and interviewers were unable to confirm with the PR that the respondent actually had no acquisitions. By using the reporting status information from the PR, we were able to distinguish between no food acquisitions and nonresponse on each of the seven days, and this is a unique feature of the present study relative to past research in this area (e.g., Bee et al., 2012; Hanley & Lippman-Hand, 1983). In this study, we combined the two confirmed statuses on a given day to indicate a valid response, while refusal and unconfirmed status are separate nonresponse categories.

To explore how the identified response patterns can inform adaptive design approaches for future diary surveys, we examined descriptive profiles of respondents assigned to each derived latent class. Specifically, we employed respondent-level information collected during the screener and initial interviews, including respondent characteristics (gender, age groups, race, education, marital status, relationship with PR, and employment status). Two household-level variables are also included in our analysis—household size and whether

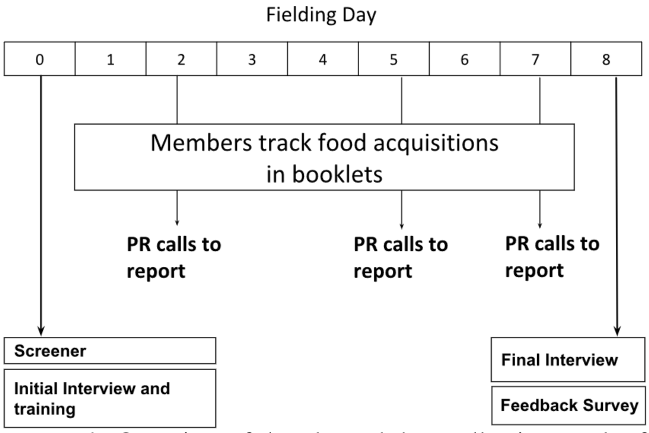


*Figure 1*. Overview of the planned data collection week of FoodAPS

[1] https://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey.aspx

the household is a Supplemantal Nutrition Assistance Program (SNAP) household.

We also employed several variables generated based on paradata. The first variable is the initial interview time adjusted for household size (total initial interview time for a household/household size), in minutes. The second variable describes whether a household ever refused to participate in the screening interview before finally agreeing to complete the diary survey. The third variable is the time gap (in days) between the initial interview and start of the diary period for each household. Although each household was instructed to start the diary survey right after the first face-to-face interview, this unfortunately did not always occur. This means that the time gap differed across households (Hu et al., 2017). The fourth variable (interview started in summer) indicates whether the interview started in June, July or August, which was used to capture possible seasonal differences in reporting patterns. The last series of variables are the number of interviewer contacts attempted or made before the diary period (e.g., for screening interviews and for the initial interviews) for each household, reflecting the amount of effort an interviewer needed to make before each household agreed to participate in the study. We explicitly did not include any variables measured in the final PR interview in these analyses, because these variables would not be available for informing adaptive survey designs in advance of the data collection.

### 2.3 Analytic Approach

To identify response patterns across respondents and across days of the diary period, we performed latent class analysis (LCA). Indicators used for the LCA were the response status for each of the seven days, measured by seven three-category variables—confirmed, refusal and unconfirmed. Because it was unclear a priori how many classes may best describe response patterns for a multi-day diary survey like FoodAPS, we fit models with different counts of classes specified. To determine the optimal number of classes, models with different counts of classes were compared against each other based on a number of model fit criteria (the Akaike Information Criterion, AIC; the Bayesian Information Criterion, BIC; and sample size–adjusted BIC, where smaller values indicate better fit in each case(Dziak, Coffman, Lanza, Li, & Jermiin, 2018); and the log-likelihood statistic). Apart from these criteria, we also considered the substantive interpretation of the models, focusing on the size and the interpretation of each class (Lugtig, 2014). We employed maximum likelihood estimation to fit the models and computed standard errors for all estimates. Because PRs were responsible for reporting food acquisitions by children under 11 years old, only respondents 11 years old and over (n = 11,552) were included in this analysis.

