

REVIEW OF THE NATIONAL HOUSEHOLD FOOD ACQUISITION AND PURCHASE SURVEY (FOODAPS) FROM A DATA USER'S PERSPECTIVE

February 2018

Parke Wilde
Mehreen Ismail

Submitted to:

U.S. Department of Agriculture
Economic Research Service
355 E Street, SW
Washington, D.C. 20024
Agency Principal Investigator: Michele Ver Ploeg
Agreement Number: 59-5000-6-0070

Submitted by:

Friedman School of Nutrition Science and Policy
Tufts University
150 Harrison Avenue
Boston, MA 02111
Non-Federal Entity Principal Investigator: Parke Wilde

TABLE OF CONTENTS

EXECUTIVE SUMMARY	i
1. INTRODUCTION	1
1.1. Goals	1
1.2. Introduction to approach and themes	1
2. METHODS	3
2.1. Data user surveys	3
2.2. FoodAPS tabulations	4
3. LITERATURE REVIEW	5
4. RESULTS	8
5. DISCUSSION AND CONCLUSION	18
5.1. Breadth	18
5.2. Depth	19
5.3. Power	20
5.4. Ease of use	20
REFERENCES	21
APPENDICES	
A. Survey instrument	24
B. Full list of recommendations for documentation	31
C. Full list of recommendations for data files	33
EXHIBITS	
1. Research literature using the National Household Food Acquisition and Purchase Survey (FoodAPS)	6
2. Comparison of FoodAPS and SNAP QC data on SNAP participant counts and selected household characteristics	9
3. Datasets comparable to the National Household Food Acquisition and Purchase Survey (FoodAPS) as reported by FoodAPS data users	10
4. How well data files met data users' needs	11
5. How well documentation met data users' needs	12
6. Comparison of FoodAPS and administrative data on WIC participant counts	13
7. Time required to get oriented to FoodAPS compared to other datasets ..	14
B-1 . Codebook-specific recommendations from data user survey	32
C-1 . Data file-specific recommendations from data user survey	34

EXECUTIVE SUMMARY

The U.S. Department of Agriculture's (USDA) National Household Food Acquisition and Purchase Survey (FoodAPS) offers an ambitious and unique view into how Americans acquire food. FoodAPS was administered to a nationally representative sample of 4,826 households between 2012 and 2013. Researchers gained restricted access to the data as early as 2014. As USDA and a contracting research firm develop the surveys next round (FoodAPS-2), multiple perspectives of the dataset's strengths and weaknesses will be informative. This report considers data user perspectives on potential tradeoffs between breadth, depth, power, and ease of use for FoodAPS-2.

We used two complementary strategies to review FoodAPS from a data user's perspective. First, we reviewed 25 publications that presented results from FoodAPS. Second, we surveyed 24 research teams that used the dataset. We organize findings from our literature review and analysis of survey results into the following topic areas:

- (a) motivations for using FoodAPS,
- (b) data strengths (including positive aspects of survey file construction and documentation),
- (c) data limitations (including barriers to use), and
- (d) directions for future rounds of data collection or future research using these data.

The literature review and data user survey offered several overlapping insights. Both information sources revealed that researchers largely were motivated to use FoodAPS for its high quality, detailed coverage of food acquisitions and purchases, the food retail environment, and Supplemental Nutrition Assistance Program (SNAP) participation. The literature review especially demonstrated how FoodAPS results have filled knowledge gaps about food access and nutritional quality of food choices. These motivations and uses of FoodAPS reflected the dataset's strengths. Both the literature review and data user survey also revealed limitations in data coverage and data quality. The survey in particular highlighted specific needs for improvement to documentation, data files, and data access. Limitations of FoodAPS offer suggestions for future rounds of data collection. While researchers expressed some very specific suggestions in their publications and survey responses, broader suggestions about survey design were most prominent.

Reflecting on data user perspectives may provide USDA with affirmation about the merits of FoodAPS and direction for continued research under FoodAPS-2. Data users recognize that FoodAPS has considerable breadth and depth in primary topic areas, including food acquisitions and purchases, the food retail environment, food security, and SNAP. Data users also acknowledge that FoodAPS allows for adequately powered analyses focused on low-income households, including SNAP participants. Deciding on how to enhance the dataset's capabilities in breadth, depth, and power must involve an assessment of tradeoffs related to cost, respondent burden, and ease of use. In assessing these tradeoffs, it may be important to incorporate data user input as USDA advances its research priorities through FoodAPS-2.

1. INTRODUCTION

1.1. Goals

The U.S. Department of Agriculture's (USDA) National Household Food Acquisition and Purchase Survey (FoodAPS) is an ambitious nationally representative household-level survey that combines information from respondents and administrative sources to offer an unusually complete picture of household food purchases and acquisitions. Data were collected for 14,317 individuals in 4,826 households between April 2012 and January 2013.

The goal of this report is **to review FoodAPS from the perspective of data users**, providing information about four topic areas:

- (a) motivations for using FoodAPS,
- (b) data strengths (including positive aspects of survey file construction and documentation),
- (c) data limitations (including barriers to use), and
- (d) directions for future rounds of data collection or future research using these data.

This report's input from the data user perspective may be combined with separate input from at least three other perspectives -- the survey production perspective, the USDA management perspective, and external stakeholder perspectives -- to inform development of a second round of the survey (FoodAPS-2) that will be fielded in 2019.

1.2. Introduction to approach and themes

We used two strategies to assess the data and research uses of FoodAPS. These complementary strategies allowed us to evaluate strengths and weaknesses of FoodAPS from a data user's perspective.

First, we reviewed 25 publications that provided results from the dataset, including articles, reports, and working papers. This literature review summarizes ways in which researchers have used the data and contributions to the broader literature made possible through FoodAPS (see Section 3).

Second, we surveyed researchers who used or are using FoodAPS. In the summer and fall of 2017, we recruited 38 individuals across 24 research teams to complete the data user survey. Of these individuals, 31 respondents returned written surveys and 7 respondents answered questions during a telephone survey. Our analysis of survey responses offers insight into the user experience and recommendations for future rounds of data collection (see Section 4).

For selected data issues that arose during the data user survey, we created empirical tabulations of FoodAPS data. These tabulations highlight: a) attractive aspects of the dataset and b) potential sampling issues. Complementing other assessments of FoodAPS data quality issues (Li, Van de Kerckhove, & Krenzke, 2016; Maitland & Li, 2016; Petraglia, Van de Kerckhove, & Krenzke, 2016), the tabulations in this report emphasize issues that were most prominent for a sample of researchers using the dataset.

One could anticipate that data users would recommend so many desirable topics, such a large sample size, and data quality reviews so thorough that the resulting survey operation would be unaffordable or infeasible. As it turned out, many of our survey respondents recognized both the importance of cost constraints and the need for tradeoffs across multiple survey design goals.

To enhance the usefulness of this report for real-world survey design decisions in FoodAPS-2, we sought to consolidate and organize the results of both the literature review and the data user survey into four themes:

- (Theme 1) **Breadth.** What topics related to food acquisition and purchase can be covered? Breadth is desirable, and yet can lengthen survey instruments, raising costs per response or threatening participation rates.
- (Theme 2) **Depth.** For which topics can the survey provide an especially high level of detail or precision? Depth can be provided either by multiple survey questions on the same topic or by careful linkages with administrative data on assistance programs or on the food retail environment.
- (Theme 3) **Power.** How large a sample size can be provided for particular subgroups or research topics? A larger sample size increases statistical power, but at a cost.
- (Theme 4) **Ease of use.** How accessible are the data and documentation from the data user perspective? Instruments that pursue breadth or depth using more complex survey skip patterns, or that attempt more complex administrative data linkages, may achieve goals for breadth, depth, or cost at the expense of ease of use.

For a given survey budget, tradeoffs are required among these four themes. This report provides information from existing research and our data user survey to help inform such tradeoffs.

2. METHODS

2.1. Data user surveys

Sample and recruitment. Our sampling frame included names of 50 principal investigators who had access to restricted-use FoodAPS data files.¹ We constructed this sampling frame by consulting with USDA Economic Research Service (ERS) staff and reviewing a list of FoodAPS research projects and publications posted to the USDA\ERS website. The resulting sampling frame excluded USDA\ERS staff but included researchers at varying stages of FoodAPS-related work sponsored through:

- University of Kentucky Center for Poverty Research grants,
- National Bureau of Economic Research grants,
- USDA\ERS cooperative agreements, and
- Independent external research projects.

To pretest our recruitment protocol and survey instrument, we selected four principal investigators representing each of the four categories listed above. We contacted these principal investigators via email and collected survey responses from three of the four research teams in March 2017. Based on pretest feedback, we modified our recruitment protocol to request that principal investigators refer us to data analysts who we could contact separately to complete sections of the survey pertaining to data usage.

