Submitted to *Manufacturing & Service Operations Management*
manuscript MSOM-12-270

Stockout-Based Substitution and Inventory Planning in Textbook Retailing

---

Online Supplement

This online supplement presents robustness tests and extensions of our choice model. Appendix A addresses the implications of the Independence of Irrelevant Alternatives (IIA) property in our choice model, the sensitivity of estimation error to the size of the data set, and the effect of product returns on estimation error. Appendix B presents likelihood functions for the benchmark models defined in Section 4.2. Appendix C shows the extension of our likelihood function to stockout-based substitution among more than two products.

Appendix A: Robustness Analysis of the Choice Model

We present three types of robustness analyses of our model and results. In A.1, we estimate a random-coefficients MNL model on the Cornell bookstore dataset, and compare its results with the fixed-coefficients MNL model used in the main text of the paper. The former model does not suffer from the restrictions of the IIA property. In A.2, we examine the sensitivity of the simulation results from Section 4.3 to the size of the data set by re-estimating the simulation for datasets of varying sizes. In A.3, we modify the simulation to assess the effect of unobserved product returns on the estimation error of the choice model.

A.1. Implications of IIA Property for the Estimation of the Choice Model

In the main text of the paper, we use a fixed-coefficients MNL model in which the consumer chooses from three choices—buying a new book, buying a used book, or not buying any book. Since this model suffers from the IIA property, in this section, we present the results for a random-coefficients MNL model with a random coefficient for the intercept to account for a flexible substitution pattern with the outside good. The utility of consumer c buying book type k for ISBN i is given by

$$Utility_{cik} = \beta_k' x_{ik} + \varepsilon_{cik} + \sigma_k \epsilon_{cik},$$

where ϵ_{cik} is an additional error term with a standard normal distribution, σ_k denotes the standard deviation of the random intercept, and the remaining terms and indices are as defined in Section 3. This specification is equivalent to the specification in which the intercept term is a random coefficient following a normal distribution with mean β_{k1} and standard deviation σ_k. The additional term introduces correlation in utility over types, and hence, allows for flexible substitution patterns even when the additional error terms are independent (Train 2009). Thus, it mitigates the IIA property and accompanying restrictive substitution patterns implied by the fixed-coefficients MNL model.

The computation of the likelihood function for the random-coefficients MNL model under stockouts poses an additional challenge. For a given ISBN i, the probability that consumer c chooses product type k from an assortment $\mathcal{K}_{i1}$ is given by

$$P_{cik}(\beta,\sigma|\mathcal{K}_{i1},X_i)=\int\left(\frac{e^{\beta_k' x_{ik}+\sigma_k\epsilon_{cik}}}{1+\sum_{l\in\mathcal{K}_{i1}} e^{\beta_l' x_{il}+\sigma_k\epsilon_{cil}}}\right)f(\epsilon_{ci})d\epsilon_{ci},$$

where each component ϵ_{cil} for $l\in\mathcal{K}_{i1}$ follows a standard normal distribution. Calculating the choice probabilities requires a numerical integration. Thus, we compute the probabilities as:

$$\breve{P}_{cik}(\beta,\sigma|\mathcal{K}_{i1},X_i)=\frac{1}{R}\sum_{r=1}^{R}\frac{e^{\beta_k' x_{ik}+\sigma_k\epsilon_{cik}^r}}{1+\sum_{l\in\mathcal{K}_{i1}} e^{\beta_l' x_{il}+\sigma_k\epsilon_{cil}^r}},$$

where ϵ_{cil}^r for $l\in\mathcal{K}_{i1}$ refers to the r-th value drawn from a standard normal distribution and $R(=10,000)$ represents the number of draws. A numerical search for the maximum of the likelihood function is possible because of the following features of the estimated probabilities: (i) $\breve{P}_{cik}$ is an unbiased estimator of P_{cik}, (ii) $\ln\breve{P}_{cik}$ is defined, and (iii) $\breve{P}_{cik}$ is twice differentiable in the parameters (Train 2009). Thus, the simulated log likelihood function can be calculated using $\breve{P}_{cik}$ in the log likelihood function (3). Similar to (4), we can then find the estimates of β and σ that maximize the simulated log likelihood function (Train 2009).

Table A.1 presents the parameters' estimates for the above model. We observe that the coefficients' estimates for the random-coefficients MNL model are statistically significant and similar to the estimates obtained for the MNL model used in the paper. We also assess if the random-coefficients MNL model has a significantly better fit than the fixed-coefficients MNL model. The latter can be considered as a restricted model of the former with two fewer degrees of freedom for the σ_k terms, so a likelihood ratio test with two degrees of freedom is applicable. Our estimates yield $2\ln(L_0/L_1)=2(40577-40519)=58$ with $p<0.001$. Thus, the random-coefficients MNL has a better statistical fit than the fixed-coefficients MNL. Finally, Table A.2 shows the average purchase probabilities for new and used books under various stocking situations across all observations in our data set, and compares them with the corresponding averages for the fixed-coefficients MNL model. We observe that there are no significant differences in the purchase probabilities estimated from the two models. These results suggest that the random-coefficients MNL model can serve as an alternative model in our context. Due to the closeness of results between the two model specifications, we use the simpler MNL model in the main text of the paper.