Table 1
*Distribution of the number of days on which individuals responded[a] for the seven-day diary period.*

| Number of days responded | % | $n$ |
|---|---|---|
| 0 | 5.9 | 683 |
| 1 | 2.4 | 279 |
| 2 | 2.3 | 267 |
| 3 | 2.0 | 227 |
| 4 | 2.1 | 245 |
| 5 | 4.9 | 566 |
| 6 | 9.0 | 1,040 |
| 7 | 71.4 | 8,245 |

[a] Responded here includes both confirmed yes and confirmed no status.

### 2.4 Examination of Class Profiles

To evaluate the profiles of respondents associated with each derived latent class based on variables available at the onset of data collection, we examined the frequency distributions of categorical variables and the means of the continuous variables across the classes. For categorical variables, we used the ML method, which produced proportions for each category (Vermunt & Magidson, 2015). For continuous variables, we used the BCH method based on the work of Bolck, Croon, and Hagenaars (2004). The ML and BCH methods take into account the uncertainty of the class membership in its estimation (Asparouhov & Muthén, 2014; Bakk, Tekle, & Vermunt, 2013; Bakk & Vermunt, 2016; Vermunt, 2010; Vermunt & Magidson, 2015). We applied a conservative Bonferroni correction when performing multiple comparisons of means and proportions across the derived classes. The Latent GOLD software (Version 5.1) was used for all model fitting in this paper.

## 3 Results

### 3.1 Descriptive Analysis

Table 1 shows the distribution of the number of days on which individuals responded during the seven-day diary period. The majority of respondents (71.4%) responded (i.e., with status being confirmed yes or confirmed no) on all seven days, and 5.9% of respondents failed to respond for all seven days.

Table 2 shows the distribution of the number of days on which individuals are unconfirmed during the seven-day diary period. The majority of respondents (76.0%) have zero unconfirmed days on all seven days, and 3.1% of respondents are unconfirmed for all seven days.

Table 2
*Distribution of the number of days on which individuals are unconfirmed for the seven-day diary period.*

| Number of days responded | % | $n$ |
|---|---|---|
| 0 | 76.0 | 8,781 |
| 1 | 9.4 | 1,089 |
| 2 | 5.0 | 573 |
| 3 | 2.0 | 225 |
| 4 | 1.5 | 177 |
| 5 | 1.7 | 198 |
| 6 | 1.4 | 156 |
| 7 | 3.1 | 352 |

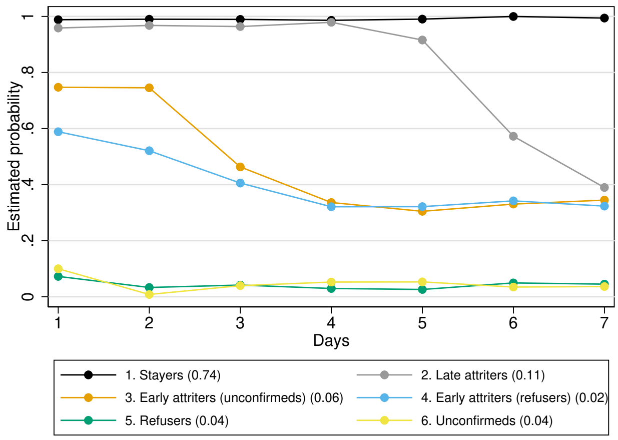


*Figure 2*. Estimated posterior probabilities of being confirmed on the seven daily response indicators for the six-class solution.

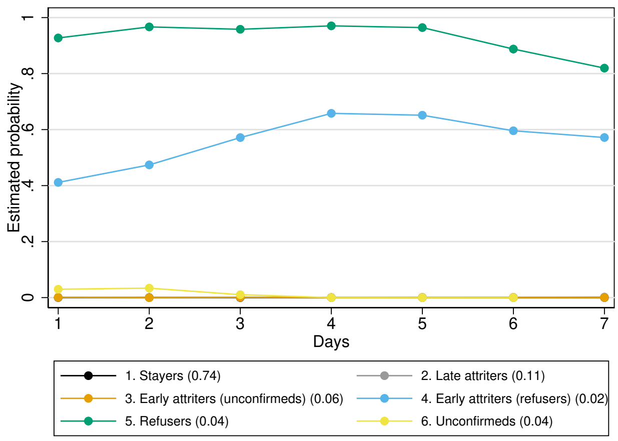


*Figure 3*. Estimated posterior probabilities of refusals on the seven daily response indicators for the six-class solution.

