MEMORANDUM

To: Sammy Eaddy
From: William Medwid, Ryan Destefano, Richa Puranik, Isabelle Estrada
Date: May 17, 2023

Re: California Polytechnic State University, Department of Kinesiology and Public Health

---

The purpose of this memo is to describe the statistical methods and findings from an analysis of your data. We hope this information helps assess the following research questions:

1. *"Does family food insecurity predict lower child diet quality for 2-year-olds?"*
2. *"Is there a cumulative effect of longer durations of food insecurity on child diet quality for 2-year-olds?"*
3. *"Are associations between food insecurity and child diet quality moderated by WIC participation and parents' use of pressuring or restrictive feeding practices?"*

The memo is organized into five sections.

If you have any addition questions regarding our recommendations and analysis following today's meeting, please feel free to reach out to either Professor Heather Smith or Professor Hunter Glanz at hsmith@calpoly.edu or hglanz@calpoly.edu to coordinate an additional meeting to address any questions.

I: <u>Abstract of Key Findings</u>

We found no statistically significant relationship between family food security and child diet quality at 2 years of age. Cumulative food security was a better predictor of child diet quality, but still not statistically significant with this current data. It is likely that more data or more explanatory factors could reveal a statistically significant relationship between these. WIC participation (or non-participation) did not significantly increase or decrease the association between family food security and child diet quality either.

II: Background and Data

Our understanding is that you want assistance in selecting an appropriate statistical method that will assist you with answering your research questions. There are three questions that we will focus on in this report: *"Does family food insecurity predict lower child diet quality for 2-year-olds?", "Is there a cumulative effect of longer durations of food insecurity on child diet quality for 2-year-olds?",* and *"Are associations between food insecurity and child diet quality moderated by WIC participation and parents' use of pressuring or restrictive feeding practices?".* Your data was collected via telephone interviews and from other databases, so it should be noted that your research is an observational study, and therefore any conclusions drawn from our analyses are not causational effects, but associations between the variables in question.

The WIC Infant and Toddler Feeding Practices Study 2 (WIC ITFPS-2) dataset provided a large but unbalanced sample of dietary information to draw conclusions from. Because the data is sourced from caregivers enrolled in WIC, there are very few observations in the "Good" diet category. Though the dataset originally included a high number of "Poor" diet observations, once we removed observations with missing values, these were less common than "Needs Improvement" observations as well. **Table 1** displays the number of observations in each diet category in the original dataset along with the number which had no missing values, passing our pruning process.

Diet Category	Original Observations	Post-Pruning Observations
Poor (0-50)	1880	323
Needs Improvement (51-80)	1980	1510
Good (81-100)	58	41

Table 1*. Count of observations by dietary quality category, before and after pruning missing values from dataset.*

Unfortunately, 52% of families in the dataset had missing data and had to be excluded from our analysis. This attrition of case counts may be problematic due to the possibility that we are removing specific types of people by removing those with missing values. We will proceed under the assumption that those who remain are still representative of the population at large, but this assumption may warrant further investigation. **Table 2** displays the number of missing values for each variable. While some variables are missing more often than others, and could therefore be excluded in order to include more observations, 1453 of the 2044 total missing observations were missing their total diet score. Because excluding those 1453 observations would come with similar ethical concerns as removing all 2044, we decided to remove all 2044 in order to retain the use of all available variables for analysis.

Variable	Label	Missing Values
Race	Race category of mother or caregiver (imputed) (SL_RaceCG) (Not used)	521
Ethnicity	Ethnicity of mother or caregiver (imputed) (SL_EthnicityCG)	521
Mother Marital Status	Marital status (imputed) (SL_MomMaritalStatus)	732
Baseline Food Security	Food security score obtained (imputed) (SL_FoodSecurity) (Not used)	521
Parity	Parity (imputed) (SL_Parity) (Not used)	732
Poverty	Poverty level 2013 (imputed) (SL_Poverty2013s)	521
Age at Birth	Age of mother or caregiver at child birth (SL_AgeatBirth_CG) (Not used)	732
13 Month Food Security	Refreshed at 13 Month: Food security score obtained (imputed) (RSL_FoodSecurity) (Not used)	898
Food Security	Refreshed at 24-Month: Food security score obtained (imputed) (RSL_FoodSecurity24m)	898
WIC Participation	Refreshed at 24-Month: Child's WIC participation status at 24-Month (imputed) (RSL_ChildWICPartStat24m)	898
Cumulative Food Security	Cumulative experiences of Food Insecurity	898
Exposure	24mo: Parent use of repeated exposure (Not used)	2025
Monitoring	24mo: Parent monitoring of child diet (Not used)	2027
Pressuring	24mo: Parent pressures child to eat	2025
Restrictive	24mo: Parent restricts child intake	2025
Diet Score	TOTAL HEI-2015 SCORE	1453

