AJAE appendix for "Food Values"

Authors: Jayson L. Lusk and Brian C. Briggeman

February 22, 2008

Note: The material contained herein is supplementary to the article named in the title and published in the *American Journal of Agricultural Economics* (AJAE).

Appendix: Count-Based Analysis

Description of Counting Approach

Although the data analytic approach described in the main body of the paper is appealing from a conceptual standpoint, it can be difficult to implement. Fortunately, there is a much simpler and straightforward (though somewhat less conceptually appealing) approach to determine the relative importance of food values. In particular an interval scale of importance can be constructed, given the aforementioned experimental design, simply by counting the number of times a person chooses a particular value as most important and subtracting it from the number of times a person chooses the value as least important across the 12 choice sets.

Because each value appears six times across the 12 sets, the highest possible score for any value is +6, and the lowest possible score is -6. Due to the construction of the experimental design, the measured importance scores across all 11 values must sum to zero. In essence this implies that the measured importance scores for a value can be interpreted as the importance of that value relative to the mean level of importance (i.e., importance measures are effects-coded). If one wants to translate the value measures to dummy-variable coding, one simply has to set one of the values as the base and subtract the importance measurement for this base value from all other 10 values. Such count measures of importance are easy to calculate and provide individual-specific measures of importance for each value.

Individual-level estimates provided by this counting procedure can be further analyzed to yield fruitful insight. In particular identifying consumer segments is of particular relevance in marketing applications. Identifying groups of consumers with similar food values can better inform advertising, pricing, product development, etc. The standard approach taken in empirical market segmentation literature is to use standard hierarchical and non-hierarchical clustering algorithms to identify groups of observations that share common characteristics within a given data set. A common criticism of these standard clustering approaches, such as k-means clustering, is that they are not based on an underlying statistical model. To circumvent this criticism, we employ a latent-class approach, which involves maximization of a likelihood function. An advantage of the likelihood function approach relative to standard clustering techniques is that the optimal number of classes can be determined through standard statistical tests. Let c_{ij} represent individual i's measure of importance for value j (as determined by the simple counting procedure described above). For K latent classes, the probability density function for a sequence of J values can be written as follows:

(A1) $$f(\lambda_i) = \sum_{k=1}^{K} P(k) \prod_{j=1}^{J} \phi\left(\frac{c_{ij} - \bar{c}_{jk}}{\sigma_{jk}^2}\right) \qquad ,$$

where ϕ is the standard normal pdf; $\bar{c}_{jk}$ is the mean level of importance for value j in cluster k; σ_{jk} is the cluster- and value-specific standard deviation; and $P(k)$ is the probability of belonging to cluster k, which we parameterize as $P(k) = \frac{e^{\alpha_k}}{\sum_{m=1}^{K} e^{\alpha_m}}$. The parameters in equation (2), α_k, $\bar{c}_{jk}$, and σ_{jk}, are estimated by maximizing the log value of (A1) summed across the N individuals in the sample.

Count-Based Results

Table A1 shows the relative importance of each of the 11 food values as estimated by the counting method. Overall, results are similar to the MNL and RPL shown in the main body of the paper. The last columns in table A1 show the effects-coded importance of each of the food values measured relative to the mean level of importance. These effects-coded count estimates imply, for example, that on average across respondents the difference in the number of times that food safety was picked as the most important food value and the number of times that food safety was picked as least important out of the 6 total times the value appeared in the survey was 2.824. This implies that respondents picked food safety as the most important food safety issue many more times than they picked it as the least important food safety issue.

Table A2 reports the correlation between food values, as determined by the count estimates. The count estimates reveal that none of the value estimates exhibit correlations above 0.5, indicating that each of the values represents a unique construct. Safety and taste exhibit negative correlations, meaning people who believe safety is important are less likely to believe taste is important.

Table A3 reports Pearson and Spearman-rank correlation coefficients between each person's measured food value and the person's willingness-to-pay premium for organic bread.

Table A4 shows how the count-estimated food values differ across purchasers and non-purchasers of organic food.

Results of Cluster Analysis on Count-Based Value Estimates

We sought to identify whether people could be grouped into similar segments or clusters in terms of the importance placed on food values. To investigate this issue, equation (A1) was used to classify people into different groups. To determine the optimal number of clusters or latent classes, the Bayesian information criterion (BIC) was used. The BIC for 1, 2, 3, 4, 5, and 6 latent classes was 8,477.0, 8,259.1, 8,214.8, 8,207.9, 8,226.8 and 8,246.2, respectively. Therefore, the 4-class model was deemed optimal as it yielded the lowest BIC measure.

The top portion of table A5 presents the results of the latent class cluster analysis, including the estimated size of each class and associated means and standard errors. The first cluster is the largest (representing 40% of the sample) and represents a group of people who find personal or self-centered values to be most important. Mean values for taste (2.27), safety (2.86), price (2.90) and nutrition (2.23) are all positive and significant, whereas the origin of the consumers' food comes (origin = -2.58), distribution of profits across the supply chain (fairness = -3.03) and impact to the environment (-2.18) are relatively unimportant to this group.

