

# Consumption and Food Security: An Examination

David Tirrell, Anthony Karabenick, Ben Mitchell, Swaraj Rai

**QAC 380**

## **1 Research Question**

Our primary research goal was to investigate whether an increase in consumers' total amount in dollars spent at Food on the Move (FOTM) markets who completed the Rhode Island Public Health Institute (RIPHI) self-reported survey(s) significantly impacted their food security.

## Table of Contents

Title, Research Question.....	Page 1
Table of Contents.....	Page 2
Hypothesis, Confounding Variables, Methods.....	Page 3
Methods.....	Page 4
Methods.....	Page 5
Methods.....	Page 6
Methods.....	Page 7
More Methods.....	Page 8
Results.....	Page 9
Results.....	Page 10
Results.....	Page 11
More Results.....	Page 12
Summary, Implications.....	Page 13
Limitations.....	Page 14
Limitations.....	Page 15

## **2 Hypothesis**

We hypothesized that the increase in amount of dollars spent at FOTM markets at different survey assessments (baseline (0), first follow up (1), and second follow up (2)) would be associated with improved food security when controlling for key demographic variables.

## **3 Confounding Variables**

Confounding variables that were potentially present in this analysis included the number of people covered by a single loyalty card (how many people are in your individual household), how many weeks the average consumer's SNAP benefits lasted, the dropout rate of members in the SNAP program, and the timing between the baseline and follow up surveys.

## **4 Methods**

### **4.1 FOTM Survey Data**

Initial management of the FOTM participant data included calculating a food security raw score for each participant. The food security metric, a six-item short form of the survey module and the associated Food Security Scale, was developed by researchers at the National Center for Health Statistics and is endorsed by the USDA. The six questions used to determine food security scores of an individual/household are as follows:

*1) 'The food that (I/we) bought just didn't last, and (I/we) didn't have money to get more.' Was that often, sometimes, or never true for (you/your household) in the last 12 months?*

*2) '(I/we) couldn't afford to eat balanced meals.' Was that often, sometimes, or never true for (you/your household) in the last 12 months?*

*3) In the last 12 months, since last (name of current month), did (you/you or other adults in your household) ever cut the size of your meals or skip meals because there was not enough money for food?*

4) *[IF YES ABOVE, ASK] How often did this happen—almost every month, some months but not every month, or in only 1 or 2 months?*

5) *In the last 12 months, did you ever eat less than you felt you should because there was not enough money for food?*

6) *In the last 12 months, were you every hungry but did not eat because there was not enough money for food?*

The responses of these questions were taken from a response scale varying from “Yes/No” to “Often True/Sometimes True/ Never true/ DK or Refused” and recoded to reflect a binary variable in which affirmative responses of “Yes”, “Often True”, or “Sometimes True” were taken as affirmative and coded to reflect a 1, whereas, non-affirmative responses like “No”, “Never True”, or omissions like “DK or Refused” were taken as non-affirmative responses. After each of the survey questions were coded to act as a binary numeric variable, a new variable (*foodsecurityrawscore*) was calculated by adding all the affirmations of a respondent and assigning a score dependent on their number of affirmations. The scoring rubric was as follows:

Number of affirmatives	Scale score
0	NA
1	2.86
2	4.19
3	5.27
4	6.30
5	7.54
6 (evaluated at 5.5)	8.48

The USDA categorizes food security using these scores as follows:

- Raw score of 0-1: High or marginal food Security
- Raw score of 2-4: Low Food Security
- Raw score of 5-6: Very Low Food Security

Conclusively, across all three time periods, there were [REDACTED] total observations: [REDACTED] in the baseline, [REDACTED].

#### 4.2 POS Data

Our main independent variable in our analysis (*Totalamount*) was derived from our POS dataset described below. Total amount represents the sum of all payment methods used by FOTM consumers during all three time periods. We used Total amount to represent consumption for unique consumers in our analysis. Figure 1 below further shows the specific payment methods that contributed to the Total amount, with SNAP Discount and EBT being the major contributors.