Table 2. *Missing value counts of each column of data.*

III. Statistical Methods

We recommend using the statistical method known as Ordinal Logistic Regression with a separate model for each research question resulting in a total of three models. With these models you will be able to explore the **association between family food security and child diet quality** for 2 year olds and the **impact on this association** when **including WIC participation and parent use of pressuring or restrictive feeding practices** to the model. You will also be able to explore any effects of **cumulative food insecurity on child diet quality** for 2 year olds.

Research Question 1 Model:

$$P(Y \leq J) = \frac{e^{\alpha j + \beta 1\, x1 + \beta 2\, x2 + \beta 3\, x3 + \beta 4\, x4 + \beta 5\, x5 + \beta 6\, x6}}{1 + e^{\alpha j + \beta 1\, x1 + \beta 2\, x2 + \beta 3\, x3 + \beta 4\, x4 + \beta 5\, x5 + \beta 6\, x6}}$$

Where Y = Diet score
J = Diet score category (poor, needs improvement, or good)
P(Y≤ J) = Probability child diet score is of category J or lower
x_1 = Food security of category 1 (1 for food security = 1, 0 otherwise)
x_2 = Food security of category 2 (1 for food security = 2, 0 otherwise)
x_3 = Ethnicity (1 for ethnicity = hispanic/latino, 0 otherwise)
x_4 = Mother marital status (1 for mother marital status = married, 0 otherwise)
x_5 = Poverty category 1 (1 for poverty level = 1, 0 otherwise)
x_6 = Poverty category 2 (1 for poverty level = 2, 0 otherwise)

With this model, the statistical quantities that you will be most interested in are β_1 and β_2 because these coefficients are the ones that help to measure how much more or less likely an observation is to be in diet score category J based on food security category after accounting for the other covariates (ethnicity, mother marital status, and poverty).

Research Question 2 Model:

$$P(Y \leq J) = \frac{e^{\alpha j + \beta 1\, x1 + \beta 2\, x2 + \beta 3\, x3 + \beta 4\, x4 + \beta 5\, x5 + \beta 6\, x6}}{1 + e^{\alpha j + \beta 1\, x1 + \beta 2\, x2 + \beta 3\, x3 + \beta 4\, x4 + \beta 5\, x5 + \beta 6\, x6}}$$

Where Y = Diet score
J = Diet score category (poor, needs improvement, or good)
P(Y≤ J) = Probability child diet score is of category J or lower
x_1 = Cumulative food security of category 1 (1 for cumulative food security = 1, 0 otherwise)

x_2 = Cumulative food security of category 2 (1 for cumulative food security = 2, 0 otherwise)
x_3 = Ethnicity (1 for ethnicity = hispanic/latino, 0 otherwise)
x_4 = Mother marital status (1 for mother marital status = married, 0 otherwise)
x_5 = Poverty category 1 (1 for poverty level = 1, 0 otherwise)
x_6 = Poverty category 2 (1 for poverty level = 2, 0 otherwise)

With this model, the statistical quantities that you will be most interested in are β_1 and β_2 because these coefficients are the ones that help to measure how much more or less likely an observation is to be in diet score category J based on cumulative food security category after accounting for the other covariates (ethnicity, mother marital status, and poverty).

Research Question 3 Model:

This model is the same as the model for Research Question 1 except it includes six more β terms, representing six individual interactions between the variables WIC participation, restrictive, and pressuring with food security.

An important note to consider is that we excluded any observations from the dataset that contained missing values. Thus, we removed all rows from the dataset that contained missing values. This brought the sample size for the analysis down to 1, 874 observations.

For all tests, we will consider a result to be statistically significant if the p-value is less than 0.05.

IV. Results

Analysis #1: Does family food insecurity predict lower child diet quality for 2-year-olds?

To investigate this, we created an ordinal logistic regression model as stated in the Statistical Methods section. This model predicts the probability of being in a specified diet score group or lower based on the explanatory variables food security, ethnicity, mother marital status, and poverty.