A.2. Variation in Parameters' Estimates with the Size of the Data Set

In Section 4.4, we concluded that the higher the heterogeneity in product attributes in the data set used to estimate the choice model, the lesser is the estimation error. Our finding may be sensitive to the number of book titles in the data set. To examine this sensitivity, we conduct additional numerical experiments with

Table A.1 Estimated parameters of the choice model for the Cornell Store data set for the period 2007-2011 using a Random-Coefficient Multinomial Logit specification

		New books		Used books	
Attributes		Estimates	Standard errors	Estimates	Standard errors
Intercept	*Intercept*	-2.002***	0.007	-0.128***	0.015
Standard deviation of random coefficient		-0.002	0.007	-0.00005	0.005
New-book price	*NP*	-0.288***	0.006	-0.132***	0.007
Number of courses	*NC*	-0.135***	0.004	-0.137***	0.009
Average number of books per course	*NB*	-0.019***	0.003	0.005***	0.002
Proportion required	*PR*	1.329***	0.008	0.099***	0.006
New ISBN	*NI*	0.431***	0.008	0.133***	0.007
Course level 1	*CL1*	0.614***	0.017	0.715***	0.011
Course level 2	*CL2*	0.557***	0.009	0.406***	0.012
Course level 3	*CL3*	0.207***	0	0.382***	0
Course level 4		-	-	-	-
Agriculture	*AGR*	-1.067***	0.016	-0.728***	0.014
Architecture	*ARC*	-1.395***	0.035	-1.07***	0.023
Art and Science	*AAS*	-1.405***	0.011	-0.814***	0.022
Engineering	*ENG*	-0.945***	0.018	-0.938***	0.028
Hotel Administration	*HAD*	-1.132***	0.027	-0.711***	0.047
Human Ecology	*HEC*	-0.781***	0.016	-0.577***	0.025
Industrial and Labor Relations	*ILR*	-1.272***	0	-0.168***	0
Law		-	-	-	-

, * statistically significant at $p<0.05$ and $p<0.01$, respectively, for two tailed tests.

Table A.2 Comparison of average purchase probabilities using the Random coefficient and the Multinomial Logit specification under various stocking situations

		Purchase Probability of New Books		Purchase Probability of Used Books	
		New and Used books are stocked	Only New books are stocked	New and Used books are stocked	Only Used books are stocked
Random Coefficient Model	Mean	7.54%	11.22%	31.09%	33.77%
	Std Dev	3.31%	5.38%	7.23%	8.24%
Multinomial Logit Models	Mean	8.53%	12.81%	31.92%	35.03%
	Std Dev	3.33%	5.48%	7.32%	8.38%

fewer titles (i.e. 100 and 300 titles as opposed to 1000 titles used in Table 4 of the manuscript) to estimate the model parameters. The results of these experiments are presented in Tables A.3 and A.4 below.

First consider the results from a 100-titles data set. We observe that, at each stocking level, as the coefficient of variation of new book prices increases, the standard errors of the estimates decline quickly and the estimates of the parameters get closer to the true values. Further, for each value of the coefficient of variation of new book prices, as the stocking level increases, the standard errors of estimates again decline and the parameters' estimates become more accurate. The same inference is obtained from the results from the 300-titles data set. Thus, the inferences shown in the paper about the effects of stocking level and heterogeneity of product attributes are supported for smaller estimation samples as well.

It is also instructive to compare the standard errors across Tables 4, A.3, and A.4. We observe that as the size of the estimation sample increases, the standard errors of the estimates decline. In particular, the standard errors for a 1000-titles sample are about one-fourth and one-half, respectively, of the standard errors for the 100-titles and the 300-titles samples. This shows that larger data sets are beneficial, but even

Table A.3 The impact of heterogeneity in new-book price NP on the estimation error at different stocking levels (Results estimated with 100 titles)

			Estimates						Standard error in estimates					
			New books			Used books			New books			Used books		
Stocking level	Stockout rate	Coefficient of Variation of New-book price	Intercept	New-book price	Course level	Intercept	New-book price	Course level	Intercept	New-book price	Course level	Intercept	New-book price	Course level
0.5	95.0%	0.25	-3.57	2.09	0.43	2.70	0.15	-3.99	3.800	3.790	0.368	4.414	4.387	0.346
	95.0%	0.50	0.09	-1.52	0.38	7.07	-4.29	-3.42	1.475	1.413	0.360	1.622	1.582	0.396
	95.0%	0.75	-1.22	-0.27	0.40	4.51	-1.72	-3.85	1.060	1.081	0.362	1.140	1.183	0.384
	96.0%	1.00	-1.08	-0.31	0.30	5.54	-2.03	-4.52	1.025	1.027	0.363	1.074	1.068	0.354
	97.0%	1.25	-1.20	-0.20	5.22	3.54	0.00	-4.69	0.800	0.795	0.543	0.920	0.909	0.351
1	62.0%	0.25	-2.10	0.65	0.76	-1.61	0.49	0.33	0.959	0.955	0.101	0.829	0.823	0.093
	61.0%	0.50	-0.90	-0.55	0.77	-1.04	-0.09	0.41	0.412	0.412	0.101	0.374	0.377	0.096
	62.0%	0.75	-1.28	-0.16	0.82	-1.24	0.13	0.37	0.298	0.295	0.102	0.263	0.263	0.095
	59.0%	1.00	-1.12	-0.32	0.77	-0.30	-0.78	0.30	0.253	0.250	0.100	0.247	0.241	0.093
	61.0%	1.25	-1.38	-0.05	0.79	-0.84	-0.26	0.37	0.216	0.205	0.102	0.202	0.194	0.094
1.5	14.0%	0.25	-1.87	0.45	0.70	-1.52	0.41	0.42	0.831	0.826	0.087	0.747	0.742	0.084
	14.0%	0.50	-0.88	-0.53	0.71	-1.10	-0.01	0.41	0.364	0.365	0.087	0.338	0.339	0.084
	13.0%	0.75	-1.20	-0.21	0.72	-1.04	-0.06	0.40	0.254	0.253	0.087	0.232	0.231	0.085
	13.0%	1.00	-0.96	-0.45	0.71	-0.67	-0.43	0.37	0.226	0.222	0.086	0.208	0.205	0.084
	14.0%	1.25	-1.23	-0.17	0.71	-0.90	-0.19	0.40	0.188	0.180	0.087	0.174	0.167	0.084
2	2.0%	0.25	-1.68	0.26	0.70	-1.46	0.37	0.41	0.822	0.817	0.086	0.745	0.740	0.084
	2.0%	0.50	-0.79	-0.62	0.70	-1.08	-0.01	0.40	0.362	0.363	0.087	0.335	0.336	0.084
	2.0%	0.75	-1.23	-0.18	0.72	-0.96	-0.13	0.38	0.251	0.251	0.087	0.231	0.230	0.084
	2.0%	1.00	-0.99	-0.42	0.70	-0.68	-0.41	0.36	0.224	0.221	0.086	0.207	0.204	0.084
	2.0%	1.25	-1.25	-0.17	0.71	-0.85	-0.22	0.38	0.187	0.179	0.086	0.173	0.166	0.084