We rolled out our recruitment protocol to the entire sampling frame of principal investigators in July 2017. As we received responses, we added to the sampling frame names of data analysts, whom we contacted via email between July and September 2017. We made up to four attempts to contact principal investigators by following up via email and telephone. Our final sample included 38 individuals across 24 research teams, reflecting a 47 percent response rate. We received approval to administer the data user survey from the Office of Management and Budget and Tufts University Institutional Review Board.

Survey instrument. The survey instrument included open and closed-ended questions on: 1) motivations for using FoodAPS, 2) data usage, and 3) recommendations for FoodAPS-2 (see Appendix 1). Based on pretest feedback, we revised the original survey instrument to include clearer signposts for questions that were either broader or more specific in focus. While we encouraged respondents to complete as much of the survey as possible, we targeted broader questions on motivation to principal investigators and more specific questions on data usage to data analysts.

Survey data collection. We distributed the survey via email and gave participants the option to respond via email or telephone. Thirty-one respondents elected to send responses via email, and

¹ One research team using the public use files contacted us directly about participating in the survey, so we included this research team in our sampling frame. The final sampling frame was comprised of 51 research teams.

eight respondents elected to complete a telephone or in-person survey. We audio recorded these survey responses and used recordings and interviewer notes to generate detailed summaries.

Survey data analysis. We used quantitative and qualitative approaches to analyze survey responses. For closed-ended responses, we determined response frequencies of numerical ratings, which respondents provided on data usage questions. For ratings on documentation and data files, we determined mean, within-respondent ratings for codebooks and data files of the same category (e.g., User's Guide, Household-Level, Geography Component). In a small number of cases, two research team members rated individual codebooks or data files differently; in these cases, we averaged the ratings for the codebook or data file and used this average to determine the mean rating for the category. We then converted numerical, within-respondent means to qualitative response options (e.g., poorly, neutral, well) for ease of interpretation.

For open-ended responses, we carried out a thematic analysis. As part of an iterative process, we reviewed open-ended responses to accomplish these tasks: 1) create and apply codes, 2) identify themes, and 3) identify negative or contrasting cases.

2.2. FoodAPS tabulations

As part of our thematic analysis of survey responses, we identified features of the FoodAPS dataset that warranted further exploration. Of these features, we selected two frequently mentioned characteristics of the dataset to explore, including 1) oversampling of Supplemental Nutrition Assistance Program (SNAP) participants and 2) representation of non-SNAP nutrition assistance program participants.

Oversampling of SNAP participants. We assessed oversampling of SNAP participants by tabulating participant counts and selected household characteristics using FoodAPS and SNAP Quality Control (QC) data. We linked SNAP QC data from Fiscal Years 2012 and 2013 and performed analysis on data collected between April 2012 and January 2013 to match the timeframe of FoodAPS data collection. Our tabulations accounted for the complex survey design of each dataset.

Representation of non-SNAP nutrition assistance program participants. We assessed representation of Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) participants in FoodAPS by tabulating participant counts and compared to WIC monthly participation summaries for April 2012 – January 2013, which are available through USDA Food and Nutrition Service. It would be useful in future research to do a similar comparison for food pantry clients represented both in FoodAPS and an authoritative source on charitable food program participation.

3. LITERATURE REVIEW

Researchers have used FoodAPS in a rapidly growing literature with important implications for food retail access, nutrition assistance programs, the economics of food choices, determinants of dietary quality, and many other topics. We reviewed 25 sources (articles, reports, and working papers) for insight into motivations for using FoodAPS, data strengths for particular research purposes, data limitations, and directions for future research using these data (Exhibit 1). Of these sources, 10 were published in refereed journals, 7 were USDA reports, and the remaining 8 were working papers or unpublished papers. We also reviewed existing summaries of the research experience from the first round of FoodAPS (Ziliak & Gundersen, 2016; Kirlin & Denbaly, 2017).

Major topics of empirical research in the 25 sources included food access (16 sources), SNAP (10 sources), nutritional quality (9 sources), food away from home (7 sources), food security (5 sources), and food prices (4 sources). In addition to these frequently studied topics, there was a long tail with many more occasionally studied topics such as housing (2 sources), WIC (1 source), and many others. One source may cover multiple topics (Exhibit 1). The research topics reflected the distinctive strengths of FoodAPS, including detailed information about the food retail environment and the high-quality combination of survey data and administrative data about SNAP participation status.

FoodAPS already is changing how researchers understand important policy-relevant issues related to food purchases and acquisitions. For example, related to the longstanding and active research thread on food deserts and food retail access, new FoodAPS research included some studies that confirmed in part the concern about lack of retailers within a close distance from home (for example, Gustafson et al., 2016). Other FoodAPS research offered empirical evidence that partly challenged previous conventional wisdom, including evidence that suggested that retailer proximity was less strongly associated with at least some important outcomes (Bowen et al., 2016; Downing & Laraia, 2016; Wilde et al., 2014; Wilde et al., 2017; Ver Ploeg & Wilde, 2017). For example, in this vein, Handbury et al. (2016) analyzed scanner data, finding that less than one third of disparities in nutritional quality of food acquisitions appeared to be related to variation in proximity to retailers (this study used FoodAPS to corroborate its nutrition quality measure).

The lively empirical literature using FoodAPS provided much insight into data quality for both FoodAPS and alternative data sources. Clay et al. (2016) compared descriptive statistics for FoodAPS to other survey sources. Kirlin and Denbaly (2017) offered a thoughtful reflection on lessons learned, as well as advantages of FoodAPS. Hu et al. (2017) considered possible symptoms of respondent fatigue and reduced response over the course of the reference week. The research firm Westat produced a series of reports that also consider FoodAPS data quality issues including non-response bias (Petraglia, Van de Kerckhove, & Krenzke, 2016) and data completeness and accuracy (Maitland & Li, 2016).

Along with the distinctive strengths that motivated selection of FoodAPS as a data source, some of the literature noted limitations that could be considered in a future round of data collection. Taylor and Villas-Boas (2016) observed that free foods had zero weight in aggregate statistics,

Source	Published	Topics
Allard & Ruggles, 2015		Food access.
Basu et al., 2016	Journal article.	SNAP; nutritional quality; housing.
Brewster et al., 2017	Journal article.	Nutritional quality.
Bowen et al., 2016		Food access; nutritional quality; housing.
Chang et al., 2017	Journal article.	SNAP; food security; knowledge.
Chenarides & Jaenicke, 2016		Food access; retailer choice; retailer supply.
Clay et al., 2016	USDA report.	SNAP; nutritional quality; food security; food away from home.
Dong et al., 2016	USDA report.	WIC; food prices.
Downing & Laraia, 2016		Food prices; food access; retailer choice; food security; nutritional quality (obesity).
Fan et al., 2015		Food prices; food access; nutritional quality (TFP).
Gustafson et al., 2016		SNAP; food access; nutritional quality (food categories).
Handbury et al., 2016		Food access; nutritional quality (disparities).
Hillier et al., 2017	Journal article.	SNAP; food access; retailer choice.
Hu et al., 2017	Journal article.	SNAP; food away from home.
Jo, 2017	USDA report.	Food access; nutritional quality (childhood obesity and HEI).
Kirlin & Denbaly, 2017	Journal article.	Data quality; review of FoodAPS literature.
Lyford et al., 2016		SNAP; food prices; food access.
Smith et al., 2016	Journal article.	SNAP; food security; food away from home.
Taylor & Villas-Boas, 2016	Journal article.	SNAP; food access; retailer choice; food away from home.
Tiehen et al., 2016	USDA report.	SNAP; retailer choice; food away from home.
Todd & Scharadin, 2017	USDA report.	Retailer choice; food away from home.
Ver Ploeg et al., 2015	USDA report.	Food access; retailer choice.
Ver Ploeg et al., 2017	USDA report.	Food access; retailer choice; food away from home.
Ver Ploeg & Wilde, 2017	Journal article.	Food access.
Wilde et al., 2014	Journal article.	Food access.
Wilde et al., 2017	Journal article.	Food access; food security.
Ziliak & Gundersen, 2016		Review of FoodAPS literature.

Exhibit 1. Research literature using the National Household Food Acquisition and Purchase Survey (FoodAPS)

Note: Thrifty Food Plan (TFP); Healthy Eating Index (HEI); Supplemental Nutrition Assistance Program (SNAP); Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).

and outlet-level price data was missing in some cases. Fan et al. (2015) said some census tracts did not have any stores for which there was price data for a TFP market basket. Some studies noted that FoodAPS permitted measuring distance from retailers to home, but not from retailers to other locations to which household members may travel, such as places of employment (Hillier et al., 2017).

