Westat

Improving lives through research™

Review of the Completeness and Accuracy of FoodAPS 2012 Data

Authors

Aaron Maitland
Lin Li

December 19, 2016

Prepared for:
Economic Research Service
U.S. Department of Agriculture
355 E Street, SW
Washington, DC 20024-3221

Prepared by:
Westat
An Employee-Owned Research Corporation®
1600 Research Boulevard
Rockville, Maryland 20850-3129
(301) 251-1500

Table of Contents

The preferred citation is:

Maitland, A., and Li, L. (2016). *Review of the Completeness and Accuracy of FoodAPS 2012 Data.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

This report is part of a series of five reports. The citations for the other reports are as follows:

Krenzke, T., and Kali, J. (2016). *Review of the FoodAPS 2012 Sample Design.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Li, J., Van de Kerckhove, W., and Krenzke, T. (2016). *Review of the FoodAPS 2012 Imputation Approaches for Income and Price Data.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Petraglia, E., Van de Kerckhove, W., and Krenzke, T. (2016). *Review of the Potential for Nonresponse Bias in FoodAPS 2012.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Yan, T., and Maitland, A. (2016). *Review of the FoodAPS 2012 Instrument Design, Response Burden, Use of Incentives, and Response Rates.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Executive Summary

The 2012 National Household Food Acquisition and Purchase Survey (FoodAPS) (hereafter referred to as "FoodAPS-1") is a household survey fielded primarily in 2012 and designed to capture detailed information on the food acquisitions of U.S. households. FoodAPS-1 was sponsored by the U.S. Department of Agriculture (USDA) and managed by its Economic Research Service (ERS). In 2015, ERS contracted with Westat to conduct an independent assessment of the quality of the FoodAPS-1 sample design, instrumentation, data collection procedures, and resulting data. This report is part of a series of five reports that constitute that assessment.

This report summarizes the findings from several analyses that Westat conducted to assess the completeness and accuracy of the data from FoodAPS-1. Specifically, we investigate the effect of several changes to the data procedures as well as assess the consistency of variables that were measured twice during the data collection period, the consistency of the administrative data as compared to self-reports, and issues of respondent fatigue.

During data collection, several changes were made to how screener data were collected. Chapter 1 examines the effect the changes to the household screener procedures had on screener completion, screener results, and data collected as part of the FoodAPS study. The most important changes to the screener procedures were (1) changing from offering a $5 incentive at the start of recruiting to offering after an initial refusal, or upon completion, and (2) shortening the introduction. We found an 8 percent decline in screener completion after these procedural changes were made. The decrease in response from offering an unconditional incentive versus a conditional incentive is consistent with much of the literature on incentives suggesting that prepaid unconditional incentives are the most effective at encouraging participation in surveys (Mercer et al., 2015; Singer et al., 1999). It is also possible that the change to the introduction may have had some effect on the differences between the screeners. Therefore, we recommend an unconditional incentive in the future. Despite the decline in screener completion, we did not find large effects of screener version on any of the resulting data from FoodAPS-1.

Chapter 2 present results on differences in reporting income between the screener and Final Interview. We found that income measured at the screener and Final Interview were in the same income group 58 percent of the time. Measuring income with screener questions is challenging, and other surveys such as the National Survey of American Families have faced similar issues with similar levels of agreement (see Cantor and Wang, 2001). Regression analyses revealed a few factors that may influence the consistency in income reports between the screener and Final Interview. For example, households who have experienced recent economic setbacks such as job changes and households with certain income sources such as unemployment compensation are more likely to have at least minor mismatches. Household size is one of the strongest predictors of both minor and major mismatches between interviews. This suggests that emphasizing that the respondent needs to include income from all persons in the household is important during both interviews.

Chapter 3 examines the accuracy of reporting Supplemental Nutrition Assistance Program (SNAP) participation by comparing survey reports with the records of SNAP participation. We conducted a latent class analysis using different indicators of the household's SNAP status from the screener interview, Final Interview, and administrative records. There were some limitations to the analysis since it was not possible to match all households to administrative records. Hence, we imputed values for these cases. The latent class model found low false negative rates for the Final Interview reports, but the screener reports had a false negative rate of approximately 10 percent. We also examined the effect of using different indicators of SNAP participation on some common analyses. Importantly, we did not find any significant changes to the analyses depending on whether administrative records were used to supplement interview reports to identify SNAP households. This suggests that attempting to link all respondents to administrative records may not be all that beneficial in future FoodAPS surveys. An alternative approach that may require fewer resources is to use a model-based strategy that uses survey data to predict some validated subset of cases that could be linked. Then use the model to correct for survey data that are not able to be linked.[1]

Chapter 4 examines the effect of declining participation within the household over the course of the study week on collected food expenditure data. We found primarily that the percentage of households that do not report any daily expenditure on food increases over the course of the week. The size of this increase in households not reporting daily food expenditures varies depending on household size. We did not see a significant decline in the amount of daily expenditures reported over the week. We also examined correlates of refusing to report and unconfirmed reports of no expenditures. Several socio-demographic factors, including race, ethnicity, and education, are related to the likelihood of providing these types of reports. In addition, those who have weaker connection to the household, such as nonrelatives of the primary respondents, are more likely to refuse or provide unconfirmed reports. Renters and members of SNAP households are also marginally more likely to provide these types of reports. The burden imposed by a data collection effort is similar to other surveys such as the Consumer Expenditure Survey (CE). A key challenge for FoodAPS will be

[1] See http://sites.usa.gov/fcsm/files/2016/03/H1_Davern_2015FCSM.pdf for an example.

finding ways to minimize burden and increase reporting over the study week, perhaps through the use of new technology such as web-based reporting, scanning, and other technology.

Finally, in Chapter 5 we conducted an analysis of outliers on a few key variables in the study. Univariate analyses of outliers were conducted for height, weight, person income, household income, and household expenditures. The analyses revealed a tiny fraction of outliers (~0.1%) for most of the examined variables except weight and height. For weight and height respectively, about 0.7 percent and 0.3 percent of sampled persons reported values that are significant outliers. And the percentages of outliers for height and weight are in general higher for infants under 2 than the other age groups. To improve the quality of height and weight data, we recommend incorporating edits (both soft and hard) for them into the computer-assisted personal interviewing (CAPI) system (e.g., by defining lower and upper bounds). In addition, a regression model on household income was run to explore the effect of outliers on multivariate analysis. The regression analysis showed a considerable impact of outliers on model conclusions. It also showed the large influence from a case with both extreme sample weight and an extreme data value.

The sampling weights of FoodAPS are quite variable as a result of oversampling some domains and weighting adjustments. Analysts should be aware of the potential influence that cases with large sampling weights can have on their analyses. The impact of outliers would be greater when outliers are associated with extreme sampling weights. Finally, we note that data should not be rejected just because it is unusually extreme. An investigation of why the extreme data occurred is always recommended before determining whether to keep, edit, or drop the outliers.

Analysis of Household Screener 1

The 2012 National Household Food Acquisition and Purchase Survey (hereafter referred to as "FoodAPS-1") gathered detailed information about household food acquisitions from April 2012 to mid-January 2013. The survey was sponsored by the U.S. Department of Agriculture (USDA) and developed and fielded by Mathematica Policy Research (Mathematica). The nationally representative sample consisted of nearly 5,000 households that completed the FoodAPS-1 Final Interview. FoodAPS collects comprehensive data on American households' food acquisition, factors influencing food choices, and household well-being. In 2015, the Economic Research Service (ERS) contracted with Westat to conduct an independent assessment of the quality of the FoodAPS-1 sample design, instrumentation, data collection procedures, and resulting data. This document is part of a series of five reports that constitute that assessment, specifically reporting on Westat's review of the completeness and accuracy of the FoodAPS-1.

FoodAPS-1 conducted screening interviews to determine a household's eligibility for the study. The survey had four target groups that are defined in terms of the household's participation in the Supplemental Nutrition Assistance Program (SNAP) and the poverty status of the household. The household screener collected information about household size, household income, and SNAP participation to identify these four target groups:

- Households participating in SNAP;
- Non-SNAP households with income below the poverty guideline that may be eligible for SNAP but do not participate;
- Non-SNAP households with income at or above 100 percent and less than 185 percent of the poverty guidelines that may be eligible for SNAP but do not participate; and
- Non-SNAP households with income equal to or greater than 185 percent of the poverty guideline.

Several questions were used in the screener to identify the above subgroups. The study defines household size as the number of people "who live together and share food." The screener respondent was asked to report income in two steps. First, the respondent identified all types of income received by the household. The interviewer showed the respondent a list of potential sources of income on a showcard as an aid to remember and identify all types of income. Next, the interviewer referred the respondent to a showcard that included ranges of household income corresponding to poverty thresholds of 100 and 185 percent and asked the respondent to identify "which group corresponds to your household total income before taxes."[2]

In addition to screening households, the screener also included questions to identify the primary meal planner and food shopper in the household who was asked to participate as the primary respondent for his or her household during the duration of the study.

In the first week of August 2012, or nearly mid-way through data collection, the wording of the household screener was modified slightly. For example, some new questions and observations were added to provide more information for the analysis of patterns of nonresponse to the study. The changes to the screener[3] are summarized below:

- A shortened introduction that revised the text and eliminated a question about receipt of the study's advance postcard;
- The timing of the offer of a $5 incentive was changed from being offered immediately to the respondent to being offered after an initial refusal (the $5 incentive was provided at the end of a version 2 screener even if the household did not answer any screening questions);
- Four questions were added for households that refused to complete the full screener;
- The questions were changed for households that were eligible for the study but refused to participate; and
- Space was added to the form for interviewer observations about the gender, age, race/ethnicity, and language of the person contacted.[4]

The goals of this analysis are to determine if these changes improved response to the screener, changed the composition of the households that completed the screener, or affected the resulting data collected as part of the main study. This analysis includes 13,445 cases for which screener data are available, of which 8,298 (61.7%, unweighted) responded to the revised screener.