Parameter	Level	Coefficient	P-value
Intercept (diet score)	1	-1.418	<.0001
Intercept (diet score)	2	4.016	<.0001
Food security	1	-0.088	.628
Food security	2	0.049	.820
Ethnicity	1	-0.555	<.0001
Mother marital status	1	0.094	.466
Poverty	1	0.467	.822
Poverty	2	0.129	.552

Table 3. *Food Security vs Diet Score Model coefficient estimates and P-values*

Looking at **Table 3**, we see that the coefficients for the variables food security (level 1) and food security (level 2) are **not significant predictors** of diet score. This is evidenced by the large p-values for both of these variables, respectively. We must note that this result is adjusted for the covariates of ethnicity, mother marital status, and poverty. So, in the presence of these covariates, while there is a small association between food security at 24 months and diet score, it is not large enough to be considered statistically significant.

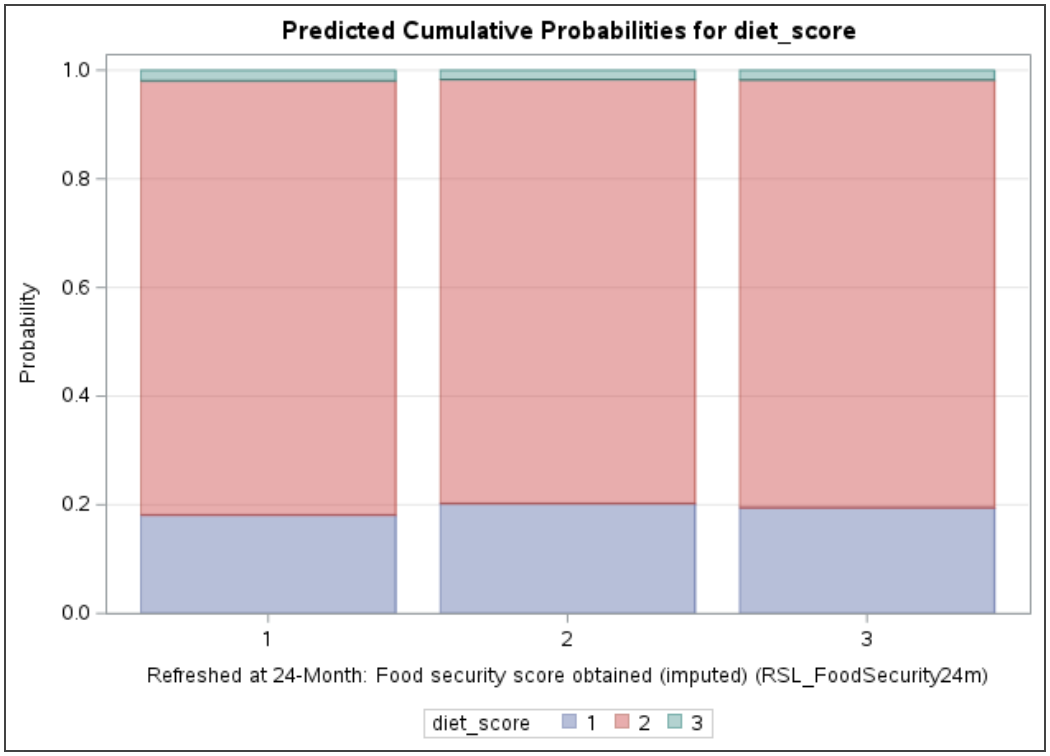

Figure 1. *Cumulative predicted probability for diet score, by 24-Month Food Security Category.*
Food Security Labels: 1 = High or Marginal Food Security, 2 = Low Food Security, 3 = Very Low Food Security
Diet Score Labels: 1 = Poor, 2 = Needs Improvement, 3 = Good

This lack of significant association between food security level and diet score for two year olds is evidenced by **Figure 1**. The cumulative predicted probabilities for diet score barely differ at each food security level. For example, looking at the cutoff for a diet score of level 1 across all food securities which is at about a probability of 0.2, tells us that regardless of food security level about 20% of children will be predicted to be in diet score category 1.