For a promising line of research into nutritional quality of food choices, several studies discussed the critical early data collection decision in FoodAPS to focus on purchases and acquisitions rather than intake (Basu et al., 2016; Chang et al., 2017; Gustafson et al., 2016). Some studies measured an adaptation of the Healthy Eating Index (HEI) suitable for use with spending data instead of just intake data (Bowen et al., 2016; Clay et al., 2016; Jo, 2017). One study described validation of a new index developed specifically for assessing nutritional quality of grocery purchases (Brewster et al., 2017). Others measured spending on particular food groups with nutritional implications, such as fruits and vegetables (Gustafson et al., 2016). Still others addressed nutrition-related outcomes, such as childhood obesity (Jo, 2017).

Finally, researchers recognized the limitations inherent in analysis of cross-sectional data, including the challenges for determining cause and effect when measuring the impact of the food retail environment (Taylor & Villas-Boas, 2016; Wilde et al., 2017) and nutrition assistance program participation (Basu et al., 2016). In future research with a new round of FoodAPS, much will depend not just on the data collection itself but also on anticipating the data requirements for supporting continued advances in research design for measuring the effects of the food retail environment, nutrition program participation, and other explanatory variables of interest.

4. RESULTS

This section organizes results from our analysis of data user survey responses by four key topic areas. We present tabulations of FoodAPS data and other data sources where relevant. Although results from our thematic and descriptive analyses do not reflect experiences of all data users, results still offer insight into experiences of researchers who were at varying project stages and used FoodAPS data for varying purposes. Data user survey results corroborate findings from the literature review and also provide new information on data users' perspectives.

Motivations for Using FoodAPS. Out of 24 respondents,² 22 offered their motivations for using FoodAPS data. Motivating factors generally were related to the dataset's unique design and content. In terms of design, five respondents commented on the sample's national representativeness, while eight respondents remarked that oversampling of SNAP participants and low-income non-participants was an especially appealing feature.

We compared weighted counts of SNAP-participating households using FoodAPS and SNAP QC data (Exhibit 2). Our estimates of weighted household counts in FoodAPS were lower than our estimates from SNAP QC. FoodAPS captured information on a larger number of SNAP-participating households in higher income categories. FoodAPS also captured information on households that were larger in size, on average. While there were differences between the two datasets in terms of weighted household counts and mean household size, weighted counts of SNAP-participating individuals were similar (results not shown).

We also compared means for gross monthly income and SNAP benefits. Our estimates of mean gross monthly income from FoodAPS were higher than our estimates from SNAP QC data. FoodAPS data generally aligned with SNAP QC data on mean SNAP benefits, particularly for households overall and for households with gross monthly income \leq 100 percent of the Federal Poverty Level.

Although FoodAPS data do not fully replicate certain household-level characteristics of SNAP units, SNAP participants in the FoodAPS sample still are representative of the underlying participant population. Estimated household-level characteristics might have been more comparable across FoodAPS and SNAP QC data if FoodAPS had been able to identify 1) SNAP-participating individuals within households and 2) multiple SNAP units within the same surveyed household. None of our survey respondents made this observation though, and only one respondent noted that gross monthly income in FoodAPS was higher than what would be expected.

Oversampling of SNAP participants is a feature of FoodAPS that data users already value. Our comparison of FoodAPS and SNAP QC data points to at least two important survey design considerations for FoodAPS-2. First, it may be important to understand potential differences between survey households and corresponding SNAP units. Second, it may be important to clarify the reference period for questions about income.

² Our use of the term "respondent" refers to the entire research team. In some cases, "respondent" may refer to multiple individuals.

	Mean Monthly Number of Households (thousands)		Mean Household Size (persons)		Mean Gross Monthly Income (\$)		Mean Monthly SNAP Benefit (\$)	
	FoodAPS	QC	FoodAPS	QC	FoodAPS	QC	FoodAPS	QC
Gross Income to Poverty Category								
≤ 50%	2,979 (360)	9,518 (102)	2.73 (0.10)	2.08 (0.02)	336.65 (17.18)	211.07 (3.18)	332.53 (11.78)	359.24 (2.27)
51 – 100%	6,040 (587)	9,011 (95)	2.69 (0.09)	2.04 (0.02)	1,112.04 (17.98)	956.62 (4.24)	244.56 (14.04)	243.24 (1.84)
101 – 130%	2,590 (346)	2,718 (54)	2.85 (0.18)	2.23 (0.03)	1,732.05 (68.26)	1,497.52 (10.25)	219.11 (20.30)	153.69 (2.36)
130 - 185%	2,139 (193)	1,088 (35)	3.44 (0.12)	1.95 (0.04)	2,638.69 (65.47)	1,819.77 (18.79)	273.31 (16.05)	80.19 (2.86)
> 185%	2,952 (322)	107 (12)	3.16 (0.15)	1.39 (0.06)	5,559.57 (334.83)	2,279.92 (70.50)	248.82 (13.66)	26.13 (3.27)
Total	16,700 (1,143)	22,442 (116)	2.90 (0.08)	2.08 (0.01)	2,051.73 (106.00)	754.07 (4.58)	260.04 (9.15)	272.65 (1.39)

Exhibit 2. Comparison of FoodAPS and SNAP QC data on SNAP participant counts and selected household characteristics

Source: Authors' tabulations of FoodAPS and SNAP QC data

Note: National Household Food Acquisition and Purchase Survey (FoodAPS); Quality Control (QC); Supplemental Nutrition Assistance Program (SNAP).

All estimates are weighted, with standard errors in parentheses. For estimates from FoodAPS, we present linearized standard errors. For both datasets, the data collection period was April 2012 – January 2013. Mean monthly SNAP benefits estimated using FoodAPS were based on administratively matched and self-reported data. For 11 SNAP-participating households in the FoodAPS sample, SNAP benefit level information was missing.

In terms of content, respondents noted that FoodAPS facilitates pursuit of both previously asked and new research questions related to: 1) food purchasing and acquisition behaviors, 2) food retail environment features, and 3) federal nutrition assistance programs. While most research teams commented on only one of these content areas, at least five research teams valued how FoodAPS covers all of these topic areas in one dataset. Fourteen respondents provided examples of other existing datasets that could have provided comparable information, yet seven other respondents reported that no other dataset apart from FoodAPS would have been suitable for their research teams' purposes (see Exhibit 3).

- National Health and Nutrition Examination Survey (NHANES)
- Nielsen datasets
- IRI datasets
- Supplemental Nutrition Assistance Program (SNAP) administrative data
- Current Population Survey (CPS)
- Survey of Income and Program Participation (SIPP)
- American Community Survey (ACS)
- Behavioral Risk Factor Surveillance System (BRFSS)
- Consumer Expenditure Survey (CEX)
- Nationwide Food Consumption Survey
- Panel Study of Income Dynamics (PSID)
- State and local-level datasets with information on low-income households

Exhibit 3. Datasets comparable to the National Household Food Acquisition and Purchase Survey (FoodAPS) as reported by FoodAPS data users

Note: Ordered by highest to lowest frequency of being reported

On food purchases and acquisitions, 13 respondents viewed FoodAPS as being unique for including objective measures of all foods purchased or acquired from all sources. Disaggregated, item-level data on food at home (FAH) and food away from home (FAFH) purchases and acquisitions were critical inputs for certain research questions. Respondents reported using this type of information to study topics, ranging from nutritional quality to environmental impacts of food purchases and acquisitions.

On food retail environment features, eight respondents appreciated that FoodAPS had detailed information on features, including food prices and proximity to food retailers. Possible applications of this type of data seemed to be diverse. One research team expressed interest in using price data to construct indices of local food prices. Three other research teams discussed using measures of food retailer proximity to understand associations between the local food environment and food and nutrition outcomes, such as obesity and food insecurity.

On federal nutrition assistance programs, five respondents commented on the suitability of FoodAPS data to pursue research questions that focus either on program participants or on

program features. At least one respondent reported using administratively matched data on SNAP participation and benefit levels to compare food purchases and acquisitions of SNAP participants and non-participants. At least two respondents also have used information on WIC and school meals to study effects of child nutrition program participation. Administratively matched and geocoded data have provided at least two research teams with the opportunity to study misreporting of SNAP participation and participation in the context of state-level policies.

While nearly all respondents were motivated to use FoodAPS based on its unique content for particular research questions, other factors motivated at least three respondents. One respondent was invited to work on developing a component of the dataset. Two other respondents were motivated by funding opportunities, which provided grant support and a built-in network of other researchers using the same dataset.