Table 1-1 shows the outcome of the screener by version. Panel A of Table 2-1 shows that overall response defined as the number of households that completed the study out of the number of

[2] The showcard used depended on household size.

[3] Screener 1 can be found here: http://www.ers.usda.gov/media/8618/screeninginterview.pdf; Screener 2 can be found here: http://www.ers.usda.gov/media/8616/householdscreener2.pdf

[4] These questions were analyzed as part of Westat's nonresponse bias analysis.

households that were screened was significantly (p<.0001) higher for screener version 1 (39.0%) compared to screener version 2 (27.7%). Next, the table shows a breakdown of the response process in steps. Panel B shows that the completion rate is significantly higher (p<.0001) for screener 1 (93.7%) compared to screener 2 (85.3%). Panel C shows that the eligibility rate, conditional on completion of the screener, is significantly higher (p<.0001) for screener version 1 (67.2%) compared to screener version 2 (54.2%).[5] There were no statistically significant differences in the percentage that agreed to participate conditional on eligibility (Panel D) or the percentage that completed the study conditional upon agreeing to participate at the screener (Panel E).[6]

The completion rates and group composition of the released sample differed over the course of the field period. We, therefore, controlled for differences in sample composition using logistic regression models. The results are shown in Table 1-2. We ran models predicting screener completion, eligibility, and overall response to understand if the sample composition could explain any of the significant differences found in Table 1-2. These models include an indicator for screener version and controls, including sampling frame source (SNAP versus Non-SNAP), address type (single unit versus multi-unit), and sample release number.

Table 1-1. Comparison of screener outcome by version

(A) Overall study response		Screener #1	Screener #2
Completed entire study		39.0%	27.7%
Did not complete entire study		61.0	72.3
		100	100
	(n)	(5,139)	(8,285)
*Rao-Scott Chi-Square = 57.68, 1 DF, p<.0001			
(B) Screener completion			
Completed screener		93.7%	85.3%
Refused screener		6.3	1.7
		100	100
	(n)	(5,139)	(8,285)
*Rao-Scott Chi-Square = 17.81, 1 DF, p<.0001			

[5] There is little reason to expect that the changes made to the screener would affect the survey eligibility rate conditional on completing the screener. This suggests that the samples of households subject to the two versions of the screener were not similar.

[6] We used the variables CATEG1 and CATEG2 on the screener file to determine screener and study outcomes. Some inconsistencies were found between these variables and a revised version of RSTATUS. We adjusted CATEG1 and CATEG2 to be consistent with RSTATUS. The changes to CATEG1 are as follows: 80 cases originally coded as study ineligibles were recoded as eligible and agrees, 10 cases originally coded as study refusals were recoded as eligible and agrees, 29 cases originally coded as study refusals were recoded as eligible and refuses, 11 cases originally coded as refused screening were recoded as ineligible, and 21 without weights were coded as missing. The changes to CATEG2 are as follows: 39 screener refusals were recoded as study refusals, 11 screener refusals were recoded as study ineligible.

Table 1-1. Comparison of screener outcome by version (continued)

(C) Eligibility determination conditional on screener completion			
Eligible		67.2%	54.2%
Ineligible		32.8	45.8
		100	100
	(n)	(4,844)	(7,217)
*Rao-Scott Chi-Square = 57.36, 1 DF, p<.0001			
(D) Agree to participate in study conditional on eligibility			
Agree to participate		80.8%	79.8%
Refused to participate		19.9	20.2
		100	100
	(n)	(3,467)	(4,076)
*Rao-Scott Chi-Square = .03, 1 DF, p=.85			
(E) Complete study conditional on agreeing to participate			
Completed study		76.6%	75.0%
Did not complete study		23.4	25.0
		100	100
	(n)	(2,853)	(3,321)
*Rao-Scott Chi-Square = .80, 1 DF, p=.37			

Note: Weighted percentages were computed using weights that incorporate a base weight and nonresponse at different stages of the response process. Standard errors were computed using Taylor Series linearization with TSPSU defined as the clustering variable and TSSTRATA defined as the strata variable. There were 21 cases dropped from the analysis who did not complete the study and had missing weights.

Models 1a and 1b show the results of predicting screener completion. The coefficient for screener 1 is significant in Model 1a with no controls for sample composition. This coefficient is consistent with the results from Table 1-1 and indicates that screener completion is higher for screener version 1. Adding the controls in the model does not explain the differences in screener completion between screener versions. The coefficient for screener version 1 is still significant and actually larger in Model 1b compared to Model 1a.

Models 2a and 2b show the results of predicting eligibility conditional upon screener completion. The coefficient for screener 1 is significant in Model 2a with no controls for sample composition; however, the coefficient is no longer significant in Model 2b, which includes the controls. This change in significance indicates that differences in eligibility rates between screener versions can be largely attributed to differences in the sample that was released between versions.

Models 3a and 3b show the results of predicting overall response. The coefficient for screener type remains significant in both models. This finding indicates that differences in overall response across screener versions cannot be attributed to solely to differences in sample composition.

Table 1-2. Logistic regression models prediction screener completion, study eligibility, and overall response

Predictor	Screener completion (n= 13,424)				Eligibility (n=12,061)				Overall response (n=13,424)			
	Model 1a		Model 1b		Model 2a		Model 2b		Model 3a		Model 3b	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Intercept	2.23*	0.15	2.83*	0.16	0.44*	0.06	1.03*	0.07	-0.70*	0.05	-0.18*	0.05
Screener version 1	0.47*	0.10	0.75*	0.13	0.28*	0.03	0.20	0.11	0.26*	0.03	0.29*	0.05
Screener version 2 (ref.)												
SNAP Frame			0.43*	0.06			1.08*	0.08			0.44*	0.04
Non-SNAP frame (ref.)												
Single unit structure			-0.04	0.07			-0.32*	0.07			-0.17*	0.04
Multi-unit structure (ref.)												
Release number 0			-0.67*	0.22			3.03*	0.56			0.46*	0.10
Release number 1			-0.75*	0.17			-0.20	0.14			-0.08	0.10
Release number 2			-0.66*	0.17			-0.25	0.13			-0.10	0.10
Release number 3			-0.24	0.16			-0.58*	0.14			-0.21*	0.09
Release number 4			-0.24*	0.12			-0.60*	0.11			-0.22*	0.10
Release number 5			-0.01	0.13			-0.65*	0.11			-0.11	0.08
Release number 6			0.50*	0.16			-0.65*	0.11			-0.01	0.07
Release number 7			0.20	0.16			-0.87*	0.10			-0.08	0.12
Release number 8			0.20	0.18			-0.72*	0.11			-0.05	0.15
Release number 9			0.69*	0.19			-0.94*	0.15			0.06	0.10
Release number 10			0.19	0.22			-1.23*	0.15			-0.22	0.15
Release number 11			0.22	0.14			-0.85*	0.13			0.09	0.09
Release number 12			0.12	0.26			0.25	0.41			0.19	0.22
Release number 13 (ref.)												

*p<.05

Note: Models estimated using weights that incorporate base weights and nonresponse at each stage of the response process. Standard errors were computed using Taylor Series linearization with TSPSU defined as the clustering variable and TSSTRATA defined as the strata variable.

The findings from Table 1-1 and Table 1-2 suggest that the change in the timing of the offer of a $5 unconditional incentive (i.e., from being offered immediately to being offered after an initial refusal) had a negative effect on response to the screener and subsequently to overall response. This is consistent with much of the literature on incentives suggesting that prepaid unconditional incentives are the most effective at encouraging participation in surveys (Mercer et al., 2015; Singer et al., 1999).

It is also possible, however, that the change to the introduction or a correlate of seasonality (like schools being in session or the impending presidential election[7]) may have had some effect on the differences in response rates between the two screener versions. We do not have the data needed to test these alternative hypotheses.

We also examined the screener data and the interview data to see if there were any significant differences in reporting across versions of the screener. We first examined reports made at the screener. Table 1-3 shows that there were no differences across screener versions in reported SNAP participation and the income categories that respondents were grouped into during the screener. There was a statistically significant difference in household size, but examination of the distributions shows that the differences are relatively small. Screener version 2 had a slightly higher percentage of single-person households. We conducted a followup analysis treating household size as a continuous variable and found that the mean household size did not differ between screener versions ($p=.12$).

Next, we examined the effect of the screener on reporting in the interview and food expenditure data collection. The results are shown in Table 1-4. We conducted these models using the full sample weight and design information. There were no significant differences in household size or the number of children that households with children reported. There was also no significant difference in reporting of food-away-from-home (FAFH) or total food expenditures. There was a significant difference in the reporting of food-at-home (FAH) expenditures, with respondents who were given screener 1 reporting higher levels of FAH expenditures; however, this difference is still relatively small (e.g., $9, or a 19% increase) and barely reaches significance at the .05 level.

[7] The survey contractor has suggested that the door-to-door campaigning by candidates and volunteers, which certainly intensified after the screener was revised, made it more difficult for interviewers to successfully engage with households to conduct the screener interview.

Table 1-3. Comparison of reporting to screener items between screener versions

(A) SNAP Participation		Screener #1	Screener #2
Yes		13.9%	13.6%
No		86.1	86.4
		100	100
	(n)	(4,735)	(7,096)
*Rao-Scott Chi-Square = .09, 1 DF, p=.77			
(B) Income Category			
Income below 100% poverty		17.9%	19.2%
Income between 100-185% poverty		24.6	25.4
Income above 185% poverty		57.5	55.4
	(n)	100	100
		(4,611)	(7,002)
*Rao-Scott Chi-Square = 1.47, 2 DF, p=.48			
(C) Household Size			
1 person		19.4%	2.5%
2 person		34.8	33.5
3 person		17.3	17.0
4 person		28.6	27.0
		100	100
	(n)	(4,868)	(7,269)
*Rao-Scott Chi-Square = 9.15, 3 DF, p<.05			

Note: Weighted percentages were computed using base weights, and standard errors were computed using Taylor Series linearization with TSPSU defined as the clustering variable and TSSTRATA defined as the strata variable. Sample sizes vary slightly due to missing data.