## 3.2 Latent Class Analysis

Table 3 presents the model fit criteria for the tested models with anywhere from 1 to 7 response pattern classes. Both the AIC and BIC decreased as more classes were added in the model, indicating that respondents do in fact have different response patterns. Although most of the criteria and tests favor the seven-class solution, we chose the model with six latent classes as our final model for the following two reasons. First, the seven-class solution produced a class that is only 0.5% of the sample and not fully distinguishable from existing classes in a six-class solution (see the supplementary materials for details related to the seven-class solution). Second, and more importantly, the six-class solution provides more interpretable results, making it the more practical choice.

Figures 2 to 4 show the estimated posterior probabilities of falling into the confirmed, refusal and unconfirmed categories for each of the six derived classes. The first class is the "stayer" class, where respondents' probabilities of being in the confirmed category are almost 1.0 for all 7 days. The predicted probability of membership in this class was by far the highest (0.74). Class 2 is the "late attriter" class, where the predicted probability of membership is 0.11. For Class 2, the respondents' probabilities of being in the confirmed category are close to 1 for the first 4 days and then start to decline. Respondents' probabilities of being in the unconfirmed category increase for later diary days. Classes 3 and 4 both represent "early attriter" classes, where Class 3 are early attriters who tend to be unconfirmed and Class 4 are early attriters who tend to refuse later in the diary period. Specifically, for Class 3, the probabilities of confirming for the first two days are slightly lower than Class 2 and then quickly start to decline, while the probability of being unconfirmed increases over time. The predicted probability of membership in Class 3 is about 0.06. For Class 4, respondents' probabilities of being in the confirmed category decrease over time, while their probabilities of refusal increase over the diary period. The predicted probability of membership in Class 4 is about 0.02. Class 5 is the "refusal" class, where the probabilities of confirming or being unconfirmed for all seven days are at or below 0.2, and the probabilities of refusing are above 0.8 for all seven days. The predicted probability of membership in Class 5 is about 0.04. Class 6 is the "unconfirmed" nonresponse class, where the probabilities to be in the unconfirmed category are above 0.8 for all seven days. The predicted probability of membership in this class is 0.04.

## 3.3 Profiles of respondents in each class

Table 4 presents the descriptive profiles of respondents in each class. As indicated by the Wald tests, for the majority of the respondents' characteristics, there are no significant differences across the six classes. Results for the household SNAP status suggest that the "early attriter (un-

Table 3
*Measures of model fit by number of response pattern classes*

| | 1 Class | 2 Class | 3 Class | 4 Class | 5 Class | 6 Class | 7 Class |
|---|---|---|---|---|---|---|---|
| AIC | 80223 | 52123 | 43113 | 41568 | 41222 | 40894 | 40627 |
| Δ AIC | | 28100 | 9010 | 1545 | 346 | 327 | 267 |
| BIC | 80326 | 52336 | 43436 | 42002 | 41766 | 41549 | 41392 |
| Δ BIC | | 27990 | 8900 | 1435 | 236 | 217 | 157 |
| SABIC[a] | 80281 | 52244 | 43296 | 41814 | 41534 | 41266 | 41062 |
| Δ SABIC | | 28038 | 8947 | 1482 | 280 | 268 | 204 |
| Log-likelihood | −40097 | −26032 | −21512 | −20725 | −20538 | −20358 | −20210 |

[a] Sample size-adjusted BIC

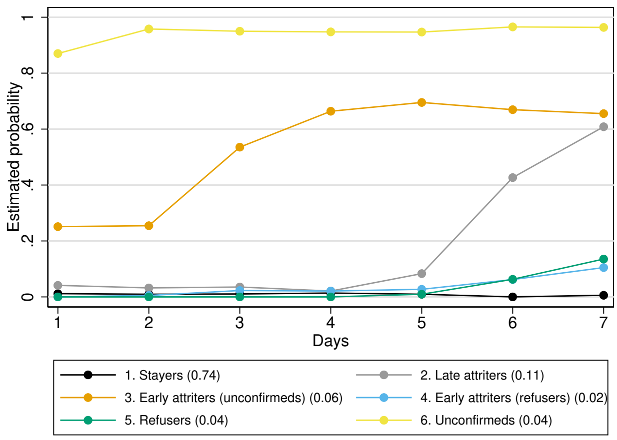


*Figure 4*. Estimated posterior probabilities of being unconfirmed on the seven daily response indicators for the six-class solution.

