

# ANALYSIS OF CLASSIFICATION DECISIONS ON THE CONSUMER EXPENDITURE SURVEY

David Cantor, Kerry Levin, Westat, Inc., and Leslie Miller, Paul Hsen, Bureau of Labor Statistics  
David Cantor, 1650 Research Blvd., Rockville, MD 20850

## 1. Introduction

The Consumer Expenditure Survey (CE) collects data on expenditures for a random sample of U.S. households. The data are used as input into the Consumer Price Index to trace inflation, as well as analysis of more detailed consumer patterns. The CE is split into two components: (1) an in-person survey, and (2) a diary that is filled in over a pair of one week periods.

The purpose of this paper is to describe research that addresses two questions regarding the diary for the CE:

1. How do diary respondents determine whether the classification of a food item should be food for home consumption (FFHC) or food away from home (FAFH)?; and
2. Does the type of diary used affect the way respondents make the FFHC vs. FAFH decision?

Addressing the issue of classification into the FAFH and FFHC categories is important for measuring variation in food purchases. The price of food is highly influenced by where it is bought and the amount of preparation involved. Food that is prepared and eaten at home is less expensive than food bought at a restaurant. For this reason, the decision to eat out or eat at home serves as an important indicator of trends in discretionary spending.

Making the FFHC vs. FAFH distinction in the CE diary is not straightforward. The labels FAFH and FFHC are ambiguous with respect to whether the location of purchase or consumption should be used in the classification. Furthermore, there are other factors that respondents may consider when classifying a food item. This includes, for example, the purchaser's intent and whether the food has to be prepared or not. Even if respondents use the same rules, it is unclear whether the rule conforms to the correct BLS definition. For example, from the respondent's point of view, purchasing food in bulk but eating it at work is not clearly either FAFH or FFHC. By BLS definition, however, it should be classified as FFHC.

To examine these issues, this project performed an experiment that examined factors considered important when individuals make classification decisions. In particular, a number of issues were addressed, including:

- How much agreement is there in the classification decision across a wide range of situations?;
- What are the important criteria used in making a classification decision?;
- How do these criteria conform to the "correct" BLS definition?;
- Do these decisions differ by the type of diary form used?; and
- Do these criteria differ by the demographic characteristics of the respondent?

Once these questions were examined, recommendations were formulated with respect to the diary form and the current CE diaries.

## 2. Experimental Design

The design of the experiment consisted of showing each respondent 43 vignettes which described a situation in which food was purchased and consumed. After reading each vignette, the respondent was asked to classify the situation as either FFHC or FAFH. Once all 43 vignettes were completed, the respondent was asked to describe (in writing) the decision rules used to make his/her classifications.

Vignettes were presented using a personal, laptop computer. The order of the vignettes and the sex of the purchaser in each vignette was randomized for each respondent. Each respondent was randomly assigned to one of three diary conditions: (1) no form; (2) production diary (diary A); and (3) revised diary (diary B). Groups of six respondents were administered the 43 vignettes. Each respondent had his/her own personal computer to read the vignettes and record his/her answers. The same type of diary was used for all respondents in each group of six.

Each of the 43 vignettes represented a cell in an experimental design that crossed two different factors. Factors were selected based upon their presumed importance to the decision making process. One of the factors represented where the food was eaten. This factor had four levels:

- At home;
- At the establishment where it was bought;
- In transit; and
- At work.

The second factor represented where food was purchased. This factor had six levels:

- At home for pickup;
- At home for delivery;
- Grocery store;
- Fast food restaurant;
- Table service restaurant; and
- Vending machine.

Nested within this second factor were two additional variables. The first variable made the distinction between whether the purchase was primary or secondary to the shopping. This was nested within the grocery store, fast food and table service restaurant conditions. The second nested variable was whether the food was prepared or not prepared. This was nested within the grocery store condition.

To assist in clarifying the design, the following provides the text of one vignette:

"(Tom/Theresa) went to (NAME OF GROCERY STORE) to do some shopping. Besides (his/her) regular groceries, (he/she) purchased some prepared potato salad from the deli and took it home and ate it"

This vignette represents a secondary purchase of prepared food made in a grocery store and eaten at home. The secondary purchase refers to the fact that the food consumed was not the intended part of the shopping trip. Prepared food refers to those involving no preparation by the purchaser, in this case, potato salad.

A design that crosses all of these conditions results in 44 potential cells (or vignettes). The 43 used in the experiment resulted from eliminating one cell in the design because it was not logical (i.e., purchasing food from home by phone for delivery and eaten at the establishment).

### 3. Experimental Procedures

Prior to conducting the testing session, the facilitators set up the testing room for the respondents. Regardless of location, the testing set up was identical. Each testing session included only one of the form conditions.

