Online Supplement for

"A Discrete/Continuous Model for Multi-Category Purchase Behavior of Households"

Inseong Song
mksong@ust.hk
Department of Marketing
Hong Kong University of Science and Technology

and

Pradeep K. Chintagunta
pradeep.chintagunta@chicagogsb.edu
Graduate School of Business
University of Chicago

I. Technical Appendix

In the technical appendices, we provides the proof for the corner solution property and the derivations of marginal and joint category purchase probability, the conditional brand choice probability, and the conditional purchase quantity. We also provide derivation of the conditional distributions of error terms.

Appendix A: Corner Solution

First, we demonstrate how the model formulation leads to a corner solution within each category. The utility maximization problem in (4) of the paper is a constrained maximization problem whose Kuhn-Tucker conditions are given as follows:

$$\frac{\partial U^*}{\partial Q_{cj}} + \lambda_{cj} = U_c \psi_{cj} - U_{M+1} \psi_O P_{cj} + \lambda_{cj} = 0 \text{ and } \lambda_{cj} \geq 0, \; -\lambda_{cj} Q_{cj} = 0 \;\; \forall c, j$$

where λ_{cj} is the Lagrange multiplier associated with the nonnegativity condition for Q_{cj}. Therefore, if Q_{cj}>0, then λ_{cj} must be zero. Suppose, at the optimal demand, there are two brands chosen for category c, say Q_{cj} and Q_{ci}. This implies that $\lambda_{cj}=\lambda_{ci}=0$. Then the Kuhn-Tucker conditions become $U_c\psi_{cj}=U_{M+1}\psi_O P_{cj}$ and $U_c\psi_{ci}=U_{M+1}\psi_O P_{ci}$, which in turn imply that $P_{cj}/\psi_{cj}=P_{ci}/\psi_{ci}$. Since ε_{cj}'s are continuous, $\Pr(P_{cj}/\psi_{cj}=P_{ci}/\psi_{ci})=0$. What if $P_{cj}/\psi_{cj}>P_{ci}/\psi_{ci}$? Note that

$U_c = U_{M+1}\psi_O P_{cj}/\psi_{cj}-\lambda_{cj}/\psi_{cj} = U_{M+1}\psi_O P_{ci}/\psi_{ci}-\lambda_{ci}/\psi_{ci}$. Therefore, $U_{M+1}\psi_O(P_{cj}/\psi_{cj}-P_{ci}/\psi_{ci}) = \lambda_{cj}/\psi_{cj}-\lambda_{ci}/\psi_{ci} > 0$ implies $\lambda_{cj} > \lambda_{ci}\psi_{cj}/\psi_{ci}$. Under this condition, λ_{cj} cannot be zero since it should be larger than a nonnegative number. So Q_{cj} must be zero. Therefore, at most one brand will be chosen.

Appendix B: Marginal Category Purchase Probability

Using the definition $V_{cj}\equiv\rho_{cj}+Z_{cj}'\theta_{cj}-\log P_{cj}$ and the condition (10) of the paper, the no-purchase probability is given by:

$$\Pr(I_c=0)=\Pr\left(\xi_c\le\min_j\frac{P_{cj}}{\psi_{cj}}\right)=\Pr\left(\xi_c\le\frac{P_{cj}}{\psi_{cj}}\forall j=1,\ldots J_c\right)=\Pr\left(\log\xi_c\le\log P_{cj}-\rho_{cj}-Z_{cj}'\theta_c-\varepsilon_{cj}\right)$$

$$=\Pr\left(\gamma_c-\varepsilon_O\le-V_{cj}-\varepsilon_{cj},\forall j=1,\ldots J_c\right)=E_{\varepsilon_O}\left[\prod_{j=1}^{Jc}\Pr\left(\varepsilon_{cj}\le\varepsilon_O-\gamma_c-V_{cj}\mid\varepsilon_O\right)\right]$$

$$=E_{\varepsilon_O}\left[\prod_{j=1}^{Jc}\exp\left(-\exp\left(-\tau\left(\varepsilon_O-\gamma_c-V_{cj}\right)\right)\right)\right]=E_{\varepsilon_O}\left[\exp\left(-\sum_{j=1}^{J_c}\exp\left(-\tau\left(\varepsilon_O-\gamma_c-V_{cj}\right)\right)\right)\right]$$

$$=E_{\varepsilon_O}\left[\exp\left(-\exp\left(\tau\left(\gamma_c-\varepsilon_O\right)\right)\sum_{j=1}^{J_c}\exp\left(\tau V_{cj}\right)\right)\right]$$