confirmeds)" class has a significantly higher proportion of non-SNAP households, compared to the "stayers" class (p-value = 0.001), suggesting that food security (important to SNAP households) may motivate those individuals to participate more frequently. When focusing on the indicator of whether a household ever refused to participate the screener interview before finally agreeing to complete the diary survey, the results suggest that "stayers" and the "late-attriter" class have lower proportions of households that ever refused the screener, compared to the "early attriter (refusers)" class (p-values = 0.006 and 0.004, respectively). These results suggest that the "early attriter (refusers)" class may need additional motivation or incentives in the early stages of data collection in an effort to encourage consistent participation.

## 4 Discussion

The current study presented evidence of different response patterns across respondents and across time in a seven-day national food acquisition diary survey. The majority of respondents (74%) had very high (close to 1) probabilities of responding to the survey for all 7 days. About 8% of respondents had very low probabilities (below 0.1) of responding for all seven fielding days. Clearly, their motivation to report was lower than other respondents, and they may have developed habits of non-reporting. About 8% of respondents tended to drop out as early as after the first two fielding days (i.e., "early attriters"). About 11% dropped out at a much later time (e.g., for the last two fielding days, or "late attriters"). These dropouts are likely caused by increased fatigue, and the breaking of reporting habits that are initially good. "Early attriters" tend to drop out in different ways—about 2% of respondents start to drop out after the first two days and then refuse on later fielding days, while about 6% of respondents start to drop out and then become unconfirmed on later fielding days. For "late attriters" who drop out and become unconfirmed mainly for the last two fielding days, it is possible that the PRs from those households simply did not or could not complete the final phone call. For this specific class, additional follow-up via different modes of data collection may prove useful in future diary surveys.

This study has important implications for the future designs of multi-day diary surveys like FoodAPS. Analyses of the response patterns indicate that respondents with similar nonresponse patterns may differ in terms of the types of nonresponse—refusal or unconfirmed. For the 8% of respondents who have very low probabilities (below 0.1) of responding for all fielding days, about half of them are refusers who tend to refuse all seven days, and the other half have an unconfirmed status for the seven days. This important difference may have critical implications for future adaptive designs. FoodAPS provided an engaging introductory video and requested that the PR show it to the household members for orientation purposes. For respondents who tend to refuse consistently across the diary period, reminders (e.g., telephone calls or text reminders) could be sent to the PRs of those households or to the household members directly to make sure that all members watch the video. Given cost constraints, this approach might only be applied for respondents predicted to be in the "refusal" class identified here.