Table 1-4. Comparison of reporting in the interview and food expenditures between screener versions

Variable	Screener #1	Screener #2	F Statistic	P-value
Household size	2.45 (.03)	2.39 (.04)	.53	.47
	(n = 2,238)	(n=2,555)		
Number of kids	1.88 (.03)	1.88 (.04)	.00	.95
	(n=1,026)	(n=1,125)		
Log (Weekly FAH expenditure + 1)	4.06 (.05)	3.89 (.07)	4.11	.05
	(n=2,170)	(n=2,496)		
Log (Weekly FAFH expenditure + 1)	3.17 (.05)	3.10 (.05)	.94	.34
	(n=2,160)	(n=2,467)		
Log (Weekly total food expenditure + 1)	4.70 (.04)	4.60 (.04)	3.05	.09
	(n=2,096)	(n=2,415)		

Note: Weighted percentages were computed using household weights, and standard errors were computed using replicate weights. Sample sizes vary slightly due to missing data.

Income Reporting in Screener and Final Interview 2

Household income was an important piece of information collected by FoodAPS for two primary reasons. First, as noted in Chapter 1, income was used to screen households to participate in the surveys. The screener has to balance the potential intrusiveness of asking about income with the need to accurately screen households. Therefore, a minimal amount of information is asked about income on the screener.

A second purpose for collecting income information is for analysis. The amount of income available to household members is critical for understanding how much food a household purchases. A significant amount of detail is necessary for creating an adequate measure of household income for analysis. Hence, the income questions in the Final Interview ask respondents to report income amounts within detailed categories for all members of the household age 16 or more. The primary respondent was asked to complete an Income Worksheet about income for all household members during the data collection week to help report the income of the household during the Final Interview.

There is often disagreement between less detailed income questions on screeners compared to more detailed income questions in an interview, and this can have implications for survey error (see Cantor and Wang, 2001). For example, it can lead to some households being recruited into one group (e.g., below poverty) but analyzed in another group (e.g., above poverty). This results in inefficiencies because it increases the variability in the survey weights when analyzing particular groups. The purpose of this analysis is to understand the factors that contribute to mismatches in income reporting between the FoodAPS screener and Final Interview.

We begin by showing the rate of disagreement between the screener and Final Interview. We used the income/poverty category that the household was classified into from the screener and the reported household income from the Final Interview for this analysis.[8] The household income variable from the Final Interview includes both reported and imputed income values.[9] Table 2-1

[8] We included possible outliers on household income in this analysis. The small number of outliers is unlikely to affect any analyses in this section (see Table A-3).

[9] There were 15 cases that were on the household level data file but not in the household income imputation file. They had a value of 0 on INCHH in the household data file. We used this value for these 15 cases.

shows the cross-tabulation of the two variables. There are 178 cases that have missing values for poverty status on the screener. The rate of agreement between the screener and Final Interview for nonmissing cases is approximately 58 percent.[10] We computed several measures of agreement among the nonmissing cases including kappa (.36), weighted kappa (.44), the Spearman correlation (.56), and the polychoric correlation (.69). In general, these measures suggest only a low to moderate level of agreement. The overall agreement is about 56 percent including all cases. Proportionally, Table 2-1 shows that it is more likely for respondents to move into a higher poverty category on the Final Interview compared to the screener than a lower poverty category. This is expected since respondents are asked to report in detailed income categories in the Final Interview, but not in the screener. Hence, respondents might better recall the amount of income within each category of the Final Interview rather than thinking only briefly about the categories in the screener.

Table 2-1. Comparison of household income reported in the screener with household income reported in the Final Interview

	Final Interview report		
Screener report	Income below 100% poverty	Income between 100-185% poverty	Income above 185% poverty
Missing	63	45	70
Income below 100% poverty	983	418	186
Income between 100-185% poverty	334	733	802
Income above 185% poverty	128	85	979

The next step in the analysis was to determine some of the causes of mismatching between the screener and Final Interview. We conducted a logistic regression analysis with a 0-1 dependent variable with 1 indicating that the household was categorized in a different poverty group in the screener as they were in the Final Interview (mismatch) and 0 indicating that the household was categorized in the same poverty groups (match). We counted the cases missing income on the screener as mismatches for the purpose of this analysis.[11]

Our predictors in the model included demographic factors such as reported income types, education level, age, job changes in the past 3 months, changes in household size in the past 3 months, and race of the respondent. We also included household size measured with a count of the number of persons age 16 and above since, all else being equal, we would expect to see more mismatches between interviews with larger households as the reporting task becomes more difficult when the

[10]Calculated by summing the cases in the diagonal in Table 2-1 and dividing by the total number of nonmissing cases.

[11]We ran the models excluding these cases from the analysis to see how sensitive the conclusion are to this choice, and it did not alter the findings.

respondent has to report for multiple household members. We also included the type of reported income in the household from the screener. The presence of certain types of income is likely to cause some discrepancies between interviews. For example, respondents might be uncertain about reporting how much other members of the household earn from work or other types of income. Other types of income such as investment income could be indicators of households that are more financially savvy and more likely to report income consistently. We examined factors such as education and age to determine if cognitive ability or memory could be an explanation for some of the discrepancy. Race is included as an indicator of cultural differences in terms of access to and knowledge of different sources of income. Finally, indicators of changes in jobs or the size of a household are included as measures of the stability of the household.

The model also included survey method factors such as version of the screener, whether the respondent completed at least part of the Income Worksheet, the final language of the interview, and whether there were different respondents to the screener and Final Interview. We initially included indicators for the version of the screener that was used in household screening to understand if the different composition of the households screened in by version had an effect on reporting. This variable was not significant in any of our models and was dropped from the final models. We also included an indicator for whether the household completed the Income Worksheet during the study week. The Income Worksheet was designed for the primary respondent to use in the Final Interview so that they did not have to recall all sources of income in the Final Interview from memory. Approximately 40 percent (unweighted) of household respondents completed the Income Worksheet in whole or in part. The language of the interview indicates whether the discrepancies could be due to how the interviewer was administered in different languages. Finally, whether the screener respondent was the primary food shopper indicates the effect of having different people who have access to different information report to the two interviews. Table 2-2 shows the results from our final model.

There were several demographic factors that predict mismatch in poverty categories between interviews. As expected, the probability of a mismatch increases with household size, demonstrating that having more potential members with income increases the likelihood of a mismatch. The presence of earnings from work, retirement and disability income, and investment income decrease the probability of a mismatch in poverty categories between interviews. Retirement and disability income are relatively stable sources of income and are perhaps relatively easy to report accurately. The presence of investment income likely suggests a household that is financially savvy and can report accurately about income. It is also likely that individuals with investment income would have income substantially higher than 185 percent of poverty, thus making a mismatch error unlikely in this context. The probability of a mismatch is lower for those with a college degree. Nonwhites are more likely to have a mismatch in poverty categories between interviews. Households that experienced a job change in the past 3 months are also more likely to have a mismatch in poverty categories between the screener and Final Interviews.

Table 2-2. Logistic regression model predicting likelihood of any mismatch in income categories between the screener and Final Interview

Parameter	Estimate	Standard error	t-value	p-value
Intercept	.03	.17	.18	.85
Household (HH) size	.18*	.04	4.08	<.001
Earnings from work	-.79*	.14	-5.60	<.0001
Unemployment compensation	.20	.22	.93	.36
Welfare, child support, alimony	.09	.21	.42	.68
Retirement and disability income	-.64*	.17	-3.76	<.001
Investment income	-.57*	.19	-2.94	<.01
Other income	.25	.20	1.23	.22
Completed Income Worksheet	-.41*	.10	-4.00	<.001
Less than high school (ref.)				
High school	-.06	.14	-.45	.64
Some college	-.28	.18	-1.58	.12
BA and above	-.58*	.20	-2.99	<.01
English interview (ref.)				
Spanish interview	.36	.19	1.81	.08
Korean interview	.33	9.30	.04	.97
Job change in HH last 3 months	.36*	.15	2.38	.02
Screener R is not primary food shopper (ref.)				
Screener R primary food shopper	-.14	.17	-.86	.39
Unknown if screener R is primary food shopper	.06	.19	.30	.77
Nonwhite	.33*	.12	2.76	<.01
N = 4,633				
*p<.05				

Note. Models estimated using household weights, and standard errors were computed using replicate weights. Sample sizes vary slightly due to missing data.

Some of the survey method factors also are predictive of the probability of matching poverty categories among non-SNAP households. Those who complete the Income Worksheet are less likely to report a mismatch in poverty categories between interviews. It is likely that the same factors that lead to accurate reporting also lead individuals to complete the Income Worksheet (e.g., financial awareness). There was a marginally significant higher likelihood of a mismatch in poverty categories for interviews conducted in Spanish compared to English. Finally, whether or not the screener respondent was the primary food shopper had no effect on the probability of matching poverty categories.

Next, we followed up the analyses in Table 2-2 with a multinomial logistic regression model predicting the likelihoods of major and minor mismatches between the screener and Final Interview relative to a match. A major mismatch is defined as reports that are two categories apart between the screener and Final Interview (i.e., from below poverty to above 185 percent poverty, or vice versa), whereas a minor mismatch is defined as reports that are only one category apart (i.e., into or out of the middle category of 100-185 percent poverty). We used the same predictors as were used in the logistic regression model from Table 2-2.

Table 2-3 shows the results from a multinomial logistic regression model predicting the likelihood of a major mismatch relative to a match and a minor mismatch relative to a match. The table shows that the likelihood of a major mismatch relative to a match is higher for larger households. The likelihood of a major mismatch is lower for households reporting income from earnings from work; welfare, child support, or alimony; retirement and disability income; and investment income.