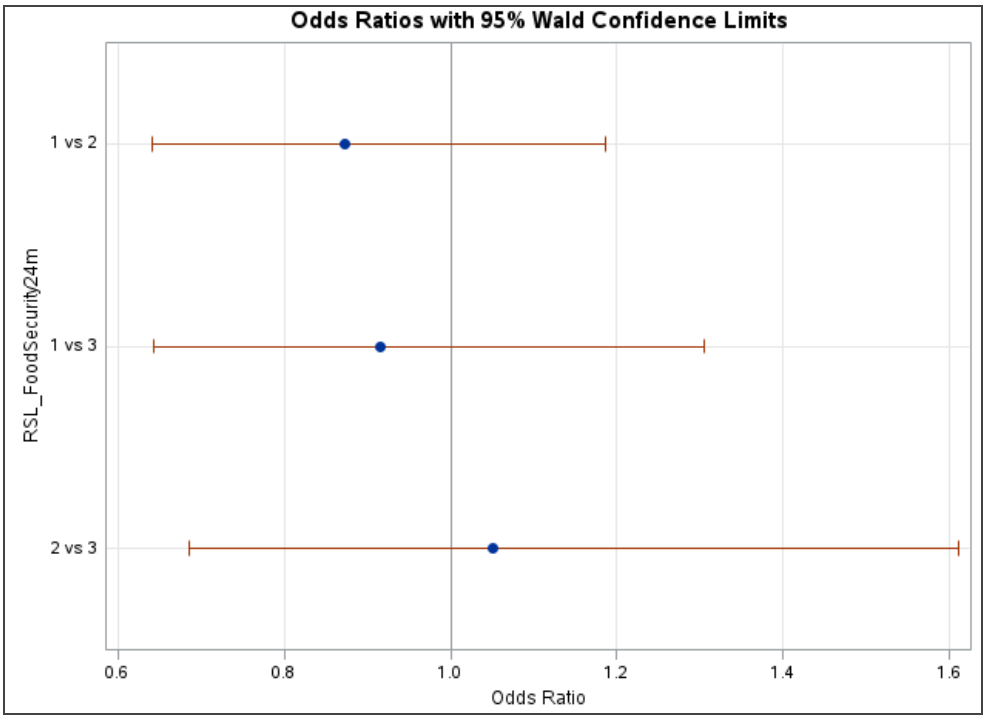

Figure 2. *Odds Ratio comparisons of the effects of 24-Month Food Security on Diet Score. 95% Wald Confidence Limits included.*
Food Security Labels: 1 = High or Marginal Food Security, 2 = Low Food Security, 3 = Very Low Food Security

This lack of significant association is also evidenced by **Figure 2**. In this figure, an odds ratio equal to 1 is contained within all the 95% confidence intervals for comparisons between different food security levels. The odds ratio comparison in Figure 2 also allows us to see the difference between food security level 1 and level 2.

The takeaway from this research question is that there is **not a statistically significant association** between the variables food security and diet score.

Analysis #2: Is there a cumulative effect of longer durations of food insecurity on child diet quality for 2-year-olds?

This model predicts the probability of being in a specified diet score group or lower based on the explanatory variables cumulative food security, ethnicity, mother marital status, and poverty.

Parameter	Level	Coefficient	P-value
Intercept (diet score)	1	-1.507	<.0001
Intercept (diet score)	2	3.934	<.0001
Cumulative food security	1	-0.011	.929

Parameter	Level	Coefficient	P-value
Cumulative food security	2	0.463	.051
Ethnicity	1	-0.555	<.0001
Mother marital status	1	0.101	.431
Poverty	1	0.057	.784
Poverty	2	0.131	.546

Table 4. *Cumulative Food Security vs Diet Score Model coefficient estimates and P-values*

In **Table 4**, we see that the coefficients for the variables cumulative food security (level 1) and cumulative food security (level 2) are **not significant predictors** of diet score. This is evidenced by the large p-values for level 1 and the p-value of .051 for level 2. We must note that again, this result is adjusted for the covariates of ethnicity, mother marital status, and poverty. It is worth noting that a cumulative food security level of 2 is very close to being a significant predictor of diet score for 2 year olds.