The second cluster represents about 24% of the sample, and relative to the other groups, this group finds social or society-centered values to be more important. Cluster 2 represents a group of consumers who, compared to the other clusters, is generally more benevolent and more concerned about food attributes that benefit other participants in the value chain. Across all of the clusters, this group places the highest level of importance on origin, fairness and natural and places the second highest level of importance on the environment.

The third cluster, representing about 20% of the sample, believes food safety (5.91) is of utmost importance. It is also interesting to note that the importance of preserving traditional consumption patterns for cluster 3 is the lowest of all clusters. Potentially, these individuals are willing to break traditional consumption patterns to ensure food safety. Convenience and appearance are not important to this group (-1.54 and -1.77, respectively).

The last cluster of people find taste (3.89) and price (2.12) to be of primary importance. Cluster 4 is similar to cluster 1; however, taste and price more clearly dominate the other values for cluster 4. For example, cluster 4 does not value nutrition (0.84 vs. 2.22) or safety (0.33 vs. 2.86) as highly as cluster 1. Because taste and price are more valuable to consumers in cluster 4, they are potentially willing to forgo nutrition and safety for those two attributes. The bottom line for cluster 4 is that these consumers want a food product that tastes good and is low-priced. It is interesting that this cluster is the smallest cluster identified since many promotional campaigns for food products focus on these two attributes.

The bottom portion of table A5 reports means of several demographic variables and the questions related to stated and revealed preferences for organic food by cluster membership (note: cluster membership was determined by assigning people to the cluster for which they had the largest posterior probability of belonging). Significance tests reveal no significant differences in gender, age, education or income across cluster membership. However, cluster membership was significantly influenced by location and the presence of children in the household. People in cluster 4 were much more likely to have children in the household than the other clusters (i.e., 34% of cluster 4 members had children in the home as compared to 11%-17% for the other clusters). It is perhaps not too surprising that households with children would place price and taste of higher importance than households without children. There were also significant differences in cluster membership across the U.S.

The last two rows of table A5 demonstrate that stated and revealed preferences for organic food are significantly influenced by cluster. In particular, the cluster 2 (the society-centered group) was willing to pay on average about $0.63 more for organic bread than non-organic bread, whereas this value was only $0.29 for cluster 1 (the self-centered group). Likewise, 82% of society-centered members had previously bought organic food, whereas only 51% of self-centered members had bought organic food.

Table A1. Importance of Food Values Estimated by Counting of Best-Worst Choices

Value	Dummy-Coded[a]	Effects-Coded[a]
Safety	4.580[*b] (0.269)[b] [3.562][d]	2.824
Nutrition	3.926* (0.214) [2.833]	2.170
Taste	3.625* (0.249) [3.301]	1.869
Price	3.756* (0.282) [3.747]	2.000
Natural	1.193* (0.228) [3.025]	-0.563
Convenience	0.938* (0.248) [3.288]	-0.818
Appearance	0.886* (0.253) [3.358]	-0.870
Environment	0.778* (0.226) [3.003]	-0.978
Fairness	0.074 (0.212) [2.816]	-1.682
Tradition	-0.443 (0.292) [3.879]	-2.199
Origin	0.000	-1.756

[a]Dummy-coded count estimates show the importance of each value relative to origin; effects-coded count estimates show the importance of each value relative to the mean level of importance.
[b]One asterisk (*) implies importance of value is statistically different from the value of Origin at p=0.05 level.
[c]Numbers in parentheses ( ) are standard errors of the mean
[d]Numbers in brackets [ ] are standard deviations of importance

Table A2. Pearson Correlations between Food Values (Effects-Coded)

Value	1	2	3	4	5	6	7	8	9	10
Count Estimates										
Safety (1)	1.00									
Nutrition (2)	-0.04	1.00								
Taste (3)	-0.43	-0.36	1.00							
Price (4)	-0.22	-0.15	0.12	1.00						
Natural (5)	-0.05	-0.03	-0.22	-0.32	1.00					
Convenience (6)	-0.22	0.00	0.19	0.27	-0.30	1.00				
Appearance (7)	-0.07	-0.14	0.25	0.06	-0.25	0.23	1.00			
Environment (8)	0.17	0.08	-0.30	-0.27	0.07	-0.41	-0.47	1.00		
Fairness (9)	0.03	0.01	-0.27	-0.25	0.12	-0.34	-0.47	0.33	1.00	
Tradition (10)	-0.29	-0.22	0.40	-0.08	-0.11	0.09	0.34	-0.42	-0.39	1.00
Origin (11)	-0.09	0.03	-0.24	-0.22	0.06	-0.29	-0.32	0.11	0.18	-0.31