Two facilitators were present at each session. One of the facilitators formally presented the instructions to the respondents. The other facilitator gave individual attention when necessary and also gave instructions to the latecomers. Upon arrival, the facilitator read an introduction to the respondents. For diary conditions A and B, the introduction to the study also included a discussion of how respondents should use the diary form during this testing session. Respondents in the "no form" condition were given a form that simply asked them to classify the food purchased into column labeled FAFH or FFHC. No other information was presented on this form.

Following the introduction to the study, respondents were given the opportunity to ask questions regarding the purposes/objective of the session.

Respondents were then asked to classify the 43 vignettes. This consisted of:

- Reading each vignette;
- Writing the vignette number onto the form; and
- Entering the FFHC or FAFH classification into the computer.

Following completion of the series of vignettes, respondents were asked to complete a debriefing form.

### 4. Results

#### 4.1 Description of Study Subjects

A total of 151 subjects participated in the study. Almost three-fourths of the subjects were female and the majority (82%) were white. Respondents were well-mixed in terms of age, being evenly distributed across the four age categories. In addition, almost two-thirds of the respondents (61%) were married.

#### 4.2 Rate of Agreement Among Respondents

The overall rate of agreement was quite low (31%). This varied, however, by the type of vignette. Table 1 provides a more detailed picture of the rate

of agreement using a measure that converts the "percent agreement" for each cell to an "agreement index" (AI). This is defined as:

$$AI = \frac{FFHC\% - 50}{50}$$

where FFHC% is the percent of responses classified as FFHC. This measure ranges from a value of -100 (when the percentage of respondents assigning a vignette to FAFH is 100%), to a value of 100 (when the percentage of respondents assigning a vignette to FFHC is 100%). Zero represents the situation where 50% of the respondents assigned a vignette to FAFH and 50% of the respondents assigned the same vignette to FFHC.

Table 1 presents the AI for each vignette. Concentrating on individual cells within the table, it seems that respondents had an easier time assigning vignettes to FAFH than to FFHC. This is indicated by several patterns. First, the column and rows having the highest negative numbers represent FAFH vignette assignments (see row 2, columns 7-10). Second, of the 10 situations involving an agreement index greater than positive or negative 90, nine are negatives.

There is a particular concentration of high agreement in vignette assignment when the food was purchased at a fast food/table service restaurant and eaten either at the establishment or in transit. The only vignette with a value greater than +90 that was primarily classified as FFHC involved a prepared food purchased as a primary purchase at the grocery store and eaten at home.

Low rates of agreement can be traced to difficulties in interpreting the meaning of the words "home" and "consumption" in the FFHC and FAFH labels. This is illustrated by focusing on the row "at home" and the column for grocery store purchases. One would expect relatively high agreement for the situation involving "a primary purchase of not prepared food in a grocery store." This is the most common purchase of food that is eventually consumed at home. As one might expect, there is high agreement for this situation when the food is eaten at home (88.5%). However, when moving down the rows of Table 1 for the other categories in the not prepared column, the agreement drops dramatically to 22.5, 22.9, and 38.9, for establishment, transit, and work, respectively. The overall agreement for this column is 43.2. When one moves across the "at home" row outside of the grocery store column, one can see similar ambiguity. If the food is bought at a table service restaurant, fast

food restaurant, or vending machine, the level of agreement drops below 20. If food is ordered from home, the level of agreement drops to below 60.

### 4.3 How Did Respondents Make Decisions?

Where the food is eaten (EATEN) is an extremely important factor when making the decision to classify. Specifically, if the food was eaten at home, respondents tended to classify the food as FFHC. If the food was eaten somewhere else, it was classified as FAFH. For the factor where the food was purchased (PURCHASED), the important distinction was whether the item was purchased in a fast food restaurant, table service restaurant, or vending machine. If the food was purchased in one of these three locations, respondents tended to classify the item into FAFH. Items purchased in a grocery store or from home had a tendency to be classified in FFHC, but this significantly interacted with the particular EATEN situation.

The other factors that were considered in the experiment, including primary vs. secondary and prepared vs. not prepared did not seem important. There are small differences among these cells and the pattern of these effects is inconsistent. For example, when scanning the prepared vs. not prepared columns, neither is uniformly greater than the other. Rather, the magnitude and direction of the effects differ by the type of establishment and by where the food was eaten. Apparently, there is a relatively complex set of decisions that were made by subjects when evaluating these vignettes. This resulted in a number of higher order interactions among the factors included in the experiment.

Despite these complications, it is interesting to note that when ordering food from home, food delivered to the house was more likely to be classified as FFHC than if it was not delivered. This is consistent with the idea that any type of purchase that is related to home activity will be classified as FFHC. There also seems to be an isolated effect of preparation within the primary purchase column for grocery stores. That is, when the item is prepared and consumed outside the home, there is more agreement that the purchase should be classified as FAFH when compared to the not prepared items.