Table 2-3. Multinomial logistic regression model predicting likelihood of major mismatches and minor mismatches in income categories between the screener and Final Interview

Parameter	Major mismatch		Minor mismatch	
	Estimate	Standard error	Estimate	Standard error
Intercept	-.66	.51	-.70*	.18
Household (HH) size	.33*	.08	.10*	.04
Earnings from work	-2.41*	.29	-.14	.13
Unemployment compensation	-.97	.58	.49*	.24
Welfare, child support, alimony	-.74*	.35	.25	.20
Retirement and disability income	-2.28*	.32	-.12	.14
Investment income	-1.20*	.54	.38	.20
Other income	.04	.66	.28	.23
Completed Income Worksheet	-.32	.17	-.37*	.10
Less than high school (ref.)				
High school	.07	.40	-.12	.14
Some college	.19	.37	-.42*	.17
BA and above	.42	.43	-.99*	.19
English interview (ref.)				
Spanish interview	.00	.41	.46*	.19
Korean interview	.24	9.20	.24	9.15
Job change in HH last 3 months	.22	.31	.45*	.15
Screener R is not primary food shopper (ref.)				
Screener R primary food shopper	-.25	.29	-.10	.20
Unknown if screener R is primary food shopper	.01	.30	.11	.24
Nonwhite	-.01	.29	.45*	.10
N = 4,633				
*p<.05				

Note. Model estimated using household weights, and standard errors were computed using replicate weights.

Table 2-3 shows that the likelihood of a minor mismatch relative to a match is higher for larger households, but with a smaller coefficient than for a major mismatch. Those receiving unemployment compensation also have a higher likelihood of a minor mismatch relative to a match. Those who have had a job change in the last 3 months have a higher likelihood of a minor mismatch. Nonwhites and those who complete the interview in Spanish also have a higher likelihood of a minor mismatch. Those with higher levels of education and who completed the Income Worksheet have a lower likelihood of a minor mismatch.

These analyses reveal a few different factors that may influence the consistency in income reports between the screener and Final Interview. Some of the sources are socio-demographic factors that do not suggest any obvious ways to improve consistencies. In addition, households that have experienced recent economic setbacks in their family members' lives (such as job changes) and households with certain income sources (such as unemployment compensation) are more likely to have at least minor mismatches. However, household size was one of the strongest predictors of both minor and major mismatches between interviews. This suggests that it is important during both interviews to emphasize that the screener respondent needs to include income from all persons in the household.

Accuracy of Reporting SNAP Participation 3

Accurate measurement of a household's participation in SNAP is important for screening households for FoodAPS and for analysis purposes. We conducted analyses to determine the accuracy of SNAP measurement and the effect of different operationalizations of SNAP participation on some analyses of interest to FoodAPS data users.

We conducted a latent class analysis (LCA) to examine the accuracy of different SNAP indicators on the FoodAPS data file. LCA is a technique that models the relationship between multiple indicators of a single concept and the underlying latent variable. The goal of LCA is to partition respondents into a set of latent or unobserved classes that make up a latent variable so that within this set of latent classes the observed indicators will be independent of one another (McCutcheon, 1987). In the case of SNAP measurement in FoodAPS, there are three observed indicators of interest: the screener report of SNAP participation, the interview report of SNAP participation, and the results from the match with SNAP administrative and ALERT data. Pseudo maximum likelihood estimation with the survey weights and design variables is used to find values of the latent class probabilities and conditional probabilities that minimize deviations between the model-predicted expectations and the observed data.

The screener and interview variables for self-reported SNAP participation include values of 0 if the household is not receiving SNAP benefits and 1 if the household is receiving SNAP benefits. Both of these variables also have some missing data. A variable summarizing the administrative match results could have four values: 3,252 cases have a value of 0 or no match; 1,316 cases have a value of 1 or match confirms SNAP participation; 136 cases have a value of 2 or match confirms non-participation; and 122 cases have a value of "v" for consent for data matching not given. A value of 0 (no match) could occur for three different reasons. First, the household could actually be receiving SNAP, but the probabilistic matching procedure did not find a match. Second, the household may actually not be receiving SNAP benefits. Third, the household may be from a state that did not provide match files or provided match files that did not provide the dates that households participated in SNAP.

We took a couple of steps in our analysis to minimize the ambiguity in the nonmatched cases. First, we excluded 952 cases from states that did not provide match files. Second, we excluded 209 cases from a state that provided a match file covering the survey field period but did not include dates of participation in SNAP. Third, we imputed SNAP administrative records for all of the remaining no match cases. For households that did not match the available administrative data, their SNAP administration record was set to missing and imputed to a value of 1 or 2. Any households that did not provide consent for administrative data match or did not match administrative data due to the unavailability of either state data or dates for confirming participation were excluded from the imputation process. This left 3,665 cases in the imputation process to have the value of their SNAP administration record imputed. We also used a simulated measure of SNAP eligibility to identify cases that did not have a match and were unlikely to be SNAP participants.[12] Households that were not considered eligible on any of the four simulated eligibility models were coded as non-participants, and the remaining 963 no-match cases were imputed. The imputation rate was about 26 percent (26% = 963/3,665).

The imputation of SNAP administration records was done through WESDECK, a Westat proprietary SAS macro. WESDECK uses the hot-deck approach, where the missing values are filled in with reported values randomly selected within imputation cells. Each reported value was allowed to be used up to 45 times. Imputation cells were formed by reported participation in SNAP in the past 12 months, sampling target group used for weight construction, income category (derived from both household income and household size), categorized imputed household income (10 categories were derived from the five imputed household income variables based on deciles), and reported SNAP participation in lifetime. For each missing value, WESDECK first searched for a reported value within the same cell of the missing value. If multiple reported values were found, one of them was randomly selected to impute the missing value. If a reported value was not available (either did not exist or had been used 45 times already) within a cell, the algorithm crossed the cell boundary and searched in the neighboring cell for a reported value. The priority of boundary crossing was such that reported SNAP participation in lifetime was the first to be crossed and reported participation in SNAP in the past 12 months was the last to be crossed.

The imputation procedure was conducted five times, each time using a different random seed and a different imputed household income variable (i.e., the categorized version of the first imputed household income was used for the first time, the categorized version of the second imputed

[12]The simulated measure of SNAP eligibility was created by the survey contractor Mathematica Policy Research (Mathematica).The SNAP eligibility estimates were conducted using the Microanalysis of Transfers to Households (MATH) SIPP+ Microsimulation Model and data collected during the initial and final interviews with the primary respondent. SNAP eligibility was estimated four times using different assumptions about income and composition of the SNAP unit.

household income was used for the second time, and so on). Five imputed variables indicating SNAP participation status for 3,665 households were created as a result.

Table 3-1 summarizes the results of the LCA for the 3,575 cases from states that provided complete records and also had valid data for both the screener and interviewer SNAP reports.[13] The LCA models were run in MPlus version 7.11 using the survey weights and design variables. The models were run five times by replacing the imputed values for SNAP administrative records with each run. Table 3-1 shows the average latent class and conditional probabilities across the five model runs. The latent class probabilities illustrate that the model assigned approximately 11 percent of cases as SNAP participants and approximately 89 percent as non-SNAP participants. The conditional probabilities in the body of the table illustrate the accuracy of the observed indicators according to the latent class model. For example, 89.7 percent of households that the model classifies as SNAP participants answered "Yes" to the screener item and 10.3 percent answered "No" to the screener item. That is, the model estimates a 10.3 percent false negative rate on the screener. The false negative rates are only around 1 percent for the interview reports and the administrative records. The final column of the table shows that the false positive rates vary between 0.1 percent for the SNAP report from the interview to 1.4 percent for the administrative records.[14]

Table 3-1. Results of latent class analysis of indicators of SNAP participation

Observed variables	SNAP participation	
	Yes	No
	Screener report	
Yes	.897	.008
No	.103	.992
	Interview report	
Yes	.987	.001
No	.013	.999
	Administrative records	
Yes	.982	.014
No	.018	.986
Latent class probabilities	**.114**	**.886**
N = 3575.		

13 An additional 90 cases were dropped from the latent class modeling due to missing data on the screener.

14 Mathematica used the monthly administrative record corresponding to the month of the data collection week to determine SNAP status. Thus, if the screener occurred in one month and the data collection week in the next month, the possibility of a change in true SNAP status is present. There were occasional lags between screener and data collection.

The latent class analysis has some noteworthy limitations as implemented. First, a latent class model with three binary indicators is a just-identified model. Therefore, we are unable to compute model fit statistics. Second, approximately one-fourth of the cases (cases with an initial value of 0) were imputed for the administrative records and the simulated eligibility variables were also used to help identify cases that were unlikely to be on SNAP. There is potential error associated with both of these modeling procedures that may not be represented in this analysis.

We also examined how using different indicators of SNAP participation may affect relationships between variables in the FoodAPS data file. We first looked at a regression model using different indicators of SNAP participation to predict food-at-home expenditures, as shown in Table 3-2. In general, the two regression models show very similar results. There is a slightly larger coefficient (in terms of absolute value) for SNAP participation when using self-reported SNAP participation compared to SNAPNOWHH, which combines input from both self-reports and administrative records.

Table 3-2. Regression model predicting FAH expenditures using different SNAP indicators

	Using self-reports and administrative records	Using self-reported SNAP participation only
Intercept	3.22*(.08)	3.23*(.08)
SNAP participation	-.29*(.011)	-.35*(.010)
Household income	.000037*(.000013)	.000036*(.000013)
Household size	.236*(.023)	.236*(.023)
Rural	.147(.095)	.147(.095)
	n= 4,697	n = 4,695

Note: Dependent variable is log (Weekly FAH expenditures + 1). Models estimated using household weights, and standard errors were computed using replicate weights. Sample sizes vary slightly due to missing data.

*p<.01.

We also examined the relationship between SNAP participation and food security using the two different indicators of SNAP participation. The dependent variable in Table 3-3 is the raw score from the 30-day adult food security measure. The results from the two models using the different SNAP indicators are nearly identical. On average, households with SNAP participants had scores on the 30-day adult food security measure approximately two points higher than non-SNAP participants for both indicators.