Table A.4 The impact of heterogeneity in new-book price NP on the estimation error at different stocking levels (Results estimated with 300 titles)

			Estimates						Standard error in estimates					
			New books			Used books			New books			Used books		
Stocking level	Stockout rate	Coefficient of Variation of New-book price	Intercept	New-book price	Course level	Intercept	New-book price	Course level	Intercept	New-book price	Course level	Intercept	New-book price	Course level
0.5	96.7%	0.25	-3.63	2.20	0.74	-4.02	2.88	0.31	1.808	1.816	0.265	1.586	1.595	0.240
	96.7%	0.50	-0.26	-1.18	0.75	-0.45	-0.73	0.27	0.913	0.899	0.264	1.065	1.056	0.238
	96.7%	0.75	-1.00	-0.44	0.71	-1.65	0.47	0.28	0.620	0.601	0.268	0.561	0.559	0.235
	97.0%	1.00	-1.85	0.46	0.70	-1.45	0.28	0.23	0.465	0.459	0.263	0.444	0.453	0.249
	96.7%	1.25	-1.12	-0.34	0.73	-1.35	0.29	0.17	0.400	0.385	0.267	0.408	0.394	0.241
1	60.0%	0.25	-1.09	-0.36	0.81	-1.74	0.63	0.47	0.586	0.583	0.058	0.547	0.546	0.056
	61.0%	0.50	-0.88	-0.57	0.80	-1.03	-0.07	0.47	0.284	0.282	0.058	0.260	0.259	0.056
	60.7%	0.75	-1.25	-0.19	0.80	-1.01	-0.11	0.50	0.161	0.156	0.058	0.149	0.144	0.056
	61.3%	1.00	-1.25	-0.20	0.82	-1.10	-0.12	0.46	0.148	0.143	0.058	0.133	0.129	0.056
	60.3%	1.25	-1.29	-0.15	0.80	-1.05	-0.07	0.45	0.107	0.104	0.057	0.098	0.095	0.055
1.5	15.7%	0.25	-1.28	-0.14	0.75	-0.88	-0.22	0.47	0.520	0.519	0.050	0.474	0.472	0.048
	16.0%	0.50	-1.20	-0.22	0.75	-0.98	-0.12	0.46	0.245	0.244	0.050	0.226	0.225	0.048
	16.0%	0.75	-1.21	-0.20	0.76	-0.86	-0.25	0.47	0.143	0.139	0.050	0.131	0.127	0.048
	16.0%	1.00	-1.24	-0.17	0.76	-1.07	-0.14	0.45	0.128	0.124	0.050	0.117	0.113	0.048
	16.0%	1.25	-1.23	-0.19	0.76	-1.09	-0.11	0.45	0.093	0.090	0.050	0.086	0.084	0.048
2	3.7%	0.25	-1.21	-0.21	0.75	-0.94	-0.17	0.47	0.516	0.514	0.050	0.471	0.470	0.048
	3.0%	0.50	-1.21	-0.21	0.75	-0.96	-0.15	0.47	0.243	0.243	0.050	0.224	0.224	0.048
	3.7%	0.75	-1.21	-0.21	0.75	-0.86	-0.25	0.47	0.143	0.138	0.050	0.131	0.126	0.048
	4.3%	1.00	-1.26	-0.16	0.76	-1.03	-0.11	0.44	0.128	0.123	0.050	0.117	0.113	0.048
	2.7%	1.25	-1.24	-0.17	0.75	-1.10	-0.10	0.45	0.092	0.089	0.049	0.086	0.083	0.048

smaller data sets with higher stocking levels or more heterogeneous products can yield accurate estimates of demand.

A.3. Implications of Product Returns

Product returns at the Cornell Store averaged 10% over 2010-2012. Returns have implications for the specification of the choice process because an accurate specification of returns will require a nested choice model as well as a model for the time duration between the purchase epoch and the return epoch. Returns also affect the sample path-based estimation of parameters because they affect the availability of books for other customers. In contrast, our model presented in the previous sections assumes zero returns. Understandably, incorporating returns in choice specification and estimation will significantly increase computational complexity. However, our analysis reveals that returns do not adversely affect the accuracy of demand estimation.