Data Strengths. Respondents viewed FoodAPS as a high quality dataset. For all categories of data files, at least 73 percent of respondents reporting data file usage rated the data as meeting their needs at least somewhat well (Exhibit 4). Most respondents in our sample of 24 research teams noted that the dataset has strong coverage both across and within topic areas. Based on survey responses, the level of within-topic detail seemed to be more salient for respondents compared to the sheer number of topics covered in FoodAPS. The most frequently mentioned topic areas, with their corresponding levels of within-topic detail, included:

- Food purchases and acquisitions (mentioned by five respondents), with objectively assessed information on all FAH and FAFH from all sources (mentioned by 10 respondents);
- Food retail environment features (noted by four respondents), with geocoded information on retailer locations (noted by five respondents); and
- Federal nutrition assistance programs (mentioned by three respondents), with administratively matched information on SNAP participation (mentioned by five respondents). On this topic, two respondents explicitly mentioned availability of WIC and school meals participation data as a benefit of using FoodAPS.

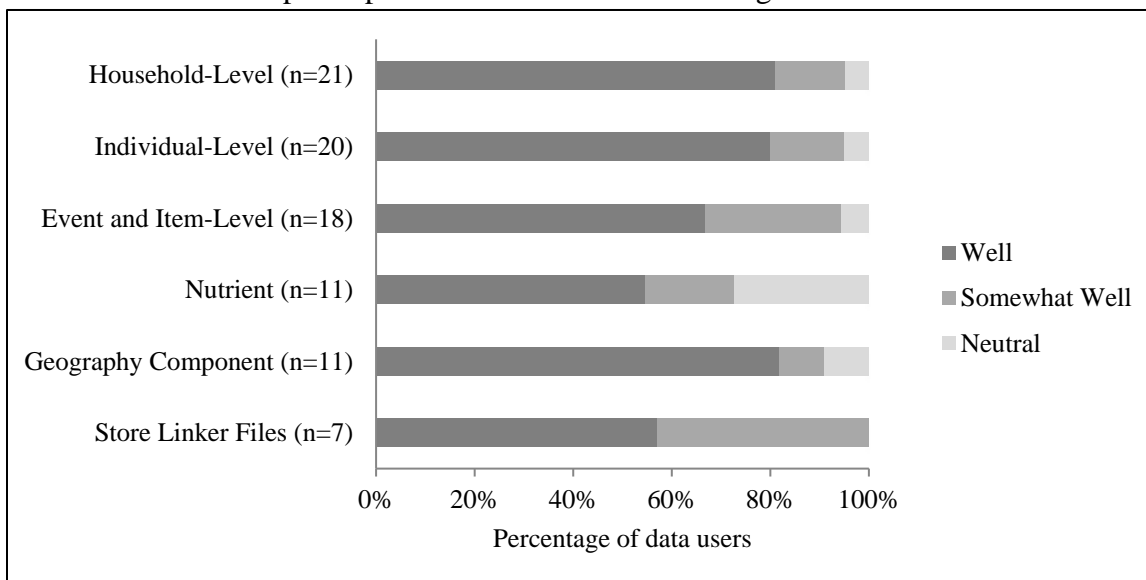


Exhibit 4. How well data files met data users' needs

Apart from generally positive views about the dataset’s content, respondents also held favorable opinions about overall support available for data users. For each category of documentation, over 80% of respondents reporting codebook usage rated the documentation as meeting their needs at least somewhat well (Exhibit 5). Thirteen respondents offered additional positive remarks, stating that documentation generally was clear, thorough, and well organized.

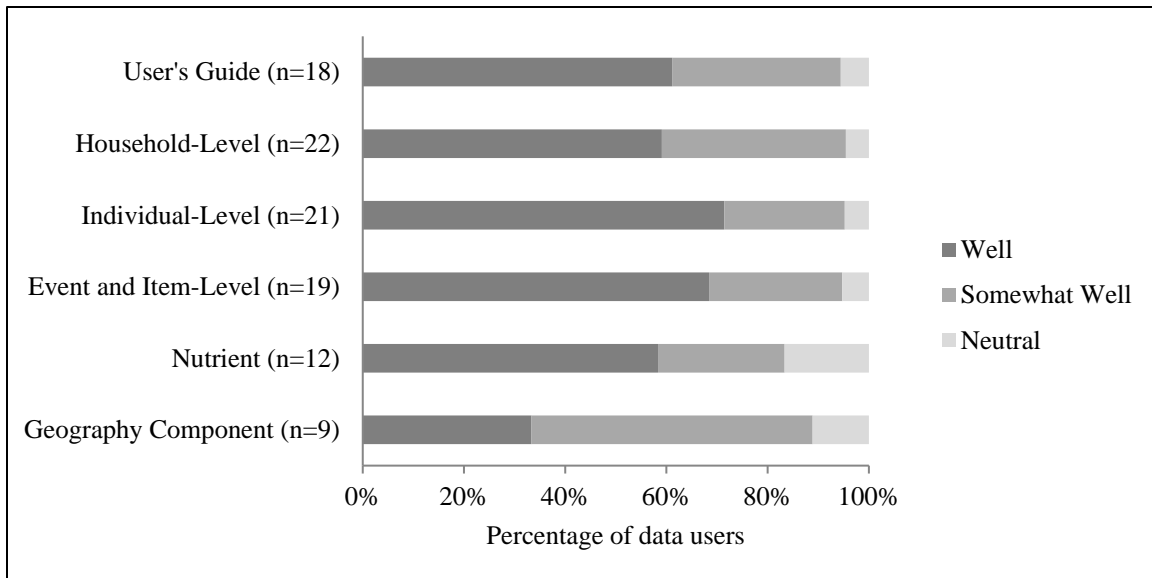


Exhibit 5. How well documentation met data users’ needs

Five respondents also appreciated the one-on-one support they received from USDA and NORC staff members. These respondents noted that staff members were very responsive to questions about the data files documentation, and the Data Enclave. One respondent also found a workshop that convened early users of FoodAPS to be useful. According to this respondent, the workshop provided helpful guidance around data usage and beneficial connections to other researchers using FoodAPS.

Data Limitations. Out of 24 respondents, several raised concerns about: 1) data quality, 2) sample size, and 3) detail on existing variables. Concerns about data quality varied, with 11 respondents mentioning issues that included:

- Missing food item-level information, including cost and weight, even when the individual reported a FAH or FAFH event (noted by four respondents);
- Inconsistent classification of supermarket or supercenter chains (noted by two respondents);
- Implausible values for individuals’ heights and weights (noted by two respondents);
- Potential underreporting of food pantry usage (noted by two respondents);
- Cumbersome SNAP verification procedures, which generated many cases of households that could not be matched (noted by one respondent); and
- Erroneous categorization of FAFH items into food groups (noted by one respondent).

Concerns about inadequate sample size were slightly more uniform, with five respondents raising this issue. Underrepresentation of specific subgroups, including WIC participants and poor and very poor SNAP participants, was of particular concern for two respondents. One

respondent also noted that FoodAPS lacks sufficient sample size to analyze consumer demand for less frequently purchased food items.

We compared FoodAPS to SNAP QC data to understand the income distribution of SNAP-participating households in both datasets. Weighted household counts from FoodAPS were lower than those from SNAP QC for households ≤ 100 percent of the Federal Poverty Level (see Exhibit 2). This difference may have arisen partly because of challenges in obtaining SNAP administrative data from all States with primary sampling units for FoodAPS.

We also compared weighted counts of WIC-participating individuals from FoodAPS to administrative participant counts (Exhibit 6). Although WIC participants were not a group targeted for oversampling, FoodAPS still included a substantial number of households with WIC-participating individuals, perhaps because groups that were oversampled were likely to include WIC participants. Compared to WIC administrative data though, weighted participant counts from FoodAPS data were lower.

Mean Monthly Participant Counts (thousands)		
	FoodAPS	Administrative Data
Total	6,381	8,844

Exhibit 6. Comparison of FoodAPS and administrative data on WIC participant counts

Source. Authors' tabulations of FoodAPS and USDA Food and Nutrition Service program data Note. National Household Food Acquisition and Purchase Survey (FoodAPS); Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).

Estimates from FoodAPS are weighted. For both estimates, the data collection period was April 2012 – January 2013.

Concerns about detail on existing variables often were related to information on food purchases and acquisitions. Four respondents noted the following distinct issues about lacking or limited information on:

- Shopping frequency (mentioned by one respondent);
- Household food inventories (mentioned by one respondent);
- Individual foods purchased using WIC and the WIC Cash Value Voucher (mentioned by one respondent); and
- Foods purchased from either food trucks or vending machines (mentioned by one respondent).

Apart from identifying issues with the data themselves, the data user survey pinpointed limitations with FoodAPS documentation and data usage logistics. Eleven respondents reported that certain codebooks contained limited descriptions of the data, with some specific issues being lack of detail about:

- Food retailer categorization (noted by three respondents);
- Data sources that were used to determine measures of distance (noted by one respondent);
- Rationale for selecting income category cutoff points to define target groups (noted by one respondent);

- Reference periods for questions about SNAP participation (noted by one respondent);
- Coding of portion sizes for food acquired by children at school (noted by one respondent); and
- Linkage of IRI data to information on FAH items (noted by one respondent).