Table 3-3. Regression model predicting FAH expenditures using different indicators of SNAP participation

	Using self-reports and administrative records	Using self-reported SNAP participation only
Intercept	.733*(.034)	.747*(.038)
SNAP participation	2.02*(.13)	2.09*(.15)
	n= 4,824	n= 4,822

Note: Dependent variable is log (Weekly FAH expenditures + 1). Models estimated using household weights, and standard errors were computed using replicate weights. Sample sizes vary slightly due to missing data.

*p<.01.

Table 3-4 compares the distribution of food security for SNAP and non-SNAP households using different indicators of SNAP participation and found similar results. SNAP participants show lower levels of food security in both cases with remarkably similar distributions across the two sets of SNAP indicators.

Table 3-4. Distribution of food security for SNAP and non-SNAP households using different indicators of SNAP participation

Food security level	Using self-reports and administrative records		Using self-reported SNAP participation only	
	SNAP	Non-SNAP	SNAP	Non-SNAP
High	33.0%	74.9%	32.5%	74.5%
Marginal	21.7	13.70	21.5	13.9
Low	25.2	6.9	24.6	7.2
Very low	20.0	4.4	21.4	4.4
	100	100	100	100
	n = 4,824		n = 4,822	
	*Rao-Scott Chi-Square = 432.47, 3 DF, p<.0001		*Rao-Scott Chi-Square = 356.00, 3 DF, p<.0001	

Note: Weighted percentages were computed using household weights, and standard errors were computed using replicate weights. Sample sizes vary due to missing data.

Next, we examined how different SNAP indicators may influence conclusions about the relationship between diet and SNAP participation. Table 3-5 shows that conclusions about the relationship between SNAP participation and assessment of a person's own diet are unaffected by the choice of SNAP indicator.

Table 3-5. Distribution of assessment of own diet for SNAP and non-SNAP households using different indicators of SNAP participation

Assessment of diet	Using self-reports and administrative records		Using self-reported SNAP participation only	
	SNAP	Non-SNAP	SNAP	Non-SNAP
Excellent	6.3	9.9	5.5	9.9
Very good	14.9	30.2	15.1	29.9
Good	34.6	37.2	35.4	37.1
Fair	34.6	18.3	34.3	18.5
Poor	9.5	4.4	9.8	4.5
	100	100	100	100
	n = 4,82 *Rao-Scott Chi-Square = 104.69, 4 DF, p<.0001		n = 4,822 *Rao-Scott Chi-Square = 105.90, 3 DF, p<.0001	

Note: Weighted percentages were computed using household weights, and standard errors were computed using replicate weights. Sample sizes vary due to missing data.

In summary, these results do not show any significant changes to the analyses depending on whether administrative records were used to supplement interview reports to identify SNAP households. This suggests that attempting to link all respondents to administrative records may not be all that beneficial in future FoodAPS surveys. An alternative approach that may require fewer resources is to use a model-based strategy that uses survey data to predict some validated subset of cases that could be linked. Then use the model to correct for survey data that are not able to be linked.[15]

[15]See http://sites.usa.gov/fcsm/files/2016/03/H1_Davern_2015FCSM.pdf for an example.

Analyses of Collected Food Data 4

One of the main areas of concern with the collection of food data is the falling participation of household members throughout the week of the data collection period. This is likely due to the burden of reporting for an entire week. Paradata from the survey include indicators of whether household members reported food acquisition each day, refused to provide information, reported no food acquisitions with interviewer confirmation, or reported no food acquisitions without interviewer confirmation that no acquisitions were made. Table 1 in the FoodAPS User Guide demonstrates that the number of refusals and unconfirmed nonreporters in this dataset generally increases from day 1 to day 7. The user guide also mentions that this effect is weaker when controlling for day of the week. In addition, the user guide mentions that the number of members refusing to provide information remained relatively steady throughout the week with less than 4 percent of all members refusing; however, the number of members with no reported acquisitions but for whom the absence of acquisitions could not be confirmed by the primary respondent or by reviewing food books increases over the study week.

The purpose of this analysis is to examine whether the decline in household reporting over the course of the week affects the estimates of food expenditures. One would expect the number of days in the study to have no effect on the amount of food expenditures reported per day if the number of days in the study does not affect reporting. A drop in predicted expenditures across the week could indicate some bias due to the burden for households participating in the study.

The analysis is conducted in two steps. First, we ran logistic regression models to predict the probability that a respondent did not report any expenditure on a given day. Next, we estimated mean expenditures across days in the study after controlling for factors such as the day of the week, household size, and other factors related to food expenditures. To do this, we created a dataset that includes the total food expenditures for each day of the week for each household. We then began by estimating the regression model shown in Equation (1).[16] The regression model shown in Equation (1) predicts the amount of food expenditures that one would expect for a specific household (i) on a

[16]We used three-level multilevel models to account for the correlation between households within clusters. The model includes measures nested within households and households nested within clusters. Three-level logistic regression models predicting days of no expenditures were estimated using PROC GLIMMIX, and PROC MIXED was used to estimate models predicting expenditures. The household weight was included as a covariate in the models.

specific day (j) given the household's size, day of the week, and number of days that the household has been in the study. Day of the week is included in the model since the FoodAPS-1 sample is not distributed equally across days of the week and purchasing patterns may be related to day of the week.

(1) $FoodExpenditures_{ij} = \beta_0 + \beta_1 HHSize + \beta_2 Day1 + \beta_3 Day2 + \beta_4 Day3 + \beta_5 Day4 + \beta_6 Day5 + \beta_7 Day6 + \beta_8 Monday + \beta_9 Tuesday + \beta_{10} Wednesday + \beta_{11} Thursday + \beta_{12} Friday + \beta_{13} Saturday + \varepsilon_{ij}$

Our final model expanded on the model in Equation (1) by adding other factors related to food expenditures such as whether a day was a holiday, household income, urban/rural status, and the number of adult males in the household. The coefficients from the regression models are shown in Table 4-1. The coefficients in the left-hand panel are logistic regression coefficients modeling the log odds of not reporting any expenses on a given day. The coefficients in the right-hand panel are predicting the amount of expenditures and days when households report any expenditure.

Table 4-1. Regression analysis predicting not reporting any expenditures and log food expenditures

Parameter	Prediction of not reporting any expenditures			Prediction of log expenditures		
	FAH	FAFH	Total	FAH	FAFH	Total
Fixed Effects						
Intercept	.62*(.06)	.73*(.08)	-.38*(.07)	2.83*(.05)	1.76*(.05)	2.55*(.04)
Household size=2	-.31*(.04)	-.49*(.07)	-.46*(.05)	.22*(.04)	.25*(.04)	.28*(.03)
Household size=3	-.32*(.05)	-.75*(.07)	-.57*(.06)	.29*(.04)	.21*(.04)	.29*(.04)
Household size=4	-.55*(.05)	-1.04*(.07)	-.88*(.06)	.45*(.04)	.26*(.04)	.44*(.03)
Day 2	.29*(.04)	.21*(.05)	.30*(.05)	-.06(.04)	-.05(.03)	-.13*(.03)
Day 3	.35*(.04)	.40*(.05)	.46*(.05)	-.17*(.04)	-.02(.03)	-.17*(.03)
Day 4	.47*(.04)	49*(.05)	.60*(.05)	-.13*(.04)	-.01(.03)	-.17*(.03)
Day 5	.57*(.04)	.53*(.05)	.69*(.05)	-.11*(.04)	-.07(.03)	-.20*(.03)
Day 6	.66*(.04)	.64*(.05)	.79*(.05)	-.11*(.04)	-.04(.03)	-.19*(.03)
Day 7	.63*(.04)	.70*(.05)	.79*(.05)	-.03(.04)	-.03(.03)	-.16*(.03)
Tuesday	.11*(.04)	-.04(05)	.06(.04)	-.04(.04)	.05(.03)	-.03(.03)
Wednesday	.082(.044)	-.09(.05)	-.04(.04)	-.09*(.04)	.07*(.03)	-.06(.03)
Thursday	.16*(.04)	-.09(.05)	.05(.04)	-.04(.04)	.10*(.03)	-.02(.03)
Friday	.09*(.04)	-.27*(.05)	-.10*(.05)	.04(.04)	.32*(.03)	.15*(.03)
Saturday	.05(.04)	.08(.05)	.10*(.04)	.18*(.04)	.43*(.03)	.32*(.03)
Sunday	.15*(.04)	.33*(.05)	.29*(.04)	.06(.04)	.37*(.03)	.21*(.03)
Holiday	.137(.072)	.48*(.08)	.39*(.07)	.07(.07)	.11(.06)	.11*(.05)
HH Income*1000	.002(.003)	-.003(.005)	-.02*(.005)	.024(.003)	.016*(.003)	.018*(.003)
SNAP participation	-.05(.03)	.53*(.05)	33*(.04)	-.06*(.03)	-.12*(.03)	-.005(.024)
Rural	.03(.04)	.110*(.055)	.13*(.05)	.09*(.03)	-.05(.03)	.06*(.03)
# adult males	-.05*(.02)	-.07*(.03)	-.07*(.03)	-.02(.02)	.08*(.02)	.033*(.015)
PSU	.02*(.01)	.04*(.01)	.03*(.01)	.008*(.004)	.008*(.003)	.014*(.004)
HH within PSU	.26*(.02)	1.33*(.05)	.84*(.03)	.15*(.02)	.27*(.01)	.14*(.01)
Residual				1.25*(.02)	.84*(.01)	1.29*(.01)
N	33,768	33,768	33,768	11,237	13,738	19,806

*p<.05. Standard errors shown in parentheses.