In this section, we evaluate the impact of returns on the estimation error of our model by modifying the simulation study presented in Section 4. We follow Step 1 of the simulation study as before to generate information on a random set of 10,000 textbooks. Then, we modify Step 2 by incorporating returns in the sales and stockout generation process. Suppose that each arriving consumer who makes a purchase has probability r of returning that book. The bookstore allows customers a grace period of one week within which sold books can be returned. Since the time dimension is captured by the sequence of customer arrivals in our analysis, we operationalize this constraint by assuming that the return must occur before the n-th subsequent customer arrival. We vary n between 0.25 and 0.5 times the enrollment for the course, subject to the constraint that n is less than the total number of remaining students who are yet to arrive. This range is chosen by comparing the grace period with the average length of the peak selling season. We experiment with different probability distributions of the time of return: (i) a uniform distribution in the range $1 - n$; (ii) a triangle distribution with increasing probability of return in the range $1 - n$. The triangle distribution is more conservative as it enables a larger fraction of customers to return books towards the end of the grace period. The returned book is added to the inventory at the bookstore, and immediately becomes available for purchase by the next arriving customer.

We use this modified process to generate sales net of returns for each ISBN in our simulation data set. For this experiment, we vary r between 0.05 and 0.15. Note that this range includes the 10% book return rate at the Cornell Store. The initial stocking levels can take seven values as before, which enables us to examine possible interaction between stockout rate and return rate. Thus, we repeat the simulation experiment described in Section 4 for this data set. Let p_j denote the initial purchase probability of book type j. Then, the effective probability of purchase of that book is $(1 - r)p_j$. We compare this number with the purchase probability estimated from our model to determine estimation error.

Table A.5 shows the estimation error in demand from this analysis. For brevity, we present results only for the triangle distribution of time of return as it is more conservative. Observe that the estimation error in demand does not get worse due to returns. For example, when stocking level is equal to demand and n is 0.25 times enrollment, the value of MPE is -0.24% when return probability is r=10% and the value of MPE is 0.50% when r=0. Surprisingly, in every case, the value of MPE with returns is smaller than that without returns.

We reason that the effect of returns on demand estimation can be decomposed as follows:

1. *Effect on product availability.* A book that is bought and later returned is unavailable for purchase by other customers who arrive during the intervening period. A stockout in the intervening period could lead to a lost sale. The longer the book is kept, the greater the chance of stockout, and consequently, the larger the effect of returns. We expect that our model, by not incorporating returns, will overstate inventory during this period, and thus, underestimate demand. However, we expect the total effect of returns to be small because only those returns that occur after a stockout should affect the assortment of book types available to customers.

2. *Effect on sample paths used in estimation.* When a book is returned on the last day of the selling period, the resulting data for this book will show that this book did not stockout. This condition, when used to generate feasible sample paths, will cause our model to underestimate the occurrence of stockouts, and thus, underestimate demand.

3. *Implications of initial stocking level.* When initial stocking levels are high, returns are less likely to cause stockouts. Thus, they will have relatively less impact on the estimation of demand. When initial stocking levels are low, then returns are more likely to lead to stockouts. However, this effect is compensated by the fact that stockouts will occur earlier so that there is a greater chance that returned books will be repurchased by later arriving customers. Thus, the effect of returns should be mitigated even for low stocking levels.

The net result of these factors is an underestimation of demand. This underestimation compensates for the MPE of the model for the scenario in which there are no product returns ($r = 0$). Therefore, returns do not adversely affect the accuracy of demand estimation. Indeed, a 10% return rate does not introduce substantial estimation error in our model.

Table A.5 The impact of product returns on the MPE in demand at different return rates and stocking levels (for triangular return distribution with mode=n)

	n=0.25*Enrollment				n=0.50*Enrollment			
Stocking	Return Probability				Return Probability			
Level	r=0	r=0.05	r=0.10	r=0.15	r=0	r=0.05	r=0.10	r=0.15
0.5	7.40%	3.18%	4.90%	6.26%	7.40%	0.00%	1.81%	2.76%
0.75	0.80%	0.09%	0.61%	0.17%	0.80%	-0.72%	-0.23%	-0.49%
1	0.50%	-0.08%	-0.24%	-0.31%	0.50%	-0.50%	-0.91%	-1.07%
1.25	0.30%	0.19%	-0.04%	-0.06%	0.30%	0.09%	-0.19%	-0.24%
1.5	0.30%	0.11%	-0.10%	-0.16%	0.30%	0.09%	-0.15%	-0.21%
1.75	0.40%	0.08%	-0.07%	-0.12%	0.40%	0.08%	-0.10%	-0.17%
2	0.30%	0.14%	-0.05%	-0.11%	0.30%	0.12%	-0.07%	-0.14%

Appendix B: Likelihood Functions for Benchmark Models

We enumerate likelihood for textbook i with two product types (new and used books) according to different stockout situations. Subscripts n and u denote new books and used books respectively, and s and E denote the sales and enrollment information respectively. For simplicity of exposition, we suppress β and X_i in the choice probability function. For a given assortment $\mathcal{A}$, let $P_{k,\text{elements}\in\mathcal{A}}$ denote $P_k(\beta_n, \beta_u \mid \mathcal{A}, X_i)$ except in benchmark model 4.