Note that the above probability within the expectation bracket is equal to the no-purchase probability in Chintagunta (1993) if we set ε_O at a constant, say 0. The reservation price, R, in Chintagunta (1993) corresponds to $R=\exp(\tau(\gamma_c-\varepsilon_O))$ in our model. The difference in our model is that we explicitly recognize the existence of common shock. For notational simplicity, temporarily, define $A\equiv\sum_{j=1}^{J_c}e^{\tau V_{cj}+\tau\gamma_c}$.

$$E_{\varepsilon_O}\left[\exp\left(-\exp\left(\tau\left(\gamma_c-\varepsilon_O\right)\right)\sum_{j=1}^{J_c}\exp\left(\tau V_{cj}\right)\right)\right]=\int\exp\left(-e^{-\tau\varepsilon_O}\sum_{j=1}^{J_c}e^{\tau V_{cj}+\tau\gamma_c}\right)f(\varepsilon_O)d\varepsilon_O$$

$$=\int\exp\left(-Ae^{-\tau\varepsilon_O}\right)\exp\left(-e^{-\tau\varepsilon_O}\right)\tau e^{-\tau\varepsilon_O}d\varepsilon_O=\int\exp\left(-e^{-\tau\varepsilon_O}(A+1)\right)\tau e^{-\tau\varepsilon_O}d\varepsilon_O$$

$$=\int\exp\left(-e^{-\tau\varepsilon_O}e^{\log(A+1)}\right)\tau e^{-\tau\varepsilon_O}d\varepsilon_O=\int\exp\left(-e^{-\tau\left(\varepsilon_O-\frac{1}{\tau}\log(A+1)\right)}\right)\tau e^{-\tau\left(\varepsilon_O-\frac{1}{\tau}\log(A+1)\right)}e^{-\tau\frac{1}{\tau}\log(A+1)}d\varepsilon_O$$

$$=e^{-\log(A+1)}=\frac{1}{1+A}=\frac{1}{1+\sum_{j=1}^{J_c}e^{\tau V_{cj}+\tau\gamma_c}}.$$

So the category purchase probability is given by

$$\Pr(I_c = 1) = 1 - \Pr(I_c = 0) = \frac{\sum_{j=1}^{J_c} e^{\tau V_{cj} + \tau\gamma_c}}{1 + \sum_{j=1}^{J_c} e^{\tau V_{cj} + \tau\gamma_c}}.$$

From the definition of V_c in (15) of the paper, the category purchase probability is expressed as

$$\Pr(I_c = 1) = \frac{e^{V_c}}{1 + e^{V_c}}.$$

Appendix C: Joint Category Purchase Probability

We first compute the joint probability conditional on ε_O and then take expectation over the distribution of ε_O, i.e.,

$$\Pr\left(I_1 = 1, ..., I_m = 1, I_{m+1} = 0, ..., I_M = 0\right) = E_{\varepsilon_O}\left[\Pr\left(I_1 = 1, ..., I_m = 1, I_{m+1} = 0, ..., I_M = 0 | \varepsilon_O\right)\right].$$

Given ε_O, the conditional category purchase probabilities are independent each other. So

$$\Pr\left(I_1 = 1, ..., I_m = 1, I_{m+1} = 0, ..., I_M = 0 | \varepsilon_O\right) = \prod_{c=1}^{m} \Pr(I_c = 1 | \varepsilon_O) * \prod_{d=m+1}^{M} \Pr(I_d = 0 | \varepsilon_O)$$

$$= \prod_{c=1}^{m}\left(1 - \exp\left(-e^{\tau(\gamma_c - \varepsilon_O)} \sum_{j=1}^{J_c} e^{\tau V_{cj}}\right)\right) * \prod_{d=m+1}^{M} \exp\left(-e^{\tau(\gamma_d - \varepsilon_O)} \sum_{j=1}^{J_d} e^{\tau V_{dj}}\right)$$

Now, recognize that

$$\prod_{i=1}^{m}(1 - x_i) = 1 - \sum_{i_1} x_{i_1} + \sum_{i_1}\sum_{i_2 > i_1} x_{i_1} x_{i_2} + ... + (-1)^m \sum_{i_1} ... \sum_{i_m > i_{m-1}} x_{i_1} ... x_{i_m}.$$

And that the last term in the above equation has only one term, i.e.,

$$(-1)^m \sum_{i_1} ... \sum_{i_m > i_{m-1}} x_{i_1} ... x_{i_m} = (-1)^m x_1 ... x_m.$$