Although this feedback suggests that data users would prefer more detailed documentation, at least three respondents noted that there were either too many codebooks in general or too many repeated details across codebooks. At least five respondents mentioned that the volume of information to understand about the data contributed to FoodAPS requiring a greater time investment relative to that for other datasets (Exhibit 7). Getting oriented to the documentation, however, was not the only factor informing respondents’ ratings about time investment. The time required to complete institutional paperwork, gain data access, become oriented with the Data Enclave, and overcome software issues all contributed to the greater than usual time investment needed to use FoodAPS.

Several respondents also named issues with data access and usage to be particularly challenging. Data users who had relatively early access to the data acknowledged improvements, including the shift from needing to use a thin client to being able to use the NORC Data Enclave. Even with such improvements, barriers to data access and usage included:

- Technical issues with the thin client and the Data Enclave (mentioned by eight respondents);
- Restricted nature of the dataset, (mentioned by six respondents);
- Delayed release of certain data files (mentioned by four respondents); and
- High costs associated with data access (mentioned by two respondents).

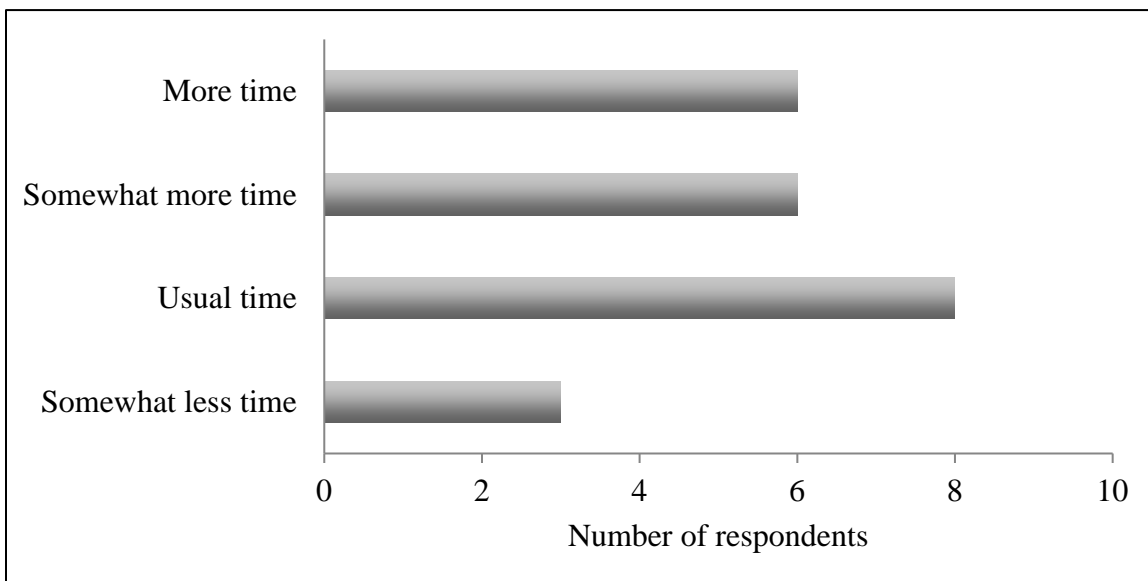


Exhibit 7. Time required to get oriented to FoodAPS compared to other datasets

Regarding the restricted nature of the dataset in particular, respondents identified three key issues. First, completing institutional paperwork to gain data access could be time intensive, as previously discussed. Second, proposing a research project without having a firmer grasp of the dataset’s contents could be difficult. Third, collaborating with research team members who did

not have data access permissions could be challenging. Although the restrictiveness of FoodAPS was a commonly reported barrier, only two of 24 respondents explicitly acknowledged tradeoffs between challenges of working with restricted data and benefits of having access to sensitive information while protecting study participant confidentiality. Placing these challenges in perspective, another respondent expressed appreciation for being able to access FoodAPS through the Data Enclave rather than through a Federal Statistical Research Data Center.

While not characterized as a limitation of the actual data themselves, two respondents disapproved of FoodAPS research that conflates food purchases and acquisitions with dietary consumption. These respondents described making inferences about dietary consumption with FoodAPS as a misuse of the data. One respondent explicitly called for USDA\ERS to educate both researchers and grant reviewers about appropriate uses. Another respondent characterized the lack of documentation around conversion of food items in as-purchased to as-consumed forms as problematic.

Recommendations for Future Rounds of Data Collection. Within our sample of 24 respondents, many researchers were enthusiastic about a future round of FoodAPS data collection. Along with recommendations for improving documentation and data files (see Appendices B and C, respectively), respondents offered recommendations related to: 1) modifying or building upon study design, 2) introducing new survey topics, 3) collecting more detail on existing survey topics, and 4) enhancing support and data access.

In terms of study design, six respondents recommended for FoodAPS to take on a panel data structure. Some of these respondents recognized the challenge of reconnecting with participants who responded to the initial round of FoodAPS data collection; to mitigate this challenge, one recommendation was to sample from similar geographic areas covered in the first round of FoodAPS. Eight respondents shared other suggestions about sampling, with two of these respondents calling for an increase in overall sample size. Other respondents recommended oversampling of the following groups:

- Low-income households who may be more likely to use food pantries (mentioned by three respondents);
- WIC participants (mentioned by two respondents);
- Farmers' market shoppers (mentioned by two respondents); and
- Vegetarians (mentioned by one respondent).

Also related to study design, six respondents suggested improvements in survey methods used to assess food purchases and acquisitions. These suggestions included:

- Changing the length of the study period either to two weeks (mentioned by two respondents) or to two shorter time periods (mentioned by one respondent);
- Maintaining the one-week study period but prompting participants to recall purchases and acquisitions for the past month (mentioned by one respondent);
- Advising households to use one form to streamline collection of food purchase and acquisition data from all household members (mentioned by one respondent);
- Providing clearer instructions to households about reporting individual foods purchased using WIC coupons or Cash Value Vouchers (mentioned by one respondent); and
- Including more probes to ascertain food pantry usage (mentioned by one respondent).

In terms of new survey topics, several respondents noted that it would be beneficial for the next round of FoodAPS to cover areas that were not included in the current dataset. Four respondents suggested collecting information on dietary consumption, which when coupled with information about food purchases and acquisitions could offer a measure of food waste. Three respondents suggested collecting information that might be relevant to the disproportionately low-income sample. These respondents were interested in data about government transfers, household debt, and health insurance coverage. Other suggested topic areas included:

- Physical activity (noted by two respondents);
- Home cooking (noted by one respondent);
- Environmental impact of food purchases and acquisitions (noted by one respondent); and
- Employment and workplace settings (noted by one respondent).

In terms of additional detail on existing survey topics, several respondents recommended that FoodAPS should capture more information about food purchases and acquisitions, food item features, and nutrition assistance program participation. Specific recommendations included collecting data on:

- Food purchasing and acquisition behaviors of children, especially in a context beyond the school food environment (noted by two respondents);
- Food purchasing and acquisition behaviors of seniors and people with disabilities (noted by two respondents);
- Food purchasing and acquisitions in the context of people's daily routines (one respondent);
- Factors influencing store choice (noted by one respondent);
- Food prices for all stores frequented by survey participants (noted by one respondent);
- Features related to production standards, such as organic and fair trade (noted by one respondent);
- Nutritional quality of food items, perhaps accompanied by guidance on constructing measures like the Healthy Eating Index (noted by one respondent);
- Foods prepared at home for away-from-home consumption (noted by one respondent);
- Length of SNAP participation spells (noted by one respondent);
- WIC participation, verified by a match to administrative data (noted by one respondent);
- Timing of other government transfers (noted by one respondent); and
- Historic program participation (noted by one respondent).

In terms of support and data access, six respondents offered suggestions on improving technical assistance, data access procedures, and funding mechanisms. While not necessarily related to the data themselves, implementing some of the following recommendations may be a relatively straightforward way for USDA to improve the data user experience.

- Improve the export request process (suggested by at least three respondents). Two respondents requested that details about the process, including allowable output formats and the expected timeline for the process, be explained, perhaps in a single reference document. Another respondent reported a potential lapse in communication between NORC and USDA about pending export requests, which could be remedied.

- Increase data access (suggested by at least two respondents). Both of these respondents commented on the value of releasing FoodAPS data for public use, with one respondent calling for more Geography Component variables to be incorporated into public use files. In terms of restricted data access, one respondent recommended that data access permissions be granted to entire research teams, rather than individual researchers, to facilitate collaboration within research teams.
- Broaden funding opportunities (suggested by at least two respondents). One respondent recommended that future funding opportunities be more inclusive of different types of research questions, particularly those without a SNAP focus. Another respondent called for future funding opportunities to be unlinked to FoodAPS data usage. This suggestion was based on the notion that there are diminishing returns to using FoodAPS since the data can only support a finite number of research questions.
- Foster connections between data users (suggested by at least two respondents). One respondent recommended that USDA\ERS set up a listserv linking data users. The listserv could partly serve as troubleshooting tool, potentially reducing technical assistance burden for USDA\ERS.