As expected, household size is a significant predictor of expenditures, with larger households being less likely to report days without any expenditure and reporting a higher level of expenditures on days that they report any expenditure. The probability of reporting days without any expenditure increases across days of the data collection week for both food at home (FAH) and food away from home (FAFH); however, the level of food expenditures varies only across days of the data collection week for food at home. Turning to calendar day, households spend significantly higher amounts of money on food expenditures on Friday, Saturday, and Sunday compared to other days for which they have expenditures. Households are more likely to not report any FAFH expenditures on Sunday. The likelihood of not reporting any food expenditure decreases with income. On days when food purchases are made, the level of food expenditure increases with household income. SNAP participants are more likely to not report any expenditures. Those who participate in SNAP have lower levels of FAH and FAFH expenditures compared to those who do not; however, SNAP participants do not have lower levels of total food expenditures. It is unclear why SNAP participation is not related to total food expenditures, but the significant relationship between household income and total food expenditures demonstrates that income level has the expected effect on purchasing of food. Finally, the likelihood of not reporting any expenditures decreases with the number of adult men in the household. On days when expenditures are made, the level of expenditures increases with the number of adult men in the household.

Table 4-2 presents the mean predicted probability of not reporting any expenditure by household size and after controlling for all of the other factors in the model in Table 4-1. The table shows that the mean predicted probability of not reporting any expenditure increases by 40-60 percent by the end of the study week (depending on household size).

Table 4-2. Mean predicted probability of not reporting any expenditure by household size and day

HH size	Day in study						
Total							
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
1	.41	.48	.51	.54	.57	.59	.58
2	.31	.37	.41	.44	.46	.48	.48
3	.29	.35	.39	.41	.44	.46	.45
4+	.24	.30	.33	.35	.38	.40	.39
FAH							
1	.65	.71	.73	.75	.77	.78	.78
2	.58	.65	.66	.69	.71	.73	.72
3	.58	.64	.66	.68	.70	.72	.71
4+	.51	.58	.60	.62	.65	.67	.66
FAFH							
1	.63	.67	.71	.72	.73	.75	.75
2	.52	.57	.61	.63	.64	.66	.66
3	.47	.52	.56	.58	.59	.61	.61
4+	.43	.47	.52	.53	.55	.56	.57

Table 4-3 presents the mean predicted log daily expenditure on a given day for a given household size after controlling for all of the other factors in the model in Table 4-1. This table shows that the level of expenditure for the households that report any expenditure does not change much over the course of the week, dropping only by 0 to 6 percent over the course of the week.

Table 4-3. Predicted log mean food expenditures by household size and day in study

HH size	Day in study						
Total							
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
1	2.73	2.61	2.58	2.60	2.54	2.54	2.56
2	3.03	2.93	2.91	2.91	2.85	2.83	2.85
3	3.07	2.96	2.94	2.93	2.91	2.88	2.89
4+	3.21	3.10	3.07	3.08	3.04	3.01	3.04
FAH							
1	2.94	2.88	2.79	2.82	2.81	2.83	2.90
2	3.16	3.12	3.03	3.07	3.07	3.02	3.13
3	3.23	3.20	3.09	3.09	3.15	3.14	3.21
4+	3.37	3.31	3.21	3.26	3.26	3.24	3.33
FAFH							
1	2.06	2.03	2.06	2.09	1.98	2.04	2.05
2	2.34	2.33	2.38	2.37	2.29	2.31	2.29
3	2.33	2.28	2.37	2.36	2.31	2.28	2.28
4+	2.37	2.34	2.39	2.39	2.33	2.33	2.31

We supplemented the analysis with a regression model that predicted the probability of a respondent refusing or reporting no food acquisitions, and the interview did not confirm that no acquisitions were made. We call the latter unconfirmed reports. The dependent variable in the analysis is 0 if the person provided reports for all days and 1 if the respondent refused at least once or the respondent provided an unconfirmed report.

We examined a variety of demographic variables that could be related to response and food expenditures including sex, race, ethnicity, education, household size, relation to the primary respondent, marital status, age, rent/own status, SNAP participation, income, and urban/rural status. We subset this analysis to adult household members since some of the variables of interest are not measured on children (e.g., education and marital status). The most significant predictors are shown in our final model in Table 4-4.

Table 4-4. Prediction of individual level response being refusal or non-confirmed

Parameter	Estimate	Standard error	t-statistic	p-value
Intercept	-1.49	0.16	-9.26	<.0001
Non-White	0.41	0.09	4.33	<.0001
Hispanic	0.49	0.11	4.36	<.0001
High school	-0.37	0.11	-3.42	<.01
Some college	-0.38	0.10	-3.80	<.001
BA or higher	-0.52	0.17	-3.10	<.01
Household size	0.07	0.03	2.14	.04
Relative of PR	0.17	0.07	2.45	.02
Nonrelative of PR	0.82	0.28	2.98	<.01
Widowed	0.32	0.22	1.48	.14
Divorced	0.02	0.13	.16	.87
Separated	-0.06	0.19	-.35	.73
Never married	0.24	0.11	2.19	.03
Rent	0.19	0.10	1.86	.07
SNAP household	0.20	0.11	1.77	.08

N = 9812

Note: Models estimated using household weight, and standard errors were computed using replicate weights.

Race and ethnicity were significant predictors of refusals or unconfirmed reports. Non-Whites were more likely than Whites and Hispanics were more likely than non-Hispanics to refuse or provide unconfirmed reports. The probability of refusals and unconfirmed reports is less likely for those with higher levels of education. Respondents from larger households were also more likely to provide these types of reports. In addition, the probability of refusals and unconfirmed reports increases when nonrelatives live in the household. This suggests that those who are more loosely attached to the household are more difficult to obtain reports from than are related household members. Respondents who have never been married are also more likely to refuse or provide unconfirmed reports. Respondents who rent and those from SNAP households are marginally more likely to refuse or provide unconfirmed reports.

This section highlights a few findings related to underreporting in FoodAPS-1. First, we found that the percentage of households that do not report any expenditure on food increases over the course of the week (Table 4-2), regardless of which calendar day the reporting began. Among those who reported food on a given day, however, we did not see a significant decline in the amount of daily expenditures over the week (Table 4-3). We also found that those who have weaker connections to the household, such as nonrelatives of the primary respondents, are more likely to refuse or provide unconfirmed reports. Renters and members of SNAP households are also marginally more likely to provide these types of reports (Table 4-4). The burden imposed by a data collection effort is similar to other surveys such as the Consumer Expenditure Survey (CE). A key challenge for a future FoodAPS will be finding ways to minimize burden and increase reporting over the study week, perhaps through the use of new technology such as web-based reporting or scanning, for example.

Outlier Analysis 5

This section describes the analysis of outliers for height, weight, person income, household income, and household expenditure and the effect of outliers on analyses. Both univariate and multivariate analyses were conducted. The names of the variables and files used in the analyses are listed in Table 5-1 below.

Table 5-1. Variables and source files for the outlier analysis

Variable	Description
	faps_individual file:
height	Individual's reported height in inches
weight	Individual's reported weight in pound
incearnind	Individual's reported earnings last month w/o net versus gross adjustment
incunempind	Individual's reported unemployment insurance income last month
inctransferind	Individual's reported income last month from welfare, child support, and alimony payments
incretdisind	Individual's reported retirement and disability income last month
incinvestind	Individual's reported investment income last month
incotherind	Individual's reported income last month from other sources
	faps_household file:
inchh	Total monthly household income, excluding missing amounts
exprentmrtg	Household's monthly rent/mortgage expense
exphomeins	Household's monthly homeowner/rental insurance expense
expproptax	Household's monthly property taxes
exppubtrans	Household's monthly public transport expense
expelectric	Household's monthly electricity expense
expheatfuel	Household's monthly heating fuel expense
expwastedisp	Household's monthly sewer/garbage removal expense
exphealthins	Household's monthly health insurance expense
expcopay	Household's monthly health insurance copays
expdoctor	Household's monthly doctor/ hospital bills
exprx	Household's monthly prescription drug expense
expopmedical60	Monthly out-of-pocket medical expenses last month for individuals who are disabled or at least 60 years old
expmedical60com	Out-of-pocket medical expenses reported as part of another expense
expchildcare	Household's monthly child care expense
expchildsupport	Household's monthly child support expense
expadultcare	Household's monthly adult care expense

5.1 Univariate Analysis

In the univariate analysis, we first used a statistical test called generalized Extreme Studentized Deviate (ESD) test to identify outliers, then verified the outliers against box plots.

For height and weight, the analysis was done by gender for people age 16+ and by age groups (0, 1, 2, 3-4, 5, 6-7, 8-9, 10-11, 12-13, and 14-15) for those under 16. The age groups for those under 16 were chosen to be consistent with the public-use file (PUF) for FoodAPS-1, except that some categories are finer than the PUF categories. Those finer categories were used to better identify outliers for the ages when kids grow fast. Age group 0 has the smallest sample size among the age groups, which is 168 for height and 190 for weight. An in-depth data review was previously performed by ERS to identify biologically implausible values of height and weight, which led to the inclusion of two flags (BIVHGT_FLAG and BIVWGT_FLAG) in the individual data file. The outlier analysis complements the previous data review in that it examines all ages, including infants under age 2, who were not flagged for biologically implausible values. In addition, this analysis focuses on the distribution of observed data and, thus, may identify outliers that are biologically plausible.

Statistical Test

When testing for outliers, problems can occur when either too few or too many outliers are specified. For example, if we are testing for a single outlier when there are in fact two or more outliers, the additional outliers may influence the value of the test statistic enough so that no points are declared as outliers. This problem is referred to as "masking." On the other hand, if we are testing for two or more outliers when there is in fact only a single outlier, both points may be declared outliers. This problem is often referred to as "swamping."

Masking is one reason that trying to apply a single outlier test sequentially can fail (e.g., modified Thompson tau test). For example, if there are multiple outliers, masking may cause the outlier test for the first outlier to return a conclusion of no outliers and so the testing for any additional outliers is not performed. Due to the possibility of masking and swamping, many tests require that the exact number of outliers be specified. The generalized ESD test requires only an upper bound on the suspected number of outliers and is the recommended test when the exact number of outliers is not known. Therefore, we used this test in the analysis. However, the generalized ESD test does not eliminate the possibility of masking entirely; thus, it is useful to complement it with graphical methods such as box plots.