Benchmark model 1: use only uncensored observations

The likelihood function of textbook i is given by

1. $STI_i = 0$: New and used books are stocked and they do not stock out.

$$L_i^1(\beta_n^1, \beta_u^1 \mid STI_i = 0, S_i, E_i, X_i) = \frac{E_i!}{s_{in}! s_{iu}! (E_i - s_{in} - s_{iu})!} P_{n,nu}^{s_{in}} P_{u,nu}^{s_{iu}} P_{0,nu}^{E_i - s_{in} - s_{iu}}.$$

2. $STI_i = 4$: Only new books are stocked and they do not stock out.

$$L_i^1(\beta_n^1, \beta_u^1 \mid STI_i = 4, S_i, E_i, X_i) = \frac{E_i!}{s_{in}!(E_i - s_{in})!} P_{n,n}^{s_{in}} P_{0,n}^{(E_i - s_{in})}.$$

3. $STI_i = 6$: Only used books are stocked and they do not stock out.

$$L_i^1(\beta_n^1, \beta_u^1 \mid STI_i = 6, S_i, E_i, X_i) = \frac{E_i!}{s_{iu}!(E_i - s_{iu})!} P_{u,u}^{s_{iu}} P_{0,u}^{(E_i - s_{iu})}.$$

Maximum likelihood estimates $\hat{\beta}_n^1$ and $\hat{\beta}_u^1$ can be found by

$$(\hat{\beta}_n^1, \hat{\beta}_u^1) := \arg\max_{\beta_n^1, \beta_u^1} \sum_{i: i \in \mathcal{I}, \text{ and } STI_i = 0, 4, \text{ or } 6} \log L_i^1(\beta_n^1, \beta_u^1 \mid STI_i, S_i, E_i, X_i).$$

Benchmark model 2: assume that demand equals sales

The likelihood function of textbook i is given by

1. $STI_i = 0$, 1, 2, or 3: New and used books are stocked.

$$L_i^2(\beta_n^2, \beta_u^2 \mid STI_i = 0, 1, 2, \text{ or } 3, S_i, E_i, X_i) = \frac{E_i!}{s_{in}! s_{iu}! (E_i - s_{in} - s_{iu})!} P_{n,nu}^{s_{in}} P_{u,nu}^{s_{iu}} P_{0,nu}^{E_i - s_{in} - s_{iu}}.$$

2. $STI_i = 4$ or 5: Only new books are stocked.

$$L_i^2(\beta_n^2, \beta_u^2 \mid STI_i = 4 \text{ or } 5, S_i, E_i, X_i) = \frac{E_i!}{s_{in}!(E_i - s_{in})!} P_{n,n}^{s_{in}} P_{0,n}^{(E_i - s_{in})}.$$

3. $STI_i = 6$ or 7: Only used books are stocked.

$$L_i^2(\beta_n^2, \beta_u^2 \mid STI_i = 6 \text{ or } 7, S_i, E_i, X_i) = \frac{E_i!}{s_{iu}!(E_i - s_{iu})!} P_{u,u}^{s_{iu}} P_{0,u}^{(E_i - s_{iu})}.$$

Maximum likelihood estimates $\hat{\beta}_n^2$ and $\hat{\beta}_u^2$ can be found by

$$(\hat{\beta}_n^2, \hat{\beta}_u^2) := \arg\max_{\beta_n^2, \beta_u^2} \sum_{i \in \mathcal{I}} \log L_i^2(\beta_n^2, \beta_u^2 \mid STI_i, S_i, E_i, X_i).$$

Benchmark model 3: ignore substitution but account for stockout

The likelihood function of textbook i is given by

1. $STI_i = 0$: New and used books are stocked and they do not stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 0, S_i, E_i, X_i) = \frac{E_i!}{s_{in}! s_{iu}! (E_i - s_{in} - s_{iu})!} P_{n,nu}^{s_{in}} P_{u,nu}^{s_{iu}} P_{0,nu}^{E_i - s_{in} - s_{iu}}.$$

2. $STI_i = 1$: New and used books are stocked. Only new books stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 1, S_i, E_i, X_i) = \sum_{h_u=0}^{s_{iu}} \sum_{h_0=0}^{E_i - s_{in} - s_{iu}} \frac{(s_{in} - 1 + h_u + h_0)!}{(s_{in} - 1)! h_u! h_0!} P_{n,nu}^{s_{in}} P_{u,nu}^{h_u} P_{0,nu}^{h_0}$$
$$\frac{(E_i - s_{in} - h_u - h_0)!}{(s_{iu} - h_u)!(E_i - s_{in} - s_{iu} - h_0)!} P_{u,nu}^{(s_{iu} - h_u)} (1 - P_{u,nu})^{(E_i - s_{in} - s_{iu} - h_0)}.$$

3. $STI_i = 2$: New and used books are stocked. Only used books stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 2, S_i, E_i, X_i) = \sum_{h_n=0}^{s_{in}} \sum_{h_0=0}^{E_i - s_{in} - s_{iu}} \frac{(s_{iu} - 1 + h_n + h_0)!}{(s_{iu} - 1)! h_n! h_0!} P_{n,nu}^{h_n} P_{u,nu}^{s_{iu}} P_{0,nu}^{h_0} \frac{(E_i - s_{iu} - h_n - h_0)!}{(s_{in} - h_n)!(E_i - s_{in} - s_{iu} - h_0)!} P_{n,nu}^{(s_{in} - h_n)} (1 - P_{n,nu})^{(E_i - s_{in} - s_{iu} - h_0)}.$$

.