5. DISCUSSION AND CONCLUSION

Results of the literature review (Section 3) and the data user survey (Section 4) may have implications for priority-setting and future design decisions within FoodAPS as a longer-term program. These implications are organized into four broad themes, summarized in the next four sub-sections.

This summary applies the evidence from the literature review and the data user survey to these four themes. As noted in the introduction to this report (Section 1), this evidence represents just some of the considerations that USDA and the contractor must consider in implementing a new round of FoodAPS in 2019. Our input may be combined with other information about costs for survey and other data collection, and about the government's research priorities, to make sound decisions about research design and scope.

5.1. Breadth

A first theme is breadth of topic coverage. **Primary** research areas clearly are in scope as core purposes of FoodAPS.

- **Descriptive statistics for food acquisition and purchases.** A strength of FoodAPS is information about food acquisition, including both food expenditures and non-cash acquisition such as gifts, hunting, or home production. FoodAPS collects data on both FAH and FAFH acquisitions and purchases.
- **Impact of the food retail environment on food acquisition and purchase.** FoodAPS provides unique information for assessing the effects of the food retail environment.
- **Household food security.** Household food security may in part reflect general resource deprivation, and in part shortages more specifically related to food purchases and acquisition, so FoodAPS is uniquely well-suited to studying food security issues.
- **Impact of SNAP.** Whereas previous research has faced considerable difficulty accurately measuring SNAP participation status, FoodAPS offers adequate sample sizes and linkage with administrative records useful for measuring SNAP impacts.

Secondary research areas are well matched to the potential capabilities of FoodAPS, but nonetheless difficult choices may be made about covering some but not all within the feasible breadth of this survey. Some but not all secondary research areas may be covered well in a single survey. Although they are not seen as core purposes equivalent to the preceding list, these research areas are each connected in some way to a distinctive strength of FoodAPS. We see no reason why any of these topics would be rejected on their research merits. Yet, accepting some and excluding others may be wise in the context of limited research budgets.

- **Impact of WIC on food acquisition and purchases.** To address this research area, it is essential to have large samples of low-income pregnant women, postpartum mothers, infants, and young children.
- **Impact of school nutrition programs on food acquisition and purchases.** In particular, food acquired in school lunch and school breakfast programs count as part of FAFH. A related research question is to what extent other household grocery spending offsets or responds to the amount of food acquired in school nutrition programs.

- **Association between health insurance and food spending.** High rates of chronic disease and the high cost of treatment are important issues for national debates over health insurance.
- **Association between physical activity, the built environment, and weight status.** Food intake and physical activity are the two basic building blocks of the food energy balance equation, with implications for weight status and risk of obesity.

Tertiary research areas are less well matched to the distinctive strengths and core purposes of FoodAPS. Tertiary research areas appeared less connected to the distinctive strengths of FoodAPS. Each of these arose in the literature review or was mentioned by one or more of our survey respondents, and yet we recognized reasons not to base future design decisions on the desire to accommodate these topics.

- **Medicare, Medicaid, and other government transfers.** Because of the oversampling of low-income households in FoodAPS, there is considerable interest in studies of other transfer programs as their own independent topics, but FoodAPS may not be the best source for transfer program research unrelated to food acquisitions and purchases.
- **Dietary Quality.** FoodAPS variables describing the nutrition quality of food purchases, or describing respondent self-assessment of dietary quality, clearly offer useful insight into the impact of resources and the food environment on nutritional outcomes, but these outcome variables are not sufficient on their own to be used more generally for studies of individual food consumption or food intake.

5.2. Depth

A second theme is depth of detailed survey questions and administrative data effort on particular topics. For each of the following topics, it is clear that FoodAPS will have some coverage, and the open question is what level of detail should be attempted.

- **Instrumentation for FAH and FAFH.** An example of what we mean by “depth” is the decision in FoodAPS to invest in a critical scanner-based data collection approach. This does not add new topics, but instead adds detail and precision to an existing topic. The bar codes used in this approach only are available for at-home foods, not restaurant foods, cafeteria foods, or home production. This investment in depth suggests a special research focus on at-home grocery food. Yet, a key strength of FoodAPS may be the inclusion of away-from-home and home-produced food, not just grocery food.
- **SNAP administrative data.** Great care was taken in FoodAPS to ensure an adequate sample size for SNAP and corroboration using administrative sources for survey data on program participation.
- **WIC and school nutrition administrative data.** In the first round of FoodAPS, analysis files did not provide information from administrative records for other nutrition assistance programs. Variables based on self-reported participation were available, however.
- **Food intake data.** Some data users described the lack of food intake or food frequency data as a limitation. It may not be possible in a single survey, with limited respondent time and burden, to investigate food acquisition and food intake together. Over the course of a longer time period, such as a month, food acquisition likely matches food consumption by household members fairly closely. Especially for SNAP households,

who are known to experience a monthly food spending cycle, it could be valuable for a second round of FoodAPS to further explore including some questions about the timing of major grocery shopping trips, including especially whether the month's largest trip was included or not included during the reference week.

5.3. Power

FoodAPS invested heavily in ensuring adequate sample size for SNAP participants, low-income non-participants, and higher-income non-participants. Although the process of oversampling can be challenging, it may be feasible and worthwhile to ensure adequate sample size for WIC participants. WIC benefits, like SNAP benefits, are primarily redeemed through traditional retail channels, so including more WIC participants on a survey about food acquisitions and purchases may be reasonable. Federal nutrition assistance programs are especially high priority for USDA research, providing further justification for oversampling of WIC participants. We recognize though that, within the requirements of budgetary feasibility, no single survey can be well-powered for all potential research topics simultaneously.

5.4. Ease of use

There may be tradeoffs between other objectives and ease of use. For example, as noted earlier, one may pursue both greater power and greater precision in some food assistance and safety net variables by investing in linkage with administrative data. Yet, linking data from multiple sources may increase the time for data users to understand and use the data linkage. The additional complexity may be worth the trouble, and the difficulty for data users may be ameliorated by continuing to provide clear data codebooks, users' guides, and technical assistance.

REFERENCES

- Allard, S. W. & Ruggles, P. (2015). The spatial context of food shopping: Understanding how local food retailer access and pricing affect household behavior. (Paper presented at 2015 International Conference for Agricultural Economists). Retrieved from http://ageconsearch.umn.edu/record/212724/files/Allard_Ruggles_ICAE%202015%20JuneUpload.pdf.
- Basu, S., Wimer, C., & Seligman, H. (2016). Moderation of the relation of county-level cost of living to nutrition by the Supplemental Nutrition Assistance Program. *American Journal of Public Health, 106*, 2064-2070.
- Brewster, P. J., Guenther, P. M., Durward, C. M., & Hurdle, J. F. (2017). Grocery Purchase Quality Index-2016 scores are moderately correlated with Healthy Eating Index-2010 scores in the Food Acquisition and Purchase Survey, 2012–13. *The FASEB Journal, 31*(1 Supplement), 1b380-1b380.
- Bowen, S., Winkler, R., Bloom, J. D. & MacNell, L. (2016). Contextualizing family food decisions: The role of household characteristics, neighborhood deprivation, and local food environments. (University of Kentucky Center for Poverty Research Discussion Paper Series, DP 2016-08). Retrieved from http://www.ukcpr.org/sites/www.ukcpr.org/files/documents/DP2016-08_Bowen_et_al.pdf.
- Chang, Y., Kim, J., & Chatterjee, S. (2017). The association between consumer competency and Supplemental Nutrition Assistance Program participation on food insecurity. *Journal of Nutrition Education and Behavior, 49*, 657-666.
- Chenarides, L. & Jaenicke, E. C. (2016). Store choice and consumer behavior in food deserts: An empirical application of the distance metric method. (Invited paper prepared for 2017 Allied Social Sciences Association Annual Meeting). Retrieved from <https://ageconsearch.umn.edu/bitstream/250118/2/Store%20Choice%20and%20Consumer%20Behavior-2r.pdf>.
- Clay, D. M., Ver Ploeg, M., Coleman-Jensen, A., Elitzak, H., Gregory, C. A., Levin, D., ... & Rabbitt, M. P. (2016). *Comparing National Household Food Acquisition and Purchase Survey (FoodAPS) data with other national food surveys' data*. (EIB-157). Washington, DC: United States Department of Agriculture, Economic Research Service.
- Dong, D., Stewart, H., Frazao, E., Carlson, A., & Hyman, J. (2016). *WIC household food purchases using WIC benefits or paying out of pocket: A case study of cold cereal purchases*. (ERR-207). Washington, DC: US Department of Agriculture, Economic Research Service.
- Downing, J. & Laraia, B. (2016). Supermarket proximity and price: Food insecurity and obesity in the United States. (University of Kentucky Center for Poverty Research Discussion Paper Series, DP 2016-05). Retrieved from http://www.ukcpr.org/sites/www.ukcpr.org/files/documents/DP2016-05_Laraia_Downing.pdf.