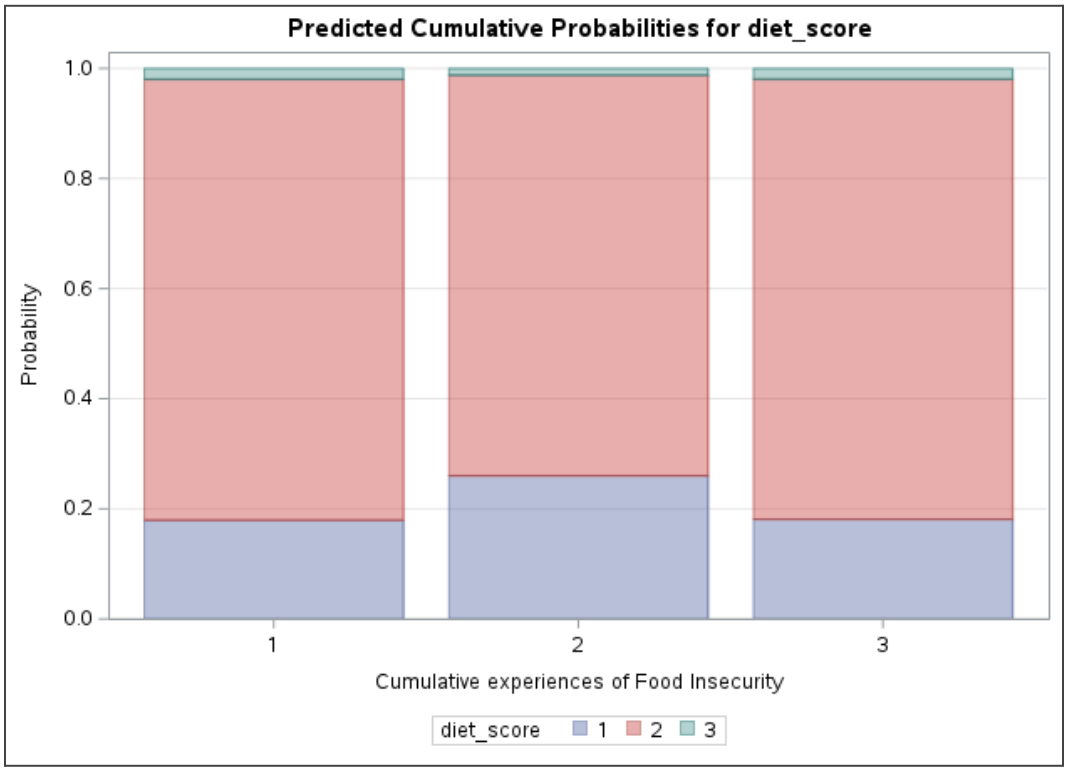

Figure 3. *Cumulative predicted probability for diet score by Cumulative experiences of Food Insecurity.*
Diet Score Labels: 1 = Poor, 2 = Needs Improvement, 3 = Good
Cumulative Food Insecurity Labels: 1 = No experiences, 2 = 2 experiences , 3 = 3 or more experiences

This lack of an overall significant association between cumulative food security level and diet score for two year olds is evidenced by **Figure 3** above. In Figure 3, we see that the cumulative predicted probabilities for diet score barely differ at each food security level. For example, looking at the cutoff for a diet score of level 1 across all food securities which is at about a probability of .2, tells us that regardless of food security level about 20% of children will be predicted to be in diet score category 1. From the graph we can see the previously noted occurrence of almost significance for cumulative food security level 2, for this level about 25% of 2 year old children will be predicted to be in diet score level 1.