For those who are unconfirmed for all seven days, it is

Table 4
*Respondent profiles in each latent class.*

| | Wald Statistic | Means/Proportions by Latent Class | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1. Stayers | 2. Late attriters | 3. Early attriters (unconfirmed) | 4. Early attriters (refusers) | 5. Refusers | 6. Un-confirmed |
| Probabilities of class membership | | 0.73 | 0.11 | 0.06 | 0.02 | 0.04 | 0.04 |
| Age | 17.79 | | | | | | |
| 11–15 | | 0.10 | 0.11 | 0.10 | 0.11 | 0.10 | 0.10 |
| 16–20 | | 0.10 | 0.11 | 0.10 | 0.16 | 0.09 | 0.08 |
| 21–30 | | 0.20 | 0.19 | 0.21 | 0.14 | 0.17 | 0.20 |
| 31–55 | | 0.38 | 0.38 | 0.39 | 0.38 | 0.44 | 0.39 |
| >55 | | 0.22 | 0.21 | 0.21 | 0.21 | 0.20 | 0.23 |
| Education | 14.40 | | | | | | |
| Below high school | | 0.31 | 0.33 | 0.32 | 0.35 | 0.32 | 0.29 |
| High school grads, below college | | 0.44 | 0.45 | 0.45 | 0.43 | 0.40 | 0.49 |
| College grads and above | | 0.25 | 0.22 | 0.23 | 0.22 | 0.29 | 0.22 |
| Marital status | 20.04 | | | | | | |
| Married | | 0.37 | 0.38 | 0.41 | 0.38 | 0.37 | 0.41 |
| Widowed | | 0.05 | 0.06 | 0.04 | 0.06 | 0.04 | 0.05 |
| Divorced | | 0.12 | 0.10 | 0.11 | 0.10 | 0.14 | 0.09 |
| Separated | | 0.03 | 0.02 | 0.04 | 0.02 | 0.04 | 0.03 |
| Never married | | 0.43 | 0.44 | 0.40 | 0.44 | 0.41 | 0.42 |
| Race | 8.26 | | | | | | |
| White | | 0.68 | 0.66 | 0.68 | 0.71 | 0.67 | 0.68 |
| Black/African American | | 0.15 | 0.14 | 0.16 | 0.15 | 0.16 | 0.15 |
| Other/Multiple races | | 0.17 | 0.21 | 0.17 | 0.14 | 0.18 | 0.16 |
| Relationship with PR | 13.25 | | | | | | |
| Primary respondent | | 0.42 | 0.41 | 0.42 | 0.40 | 0.40 | 0.41 |
| Partner or children | | 0.44 | 0.43 | 0.43 | 0.46 | 0.43 | 0.41 |
| Other relatives | | 0.10 | 0.12 | 0.12 | 0.11 | 0.12 | 0.14 |
| Non-relatives | | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.05 |
| Male | 4.15 | 0.53 | 0.54 | 0.56 | 0.47 | 0.53 | 0.54 |
| Employed | 3.67 | 0.48 | 0.45 | 0.51 | 0.48 | 0.46 | 0.47 |
| Interview started in summer | 5.83 | 0.43 | 0.41 | 0.41 | 0.47 | 0.43 | 0.47 |
| SNAP Household | 11.76* | 0.35 | 0.35 | 0.28 | 0.32 | 0.37 | 0.37 |
| Household size | 8.89 | 3.66 | 3.89 | 3.78 | 4.06 | 3.69 | 4.06 |
| Time gap | 4.82 | 1.10 | 1.04 | 1.11 | 1.05 | 1.03 | 1.11 |
| Initial interview time | 1.46 | 6.25 | 6.19 | 6.08 | 6.46 | 6.16 | 6.31 |
| Number of contacts for screener | 7.40 | 2.91 | 3.13 | 2.98 | 3.22 | 2.97 | 3.06 |
| Number of contacts for initial interview | 8.63 | 2.65 | 2.57 | 2.96 | 2.80 | 2.63 | 2.78 |
| Total number of contacts before diary | 10.47 | 5.56 | 5.69 | 5.94 | 6.03 | 5.60 | 5.84 |
| Ever refused screener | 15.25** | 0.018 | 0.013 | 0.012 | 0.050 | 0.027 | 0.027 |

* $p < 0.05$ ** $p < 0.01$

likely that the PR for those respondents did not or could not complete the three phone calls, or they could not obtain information from other household members for the seven days. For those respondents predicted to be in the "unconfirmed" class at the onset of data collection, interviewers could set up tailored telephone interview times with them by asking the PR about the best time of day and day of the week to call them before diary collection. Interviewers could also make follow-up calls to not only the PRs for these households but also household members who tend to be unconfirmed consistently. Given cost constraints, these tailored calling strategies may need to target only those who are more likely to be unconfirmed. Other design strategies, such as the use of text message surveys as a follow-up mode for those who are less likely to be reached by telephone interviewers, could also be used to follow up with those who are consistently unconfirmed. The use of mixed mode designs may encourage additional participation (De Leeuw, 2005; Dillman & Edwards, 2016).