- Fan, L., Baylis, K., Gundersen, C., & Ver Ploeg, M. (2015). Does a nutritious diet cost more in food deserts? (Selected paper prepared for 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting). Retrieved from <https://ageconsearch.umn.edu/bitstream/205424/2/Food%20Deserts%202015%20AAEAUpdated.pdf>.
- Gustafson, A., Allen, J. E. IV, Schoenberg, N., & Swanson, M. (2016). The relationship between neighborhood food environment and food store choice on purchasing habits among SNAP and lower income households, USDA FoodAPS data (University of Kentucky Center for Poverty Research Discussion Paper Series, DP 2016-09). Retrieved from https://uknowledge.uky.edu/ukcpr_papers/110.
- Handbury, J., Rahkovsky, I., Schnell, M. (2016). Is the focus on food deserts fruitless? Retail access and food purchases across the socioeconomic spectrum. Becker Friedman Institute for Research in Economics, University of Chicago (Health Economic Series No. 2016-08).
- Hillier, A., Smith, T. E., Whiteman, E. D., & Chrisinger, B. (2017). Discrete choice model of food store trips using National Household Food Acquisition and Purchase Survey (FoodAPS). *International Journal of Environmental Research and Public Health*, 14, 1133.
- 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. *Journal of Nutrition*, 147(5), 964-975.
- Jo, Y. (2017). *The differences in characteristics among households with and without obese children: Findings from USDA's FoodAPS*. (EIB-179). Washington, DC: US Department of Agriculture, Economic Research Service.
- Kirlin, J. A. & Denbaly, M. (2017). Lessons learned from the national household food acquisition and purchase survey in the United States. *Food Policy*, 72, 62-71.
- Li, J., Van de Kerckhove, W., & Krenzke, T. (2016). *Review of the FoodAPS 2012 imputation approaches for income and price data*. Washington, DC: Westat for US Department of Agriculture, Economic Research Service.
- Lyford, C., March, R. J., Carpio, C., & Boonsaeng, T. (2016). Do SNAP recipients get the best prices? (University of Kentucky Center for Poverty Research Discussion Paper Series, DP 2016-07). Retrieved from http://www.ukcpr.org/sites/www.ukcpr.org/files/documents/DP2016-07_Lyford_et_al.pdf.
- Maitland, A. & Li, L. (2016). *Review of the completeness and accuracy of FoodAPS 2012 data*. Washington, DC: Westat for US Department of Agriculture, Economic Research Service.
- Petraglia, E., Van de Kerckhove, W., & Krenzke, T. (2016). *Review of the potential for nonresponse bias in FoodAPS 2012*. Washington, DC: Westat for US Department of Agriculture, Economic Research Service.

Smith, T. A., Berning, J. P., Yang, X., Colson, G., & Dorfman, J. H. (2016). The effects of benefit timing and income fungibility on food purchasing decisions among Supplemental Nutrition Assistance Program households. *American Journal of Agricultural Economics*, 98(2), 564-580.

Taylor, R., & Villas-Boas, S. B. (2016). Food store choices of poor households: A discrete choice analysis of the National Household Food Acquisition and Purchase Survey (FoodAPS). *American Journal of Agricultural Economics*, 98(2), 513-532.

Tiehen, L., Newman, C., & Kirlin, J. A. (2017). *The food-spending patterns of households participating in the Supplemental Nutrition Assistance Program: Findings from USDA's FoodAPS*. (EIB-176). Washington, DC: US Department of Agriculture, Economic Research Service.

Todd, J. E. & Scharadin, B. *Where households get food in a typical week: Findings from USDA's FoodAPS*. (EIB-156). Washington, DC: US Department of Agriculture, Economic Research Service.

Ver Ploeg, M., Larimore, E., & Wilde, P. E. (2017). *The influence of foodstore access on grocery shopping and food spending*. (EIB-180). Washington, DC: US Department of Agriculture, Economic Research Service.

Ver Ploeg, M. & Wilde, P. (forthcoming). How do food retail choices vary within and between food retail environments? *Food Policy*.

Ver Ploeg, M., Mancino, L., Todd, J. E., Clay, D. M., & Scharadin, B. (2015). *Where do Americans usually shop for food and how they travel to get there? Initial findings from the National Household Food Acquisition and Purchase Survey* (EIB-138). Washington, DC: US Department of Agriculture, Economic Research Service.

Wilde, P., Llobrera, J., & Ver Ploeg, M. (2014). Population density, poverty, and food retail access in the United States: An empirical approach. *International Food and Agribusiness Management Review*, 17(A), 171-186.

Wilde, P. E., Steiner, A., & Ver Ploeg, M. (2017). For low-income Americans, living ≤ 1 mile (≤ 1.6 km) from the nearest supermarket is not associated with self-reported household food security. *Current Developments in Nutrition*, 1(11), e001446.

Ziliak, J. P. & Gundersen, C. (2016). Food security and geographic factors in food purchase and acquisition decisions: A compilation of research conducted under USDA Cooperative Agreement 58-5000-1-0050 and 58-5000-3-0066 (University of Kentucky Center for Poverty Research Discussion Paper Series, DP 2016-14). Retrieved from http://www.ukcpr.org/sites/www.ukcpr.org/files/documents/DP2016-14_Understanding_SNAP_research_compilation_0.pdf.

B. Data Usage (*Specific questions intended for data analysts*)

[If another research team member will answer these questions, check here:

3. Please consider the documentation materials you used or are using for your study. How well did these materials meet your needs?

	<i>Check if used</i>	<i>Poorly</i> 1	2	<i>Neutral</i> 3	4	<i>Well</i> 5	<i>Comments (optional)</i>
User's Guide	<input type="checkbox"/>						
<i>Household-Level</i>							
Household Codebook	<input type="checkbox"/>						
HH Geodata Codebook	<input type="checkbox"/>						
Access Codebook	<input type="checkbox"/>						
SNAP Eligibility Estimation Codebook	<input type="checkbox"/>						
<i>Individual-Level</i>							
Individual Codebook	<input type="checkbox"/>						
Meals Codebook	<input type="checkbox"/>						
<i>Event and Item-Level</i>							
Places Codebook	<input type="checkbox"/>						
FAH Events Codebook	<input type="checkbox"/>						
FAH Items Codebook	<input type="checkbox"/>						
FAH Item IRI Codebook	<input type="checkbox"/>						
FAFH Supplementary Documentation	<input type="checkbox"/>						
FAFH Events Codebook	<input type="checkbox"/>						
FAFH Items Codebook	<input type="checkbox"/>						
<i>Nutrient</i>							
FoodAPS Nutrient Coding Overview	<input type="checkbox"/>						
FAH Nutrient Codebook	<input type="checkbox"/>						
FAFH Nutrient Codebook	<input type="checkbox"/>						

3. (Continued) How well did these materials meet your needs?

	<i>Check if used</i>	<i>Poorly 1</i>	<i>2</i>	<i>Neutral 3</i>	<i>4</i>	<i>Well 5</i>	<i>Comments (optional)</i>
<i>Geography Component</i>							
Retail Environment Study Codebook	<input type="checkbox"/>						
Variable List: Retail Environment Study	<input type="checkbox"/>						
Construction of Weekly Store-Level Food Basket Costs: Documentation	<input type="checkbox"/>						

General comments on documentation (*optional*):

4. Please consider the data files you used or are using for your study. How well did these data files meet your needs?

	<i>Check if used</i>	<i>Poorly 1</i>	<i>2</i>	<i>Neutral 3</i>	<i>4</i>	<i>Well 5</i>	<i>Comments (optional)</i>
<i>Household-Level Data</i>							
faps_household	<input type="checkbox"/>						
faps_hhweights	<input type="checkbox"/>						
faps_hhgeodata	<input type="checkbox"/>						
faps_access	<input type="checkbox"/>						
faps_snapelig_hh	<input type="checkbox"/>						
faps_snapelig_unit	<input type="checkbox"/>						
faps_snapelig_ind	<input type="checkbox"/>						
<i>Individual-Level Data</i>							
faps_individual	<input type="checkbox"/>						
faps_reportstatus	<input type="checkbox"/>						
faps_meals	<input type="checkbox"/>						
<i>Event and Item-Level Data</i>							
faps_places	<input type="checkbox"/>						
faps_fahevent	<input type="checkbox"/>						
faps_fahitem	<input type="checkbox"/>						
faps_fahitem_iri	<input type="checkbox"/>						
faps_fafhevent	<input type="checkbox"/>						
faps_fafhitem	<input type="checkbox"/>						
<i>Nutrient Data</i>							
faps_fahnutrient	<input type="checkbox"/>						
faps_fafhnutrient	<input type="checkbox"/>						
<i>Geography Component Data</i>							
BGMain_v6	<input type="checkbox"/>						
BGRestaurant_v2	<input type="checkbox"/>						
Tract_v5	<input type="checkbox"/>						
County_v6	<input type="checkbox"/>						
basketprices	<input type="checkbox"/>						
basketprices_raw	<input type="checkbox"/>						

4. (Continued) How well did these data files meet your needs?