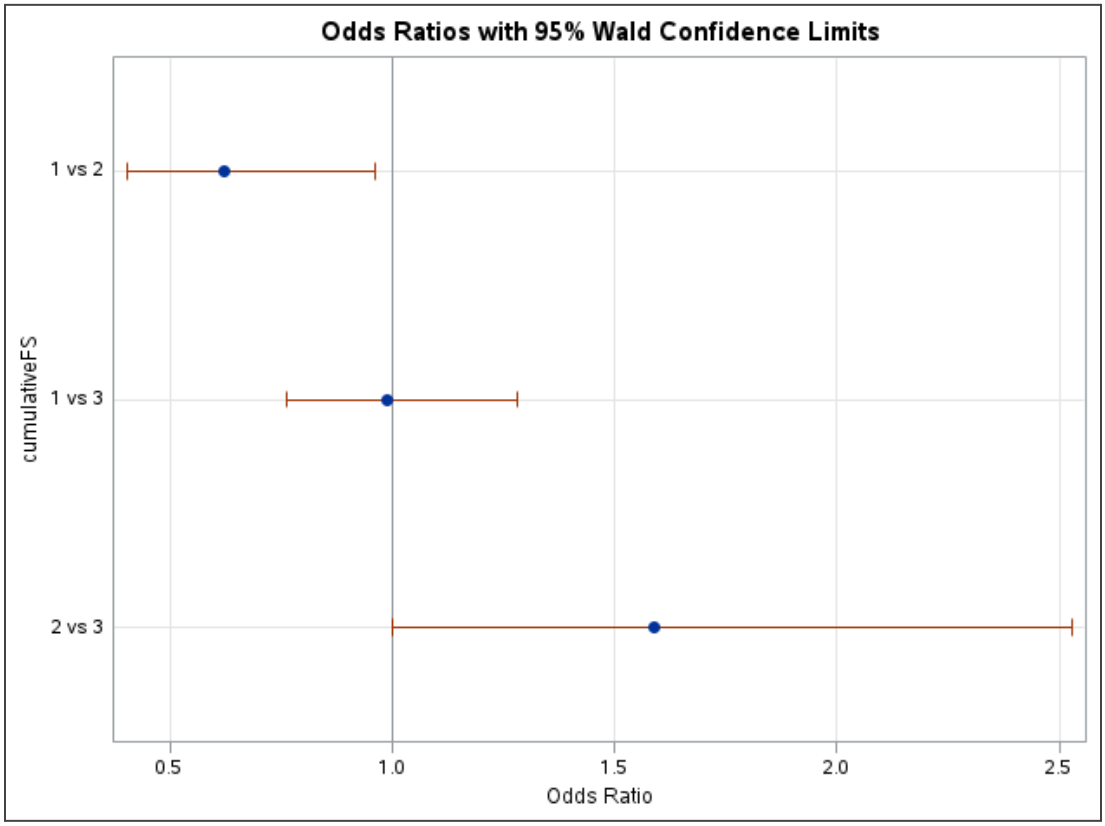

Figure 4. *Odds Ratio comparisons of the effects of Cumulative experiences of Food Insecurity on Diet Score. 95% Wald Confidence Limits included.*
Cumulative Food Insecurity Labels: 1 = No experiences, 2 = 2 experiences , 3 = 3 or more experiences

This lack of significant association is also evidenced by **Figure 4**. In this figure, an odds ratio equal to 1 is contained within the 95% confidence intervals for comparisons between cumulative food security levels (1 and 3) and (2 and 3). This figure uncovers the statistically significant result between cumulative food security levels 1 and 2, an odds ratio of 1 is not contained in this comparison's odds ratio confidence interval. Since the interval is entirely below 1, this tells us that a 2 year old child is less likely to be in a lower diet score group when the child is in cumulative food security level 1 compared to cumulative food security level 2.

The takeaway from this section is that while there is 1 significant difference between cumulative food security groups 1 and 2, the overall variable **cumulative food security does not have a significant association with diet score**.

Analysis #3: Are associations between food insecurity and child diet quality moderated by WIC participation and parents' use of pressuring or restrictive feeding practices?

This model predicts the probability of being in a specified diet score group or lower based on the explanatory variables cumulative food security, ethnicity, mother marital status, poverty, and interactions between food security and the variables WIC participation, restriction, and pressuring.

Effect	P-value
WIC participation * food security	.112
restriction * food security	.241
pressuring * food security	.296

***Table 5**. Interaction model effect tests (only displaying interaction effects)*

In **Table 5** above, none of the interaction terms have a significant effect on diet score for 2 year olds which is evidenced by the large P-values. This tells us that the association between food security level and diet score is not significantly modified by the variables WIC participation, restriction, and pressuring. Again, we must note that these results are adjusted for the variables ethnicity, mother marital status, and poverty.

V. Conclusion

We found no statistically significant results to any of the three research questions. This means that we have not found evidence that current food insecurity and cumulative food insecurity experiences increase or decrease a child's diet score for 2 year olds. That doesn't mean that the child receives enough food, only that the quality of the food that they do receive is of approximately the same nutritional quality despite their food insecurity situation. Further studies with a larger sample size or investigating diet score as a continuous variable may be able to produce more significant results.