Temporally define $x_O \equiv \prod_{d=m+1}^{M} \exp\left(-e^{\tau(\gamma_d - \varepsilon_O)} \sum_{j=1}^{J_d} e^{\tau V_{dj}}\right)$ and $x_c \equiv \exp\left(-e^{\tau(\gamma_c - \varepsilon_O)} \sum_{j=1}^{J_c} e^{\tau V_{cj}}\right)$. Now it is clear that the joint purchase probability conditional on the common shock is given by $x_O \prod_{i=1}^{m}(1 - x_i) = x_O\left(1 - \sum_{i_1} x_{i_1} + \sum_{i_1}\sum_{i_2 > i_1} x_{i_1} x_{i_2} + ... + (-1)^m \sum_{i_1} ... \sum_{i_m > i_{m-1}} x_{i_1} ... x_{i_m}\right)$. We will integrate out the common shock for each term in the above equation.

$$\int x_O (-1)^k \sum_{i_1} ... \sum_{i_k > i_{k-1}} x_{i_1} ... x_{i_k} f(\varepsilon_O) d\varepsilon_O = (-1)^k \sum_{i_1} ... \sum_{i_k > i_{k-1}} \int x_{i_1} ... x_{i_k} x_O f(\varepsilon_O) d\varepsilon_O$$

$$= (-1)^k \sum_{i_1} \cdots \sum_{i_k > i_{k-1}} \int \exp\left(-\sum_{c=i_1}^{i_k} e^{-\tau\varepsilon_O} \sum_{j=1}^{J_c} e^{\tau V_{cj} + \tau\gamma_c} - \sum_{d=m+1}^{M} e^{-\tau\varepsilon_O} \sum_{j=1}^{J_d} e^{\tau V_{dj} + \tau\gamma_d} \right) f(\varepsilon_O) d\varepsilon_O$$

$$= (-1)^k \sum_{i_1} \cdots \sum_{i_k > i_{k-1}} \int \exp\left(-\left(\sum_{c=i_1}^{i_k} e^{V_c} + \sum_{d=m+1}^{M} e^{V_d} \right) e^{-\tau\varepsilon_O} \right) f(\varepsilon_O) d\varepsilon_O$$

Again define temporarily $A \equiv \sum_{c=i_1}^{i_k} e^{V_c} + \sum_{d=m+1}^{M} e^{V_d}$. Then the above expression becomes

$$= (-1)^k \sum_{i_1} \cdots \sum_{i_k > i_{k-1}} \int \exp\left(-Ae^{-\tau\varepsilon_O} \right) f(\varepsilon_O) d\varepsilon_O = (-1)^k \sum_{i_1} \cdots \sum_{i_k > i_{k-1}} \int \exp\left(-Ae^{-\tau\varepsilon_O} \right) \exp(-e^{-\tau\varepsilon_O}) \tau e^{-\tau\varepsilon_O} d\varepsilon_O$$

$$= (-1)^k \sum_{i_1} \cdots \sum_{i_k > i_{k-1}} \int \exp\left(-e^{-\tau\left(\varepsilon_O - \frac{1}{\tau}\log(A+1)\right)} \right) \tau e^{-\tau\varepsilon_O} d\varepsilon_O = \frac{1}{1+A} = \frac{1}{1 + \sum_{c=i_1}^{i_k} e^{V_c} + \sum_{d=m+1}^{M} e^{V_d}}$$

Hence,

$$\Pr\left(I_1 = 1, \ldots, I_m = 1, I_{m+1} = 0, \ldots, I_M = 0\right) = \int \Pr\left(I_1 = 1, \ldots, I_m = 1, I_{m+1} = 0, \ldots, I_M = 0 | \varepsilon_O\right) f(\varepsilon_O) d\varepsilon_O$$

$$= \frac{1}{1 + \sum_{d=m+1}^{M} e^{V_d}} - \sum_{i_1=1}^{m} \frac{1}{1 + e^{V_{i_1}} + \sum_{d=m+1}^{M} e^{V_d}} + \sum_{i_1=1}^{m} \sum_{i_2 > i_1} \frac{1}{1 + e^{V_{i_1}} + e^{V_{i2}} + \sum_{d=m+1}^{M} e^{V_d}} + \ldots$$

$$+ (-1)^m \sum_{i_1} \sum_{i_2 > i_1} \cdots \sum_{i_m > i_{m-1}} \frac{1}{1 + e^{V_{i_1}} + \ldots + e^{V_{i_m}} + \sum_{d=m+1}^{M} e^{V_d}}.$$

Appendix D: Brand Choice Probability

Here we derive the brand choice probability conditional on making a category purchase.

$$\Pr(I_{cj} = 1 \mid I_c = 1) = \frac{\Pr(I_{cj} = 1,\ I_c = 1)}{\Pr(I_c = 1)}$$

We only need to obtain the numerator in the above equation as the denominator is derived above.