The latent class profile analysis identified two pre-diary survey variables that differed significantly across the classes. The “early attriters (unconfirmeds)” class had a higher proportion of non-SNAP households. For non-SNAP households, motivation via incentives may not be as effective as for SNAP households. Furthermore, the topic of the survey may provide more motivation for SNAP households, due to their increased food insecurity. To help increase the motivation of PRs and members in non-SNAP households (or households where the topic of the survey may not be as salient more generally), future diary surveys like FoodAPS could use modified introductions to address their unique circumstances. For example, if respondents likely to fall in this class express concerns about disclosure risk during the screener, details about how FoodAPS protects confidentiality can be provided. Researchers could also ask respondents likely to be in this class about their preferred mode of communication at the initial interview to follow up about reporting status during the dairy period.

In addition, the “early attriters (refusers)” class has a higher proportion of households whose PR ever refused the screening interview. These results suggest that those who ever refuse the screening interview may also be less cooperative during the diary survey. To encourage the participation of respondents likely to be in this class, future FoodAPS-like diary surveys could consider emphasizing the importance of the study to these respondents (e.g., sending them extra introductory brochures about the study) and providing them with increased conditional incentives (e.g., incentives depending on the number of days reported, or prize draws where each diary survey response increases the likelihood of winning), in hopes of increasing extrinsic motivation for participation.

To better understand the reasons for nonresponse in these classes and to better guide practical suggestions on how to reduce item nonresponse at the day level, the developers of a diary survey could conduct focus groups and / or small-scale pretests with particular types of respondents (e.g., non-SNAP respondents) and discuss with them what may prevent them from participating every day in a diary survey like FoodAPS. The USDA is currently planning to conduct a second iteration of FoodAPS, namely FoodAPS-2. Given the results from this study, focus groups and pretests will be conducted in FoodAPS-2 to further explore reasons for lack of participation in this setting.

Given the limited measures available in FoodAPS prior to data collection, we were constrained in terms of the variables that we could use to profile each derived class. To gain a more comprehensive picture of class profiles and to better implement adaptive designs in future iterations of FoodAPS or related diary surveys, future studies could focus on collecting and evaluating more measures, including paradata, that could describe class profiles and guide adaptive designs.

First, to gain additional insights into whether the “early attriter” classes (both the refusal and unconfirmed types) emerge due to unpleasant experiences with interviewers at the screening or initial interviews, interactions between interviewers and respondents could be audio-recorded and analyzed. Dedicated design efforts like this for future diary surveys would enable researchers to better design tailored interventions for FoodAPS-2 and other similarly-designed diary surveys, given the presence of ancillary information that was collected by design prior to the diary survey data collection (e.g., focus groups, audio recordings, interviewer observations of interactions, training time for each household, etc.). The availability of this information for different response pattern classes like those identified in this study would greatly enhance the development of future adaptive designs, and FoodAPS-2 is currently taking steps in this regard.

Second, Ong et al. (2018) examined the effects of interviewers on a number of outcomes obtained from the initial and final interviews of the FoodAPS. The authors found that interviewers may introduce error in food acquisition survey data on several variables. No studies so far have examined interviewer effects on respondents’ reporting status trajectories over time in FoodAPS, which has practical implications for interviewer training. For example, if one interviewer tends to have respondents in the “early attriters” class as the survey proceeds, their interactions could be studied more closely, and corrections could be made during data collection. Future studies could examine these types of interviewer effects on latent class membership, where the interviewer ID is another piece of ancillary information that could be used to inform adaptive survey designs based on richer profiles of the response pattern classes derived here.

Third, some information collected in the final interviews in FoodAPS may be useful in profiling respondents in these classes, e.g., household income. For example, those with higher income may be less motivated by incentives and may require additional interventions. Although it is outside the scope of this study, future studies could compare the response pattern classes in terms of data collected later in the survey and determine whether some of these measures could be collected in the initial interviews to inform adaptive designs.

Fourth, Jäckle et al. (2019) suggest that people who keep track of their expenditures and who use store loyalty cards are more likely to participate in a daily diary survey. FoodAPS-2 and other similar diary surveys could consider asking questions regarding habits of tracking daily activities (e.g., expenditures) and the use of loyalty cards in the initial interview. This would allow researchers to better understand if these measures differ across response pattern classes in FoodAPS-2, and further inform future adaptive designs.