VI. Technical Output

Our analysis will be simple to replicate because it is all contained in the attached SAS file. For your convenience, we have included relevant screenshots of the raw output from this analysis:

1. Load in the provided dataset (itfps466_updated_2023511.sas7bdat).
    a. You may need to change the folder path to match your SAS environment. This is the only change you will need to make to run our analysis.
2. Encode Diet Score as a categorical variable.
3. Remove rows with missing values from the dataset.
4. Data Exploration
    a. **Display diet score frequencies from the un-pruned data.**
    (Used for table 1)

diet_score	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1880	47.98	1880	47.98
2	1980	50.54	3860	98.52
3	58	1.48	3918	100.00

    b. **Display diet score frequencies from pruned data (no missing values).**
    (Used for table 1)

diet_score	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	323	17.24	323	17.24
2	1510	80.58	1833	97.81
3	41	2.19	1874	100.00

    c. **Display how frequently each variable is missing.**
    (Used for table 2)

Variable	Label	N Miss
SL_RaceCG	Race category of mother or caregiver (imputed) (SL_RaceCG)	521
SL_EthnicityCG	Ethnicity of mother or caregiver (imputed) (SL_EthnicityCG)	521
SL_MomMaritalStatus	Marital status (imputed) (SL_MomMaritalStatus)	732
SL_FoodSecurity	Food security score obtained (imputed) (SL_FoodSecurity)	521
SL_Parity	Parity (imputed) (SL_Parity)	732
SL_Poverty2013s	Poverty level 2013 (imputed) (SL_Poverty2013s)	521
SL_AgeatBirth_CG	Age of mother or caregiver at child birth (SL_AgeatBirth_CG)	732
RSL_FoodSecurity	Refreshed at 13 Month: Food security score obtained (imputed) (RSL_FoodSecurity)	898
RSL_FoodSecurity24m	Refreshed at 24-Month: Food security score obtained (imputed) (RSL_FoodSecurity24m)	898
RSL_ChildWICPartStat24m	Refreshed at 24-Month: Child's WIC participation status at 24-Month (imputed) (RSL_ChildWICPartStat24m)	898
cumulativeFS	Cumulative experiences of Food Insecurity	898
repeatexpos24m	24mo: Parent use of repeated exposure	2025
monitoring24m	24mo: Parent monitoring of child diet	2027
pressuring24m	24mo: Parent pressures child to eat	2025
restrictive24m	24mo: Parent restricts child intake	2025
HEI2015_TOTAL_SCORE	TOTAL HEI-2015 SCORE	1453

5. Research Questions

a. **Research Question 1**

Model information:

The LOGISTIC Procedure

Model Information	
Data Set	WORK.WIC_PRUNED
Response Variable	diet_score
Number of Response Levels	3
Model	cumulative logit
Optimization Technique	Fisher's scoring

How model variables are encoded:

Class Level Information			
Class	Value	Design Variables	
RSL_FoodSecurity24m	1	1	0
	2	0	1
	3	0	0
SL_EthnicityCG	1	1	
	2	0	
SL_MomMaritalStatus	1	1	
	2	0	
SL_Poverty2013s	1	1	0
	2	0	1
	3	0	0

P-values for each predictor:

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
RSL_FoodSecurity24m	2	0.8778	0.6448
SL_EthnicityCG	1	19.0999	<.0001
SL_MomMaritalStatus	1	0.5326	0.4655
SL_Poverty2013s	2	0.5319	0.7665

Model parameter estimates and confidence levels:

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-1.4181	0.2573	30.3751	<.0001
Intercept	2	1	4.0155	0.2991	180.2460	<.0001
RSL_FoodSecurity24m	1	1	-0.0876	0.1807	0.2348	0.6280
RSL_FoodSecurity24m	2	1	0.0495	0.2180	0.0515	0.8204
SL_EthnicityCG	1	1	-0.5546	0.1269	19.0999	<.0001
SL_MomMaritalStatus	1	1	0.0940	0.1288	0.5326	0.4655
SL_Poverty2013s	1	1	0.0465	0.2063	0.0509	0.8216
SL_Poverty2013s	2	1	0.1290	0.2166	0.3545	0.5516

Confidence intervals for differences between odds ratios:

Odds Ratio Estimates and Wald Confidence Intervals			
Odds Ratio	Estimate	95% Confidence Limits	
RSL_FoodSecurity24m 1 vs 2	0.872	0.641	1.186
RSL_FoodSecurity24m 1 vs 3	0.916	0.643	1.306
RSL_FoodSecurity24m 2 vs 3	1.051	0.685	1.611

b. **Research Question 2**

Model information:

The LOGISTIC Procedure

Model Information	
Data Set	WORK.WIC_PRUNED
Response Variable	diet_score
Number of Response Levels	3
Model	cumulative logit
Optimization Technique	Fisher's scoring

How model variables are encoded:

Class Level Information			
Class	Value	Design Variables	
cumulativeFS	1	1	0
	2	0	1
	3	0	0
SL_EthnicityCG	1	1	
	2	0	
SL_MomMaritalStatus	1	1	
	2	0	
SL_Poverty2013s	1	1	0
	2	0	1
	3	0	0

P-values for each predictor:

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
cumulativeFS	2	4.6644	0.0971
SL_EthnicityCG	1	19.2108	<.0001
SL_MomMaritalStatus	1	0.6191	0.4314
SL_Poverty2013s	2	0.4877	0.7836

Model parameter estimates and confidence levels:

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-1.5070	0.2254	44.7056	<.0001
Intercept	2	1	3.9344	0.2708	211.0697	<.0001
cumulativeFS	1	1	-0.0118	0.1329	0.0079	0.9292
cumulativeFS	2	1	0.4628	0.2372	3.8066	0.0511
SL_EthnicityCG	1	1	-0.5552	0.1267	19.2108	<.0001
SL_MomMaritalStatus	1	1	0.1014	0.1289	0.6191	0.4314
SL_Poverty2013s	1	1	0.0566	0.2064	0.0751	0.7841
SL_Poverty2013s	2	1	0.1309	0.2169	0.3644	0.5461

Confidence intervals for differences between odds ratios:

Odds Ratio Estimates and Wald Confidence Intervals			
Odds Ratio	Estimate	95% Confidence Limits	
cumulativeFS 1 vs 2	0.622	0.403	0.961
cumulativeFS 1 vs 3	0.988	0.762	1.282
cumulativeFS 2 vs 3	1.588	0.998	2.529

c. **Research Question 3**

Model information:

The LOGISTIC Procedure

Model Information	
Data Set	WORK.WIC_PRUNED
Response Variable	diet_score
Number of Response Levels	3
Model	cumulative logit
Optimization Technique	Fisher's scoring

How model variables are encoded:

Class Level Information			
Class	Value	Design Variables	
RSL_FoodSecurity24m	1	1	0
	2	0	1
	3	0	0
SL_EthnicityCG	1	1	
	2	0	
SL_MomMaritalStatus	1	1	
	2	0	
SL_Poverty2013s	1	1	0
	2	0	1
	3	0	0

P-values for each predictor:

Joint Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
RSL_FoodSecurity24m	2	9.5764	0.0083
RSL_Child*RSL_FoodSe	2	4.3718	0.1124
pressurin*RSL_FoodSe	2	2.8446	0.2412
restricti*RSL_FoodSe	2	2.4378	0.2956
SL_EthnicityCG	1	15.0874	0.0001
SL_MomMaritalStatus	1	0.1623	0.6871
SL_Poverty2013s	2	0.6288	0.7302

Model parameter estimates and confidence levels:

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-1.4718	0.2591	32.2650	<.0001
Intercept	2	1	3.9969	0.3006	176.7480	<.0001
RSL_FoodSecurity24m	1	1	-1.0175	0.3322	9.3818	0.0022
RSL_FoodSecurity24m	2	1	-0.0321	0.5961	0.0029	0.9570
RSL_Child*RSL_FoodSe	1	1	0.3140	0.1503	4.3651	0.0367
RSL_Child*RSL_FoodSe	2	1	-0.0143	0.3054	0.0022	0.9626
pressurin*RSL_FoodSe	1	1	0.1098	0.0661	2.7586	0.0967
pressurin*RSL_FoodSe	2	1	-0.0369	0.1277	0.0834	0.7727
restricti*RSL_FoodSe	1	1	0.0715	0.0520	1.8924	0.1689
restricti*RSL_FoodSe	2	1	0.0775	0.1051	0.5433	0.4611
SL_EthnicityCG	1	1	-0.5051	0.1300	15.0874	0.0001
SL_MomMaritalStatus	1	1	0.0524	0.1300	0.1623	0.6871
SL_Poverty2013s	1	1	0.1116	0.2082	0.2875	0.5918
SL_Poverty2013s	2	1	0.1696	0.2178	0.6068	0.4360