Online Appendix for "Modeling Indivisible Demand" : An Additional Empirical Study

1. Data

Six popular flavors[1] of 6-ounce Yoplait yogurt from the IRI household panel dataset described in Bronnenberg et al. (2008) are used to illustrate our methodology. The data contain 109 households, each with more than 5 shopping trips on which at least one of the six items is purchased. The data contain the trips where multiple as well as a single varieties are purchased, and the trips where none of the inside goods is purchased. Conditional on the purchase occasion, an average of 3.58 units of 6-ounce yogurt are purchased in a single shopping trip, which comprises an average of 1.81 different flavors purchased. Approximately 92% of the observed data are zeros, indicating great potential of obtaining biased estimates from assuming a continuous likelihood. In our analysis, we use 90% of the data for the model calibration, leaving 10% for prediction.

Table 1: Data Description

Flavor	Blueberry	Strawberry	Vanilla	Raspberry	Key Lime	Peach
Unit Price	0.72	0.72	0.72	0.72	0.72	0.72
Purchase Incidence	541	455	420	474	383	393
Purchase Quantity	1074	774	1025	840	813	739
Zero	5046	5132	5167	5113	5204	5194
One	243	240	157	274	170	183
Two	192	151	133	127	122	137
Three	41	33	34	25	36	27
Four and Above	65	31	96	48	55	46

* Total number of shopping trips included in the data is 5587.

2. Estimates

We estimate the model parameters both by our discrete likelihood and by the continuous likelihood. Heterogeneity across households is incorporated by a random-effect specification for household parameters:

$$\theta_h = \{\alpha^*_{1h}, \alpha^*_{2h}, \alpha^*_{3h}, \alpha^*_{4h}, \alpha^*_{5h}, \alpha^*_{6h}, \gamma^*_h\} \sim Normal\left(\bar{\theta}, V_\theta\right)$$

[1]Blueberry, Strawberry, Vanilla, Raspberry, Key Lime, Peach

where $\alpha_{ih}^* = \log(\alpha_{ih}), \gamma_h^* = \log(\gamma_h)^2$. A Bayesian MCMC method is used for estimation with a conjugate but relatively diffuse prior distribution for the hyper-parameters[3]. Out of 100,000 iterations of the MCMC chain, we discarded the first 30,000 draws as burn-in, and used the remaining 70,000 iterations to summarize the posterior distributions. Table 2 displays the estimation results, comparing the two methods.

Table 2: Estimation Results

	Model-Fit		Estimates						
	In-sample	Predictive	Blueberry	Strawberry	Vanilla	Raspberry	Key Lime	Peach	Satiation
Disc.	-8691	-1237	-3.38	-3.34	-3.76	-3.33	-3.69	-3.50	-1.66
($s_i = 6$)	[-8691]	[-1237]	(0.12)	(0.10)	(0.14)	(0.10)	(0.13)	(0.11)	(0.11)
Disc.	-11051	-1527	-3.73	-3.69	-4.09	-3.68	-4.03	-3.84	-2.42
($s_i = 3$)	[-8935]	[-1251]	(0.11)	(0.09)	(0.14)	(0.09)	(0.12)	(0.10)	(0.07)
Disc.	-12168	-1668	-3.77	-3.74	-4.13	-3.72	-4.07	-3.89	-2.54
($s_i = 2$)	[-9019]	[-1259]	(0.11)	(0.09)	(0.13)	(0.08)	(0.12)	(0.10)	(0.06)
Disc.	-13948	-1894	-3.81	-3.78	-4.19	-3.76	-4.11	-3.94	-2.65
($s_i = 1$)	[-9107]	[-1268]	(0.10)	(0.09)	(0.14)	(0.08)	(0.12)	(0.10)	(0.06)
Cont.	-9843	-1370	-3.85	-3.81	-4.23	-3.80	-4.15	-3.98	-2.74
	[-9192]	[-1278]	(0.10)	(0.09)	(0.14)	(0.08)	(0.12)	(0.10)	(0.06)

* Posterior standard deviations are given in parentheses ( ).
* True log-likelihood values are given in parentheses [ ].

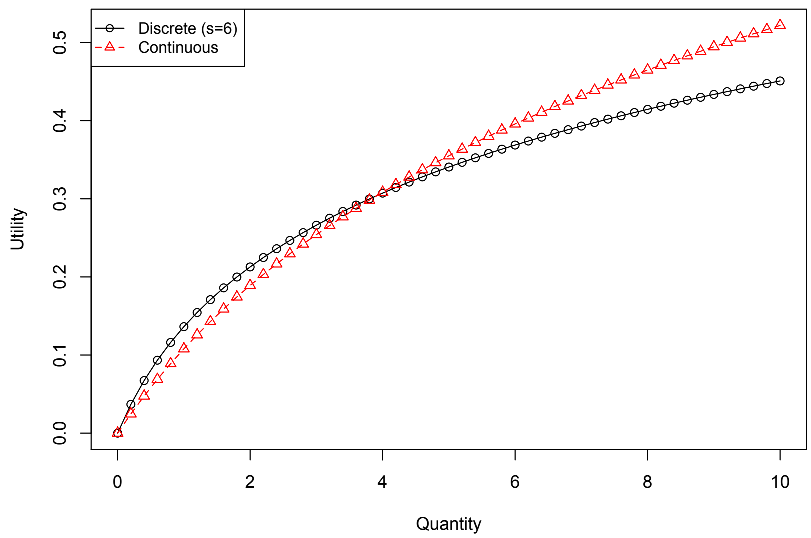
The discrete likelihood fits the in-sample data better and provides better hold-out predictions. Both the baseline and satiation parameters are underestimated with the continuous likelihood, and the magnitudes of the biases are statistically significant. As we vary the unit grid of the discrete likelihood from the true package size (i.e. 6-ounce) to smaller sizes (i.e. 3, 2, 1-ounce), the parameter estimates gradually decrease, converging toward those of the continuous likelihood. The estimates of the continuous likelihood are still quite different from those of the discrete likelihood in which we assume the smallest available package size (i.e., one). This illustrates that assuming continuously available demand can be costly when zero datum values are common.