4. $STI_i = 3$: New and used books are stocked and both of them stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 3, S_i, E_i, X_i) = \sum_{h_u=0}^{s_{iu}-1} \sum_{h_0=0}^{E_i - s_{in} - s_{iu}} \frac{(s_{in} - 1 + h_u + h_0)!}{(s_{in} - 1)! h_u! h_0!} P_{n,nu}^{s_{in}} P_{u,nu}^{h_u} P_{0,nu}^{h_0} \sum_{h_{00}=0}^{E_i - s_{in} - s_{iu} - h_0} \left[\frac{(s_{iu} - 1 - h_u + h_{00})!}{(s_{iu} - 1 - h_u)! h_{00}!} P_{u,nu}^{(s_{iu} - h_u)} (1 - P_{u,nu})^{h_{00}} \right]$$
$$+ \sum_{h_n=0}^{s_{in}-1} \sum_{h_0=0}^{E_i - s_{in} - s_{iu}} \frac{(s_{iu} - 1 + h_n + h_0)!}{(s_{iu} - 1)! h_n! h_0!} P_{n,nu}^{h_n} P_{u,nu}^{s_{iu}} P_{0,nu}^{h_0} \sum_{h_{00}=0}^{E_i - s_{in} - s_{iu} - h_0} \left[\frac{(s_{in} - 1 - h_n + h_{00})!}{(s_{in} - 1 - h_n)! h_{00}!} P_{n,nu}^{(s_{in} - h_n)} (1 - P_{n,nu})^{h_{00}} \right].$$

5. $STI_i = 4$: Only new books are stocked and they do not stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 4, S_i, E_i, X_i) = \frac{E_i!}{s_{in}!(E_i - s_{in})!} P_{n,n}^{s_{in}} P_{0,n}^{(E_i - s_{in})}.$$

6. $STI_i = 5$: Only new books are stocked but they stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid E_i, STI_i = 5, S_i, E_i, X_i) = \sum_{h_n=0}^{E_i - s_{in}} \frac{(s_{in} - 1 + h_n)!}{(s_{in} - 1)! h_n!} P_{n,n}^{s_{in}} P_{0,n}^{h_n}.$$

7. $STI_i = 6$: Only used books are stocked and they do not stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 6, S_i, E_i, X_i) = \frac{E_i!}{s_{iu}!(E_i - s_{iu})!} P_{u,u}^{s_{iu}} P_{0,u}^{(E_i - s_{iu})}.$$

8. $STI_i = 7$: Only used books are stocked but they stock out.

$$L_i^3(\beta_n^3, \beta_u^3 \mid STI_i = 7, S_i, E_i, X_i) = \sum_{h_u=0}^{E_i - s_{iu}} \frac{(s_{iu} - 1 + h_u)!}{(s_{iu} - 1)! h_u!} P_{u,u}^{s_{iu}} P_{0,u}^{h_u}.$$

Maximum likelihood estimates $\hat{\beta}_n^3$ and $\hat{\beta}_u^3$ can be found by

$$(\hat{\beta}_n^3, \hat{\beta}_u^3) := \arg\max_{\beta_n^3, \beta_u^3} \sum_{i \in \mathcal{I}} \log L_i^3(\beta_n^3, \beta_u^3 \mid STI_i, S_i, E_i, X_i).$$

Benchmark model 4: full information benchmark

For model 4, we know the number of stockout occurrences T_i for textbook i. Moreover, for each epoch $t = 1, \ldots, T_i + 1$, sales $S_{it} := (s_{int}, s_{iut}, s_{i0t})$ and available assortment $\mathcal{A}_{it}$ are known. Let $S_{iT} := (s_{i1}, \ldots, s_{iT_i})$ and $\mathcal{A}_i := (\mathcal{A}_{i1}, \ldots, \mathcal{A}_{iT_i})$. The likelihood function of textbook i is given by

$$L_i^4(\beta_n^4, \beta_u^4 \mid T_i, S_{iT}, \mathcal{A}_i, X_i) = \prod_{t=1}^{T_i} \left[\frac{(s_{int} + s_{iut} + s_{i0t})!}{s_{int}! s_{iut}! s_{i0t}!} \prod_{k \in \mathcal{A}_{it}} P_k^{s_{ikt}} (\beta_n^4, \beta_u^4 \mid \mathcal{A}_{it}, X_i) \right].$$

Maximum likelihood estimates $\hat{\beta}_n^4$ and $\hat{\beta}_u^4$ can be found by

$$(\hat{\beta}_n^4, \hat{\beta}_u^4) := \arg\max_{\beta_n^4, \beta_u^4} \sum_{i \in \mathcal{I}} \log L_i^4(\beta_n^4, \beta_u^4 \mid T_i, S_{iT}, \mathcal{A}_i, X_i).$$

Appendix C: Likelihood function for general number of product types

This section shows how to compute the likelihood function in closed form for an arbitrary number of products when sales are observed but lost sales and the time of stockout are unobserved. As we do not observe when each product type stocks out, we cannot identify the exact sequence of stockouts. Hence, we must account for all possible sequences of stockouts that yield the observed sales. We first derive the likelihood function for a given sequence of stockouts and then enumerate all possible sequences of stockouts. Let T_a and T_b denote the number of product types that have stocked out and have not stocked out respectively. Note that $T_a + T_b \leq K + 1$ since we have K substitutable types and type 0. Let $a := (a_1, a_2, \ldots, a_{T_a})$ denote the types that have stocked out with their order representing the sequence of stockouts. Let $b := (b_1, b_2, \ldots, b_{T_b})$ denote the remaining product types including type 0. The ordering of elements of b is irrelevant because those product types have not stocked out. For example, if there are five product types and types 2 and 4 have stocked out, then $K = 5$, $T_a = 2$, $T_b = 4$, the value of a can be $(2, 4)$ or $(4, 2)$, and the value of b is $(0, 1, 3, 5)$.