It should be noted that some of these interactions may be due to the content of the vignettes used for the grocery store situations. For example, the prepared vs. not prepared distinction was made using items that, under reasonable definitions, cannot clearly be put into one of these categories. For primary purchases, the "not prepared" items were sandwich meat and peanuts.

Similarly, the "not prepared" items under the secondary purchase was a candy bar. With the exception of sandwich meat, these items do not normally require any preparation when consuming. In this sense, therefore, they can be thought of as already being prepared.

#### 4.4 Individual Differences in Decision Making

Section 4.3 described the significance of the two primary experimental factors related to the FFHC or FAFH assignment. In order to understand the unique processes respondents used to arrive at different decisions, the debriefing information was used, in conjunction with the experimental results, to develop a profile of the different methods respondents used to make decisions.

The debriefing consisted of asking respondents two questions on their decision making process. In order to assist in formulating individual decision rules, these data were coded into groups according to the primary decision rules used in response to these questions. The first question on the debriefing form was an open-ended question asking respondents to provide the rules used in making classification decisions. This was the information that was primarily used to classify the response. The second question was more specific and asked the role "location" played in the decision process. Responses to the second question were used in the coding when respondents did not provide a sufficient answer to the first question.

When a unique decision rule was described on the debriefing form, a category was created and labeled. When other respondents described the same type of rule or classification scheme, they were placed in the same category. In cases where the respondent gave more than one classification rule, the project staff members classified the first rule mentioned. As a result, each respondent's debriefing form was placed into only one category.

Overall, the coding process resulted in the following eight different categories.

1. **WHERE EATEN:** Where the food was actually eaten.
2. **WHERE PURCHASED:** What type of store the food was purchased in; the differentiation occurs between a restaurant and a grocery store.
3. **WHERE PREPARED:** Where the food was prepared; home vs. restaurant/grocery store.

4. **PURCHASER'S INTENT:** Some assumption made by the respondent about how the person in the vignette intended to use the food.
5. **RESPONDENT'S FOOD USE:** Classified based upon how the respondent would use the food regardless of the situation being described.
6. **PRICE OF FOOD:** Food Away From Home is considered more expensive. This may be related to overhead costs but that specific term was never mentioned.
7. **TYPE OF FOOD:** Classified based upon (1) whether or not the food could stand alone as a "complete meal," or (b) whether or not the food is too messy to be eaten away from home.
8. **UNCLASSIFIED:** Insufficient information from respondent.

Over 50% of the respondents reported making their classifications based upon a decision rule related to where the food was eaten. Where the food was purchased, where the food was prepared, and the intent of the purchaser were the other rules identified as being used by approximately 15%, 11%, and 10% of the respondents, respectively. The other categories described above were used by a very small percent of the respondent population.

One limitation of the debriefing is that it relies on self-reports of a relatively complicated process. Respondents may have forgotten how the decision had been made or they may have deliberately simplified the process for clarity. In order to assess how closely these categories relate to the empirical patterns, a discriminant analysis was conducted. This analysis attempts to predict the criterion for assigning FFHC or FAFH (i.e., the debriefing category membership) using the classification information contained in the assignment responses for the 43 vignettes. Once the classification equations were estimated, each person was assigned a debriefing category using these equations. This predicted category was then compared to the original category for the same person.

For the purposes this analysis, several modifications to the data were necessary. First, the debriefing groups with fewer than ten respondents were combined into one group. Second, all missing

values were replaced by a value of 0.5 (i.e., to approximate a random assignment of that vignette to FFHC or FAFH) so that people with missing values would not be completely thrown out of the analysis.

The results indicated that more than half of the respondents (i.e., 77 out of the 144) were reclassified into the same debriefing group. While this is a relatively high proportion, it still indicates that the stated reasons in the debriefing do not totally capture the actual decision-making process. This is also reflected in the fact that those individuals using the simplest criterion are the most consistent. People in the WHERE EATEN category were the most consistent, with 74% being reclassified into that group on the basis of their vignette assignments. The people in the WHERE PURCHASED and WHERE PREPARED categories had a little more than one-third of the respondents reclassified into their original groups. Almost as many were reclassified into debriefing group 4, the "miscellaneous" group. This means that many of the people in these groups really used some mixture of criteria that was not captured accurately by their stated criterion. The people in debriefing group 4 were the most inconsistent. Less than one-third of them were reclassified into group 4. Interestingly, however, a very large proportion of these individuals resembled those persons who described their purchases as WHERE EATEN.