Table A3. Correlations between Food Values and Stated Willingness-to-Pay Premium for Organic Bread[a]

Value	Pearson Correlations	Spearman Rank Correlations
Safety	0.067	0.082
Nutrition	0.118	0.172*
Taste	-0.232*[b]	-0.267*
Price	-0.444*	-0.354*
Natural	0.417*	0.404*
Convenience	-0.238*	-0.241*
Appearance	-0.066	-0.128
Environment	0.166*	0.174*
Fairness	0.118	0.090
Tradition	-0.050	-0.065
Origin	0.149*	0.158*

Number of observations = 176

[a]People were assigned a willingness-to-pay for organic bread of $0, $0.25, $0.75, $1.25, $1.75 or $2.25 based on their response to an interval-censored, payment card question.

[b]One asterisk (*) implies the correlation between willingness-to-pay a premium for organic bread and the food value is significantly different from zero at p=0.05 level or lower.

Table A4. Revealed Preferences for Organic Food and Effects-Coded Mean Food Values

Value	Count Estimates: Previously Purchased Organic	Count Estimates: Have Not Previously Purchased Organic
Safety	2.911	2.672
Nutrition	2.348	1.860
Taste	1.661*[a]	2.234*
Price	1.642*	2.625*
Natural	-0.134*	-1.313*
Convenience	-1.080*	-0.359*
Appearance	-1.133	-0.406*
Environment	-0.714*	-1.438*
Fairness	-1.625	-1.781
Tradition	-2.401	-1.843
Origin	-1.437*	-2.250*
Number of Obs.	112	64

Number of observations = 176

[a]One asterisk (*) implies the hypothesis that the mean values are the same for people who have and who have not previously purchased organic is rejected at the p=0.05 level of significance or lower according to a two-tailed t-test.

Table A5. Latent Class Cluster Analysis Based on Effects-Coded Count Estimates

	Cluster			
	1	**2**	**3**	**4**
	Self-Centered	**Society-Centered**	**Safety-Centered**	**Taste and Price**
Mean Importance of Food Value by Cluster				
Safety	2.857 (0.309)[a]	1.977 (0.343)	5.910 (0.080)	0.332 (0.143)
Nutrition	2.227 (0.241)	2.817 (0.282)	2.386 (0.155)	0.835 (0.185)
Taste	2.271 (0.183)	0.955 (0.244)	0.45 (0.184)	3.886 (0.363)
Price	2.902 (0.260)	1.065 (0.494)	1.172 (0.218)	2.116 (0.598)
Natural	-1.052 (0.148)	0.535 (0.436)	-0.747 (0.261)	-0.737 (0.537)
Convenience	-0.091 (0.153)	-2.044 (0.249)	-1.543 (0.249)	0.031 (0.429)
Appearance	0.092 (0.095)	-2.309 (0.236)	-1.766 (0.294)	-0.083 (0.454)
Environment	-2.184 (0.239)	0.607 (0.219)	0.623 (0.21)	-2.201 (0.386)
Fairness	-3.032 (0.234)	0.006 (0.232)	-0.695 (0.279)	-1.978 (0.305)
Tradition	-1.406 (0.249)	-3.618 (0.300)	-4.033 (0.335)	0.080 (0.470)
Origin	-2.585 (0.218)	0.009 (0.396)	-1.757 (0.382)	-2.281 (0.410)
Cluster Size	40.3%	23.9%	19.4%	16.4%
Mean of Demographics and Willingness-to-Pay by Cluster				
Gender	0.361	0.450	0.364	0.172
Age	54.887	56.667	58.545	52.793
Degree	0.699	0.500	0.576	0.286
Child[b]	0.110	0.175	0.121	0.345
Income	75.101	70.151	72.424	78.280
West[c]	0.219	0.184	0.182	0.200
Midwest[c]	0.384	0.211	0.273	0.467
South[c]	0.301	0.368	0.182	0.233
Northeast[c]	0.096	0.237	0.363	0.100
WTP-organic[d]	0.290	0.626	0.444	0.432
Pur-organic[e]	0.507	0.825	0.727	0.600

Number of Observations = 176; Log Likelihood at maximum=-3,868.71; R^2=90.97%

[a]Numbers in parentheses ( ) are standard errors of the mean

[b]The null hypothesis that the households with children classification is independent of cluster membership is rejected at the p=0.03 level of significance according to a Chi-square test of independence.

[c]The null hypothesis that location is independent of cluster membership is rejected at the p=0.04 level of significance according to a Chi-square test of independence.

[d]The null hypothesis that mean-stated willingness-to-pay a premium for organic bread does not differ by cluster membership is rejected at the p=0.02 level of significance according to an ANOVA test.

[e]The null hypothesis that percentage of people who have previously bought organic food is independent of cluster membership is rejected at the p=0.001 level of significance according to a Chi-square test of independence.