$$\Pr(I_{cj} = 1,\ I_c = 1) = \Pr\left(\frac{P_{ck}}{\psi_{ck}} > \frac{P_{cj}}{\psi_{cj}}, \forall k \neq j,\ \ \textbf{AND},\ \xi_c > \min_{k=1,\ldots,J_c} \frac{P_{ck}}{\psi_{ck}} \right)$$

$$= \Pr\left(\frac{P_{ck}}{\psi_{ck}} > \frac{P_{cj}}{\psi_{cj}}, \forall k \neq j,\ \ \textbf{AND},\ \xi_c > \frac{P_{cj}}{\psi_{cj}} \right)$$

$$= \Pr\left(\varepsilon_{cj} + V_{cj} > \varepsilon_{ck} + V_{ck}, \forall k \neq j, \textbf{AND},\ \varepsilon_{cj} > \varepsilon_O - V_{cj} - \gamma_c \right)$$

$$= \int_{\varepsilon_O=-\infty}^{+\infty} \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \left(\prod_{k \neq j} \int_{-\infty}^{\varepsilon_{cj}+V_{cj}-V_{ck}} f(\varepsilon_{ck}) d\varepsilon_{ck} \right) f(\varepsilon_{cj}) d\varepsilon_{cj} f(\varepsilon_O) d\varepsilon_O$$

$$= \int_{\varepsilon_O=-\infty}^{+\infty} \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \left(\prod_{k \neq j} \exp\left(-e^{-\tau\left(\varepsilon_{cj}+V_{cj}-V_{ck}\right)} \right) \right) f(\varepsilon_{cj}) d\varepsilon_{cj} f(\varepsilon_O) d\varepsilon_O$$

$$= \int_{\varepsilon_O=-\infty}^{+\infty} \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-\sum_{k \neq j} e^{-\tau\left(\varepsilon_{cj}+V_{cj}-V_{ck}\right)} \right) f(\varepsilon_{cj}) d\varepsilon_{cj} f(\varepsilon_O) d\varepsilon_O$$

$$= \int_{\varepsilon_O=-\infty}^{+\infty} \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-e^{-\tau\varepsilon_{cj}} \frac{\sum_{k \neq j} e^{\tau V_{ck}}}{e^{\tau V_{cj}}} \right) f(\varepsilon_{cj}) d\varepsilon_{cj} f(\varepsilon_O) d\varepsilon_O \, .$$

Now let us compute the inner integration in the above expression. Define $A \equiv \sum_{k \neq j} \frac{\exp(\tau V_{ck})}{\exp(\tau V_{cj})}$.

Based on the definition, the inner integration becomes

$$\int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-e^{-\tau\varepsilon_{cj}} \frac{\sum_{k \neq j} e^{\tau V_{ck}}}{e^{\tau V_{cj}}} \right) \exp\left(-e^{-\tau\varepsilon_{cj}} \right) \tau e^{-\tau\varepsilon_{cj}} d\varepsilon_{cj}$$

$$= \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-e^{-\tau\varepsilon_{cj}} (A+1) \right) \tau e^{-\tau\varepsilon_{cj}} d\varepsilon_{cj}$$

$$= \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-e^{-\tau\varepsilon_{cj}} e^{\log(A+1)} \right) \tau e^{-\tau\varepsilon_{cj}} d\varepsilon_{cj}$$

$$= \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-e^{-\tau\left(\varepsilon_{cj} - \frac{1}{\tau}\log(A+1)\right)} \right) \tau e^{-\tau\left(\varepsilon_{cj} - \frac{1}{\tau}\log(A+1)\right)} e^{-\log(A+1)} d\varepsilon_{cj}$$

$$= \frac{1}{1+A} \int_{\varepsilon_{cj}=\varepsilon_O - V_{cj} - \gamma_c}^{+\infty} \exp\left(-e^{-\tau\left(\varepsilon_{cj} - \frac{1}{\tau}\log(A+1)\right)} \right) \tau e^{-\tau\left(\varepsilon_{cj} - \frac{1}{\tau}\log(A+1)\right)} d\varepsilon_{cj}$$

$$= \frac{1}{1+A} \left(1 - \exp\left(-e^{-\tau\left(\varepsilon_O - V_{cj} - \gamma_c - \frac{1}{\tau}\log(A+1)\right)} \right) \right).$$