Despite the aforementioned measures that future studies could explore to inform adaptive designs in diary surveys, our study also identifies two important directions for future research. First, given the lack of true expenditure data or ideal benchmarks, we were unable to examine the extent of nonresponse bias for each class in this paper. The presence of validation data in future studies could enable evaluation of the level of estimated bias for each response-pattern class in a diary survey. If no validation data are available, future studies might follow the imputation method outlined by Hu et al. (2019) and evaluate the extent of nonresponse bias for each class. Second, mode effects may play an important role in item nonresponse at the day level in diary surveys. Different modes in diary surveys, e.g., paper-and-pencil, web or smartphone-based mobile surveys may differ in terms of respondents' reporting patterns, given that respondents' behaviors may likely vary across modes. With the increased popularity of mobile or web diaries, the response patterns for other modes of diary data collections could also be examined.

### Acknowledgements

This research was supported via a Cooperative Assistance Grant from the Economic Research Service of the U.S. Department of Agriculture [59-5000-5-0008] and in part by the intramural research program of the Economic Research Service. The findings and conclusions in this preliminary publication have not been formally disseminated by the U. S. Department of Agriculture and should not be construed to represent any agency determination or policy.

### References

Asparouhov, T., & Muthén, B. (2014). *Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model*. Mplus Web Notes.

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology*, *43*(1), 272–311.

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 20–31.

Bee, A., Meyer, B. D., & Sullivan, J. X. (2012). *The validity of consumption data: Are the consumer expenditure interview and diary surveys informative? (No. w18308)*. National Bureau of Economic Research.

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*(1), 3–27.

Carson, J. W., Keefe, F. J., Affleck, G., Rumble, M. E., Caldwell, D. S., & Beaupre, P. M. (2006). A comparison of conventional pain coping skills training and pain coping skills training with a maintenance training component: A daily diary analysis of short-and long-term treatment effects. *The Journal of Pain*, *7*(9), 615–625.

De Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of official statistics*, *21*(5), 233–255.

Dillman, D. A. (1991). The design and administration of mail surveys. *Annual review of sociology*, *17*(1), 225–249.

Dillman, D. A., & Edwards, M. L. (2016). Designing a mixed-mode survey. *Sage Handbook of Survey Methodology*, 255–268.

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2018). Sensitivity and specificity of information criteria. *bioRxiv*, *449751*.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. A. (1998). *An analysis of sample attrition in panel data*. The Michigan Panel Study of Income Dynamics.

FoodAPS. (2016). National Household Food Acquisition and Purchase Survey: User's guide to survey design, data collection, and overview of datasets. Retrieved from https://www.ers.usda.gov/media/8804/0_foodaps-user-guide-puf.pdf

Hanley, J. A., & Lippman-Hand, A. (1983). If nothing goes wrong, is everything all right?: Interpreting zero numerators. *Jama*, *249*(13), 1743–1745.

Hu, M., Gremel, G. W., Kirlin, J. A., & West, B. T. (2017). Nonresponse and underreporting errors increase over the data collection week based on paradata from the National Household Food Acquisition and Purchase Survey. *The Journal of nutrition*, *147*(5), 964–975.

Hu, M., Kirlin, J., West, B., He, W., Ong, A., Zhang, S., & Zhang, X. (2019). Estimation of underreporting in diary surveys: An application using the National Household Food Acquisition and Purchase Survey. doi:10.1093/jssam/smz024

Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Response rates and response biases. *Survey Research Methods*, *13*(1), 23–44.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Lemay, M. (2009). Understanding the mechanism of panel attrition. PhD thesis.

Lugtig, P. (2014). Panel attrition: Separating stayers, fast attriters, gradual attriters. *Sociological Methods and Research*, *43*(4), 699–723.

Nicoletti, C., & Peracchi, F. (2005). Survey response and survey characteristics: Microlevel evidence from the European Community Household Panel. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *168*(4), 763–781.

Ong, A. R., Hu, M., West, B. T., & Kirlin, J. A. (2018). Interviewer effects in food acquisition surveys. *Public health nutrition*, *46*(2), 425–492.

Schmid, B., Balac, M., & Axhausen, K. W. (2019). Post-car world: Data collection methods and response behavior in a multi-stage travel survey. *Transportation*, *46*(2), 425–492.

Silberstein, A., & Scott, S. (1991). Expenditure diary surveys and their associated errors. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. S. (Eds.), *Measurement errors in surveys* (pp. 303–326). New York: John Wiley.