In our choice model, the purchasing probability depends on the available assortment. During the selling season, the available assortment changes after each stockout. Therefore, we partition a selling season into $T_a + 1$ epochs so that each epoch is associated with a different assortment. Define epoch t as the period after the $(t-1)$th stockout (the beginning of the selling season is $t = 1$) until the t^{th} stockout. Epoch $T_a + 1$, the last epoch, represents the period after the last stockout to the end of selling season. Let $\mathcal{A}_t(a, b) := \{a_t, a_{t+1}, \ldots, a_{T_a}\} \bigcup \{b_1, b_2, \ldots, b_{T_b}\}$ for $t = 1, \ldots, T_a + 1$ denote the set of available product types during epoch t given information a and b. Note that $\mathcal{A}_{T_a+1}(a, b) \equiv \{b_1, b_2, \ldots, b_{T_b}\}$, which means that in the last epoch the available product types are the ones that do not stock out. Given a, b, and product attribute x, the choice probability that a consumer chooses product type k during epoch t is $P_k(\beta|\mathcal{A}_t(a, b), x)$ for $k \in \mathcal{A}_t(a, b)$ and 0 otherwise.

With the above structure, Figure A.1 illustrates the components of the likelihood function for a given sequence of stockouts. We derive the likelihood of observing sales data of a product, given information on the stockout occurrence and product attributes, in three steps. In the first step, we derive the probability of observing sales $h_t := (h_{t0}, h_{t1}, \cdots, h_{tK})$ in epoch t given a and b. Note that $h_{tk} = 0$ for $k = a_1, \ldots, a_{t-1}$ since book types $a_1, \ldots, a_{t-1}$ have stocked out before epoch t. The probability of observing sales h_t during epoch t given that type a_t stocks out at the end of epoch t and given that other types do not stockout follows the p.m.f. of a negative multinomial distribution, and is given by

$$G_t(\beta|a, b, h_t, x) = \frac{(h_{ta_t} - 1 + \sum_{l=t+1}^{T_a} h_{ta_l} + \sum_{l=1}^{T_b} h_{tb_l})!}{(h_{ta_t} - 1)! \prod_{l=t+1}^{T_a} h_{ta_l}! \prod_{l=1}^{T_b} h_{tb_l}!} \prod_{l=t}^{T_a} P_{a_l}(\beta|\mathcal{A}_t(a, b), x)^{h_{ta_l}} \prod_{l=1}^{T_b} P_{b_l}(\beta|\mathcal{A}_t(a, b), x)^{h_{tb_l}}.$$

The first component of the equation is the multinomial coefficient which gives the number of all feasible sequences of sales, and the last two components represent the probability of observing each sequence of sales for given a and b.

In the second step, we derive the probability of observing given sales-to-go data from epoch t onwards using G_t. Let $H_t := (H_{t0}, H_{t1}, \cdots, H_{tK})$ denote sales-to-go, i.e., sales from epoch t to the last epoch. Note that $H_{tk} = \sum_{l=t}^{T_a+1} h_{lk}$. Different types of a product have different feasible ranges of sales in epoch t for given sales-to-go from epoch t: for types $k = a_1, \ldots, a_{t-1}$, $h_{tk} = H_{tk} = 0$, because those types have stocked out before epoch t; for type a_t, $h_{ta_t} = H_{ta_t}$, since that type stocks out at the end of epoch t; and for types $k = a_{t+1}, \ldots, a_{T_a}$, h_{tk} can range from zero to $H_{tk} - 1$ because those product types stock out in later epochs. On the other hand, for types from b, the value of sales can range from zero to the given sales-to-go. We use recursion to derive the probability $F_t(\beta|a, b, H_t, x)$ of observing sales-to-go data H_t from epoch t onwards given a and b by summing the probabilities of observing given sales in epoch t and observing the leftover sales-to-go during later epochs over the feasible ranges of sales.

$$F_t(\beta|a, b, H_t, x) =$$

$$\sum_{h_{ta_{t+1}}=0}^{H_{ta_{t+1}}-1} \sum_{h_{ta_{t+2}}=0}^{H_{ta_{t+2}}-1} \cdots \sum_{h_{ta_{T_a}}=0}^{H_{ta_{T_a}}-1} \sum_{h_{tb_1}=0}^{H_{tb_1}} \sum_{h_{tb_2}=0}^{H_{tb_2}} \cdots \sum_{h_{tb_{T_b}}=0}^{H_{tb_{T_b}}} G_t(\beta|a, b, h_t, x) F_{t+1}(\beta|a, b, H_t - h_t, x), \tag{1}$$

for $t = 1, \ldots, T_a$, where $h_{tk} = 0$ for $k = a_1, \ldots, a_{t-1}$, and $h_{ta_t} = H_{ta_t}$.

$$F_{T_a+1}(\beta|a, b, H_{T_a+1}, x) = \frac{(\sum_{l=1}^{T_b} H_{T_a+1 b_l})!}{\prod_{l=1}^{T_b} H_{T_a+1 b_l}!} \prod_{l=1}^{T_b} P_{b_l}(\beta|\mathcal{A}_{T_a+1}(a, b), x)^{H_{T_a+1 b_l}},$$

because $T_a + 1$ is the last epoch and we have to observe the remaining sales exactly. So far, we suppress the product category subscript i for simplicity because all computations were within the substitutable types for a given product category.