In order to compute the outer integral, let us define $\kappa \equiv V_{cj} + \gamma_c + \frac{1}{\tau}\log(A+1)$. Now the outer integral becomes

$$\Pr(I_{cj} = 1 \, , \; I_c = 1) = \int \frac{1}{1+A} \left(1 - \exp\left(-e^{-\tau(\varepsilon_O - \kappa)} \right) \right) f(\varepsilon_O) d\varepsilon_O$$

$$= \frac{1}{1+A} \left(1 - \int \exp\left(-e^{-\tau(\varepsilon_O - \kappa)} \right) f(\varepsilon_O) d\varepsilon_O \right) = \frac{1}{1+A} \left(1 - \int \exp\left(-e^{-\tau(\varepsilon_O - \kappa)} \right) \exp\left(-e^{-\tau\varepsilon_O} \right) \tau d\varepsilon_O \right)$$

$$= \frac{1}{1+A}\left(1 - \frac{1}{1+e^{\tau\kappa}}\right) = \frac{1}{1+A}\left(1 - \frac{1}{1+e^{\tau V_{cj}+\tau\gamma_c}(1+A)}\right) = \frac{e^{\tau V_{cj}+\tau\gamma_c}}{1+e^{\tau V_{cj}+\tau\gamma_c}(1+A)}$$

$$= \frac{e^{\tau V_{cj}+\tau\gamma_c}}{1+e^{\tau V_{cj}+\tau\gamma_c}\left(1+\sum_{k\neq j}\frac{e^{\tau V_{ck}}}{e^{\tau V_{cj}}}\right)} = \frac{e^{\tau V_{cj}+\tau\gamma_c}}{1+\sum_{k=1}^{J_c} e^{\tau V_{ck}+\tau\gamma_c}}$$

Thus, the conditional brand choice probability is given by

$$\Pr(I_{cj}=1 \mid I_c=1) = \left(\frac{e^{\tau V_{cj}+\tau\gamma_c}}{1+\sum_{k=1}^{J_c} e^{\tau V_{ck}+\tau\gamma_c}}\right) \Bigg/ \left(\frac{\sum_{k=1}^{J_c} e^{\tau V_{ck}+\tau\gamma_c}}{1+\sum_{k=1}^{J_c} e^{\tau V_{ck}+\tau\gamma_c}}\right) = \frac{e^{\tau V_{cj}}}{\sum_{k=1}^{J_c} e^{\tau V_{ck}}}.$$

Appendix E: Conditional Purchase Quantity

The expenditure share equation in (19) of the paper can be represented in a matrix form as follows:,

$$\begin{pmatrix} w_{1j_1} \\ \vdots \\ w_{mj_m} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_M \end{pmatrix} + \begin{bmatrix} \beta_{11} & \cdots & \beta_{1M} \\ \vdots & \ddots & \vdots \\ \beta_{M1} & \cdots & \beta_{MM} \end{bmatrix} \begin{pmatrix} -V_{1j_1} - \varepsilon_{1j_1} \\ \vdots \\ -V_{mj_m} - \varepsilon_{mj_m} \\ \log \xi_{m+1}(Q^*) \\ \vdots \\ \log \xi_M(Q^*) \end{pmatrix} + \begin{pmatrix} \beta_{1,M+1} \\ \vdots \\ \beta_{M,M+1} \end{pmatrix} (-\varepsilon_O).$$

Or, in a condensed form

$$\begin{pmatrix} w_{(m)} \\ 0 \end{pmatrix} = \boldsymbol{\alpha} + B \begin{pmatrix} -V_{(m)} - \varepsilon_{(m)} \\ \log \xi\left(Q^*\right)_{(M-m)} \end{pmatrix} + \left(B \cdot \mathbf{1}\right)\varepsilon_O.$$

In order to obtain a closed from expression for $w_{(m)}$, we need to solve the last M-m equations first for $\log \xi\left(Q^*\right)_{(M-m)}$. The last M-m rows of the above equation can be rewritten by

$$0 = \boldsymbol{\alpha_2} + B_{21}(-V_{(m)} - \varepsilon_{(m)}) + B_{22} \log \xi\left(Q^*\right)_{(M-m)} + \left(B_{21} \cdot \mathbf{1} + B_{22} \cdot \mathbf{1}\right)\varepsilon_O.$$

This yields the solution for $\log \xi\left(Q^*\right)_{(M-m)}$,

$$\log \xi\left(Q^*\right)_{(M-m)} = -B_{22}^{-1}\boldsymbol{\alpha_2} - B_{22}^{-1}B_{21}(-V_{(m)} - \varepsilon_{(m)}) - B_{22}^{-1}\left(B_{21} \cdot \mathbf{1} + B_{22} \cdot \mathbf{1}\right)\varepsilon_O.$$

Thus,

$$
\begin{aligned}
w_{(m)} &= \boldsymbol{\alpha_1} + B_{11}(-V_{(m)} - \varepsilon_{(m)}) + B_{12} \log \xi\left(Q^*\right)_{(M-m)} + \left(B_{11} \cdot \mathbf{1} + B_{12} \cdot \mathbf{1}\right)\varepsilon_O \\
&= (\boldsymbol{\alpha_1} - B_{12}B_{22}^{-1}\boldsymbol{\alpha_2}) + (B_{11} - B_{12}B_{22}^{-1}B_{21})(-V_{(m)} - \varepsilon_{(m)} + \mathbf{1}\varepsilon_O)
\end{aligned}
$$

Appendix F: Derivation of Conditional Distributions of Error Terms

Here we explain how to derive the conditional distributions $f\left(\varepsilon_{kj_k}\left(w_{(m)}\right) \mid \Omega, \varepsilon_O\right)$ and $f\left(\varepsilon_O \mid \Omega\right)$.