Stone, A. A., Kessler, R. C., & Haythomthwatte, J. A. (1991). Measuring daily events and experiences: Decisions for the researcher. *Journal of Personality*, *59*(3), 575–607.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, *18*(4), 450–469.

Vermunt, J. K., & Magidson, J. (2015). *Lg-syntax user's guide: Manual for latent GOLD 5.0 syntax module November 20*. Belmont: Statistical Innovations Inc.

Watson, N., & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. *Methodology of longitudinal surveys*, *1*, 157–182.

Wiseman, V., Conteh, L., & Matovu, F. (2005). Using diaries to collect data in resource-poor settings: Questions on design and implementation. *Health Policy and Planning*, *20*(6), 394–404.

Zhao, H. (2002). Toward a comprehensive hazard-based duration framework to accomodate nonresponse in panel surveys. Retrieved from https://repositories.lib.utexas.edu/bitstream/handle/2152/1080/zhaoh029.pdf?sequence=2&isAllowed=y

# Appendix
# Figures

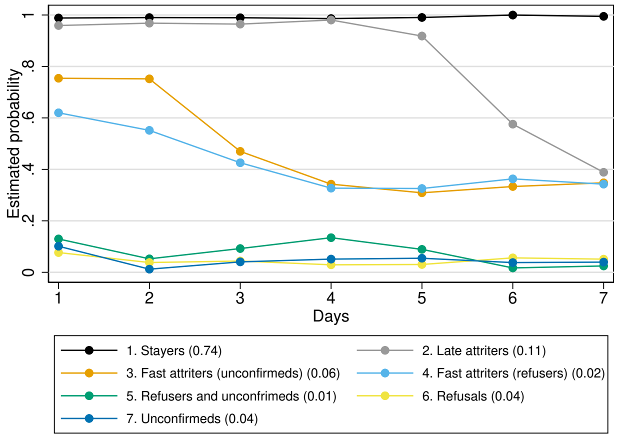


*Figure A1*. Estimated posterior probabilities for the seven-class solution predicting the confirmed category.

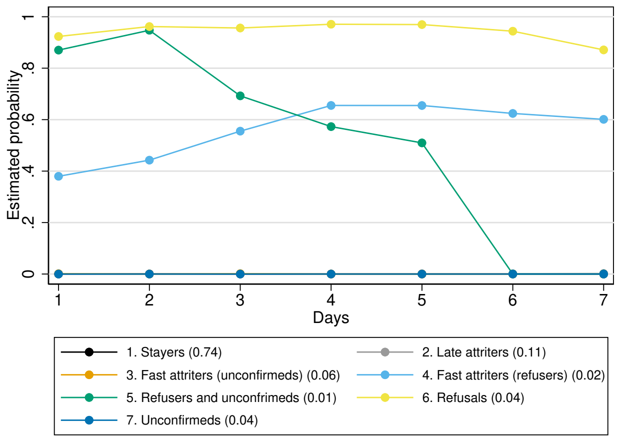


*Figure A2*. Estimated posterior probabilities for the seven-class solution predicting the refusal category.

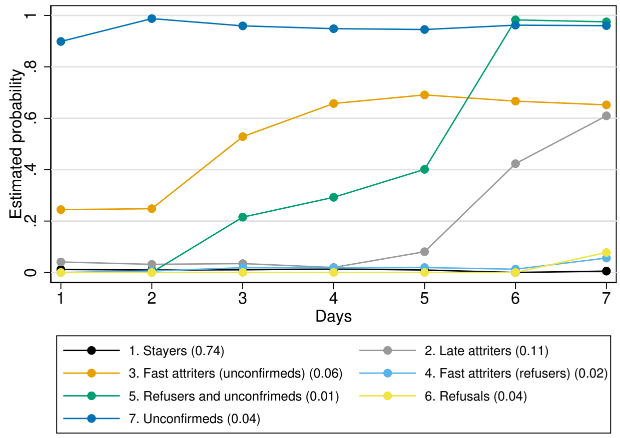


*Figure A3*. Estimated posterior probabilities for the seven-class solution predicting the unconfirmed category.