Figure 1

*\*redacted due to confidentiality\**

*Figure 1 represents the distribution of benefits that contribute to the Totalamount variable in the POS dataset.* [REDACTED]

The initial POS dataset contained all SNAP program transactions across all store locations. The total number of transactions totaled to be [REDACTED] transactions in the raw dataset. The data was imported into STATA, where we consolidated duplicates for all unique loyalty card numbers. More specifically, the data consisted of item level transactions at all FOTM Markets. In many cases, transactions of the same loyalty card number were recorded on both the same and different days. This caused the loyalty card number variable to not be a valid ID variable for our data merge (this was due to a unique identifier error in STATA when using the *merge* function). As a result, the total amount of all transactions from both same and different days were summed together to show each loyalty card number's total amount spent from baseline (t=0) to second follow up (t=2). A summing of Total amount was also necessary due to overlap in survey responses amongst FOTM consumers (inaccurate time flows between Survey and POS datasets made it impossible to distinguish consumption by different time periods). A byproduct of this was a reduced POS dataset, going from over [REDACTED] transactions to a total of [REDACTED] POS data points. This in turn allowed for the merge process to be facilitated. The cleaned POS dataset consisted of [REDACTED] unique loyalty card numbers with a corresponding sum of Total amount.

#### **4.3 Merged Data**

In STATA, after the loyalty card number variable in the FOTM dataset and the POS dataset was isolated, we were able to merge the datasets by a new unique loyalty card number variable (*uniqueloyaltycardnumber*). This variable appended the time period (*loyaltycardnumber.001*, *loyaltycardnumber.006*, *loyaltycardnumber.012*) to distinguish repeat responders across time periods in the survey dataset and allow for the merge. The cleaned merged POS and survey data included [REDACTED] observations. This resulted in [REDACTED] dropped observations in the survey dataset [REDACTED] dropped observations in the POS dataset [REDACTED]

Table 2.

Control Variable	Description	How it was coded
<i>Age</i>	Age of survey respondent	Ready for analysis (no adjustments)
<i>HHLD_SIZE_NUM</i>	Number of members in household of respondent	Converted from a string to a numeric variable (destring)
<i>RETIRED_NUM2</i>	Whether respondent is retired or not	Converted to a binary variable 1= Yes 0= No
<i>FRUIT_NUM2</i>	Frequency of fruit consumption by respondent	0= Never 1=Per month 2=Per Week 3=Per Day . = Don't Know/Refused to Respond
<i>VEG_DARK_GREEN_NUM2</i>	Frequency of dark green vegetable consumption by respondent	0= Never 1=Per month 2=Per Week 3=Per Day . = Don't Know/Refused to Respond
<i>phys_act_days_num2</i>	Number of days a week respondent	Numeric variable: 0-7

*Note: This table provides an explanation of controls that were incorporated into the merged dataset for Table 6 in Results*

#### 4.4 Additional Data Management

Additional data management included dropping missing observations and observations that were irrelevant to the premise of this project. These variables tracked the weight of the customer and other dietary habits

After basic data management, the variable which described each quartile (*quartile\_group*) was created using the Total amount variable in which the customer is categorized as either 0, 1, 2, or 3 depending on whether they were in the 0<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, or 75<sup>th</sup> percentile of total expenditure at FOTM in their given time-period. This provided room for the statistical analysis that can be found in the Results section.

## Results

Figure 2

*\*redacted due to confidentiality\**



Figure 3.

*\*redacted due to confidentiality\**



Figure 4

*\*redacted due to confidentiality\**





Table 3

*\*redacted due to confidentiality\**

Table 4

*\*redacted due to confidentiality\**

Table 5

*\*redacted due to confidentiality\**

Tables 3-5 are simple linear regressions of the Dependent Variable ( ) being regressed upon by the Independent Variable ( ).

Table 6

*\*redacted due to confidentiality\**

## Summary of Conclusions

[REDACTED]

## 5 Implications

[REDACTED]

[REDACTED]

## 6 Limitations

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]