1. $f\left(\varepsilon_{kj_k}\left(w_{(m)}\right) \mid \Omega, \varepsilon_O\right)$

The purchase condition Ω imposes two restriction on ε_{kj_k} : i) Brand Choice Condition $\varepsilon_{kj_k} + V_{kj_k} > \varepsilon_{ki} + V_{ki}, \forall i \neq j_k$ and ii) Category Purchase Condition $\max_i\left(\varepsilon_{ki} + V_{ki} + \gamma_c\right) > \varepsilon_O$.

Combining two conditions yields the following condition:

$$
\varepsilon_{kj_k} + V_{kj_k} > \varepsilon_{ki} + V_{ki}, \forall i \neq j_k, \text{AND}, \varepsilon_{kj_k} + V_{kj_k} + \gamma_c > \varepsilon_O
$$

Consider the first condition, i.e., $f(\varepsilon_{kj_k} \mid \varepsilon_{kj_k} + V_{kj_k} > \varepsilon_{ki} + V_{ki}, \forall i \neq j_k)$. Denote $A \equiv \max_{i \neq j_k}(\varepsilon_{ki} + V_{ki})$. Then A is Gumbel distributed with the scale parameter τ and the location parameter $\frac{1}{\tau}\log(\sum_{i \neq j_k} e^{\tau V_{ki}})$. The CDF of the conditional distribution given by.

$$
\Pr(\varepsilon_{kj_k} \leq x \mid \varepsilon_{kj_k} + V_{kj_k} > A) = \frac{\Pr(\varepsilon_{kj_k} \leq x, \varepsilon_{kj_k} + V_{kj_k} > A)}{\Pr(\varepsilon_{kj_k} + V_{kj_k} > A)}.
$$

$$
\begin{aligned}
&\Pr(\varepsilon_{kj_k} \leq x, \varepsilon_{kj_k} + V_{kj_k} > A) \\
&= \Pr(\varepsilon_{kj_k} \leq x, \varepsilon_{kj_k} + V_{kj_k} > A, A > V_{kj_k} + x) + \Pr(\varepsilon_{kj_k} \leq x, \varepsilon_{kj_k} + V_{kj_k} > A, A < V_{kj_k} + x) \\
&= 0 + \Pr(\varepsilon_{kj_k} \leq x, \varepsilon_{kj_k} + V_{kj_k} > A, A < V_{kj_k} + x) = \int_{-\infty}^{V_{kj_k}+x}\left(\int_{A-V_{kj_k}}^{x} f(\varepsilon_{kj_k})d\varepsilon_{kj_k}\right)f(A)dA \\
&= \int_{-\infty}^{V_{kj_k}+x}\left(e^{-e^{-\tau x}} - e^{-e^{-\tau(A-V_{kj_k})}}\right) \tau e^{-\tau(w-\frac{1}{\tau}\log(\sum_{i \neq j_k} e^{V_{ki}}))} e^{-e^{-\tau(w-\frac{1}{\tau}\log(\sum_{i \neq j_k} e^{V_{ki}}))}} dA \\
&= \exp(-\exp(-\tau(x - \frac{1}{\tau}\log\frac{\sum e^{\tau V_{ki}}}{e^{\tau V_{kj_k}}}))) \frac{e^{\tau V_{kj_k}}}{\sum_i e^{\tau V_{ki}}}
\end{aligned}
$$

$$
\Pr(\varepsilon_{kj_k} \leq x \mid \varepsilon_{kj_k} + V_{kj_k} > A) = \frac{\Pr(\varepsilon_{kj_k} \leq x, \varepsilon_{kj_k} + V_{kj_k} > A)}{\Pr(\varepsilon_{kj_k} + V_{kj_k} > A)}
$$

$$= \left\{ e^{-e^{-\tau\left(x-\frac{1}{\tau}\log\left(\frac{\sum e^{\tau V_{ki}}}{e^{\tau V_{kj_k}}}\right)\right)}} \frac{e^{\tau V_{kj_k}}}{\sum_i e^{\tau V_{ki}}} \right\} / \Pr(\varepsilon_{kj_k} + V_{kj_k} > A) = e^{-e^{-\tau\left(x-\frac{1}{\tau}\log\left(\frac{\sum e^{\tau V_{ki}}}{e^{\tau V_{kj_k}}}\right)\right)}}$$

Therefore, the distribution of the random shock of a brand conditional on the choice of the brand is Gumbel with the scale parameter τ and the location parameter $\frac{1}{\tau}\log\left(\frac{\sum_{i=1}^{J_c} e^{\tau V_{ki}}}{e^{\tau V_{kj_k}}}\right)$.