In the third step, we derive the likelihood of observing sales data of product category i by enumerating all possible sequences of stockouts. From the definition of F_t, the probability of observing sales data S_i for given a, b, and product attributes X_i is $F_1(\beta|a, b, S_i, X_i)$. Therefore, the likelihood of observing sales data S_i for product i is given by

$$L_i(\beta|\mathcal{A}_i, S_i, X_i) = \sum_{a \in \mathcal{SEQ}(\mathcal{A}_{i,stockout})} F_1(\beta|a, b, S_i, X_i), \tag{2}$$

where $\mathcal{A}_i := (\mathcal{A}_{i,stockout}, \mathcal{A}_{i,no-stockout})$, $\mathcal{A}_{i,stockout}$ and $\mathcal{A}_{i,no-stockout}$ denote the set of types of product i that have stocked out and have not respectively, and $\mathcal{SEQ}(\mathcal{A}_{i,stockout})$ denote the set of all possible permutations of $\mathcal{A}_{i,stockout}$.

Finally, the total likelihood of observing the aggregate sales data is given by

$$L(\beta|\mathcal{A}, S, X) = \prod_{i \in \mathcal{I}} L_i(\beta|\mathcal{A}_i, S_i, X_i), \tag{3}$$

where $\mathcal{A} := (\mathcal{A}_1, \ldots, \mathcal{A}_I)$, $S := (S_1, \ldots, S_I)$, and $X := (X_1, \ldots, X_I)$.

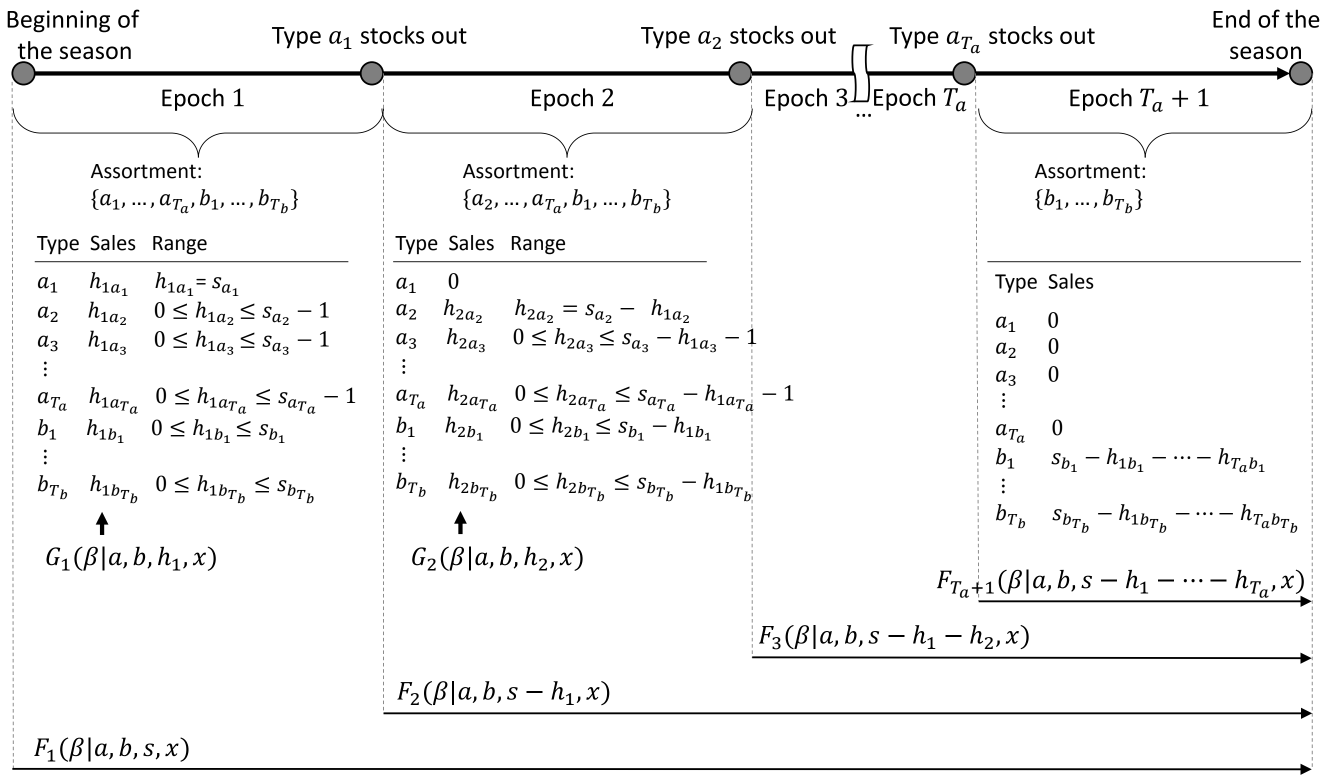
The parameters of the choice model, β, can be estimated using maximum likelihood (ML) approach when the exact likelihood is used. The parameter estimator ($\hat{\beta}$) is given by

$$\hat{\beta}_{ML} = \arg\max_{\beta} \log L(\beta|\mathcal{A}, S, X). \tag{4}$$

References

Train, K. E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Figure A.1 Illustration of the components of the likelihood function



a represents the types that stock out with their order representing the sequence of stockouts.
b represents the types that do not stock out.
h_t represents the sales of epoch t.
$G_t(\beta | a, b, h_t, x)$ is the probability of observing sales h_t during epoch t given a, b, and product attributes x .
$F_t(\beta | a, b, q, x)$ is the probability of observing sales-to-go q from epoch t onwards given a, b, and product attributes x.