Figure 1 depicts the estimated sub-utility for the blueberry-flavor yogurt from the discrete approach and compares it with one estimated from the continuous likelihood. Comparison between the two sub-utilities show two patterns of biases that stem from the

[2] γ_h can be alternative-specific as is in our simulation study. However, the γ_{ih}'s were not differentially estimated with our empirical dataset.

[3] $\bar{\theta}|V_\theta \sim N(0, V_\theta \otimes 100I), V_\theta \sim IW(10, 10I)$

Figure 1: Sub-utility for Blueberry Yogurt (Discrete vs. Continuous)



ignorance of the data indivisibility. First, the baseline utility is underestimated in the continuous approach to account for the zeros in the data, some of which are due to indivisible demand(i.e., don't like it enough), and others due to lack of preference (i.e., don't like it, period). The slope of the sub-utility at zero quantity is steeper in the discrete approach, implying that an observation of non-purchase in data represents lower-level of preference in the continuous approach. This is consistent with our theoretical prediction. Second, the degree of satiation is underestimated when the discreteness of the data is not accounted for. This can be viewed as a bi-product of the bias of the baseline utility. As the continuous approach underestimates the baseline parameter (i.e., α_i) to rationalize zeros, a smaller γ_i is required to increase the log-likelihood for the positive quantities. In other words, because the continuous approach infers lower-level of preference from non-purchase, it rationalizes the positive quantities by attributing to a lower degree of satiation.

3. Decomposition of Corner Probability

Non-purchase decisions for packaged goods can arise from two different reasons: (i) a consumer does not like it, (ii) a consumer likes it but not as much as to buy one unit of it.

Based on the full posterior distribution of individual-level parameters from the 6-ounce discrete model, we compute the probability of non-purchase (i.e. corner probability) with different package sizes[4], and decompose it into the one due to the lack of preference, and the other reflecting the restriction of demand indivisibility.

Figure 2: Impact of Package Size on Corner Probability for Blueberry Yogurt

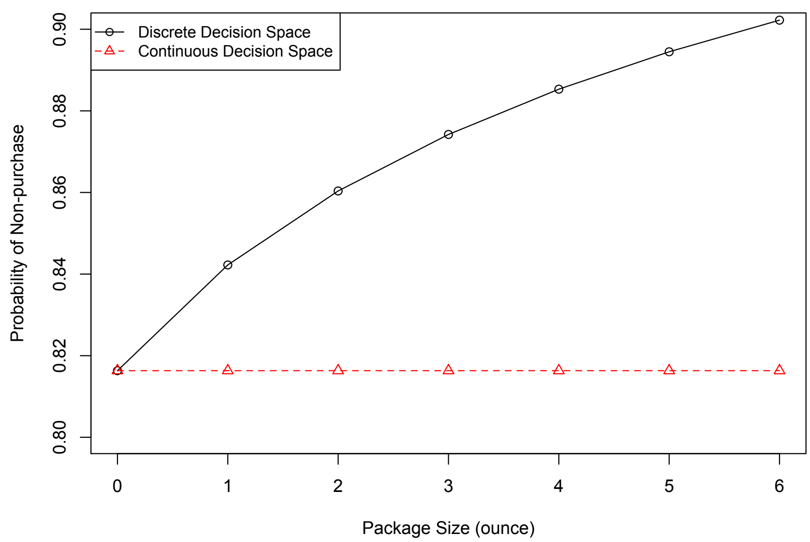


Figure 2 shows how the probability of non-purchase changes with different package sizes of the blueberry-flavor yogurt. When the true package size (=6 ounce) is assumed, the probability of a corner is 0.902, which is consistent with the frequencies of zeros in the data. While the majority of non-purchase is driven by the lack of preference (i.e. the corner probability with a continuous decision space is 0.816), there exists a significant portion of non-purchase (9.52%) that is caused by the discreteness restriction in a consumer's decision space, which is represented by the gap between the two lines in Figure 2. As the package size decreases, the corner probability also decreases and converges to the case of a continuous decision space, which reflects that the discreteness restriction in a consumer's decision space is reduced with a smaller package size.

[4] Per-volume prices are assumed to remain constant.

References

Bronnenberg, Bart J., Michael W. Kruger, Carl F. Mela. 2008. Database paper: The iri marketing data set. *Marketing Science* **27**(4) pp. 745–748.