The second condition implies a truncation. We know that if F(x) is a CDF of the untruncated distribution and x is truncated from below, i.e., x>a, then the CDF of the truncated distribution is G(x)=(F(x)-F(a))/(1-F(a)) and its PDF is given by g(x) = G'(x)=F'(x)/(1-F(a)). So

$$\Pr\left(\varepsilon_{cj_c} \le x_c \forall c = 1,\ldots,m \middle| \Omega, \varepsilon_O\right)$$

$$= \Pr\left(\varepsilon_{cj_c} \le x_c \forall c = 1,\ldots,m \;\middle|\; \begin{array}{l} \varepsilon_{cj_c} + V_{cj_c} + \gamma_c > \varepsilon_O, \forall c = 1,\ldots,m, \ \textbf{AND}, \\ \varepsilon_{cj_c} + V_{cj_c} > \varepsilon_{ci} + V_{ci}, \forall i \ne j_c \forall c = 1,\ldots,m \end{array} \right)$$

$$= \prod_{c=1}^{m} \frac{ e^{-e^{-\tau(x_c - \frac{1}{\tau}\log(\frac{\sum e^{\tau V_{ci}}}{e^{\tau V_{cjc}}}))}} - e^{-e^{-\tau\left(\varepsilon_O - V_{cjc} - \gamma_c - \frac{1}{\tau}\log\left(\frac{\sum e^{\tau V_{ci}}}{e^{\tau V_{cjc}}}\right)\right)}} }{ 1 - e^{-e^{-\tau\left(\varepsilon_O - V_{cjc} - \gamma_c - \frac{1}{\tau}\log\left(\frac{\sum e^{\tau V_{ci}}}{e^{\tau V_{cjc}}}\right)\right)}} }, \ \varepsilon_O - V_{cjc} - \gamma_c < x_c < \infty, c = 1,\ldots,m.$$

$$= \prod_{c=1}^{m} \frac{ e^{-e^{-\tau\left(x_c + \frac{1}{\tau}\log \Pr\left(I_{cj_c}=1|I_c=1\right)\right)}} - e^{-e^{-\tau\left(\varepsilon_O - \frac{1}{\tau}V_c\right)}} }{ 1 - e^{-e^{-\tau\left(\varepsilon_O - \frac{1}{\tau}V_c\right)}} }, \ \varepsilon_O - V_{cjc} - \gamma_c < x_c < \infty, c = 1,\ldots,m$$

2. $f(\varepsilon_O \mid \Omega)$

Now we show $f(\varepsilon_O \mid \Omega)$ is Gumbel with location parameter $\kappa \equiv \frac{1}{\tau}\log(1+\sum_{d=m+1}^{M} e^{V_d})$ scale parameter τ. Note that the no-purchase conditions in Ω implies $\varepsilon_O \ge \varepsilon_{dk} + V_{dk} + \gamma_d$ for k=1,…,J_d and d=m+1,…,M or equivalently $\varepsilon_O \ge \max_{\forall k, d=m+1,\ldots,M}(\varepsilon_{dk} + V_{dk} + \gamma_d)$. Define again $W \equiv \max_{\forall k, d=m+1,\ldots,M}(\varepsilon_{dk} + V_{dk} + \gamma_d)$. Then W is Gumbel with the scale parameter τ and the location parameter $\kappa \equiv \frac{1}{\tau}\log\left(\sum_{d=m+1}^{M}\sum_{k=1}^{J_d}\exp\left(\tau V_{dk} + \tau\gamma_d\right)\right) = \frac{1}{\tau}\log\left(\sum_{d=m+1}^{M} e^{V_d}\right)$. Again,

$$\Pr(\varepsilon_O \le x \mid \varepsilon_O > W) = \frac{\Pr(\varepsilon_O \le x, \varepsilon_O > W)}{\Pr(\varepsilon_O > W)}.$$

$$\Pr(\varepsilon_O \le x, \varepsilon_O > W)$$

$$= \Pr(\varepsilon_O \le x, \varepsilon_O > W, W > x) + \Pr(\varepsilon_O \le x, \varepsilon_O > W, W \le x)$$

$$= 0 + \Pr(\varepsilon_O \le x, \varepsilon_O > W, W \le x)$$

$$= \int_{-\infty}^{x} \left(\int_{w}^{x} f(\varepsilon_O) d\varepsilon_O \right) f(w) dw$$