The generalized ESD test is based on the criterion of "distance from the mean" and assumes that the data (as collected or transformed) follow an approximately normal distribution. Given the upper bound, r, the generalized ESD test performs r separate tests: a test for one outlier, a test for two outliers, and so on up to r outliers.[17]

We used the generalized ESD test function[18] in Excel to perform the analysis.[19] For income and expenditure, we specified an upper bound of 10 outliers in the ESD test. If the 10th outlier was statistically significant (α=0.05), the upper bound was then increased to test for more outliers until no outliers were found to be significant in two consecutive tests. For height and weight, an upper bound of 25 was specified. And if the 25th outlier was statistically significant (α=0.05), the upper bound was then increased to test for more outliers until no outliers were found significant in two consecutive tests. The counts and percentages of identified outliers along with the sample sizes are summarized in Appendix A for height, weight, income and expenditure, respectively. As shown in Appendix A, overall the person height and weight have a small number of significant outliers (0.3% for height and 0.7% for weight). However, the percentages of outliers vary across age groups and infants under 2 years have a much higher percentage of outliers for height than the other age groups; e.g., a 0-year-old has 6.5 percent significant outliers for height. For income and expenditure, only a tiny fraction (~0.1%) of significant outliers were found.

Plots

The ESD test is based on the assumption that the data follow an approximately normal distribution. Therefore, we examined the distribution of data using the normal quantile-quantile (Q-Q) plot before applying the outlier test. Although formal tests of normality can be done, the presence of one or more outliers may cause the tests to reject normality when it is in fact a reasonable assumption for applying the outlier test. The normal Q-Q plot compares the distribution of a variable against

[17]Details about the generalized ESD test can found at: http://www.itl.nist.gov/div898/handbook/eda/section3/eda35h3.htm .

[18]The ESD function is not in standard Excel software. It can be added to Excel by installing the Real Statistics Resource Pack downloaded from http://www.real-statistics.com/free-download/.

[19]We also verified the Excel results by comparing them with the results from the R code available for download from http://www.itl.nist.gov/div898/handbook/eda/section3/eda35h3.r.

normal distribution by plotting their quantiles against each other. A point (x, y) on the plot corresponds to one of the quantiles of a variable (y-coordinate) plotted against the same quantile of the normal distribution (x-coordinate). If the data follow normal distribution, the points in the Q–Q plot will lie on the line y = x.

The review of Q-Q plots showed that the income and expenditure data deviate greatly from a normal distribution. The income and expenditure data are censored at 0 and also have “strongly” positive skewed distribution (extraordinarily large values). Therefore, we took the natural logarithms of the nonzero data to make them as close to normal distribution as possible, and the zero values are excluded from the analysis. For height and weight, for kids under age 4 the data are highly concentrated and far from normal distribution.

If the normality assumption for the data being tested is not valid, then a determination that there is an outlier may in fact be due to the non-normality of the data rather than the presence of an outlier. Box plots can be used to help determine outliers in such cases. In the box plot, the box displays the interquartile range (IQR) and also depicts the median value with a line; the whiskers are typically set as 1.5 times IQR. Our review of the box plots found consistent patterns with the ESD test results for most of the variables. We also examined side-by-side box plots of height and weight by age groups and by gender for people age 16+. The side-by-side box plots validated our assumption of the different height and weight distributions across those age and gender groups.

Since the analysis above is based on the FoodAPS-1 sample data, we also tried to determine whether the identified outliers are true outliers by looking at the population distribution. For height and weight, we obtained the lowest percentile (3%) and highest percentile (97%) available from the Centers for Disease Control and Prevention (CDC) Growth Chart.[20] Since the CDC Growth Chart is defined by age in months and gender, we derived the values for each age group mentioned above by taking the smallest value from the months in the age group as the lowest percentile and the largest value from the months in the age group as the highest percentile. For the age groups under 16, the percentile values were obtained by taking the lower or higher value across gender. The derived range (3% - 97%) for each age group is shown in Appendix B. There are quite a few values that are far outside the range and seem biologically implausible (e.g., a 0-year-old with height at 80.7 inches). The counts and percentages of outliers that are outside the 3% - 97% range can be found in Appendix A. One of the causes for the extreme values might be reporting error due to confusion between English and metric units. Respondents were offered an option to report height and weight

[20]The growth charts can be downloaded from http://www.cdc.gov/growthcharts/clinical_charts.htm#Set2.

in either English or metric units. A single measure for height and a single measure for weight were then derived by converting any response in metric units to English units. Since the metric system and English system are different by a ratio of over 2, it could be a reason for outliers if a person chose one option but actually reported numbers in another option.

Finally, we note that even when an appropriate test for outliers is used, data should not be rejected just because they are unusually extreme. An investigation of why the extreme data occurred is always recommended before determining whether to keep, edit, or drop the outliers.

5.2 Multivariate Analysis

In the multivariate analysis, we explored the effect of outliers on a regression model that involves household income and various other survey data. We first identified outliers using an influence statistic computed from the model, then compared the model results with and without the influential observations to determine if they have an impact on the conclusions from the models.

The regression model used in the analysis was specified by ERS. The variables involved are shown in Table 5-2. FOODEXPENDITURES is the dependent variable and the rest are the independent variables.

Table 5-2. Definition of variables used in the model

Variable name	Variable definition
FOODEXPENDITURES	Total food spending at all channel types ($100)
INCOME	Total monthly household income ($1,000)
INCOME2	Total monthly household income ($1,000) squared
USEDSNAP	Used SNAP benefits during FoodAPS survey week (0/1)
CHILD	Number of household members aged 0 to 10 years
YOUTH	Number of household members aged 11 to 13 years
TEEN	Number of household members aged 14 to 18 years
ADULT	Number of household members aged more than 18 years
AGE	Age of household's main meal planner (decades)
COLLEGE	Main meal planner completed college (0/1)
FEMALE	Main meal planner is female (0/1)
BLACK	Main meal planner is Black (0/1)
HISPANIC	Main meal planner is Hispanic (0/1)
ASIAN	Main meal planner is Asian (0/1)

Identifying Outliers

There are several statistics that can be used to identify outliers and their influence on regression models. Statistics such as residuals, leverage, Cook's D, and DFFITS assess the overall impact of an observation on the regression results; and statistics such as DFBETA assess the specific impact of an observation on the regression coefficients.

We identified influential data points using the DFFITS statistic since it combines information on both the residual and leverage. It is similar to Cook's D except that they scale differently, but they yield similar answers. DFFITS is defined as the change in the predicted value for a point, obtained when that point is left out of the regression, and "Studentized" by dividing by the estimated standard deviation of the fit at that point. DFFITS can be either positive or negative, with numbers close to zero corresponding to the points with small or zero influence. The conventional cut-off point for DFFITS is $2 * \sqrt{\mathrm{p/n}}$, where p is the number of predictors and n is the number of observations in the data set. An observation with an absolute value of DFFITS greater than the cutoff is considered influential (i.e., removing the observation may substantially change the estimate of coefficients).

At present the standard statistical software packages (e.g., SAS, SPSS, STATA) cannot account for complex sample design in variance estimation of the statistics mentioned above. Therefore, we used the R program by Jane Li (Li and Valliant, 2011), which was developed specifically to compute linear regression diagnostic statistics while accounting for complex sample design. We looked at the computed DFFITS by INCOME. There are a total of 131 households above the higher cutoff point (0.103803 $=2 * \sqrt{13/4826}$) or below the lower cutoff line (-0.103803). The households with the two highest incomes have large values of DFFITS; hence, highly influential. The household with the second highest income has the largest DFFITS as a result of having both high income and a large sampling weight.

Comparing Model Results

To assess the impact of outliers, we first fit the model with the original data, then refit it with the 133 influential observations removed. The results are shown in Table 5-3. Variance was estimated using sampling weights and Taylor Series method. After removing the influential observations, the model fit statistic R^2 increased from 0.27 to 0.31. Table 5-3 shows the results from the two model runs including the estimated regression coefficients, standard errors, t-values, and p-values. The relative change in the estimated coefficients ranges from -29 percent (HISPANIC) to 60 percent (INCOME2), with two-thirds of the predictors having estimated coefficients changed by more than 10 percent. Standard errors have reduced for all of the predictors except INCOME2 and AGE. In addition, the predictor ASIAN changed from being not significant (p-value: .09) to being highly significant (p-value: .0049). This demonstrated the substantial impact that influential observations can have on the model conclusions. The analysis also showed the potential influence that cases with large sampling weights can have, especially when extreme weights are associated with extreme data points.

Finally, we reiterate that even when an appropriate test for outliers is used, a data point should not be rejected just because it is unusually extreme. An investigation of why the extreme data occurred is always recommended before determining how to treat the outliers.