	<i>Check if used</i>	<i>Poorly 1</i>	<i>2</i>	<i>Neutral 3</i>	<i>4</i>	<i>Well 5</i>	<i>Comments (optional)</i>
<i>Store Linker Files</i>							
PlaceID_IRI_TempERSID	<input type="checkbox"/>						
PlaceID_TDLinx_TempERSID	<input type="checkbox"/>						
PlaceID_TempERSID	<input type="checkbox"/>						

General comments on data files (*optional*):

5. Compared to your experience using similar datasets, how long did it take you to orient yourself with FoodAPS? Please consider the time you took to review documentation, learn about the data enclave, and any other preliminary tasks. Please highlight your response or mark it with an X.

Less time		Usual time		More time
1	2	3	4	5

Comments (optional):

6. How would you rate the ease or difficulty of linking FoodAPS data files? Please highlight your response or mark it with an X.

Very easy		Neither easy nor difficult		Very difficult
1	2	3	4	5

Comments (optional):

7. How would you rate the ease or difficulty of applying survey weights to FoodAPS data? Please highlight your response or mark it with an X.

Very easy		Neither easy nor difficult		Very difficult
1	2	3	4	5

Comments (optional):

8. What barriers, if any, did you encounter when using FoodAPS data?

C. Recommendations (*General questions intended for investigators and data analysts*)

[If another research team member will answer these questions, check here:

9. What are your suggestions, if any, for adjustments to be made for a possible next round of FoodAPS? For example, your response may consider sample size, groups for oversampling, content focus areas, questions to ask, etc.

10. Are there any other comments you wish to share?

11. Would you be willing to answer follow-up questions via telephone? The purpose of a follow-up telephone call would be to clarify answers provided in writing. Please highlight your response or mark it with an X.

Yes

No

APPENDIX B. Full list of recommendations for documentation

This list contains paraphrased comments from respondents about documentation. In some instances, we re-framed comments as recommendations if they were originally expressed as limitations. We present suggestions that either are general in scope or pertain to multiple codebooks by theme. Codebook-specific suggestions appear in Exhibit B-1.

Access to Documentation

- Allow users to access FoodAPS codebooks outside of the data enclave, even if response frequencies must be redacted.
- Include documentation for other datasets, which can be linked to FoodAPS, within the data enclave. For example, documentation on other USDA datasets like the Food and Nutrient Database for Dietary Studies (FNDDS) could be made available.
- Ensure timelier release of codebooks, especially those related to the Geography Component.

Documentation of Survey Methods

- Provide a document that researchers could use as a model when preparing manuscripts. This document should contain a high-level summary about:
 - study design,
 - sampling,
 - data collection, and
 - categories of data.
- Describe the administrative matching process, including more detail on:
 - when matching took place,
 - how administrative dates were chosen,
 - why certain records could not be matched, and
 - the probabilistic algorithm used to match survey data to administrative records.
- Provide detail on how SNAP participation was assessed during interviews, particularly with respect to the reference period.
- Describe methods used to convert items from as-purchased to as-consumed forms.

Linkage of Data Files

- Share sample SAS and Stata code for merging data files.
- Offer clearer instructions on which data files to merge and which ones to append. For data files that should be appended, standardize variable names to allow for streamlined linkage.
- Provide more guidance on how to construct a long-form dataset at the person-day level.
- Provide more guidance on how to link FoodAPS and IRI data. Making matches between stores has been challenging.

Application of Survey Weights

- Provide examples of how to use the survey weights in practice, especially in instances where multiple levels of data are being analyzed.

Addition of Detail or Documentation

- Provide more detailed descriptions of response frequencies, especially for variables that currently do not have a key for response options.
- Offer more explanation for why target groups were defined using income to poverty ratios that differ from federal income eligibility standards for SNAP.
- Improve the description of store type categorization, and include more information about STARS criteria for retailer classification. Also, explain how store types translate to other variables, including co3, cs3, mlg3, ss3, and sm3.
- Create documentation for store linker files.
- Add documentation on how to use data to construct different measures of dietary quality.

Streamlining of Documentation

- Reduce the number of codebooks.
- Make several codebooks more concise by removing background information already found in the User’s Guide.
- Consolidate information about data linkage in one document, obviating the need to refer to multiple documents.
- Create a document that cross-lists variables by data file.

	Codebook	Recommendations
Household-Level	Household Codebook	<ul style="list-style-type: none"> • Describe which road files were used to calculate distances between places of residence and retailers.
	HH Geodata Codebook	<ul style="list-style-type: none"> • Improve description of American Community Survey data, clarifying whether measures are 5-year estimates or averages of five 1-year estimates.
	Access Codebook	<ul style="list-style-type: none"> • Define variables, like the one for combination grocery/other stores, more clearly and fully.
Event and Item-Level	Places Codebook	<ul style="list-style-type: none"> • Revise categories that seem to overlap.
	FAH Events Codebook	<ul style="list-style-type: none"> • Indicate how many non-food items were included in the study more clearly.
	FAH Item IRI Codebook	<ul style="list-style-type: none"> • Provide more information on how to link FAH items with IRI data. • Provide a crosswalk between FoodAPS and IRI food categories.
	FAFH Items Codebook	<ul style="list-style-type: none"> • Include more detail on quantities of foods that children acquired at school.
Geography Component	Construction of Weekly Store-Level Food Basket Costs: Documentation	<ul style="list-style-type: none"> • Ensure consistency between descriptions of variables and corresponding data.

Exhibit B-1. Codebook-specific recommendations from data user survey

APPENDIX C. Full list of recommendations for data files

This list contains paraphrased comments from respondents about data files. In a few instances, we re-framed comments as recommendations if they were originally expressed as limitations. We group suggestions that either are general in scope or pertain to multiple data files. Data file-specific suggestions appear in Exhibit C-1.

Availability of Data

- Continue making nutrient data available.
- Provide explanations for why FAH and FAFH event data may be unavailable even if participants reported that they had purchase or acquisition events.

Addition of Data Files or Variables

- Offer several master data files, containing all variables for the same level of data (e.g. household, individual, etc.). Data users would have a smoother experience with data linkage and could clean master data files according to their needs.
- Include interviewer characteristics, including an identifier, in the data files.
- Include responses to the screener survey.
- Provide a score for administratively matched data, similar to how the ArcGIS match score is provided for the place data.
- Add variables that flag unusual values (e.g., for purchased food quantities or anthropometric information).

Streamlining and Standardization of Procedures and Variables

- Align food categories with those used in the USDA Food Plans and/or the Quarterly Food at Home Price Database. Add these food category or group identifiers to FoodAPS data files, and keep variable names consistent.
- Avoid mixing non-standard item codes with standard UPC and PLU codes, particularly for FAH items.
- Streamline the administrative data verification process. Attempt to have greater uniformity for this process so that the resulting administratively matched data is more comparable across States.

Updates and Revisions

- Indicate in advance which variables, like survey weights, are subject to change. Provide some measure of how temporary these variables are if possible.
- Make updates and revisions according to a known timeline. Create a naming convention to distinguish between different data file versions, and encourage researchers to report on which version of the data they used for their analyses.

	Data File	Recommendations
Individual-Level	faps_reportstatus	<ul style="list-style-type: none"> • Construct report status variables to differentiate between FAH and FAFH events.
Event and Item-Level Data	faps_places	<ul style="list-style-type: none"> • Provide basic geographic information, such as county name, in addition to longitude and latitude coordinates.
	faps_fahitem	<ul style="list-style-type: none"> • Provide clearer item descriptions.
	faps_fafhevent	<ul style="list-style-type: none"> • Include a variable that distinguishes fast food restaurants from other types of restaurants. For places that fall into a gray area, like fast casual establishments, provide a rationale for how this distinction was made.
	faps_fafhitem	<ul style="list-style-type: none"> • Provide clearer and more specific item descriptions.

Exhibit C-1. Data file-specific recommendations from data user survey