#### 4.5 Accuracy of Information

The results on agreement and decision-making do not address the extent to which classification decisions conform to BLS definitions of FFHC and FAFH. High agreement, for example, does not imply the information collected is "correct." In order to assess accuracy, a profile of "correct" answers was developed using definitions provided by BLS staff.

Table 2 presents, by vignette, the percentage of respondents making the "correct" assignment for each vignette. The first pattern to notice is that the factor EATEN is not nearly as important in determining the "correct" response as it is for determining assignment to FFHC. The average percent correct in the EATEN row is 46.1%. This corresponds to an average of half of the responses being correct for food eaten at home. For the PURCHASED factor, there is a large difference between the percentages correct for places other than the grocery store and the low percentages correct for vignettes involving the grocery store. This is due to the significant drop in correct responses in the grocery store for food eaten outside the home.

Many of the cells refer to situations that are not particularly common. However, some of these cells account for a significant portion of the purchases made by consumers. This includes food ordered and eaten at home (20.3%, 21.7%), food bought in a fast food restaurant and eaten at home (46.8%, 39.6% classified correctly), and unprepared food bought at a grocery store, but eaten at work (31.2%, 30.6% classified correctly).

#### 4.6 Summary of Results

The results of the experiment revealed six primary conclusions:

1. There was significant inconsistency in the classification of the vignettes.
2. The decision process is quite complicated and varied across respondents. The most common criteria used in making decisions were:
  - Whether the food was eaten at home or elsewhere; and
  - Whether the food was bought in some type of restaurant or elsewhere.

These criteria explained between 32% and 36% of the variance of the classification process.

3. Variables that were not highly significant in the decision process included:
  - Amount of food preparation;
  - Whether the purchase was primary or secondary;
  - The type of form used; and
  - Testing site, age, sex, education and income of the subject.
4. The decision rules that were most commonly used in classification were not consistent with BLS definitions of FAFH and FFHC.
5. The form did not make a difference in the classification into FFHC or FAFH categories.
6. The subcategories on revised diary form (Form B) were used much more frequently than the production diary (Form A).

**Table 1**  
**Agreement Index**

	WHERE PURCHASED										
	FROM HOME BY PHONE		AT ESTABLISHMENT								11 Machine
	1 For Delivery	2 For Pickup	Grocery Store				Fast Food		Restaurant		
			Secondary		Primary		7 Secondary	8 Primary	9 Secondary	10 Primary	
3 Prep.			4 Not Prep.	5 Prep.	6 Not Prep.						
<b>WHERE EATEN</b>											
1 At Home	59.4	56.5	78.1	58.9	94.3	88.5	6.4	20.9	11.4	32.4	29.0
2 At Establishment	NA	-83.2	-40.0	-54.9	-78.4	-22.5	-94.4	-95.8	-66.7	-90.1	-97.1
3 In Transit	-51.1	-93.1	3.6	-32.9	-72.2	-22.9	-95.7	-97.1	-94.4	-97.2	-66.0
4 At Work	-39.4	-77.6	-7.8	-37.7	-72.0	-38.9	-81.9	-70.6	-78.9	-86.0	-53.5

**Notes:** 0 = (No agreement) 50% FFHC, 50% FAFH  
 20.0 = 60% FFHC - 20 60% FAFH  
 40.0 = 70% FFHC - 40 70% FAFH  
 60.0 = 80% FFHC - 60 80% FAFH  
 80.0 = 90% FFHC - 80 90% FAFH  
 100 = 100% assigned FFHC - 100 100% assigned FAFH

**Table 2**  
**Percent of Correct Responses\***

	WHERE PURCHASED										TOTAL	
	FROM HOME BY PHONE		AT ESTABLISHMENT									
	1 For Delivery	2 For Pickup	Grocery Store				Fast Food		Restaurant			11 Machine
			Secondary		Primary		7 Secondary	8 Primary	9 Secondary	10 Primary		
3 Prep.			4 Not Prep.	5 Prep.	6 Not Prep.							
<b>WHERE EATEN</b>												
1 At Home	20.3	21.7	89.1	79.4	2.9	94.2	46.8	39.6	44.3	33.8	35.5	46.1
2 At Establishment	NA	91.6	70.0	22.5	89.2	38.7	97.2	97.9	83.3	95.1	98.6	78.4
3 In Transit	75.5	96.5	51.8	33.6	86.1	38.6	97.9	96.6	97.2	98.6	83.0	77.9
4 At Work	69.7	88.8	46.1	31.2	86.0	30.6	91.0	85.3	89.4	93.0	76.8	71.6
<b>TOTAL</b>	51.1	74.6	64.2	41.7	66.0	50.5	83.2	80.3	78.6	80.1	73.5	

\*Total sample of 144 respondents.