Table 5-3. Estimated regression coefficients for model predicting expenditures before and after removing influential observations

Predictors	Original[1]				After removing influential observations[1]				Relative change in coefficient estimate
	Estimate	Std. Error	t value	Pr(>|t|)	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	10.54	10.3226	1.021	0.319868	9.21	8.84659	1.041	0.310816	-13%
INCOME	10.36	2.09209	4.953	8.83e-05 ***	10.58	1.09604	9.655	9.24e-09 ***	2%
INCOME2	-0.15	0.03176	-4.621	0.000186 ***	-0.24	0.05297	-4.447	0.000277 ***	60%
USEDSNAP	33.00	7.36803	4.479	0.000257 ***	29.21	4.70214	6.212	5.75e-06 ***	-11%
AGE	1.98	1.21028	1.639	0.117596	1.86	1.21028	1.533	0.141683	-6%
COLLEGE	38.40	6.77394	5.669	1.83e-05 ***	31.14	5.89236	5.285	4.22e-05 ***	-19%
FEMALE	13.67	6.20857	2.202	0.040233 *	17.58	4.92562	3.57	0.002044 **	29%
BLACK	-50.99	11.90107	-4.284	0.000401 ***	-49.82	8.0698	-6.173	6.23e-06 ***	-2%
HISPANIC	-25.48	6.53167	-3.901	0.000961 ***	-18.16	4.88238	-3.718	0.001457 **	-29%
ASIAN	-29.35	16.4426	-1.785	0.090188	-25.19	7.92331	-3.18	0.004933 **	-14%
ADULT	41.73	5.61503	7.432	4.92e-07 ***	38.67	2.48273	15.578	2.83e-12 ***	-7%
TEEN	35.68	7.93471	4.496	0.000247 ***	34.16	5.67442	6.019	8.62e-06 ***	-4%
YOUTH	14.71	8.84929	1.662	0.112861	7.50	5.30952	1.413	0.173836	-49%
CHILD	17.67	4.50809	3.919	0.000921 ***	14.53	3.24403	4.478	0.000257 ***	-18%

Significant codes for α=: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

[1] The sample size for the original model and the model after removing influential observations is 4,826 and 4,695, respectively.

References

Cantor, D., and Wang, K. (2001). *Correlates of measurement error when screening on poverty status for a random-digit dial survey*. pp. 2E-1 to 2E-13, 1999 Collection of Papers, Report No. 7. Retrieved from http://www.urban.org/research/publication/1999-nsaf-collection-papers.

Li, J., and Valliant, R. (2011). Linear regression diagnostics for unclustered survey data. *Journal of Official Statistics*, *27*, 99-119.

McCutcheon, A. (1987). *Latent class analysis*. Thousand Oaks, CA: Sage.

Mercer, A., Caporaso, A., Cantor, D., and Townsend, R. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly, 79*, 105-129.

Singer, E., Gebler, N., Rhagunathan, T., Hoewyk, J. V., and McGonagle, K. (1999). The effects of incentives on interviewer-mediated surveys. *Journal of Official Statistics, 15*, 217-230.

Appendix A

Distribution of Outliers

Appendix A
Distribution of Outliers

Table A-1. Distribution of outliers for person height

Age (and sex) group	Sample size	Significant outliers*		Insignificant outliers*	
		Count	Percent	Count	Percent
0 year	168	11	6.5%	14	8.3%
1 year	225	8	3.6%	0	0.0%
2 years	236	4	1.7%	25	10.6%
3-4 years	489	1	0.2%	34	7.0%
5 years	254	2	0.8%	4	1.6%
6-7 years	428	2	0.5%	2	0.5%
8-9 years	445	2	0.4%	13	2.9%
10-11 years	451	2	0.4%	0	0.0%
12-13 years	422	1	0.2%	0	0.0%
14-15 years	480	0	0.0%	2	0.4%
16+ years male	4,715	9	0.2%	3	0.1%
16+ years female	5,544	5	0.1%	13	0.2%
Overall	13,857	47	0.3%	110	0.8%

* Outliers are values of person height that are outside the CDC Growth Chart range by more than 20 percent for the respective age (or age and sex) group. Significant outliers are statistically significant at α=0.05 level. Insignificant outliers are not significant at α=0.05 level.

Table A-2. Distribution of outliers for person weight

Age (and sex) group	Sample size	Significant outliers*		Insignificant outliers*	
		Count	Percent	Count	Percent
0 year	190	3	1.6%	1	0.5%
1 year	241	5	2.1%	3	1.2%
2 years	259	1	0.4%	6	2.3%
3-4 years	536	9	1.7%	1	0.2%
5 years	272	6	2.2%	2	0.7%
6-7 years	460	6	1.3%	1	0.2%
8-9 years	477	6	1.3%	4	0.8%
10-11 years	454	3	0.7%	9	2.0%
12-13 years	419	2	0.5%	9	2.1%
14-15 years	475	5	1.1%	5	1.1%
16+ years male	4,649	25	0.5%	263	5.7%
16+ years female	5,403	20	0.4%	353	6.5%
Overall	13,835	91	0.7%	657	4.7%

* Outliers are values of person weight that are outside the CDC Growth Chart range by more than 20 percent for the respective age (or age and sex) group. Significant outliers are statistically significant at α=0.05 level. Insignificant outliers are not significant at α=0.05 level.

Table A-3. Distribution of outliers for individual and household income

Income type	Sample size*	Significant outliers**	
		Count	Percent
Individual's reported earnings last month w/o net versus gross adjustment	9,968	12	0.1%
Individual's reported unemployment insurance income last month	10,227	5	0.0%
Individual's reported income last month from welfare, child support, and alimony payments	10,224	3	0.0%
Individual's reported retirement and disability income last month	10,156	7	0.1%
Individual's reported investment income last month	10,167	0	0.0%
Individual's reported income last month from other sources	10,151	0	0.0%
Total monthly household income, excluding imputed amounts	4,826	3	0.1%

* Sample size refers to the number of people with nonmissing value for individual income, and the number of households with nonmissing value for household income.

**Significant at α=0.05 level.

Table A-4. Distribution of outliers for household expenditure

Household expenditure	Sample size	Significant outliers*	
		Count	Percent
Household's monthly rent/mortgage expense	4,515	21	0.5%
Household's monthly rental/homeowner's insurance expense	4,563	0	0.0%
Household's monthly property taxes	4,630	0	0.0%
Household's monthly public transportation expense	4,802	2	0.0%
Household's monthly electricity expense	4,723	5	0.1%
Household's monthly heating fuel expense	4,739	0	0.0%
Household's monthly sewer/garbage removal expense	4,698	2	0.0%
Household's monthly health insurance expense	4,601	3	0.1%
Household's monthly health insurance copays	4,748	1	0.0%
Household's monthly doctor/hospital bills	4,769	0	0.0%
Household's monthly prescription drug expense	4,758	0	0.0%
Out-of-pocket medical expenses last month for those 60 and older or disabled	2,775	0	0.0%
Out-of-pocket medical expenses reported as part of another expense	2,259	0	0.0%
Household's monthly child care expense	4,808	1	0.0%
Household's monthly child support expense	4,807	1	0.0%
Household's monthly adult care expense	4,819	0	0.0%

* Significant at α=0.05 level.

Appendix B

Range of Height and Weight
(Based on CDC Growth Chart)

Appendix B
Range of Height and Weight (Based on CDC Growth Chart)

Table B-1. Range of height by age and sex groups, derived from CDC Growth Chart*

Age			CDC percentiles (centimeters)		Derived range (inches)	
Year	Month	Sex	3%	97%	Low**	High**
0	0-12.5	Boys	44.9251	82.03585	14.1	38.8
		Girls	45.09488	79.80419		
1	12.5-24.5	Boys	70.91088	94.31998	21.7	44.6
		Girls	68.77613	92.80876		
2	24.5-36.5	Boys	80.99959	102.9402	25.1	48.6
		Girls	79.55974	101.7931		
3-4	36.5-60.5	Boys	88.37864	117.8314	27.4	55.7
		Girls	86.90307	117.3552		
5	60.5-72.5	Boys	100.3318	125.1095	31.3	59.2
		Girls	99.35047	125.2536		
6-7	72.5-96.5	Boys	106.1048	139.2502	33.3	65.9
		Girls	105.7615	139.411		
8-9	96.5-120.5	Boys	117.5	151.5294	36.9	71.6
		Girls	117.2737	151.292		
10-11	120.5-144.5	Boys	126.6678	163.7185	39.7	78.0
		Girls	125.9599	165.1503		
12-13	144.5-168.5	Boys	135.6621	178.8165	42.7	84.5
		Girls	137.4381	172.8796		
14-15	168.5-192.5	Boys	148.5284	187.0941	46.7	88.4
		Girls	148.1173	174.7708		
16+	192.5+	Male	158.845	190.1943	50.0	89.9
		Female	150.4209	175.4671	47.4	82.9

* Source: CDC Growth Chart (http://www.cdc.gov/growthcharts/clinical_charts.htm)

**The low range is 20 percent less than the 3rd percentile from CDC Growth Chart for the age group; and the high range is 20 percent more than the 97th percentile from CDC Growth Chart for the age group.

Table B-2. Range of weight by age and sex groups, derived from CDC Growth Chart*

Age			CDC percentiles (kilograms)		Derived extreme values (pounds)	
Year	Months	Sex	3%	97%	Low**	High**
0	0-12.5	Boys	2.355451	12.91645	4.2	34.2
		Girls	2.414112	11.85539		
1	12.5-24.5	Boys	8.534275	15.68841	14.0	41.5
		Girls	7.93103	15.11839		
2	24.5-36.5	Boys	10.44144	18.0457	17.7	47.7
		Girls	10.04881	17.99807		
3-4	36.5-60.5	Boys	11.81842	24.46169	20.1	66.0
		Girls	11.38824	24.94401		
5	60.5-72.5	Boys	14.85692	28.27115	25.3	76.5
		Girls	14.34277	28.92162		
6-7	72.5-96.5	Boys	16.5037	37.41935	28.2	101.9
		Girls	16.01186	38.53537		
8-9	96.5-120.5	Boys	20.11467	49.41515	34.5	136.1
		Girls	19.54481	51.42913		
10-11	120.5-144.5	Boys	24.19264	63.30853	42.3	174.3
		Girls	23.99143	65.90103		
12-13	144.5-168.5	Boys	29.47257	76.9639	52.0	205.5
		Girls	30.0176	77.68817		
14-15	168.5-192.5	Boys	37.07331	88.95303	64.7	235.3
		Girls	36.70144	84.36639		
16+	192.5+	Male	45.79301	100.7784	80.8	266.6
		Female	41.82734	89.04485	73.8	235.6

*Source: CDC Growth Chart (http://www.cdc.gov/growthcharts/clinical_charts.htm)

**The low range is 20 percent less than the 3rd percentile from CDC Growth Chart for the age group; and the high range is 20 percent more than the 97th percentile from CDC Growth Chart for the age group.