$$= \int_{-\infty}^{x} \left(e^{-e^{-\tau x}} e^{-e^{-\tau w}} \right) \ \tau e^{-\tau(w-\kappa)} e^{-e^{-\tau(w-\kappa)}} dw$$

$$= \exp\left(-\exp\left(-\tau(x - \frac{1}{\tau}\log(1+e^{\tau\kappa})) \right) \right) \left(\frac{1}{1+e^{\tau\kappa}} \right)$$

$$\Pr(\varepsilon_O > W) = \int_{-\infty}^{\infty} \left(\int_{w}^{\infty} \tau e^{-\tau x} e^{-e^{-\tau x}} dx \right) \ \tau e^{-\tau(w-\kappa)} e^{-e^{-\tau(w-\kappa)}} dw = \frac{1}{1+e^{\tau\kappa}}$$

So, $\Pr(\varepsilon_z \le x \mid \varepsilon_O > W) = \exp\left(-\exp\left(-\tau(x - \frac{1}{\tau}\log(1+e^{\tau\kappa})) \right) \right)$ which is the CDF of the Gumbel with location parameter $\frac{1}{\tau}\log(1+e^{\tau\kappa}) = \frac{1}{\tau}\log(1+\sum_{d=m+1}^{M} e^{V_d})$.

II. Simulation: The impact of item aggregation on price elasticity

We conducted two small scale simulations to illustrate the effect of item aggregation on price elasticity estimates.

Logit Model

In the first simulation, we use a homogeneous logit model. We generate individual choice data. Consider a SKU level choice model with 4 products whose indirect utilities are given by $U_{jt}=a_j+b*Price_{jt}+e_{jt}$ for j=1,2,3,4 and t=1,2,.., T where e_{jt} is iid Type I extreme value distributed. The price vector is generated by $Price_{jt}=c_j+u_{jt}$, where u_{jt} is a iid draw from U(0,1). We use the following parameters for the simulation: a={0.5, 0.6, 0.7, 0.8}, b=-2.0, and c={0.5, 0.7, 0.8, 1.0}. We first generate the indirect utility for each SKU and generate the choice outcome by comparing the utility values. For each run of the simulation, we generate 10000 observations (T=10000). For each run, we first estimate the SKU level choice model and compute the elasticity matrix (averaged across observations). Then, we generate brand level data by aggregating SKU data into brand level data. We assume there are two brands and also assume that the first two SKUs (j=1,2) belong to brand 1 and the rest (j=3,4) belong to brand 2. The brand level prices are weighted averages of SKU prices. We use the choice share as the weights. Then we estimate the brand level choice model and compute the elasticity. We repeat this process 50 times.

The estimates of own elasticities from the SKU level model are -1.39, -1.83, -2.03, -2.49 respectively. The estimates of own elasticities from the brand level model are -0.88 and -1.73. That is, the own elasticity of brand 1 (-0.88) is smaller than its SKU level elasticities (-1.39, and -1.83). Also the own elasticity of brand 2 is -1.73 while its SKU elasticities are -2.03 and -2.49.

Log-log Model

In the second simulation, we use a multiplicative demand model to illustrate the item aggregation effect is not limited to the logit model. Consider a demand system with 4 SKUs whose demand function is given by $\log Q_{jt} = a_j + \sum_{k=1}^{4} b_{jk} \log P_{jt} + e_{jt}$ where e_{jt} is a draw from iid standard normal distribution. Again, the price vector is generated by $Price_{jt}=c_j+u_{jt}$, where u_{jt} is a iid draw from U(0,1). We use the following parameters for the simulation: a={0.5, 0.6, 0.7, 0.8}, c={0.5, 0.7, 0.8, 1.0}, and the elasticity matrix {b}

$\{b_{kj}\}=$				
	-1.5	0.8	0.3	0.3
	0.7	-2	0.4	0.4
	0.3	0.4	-2.5	0.7
	0.5	0.4	0.8	-3.0.

For each run of the simulation, we generate 10000 observations (T=10000). For each run, we first estimate the SKU level demand model. Then, we generate brand level data by aggregating SKU data into brand level data. We assume there are two brands and also assume that the first two SKUs (j=1,2) belong to brand 1 and the rest (j=3,4) belong to brand 2. The brand level prices are weighted averages of SKU prices. We use the long term market shares as the weights. Then we estimate the brand level demand model. We repeat this process 50 times.

The elasticity estimates from the SKU level demand model are given by

-1.4984803	0.79825247	0.30014613	0.29869350
0.70008624	-2.0008275	0.39953119	0.40174906

0.29911024	0.40102242	-2.5015824	0.69653448
0.50140490	0.40038450	0.79868990	-3.0027447.

And the elasticity estimates from the brand level model are given by

-1.0454643	0.66115562
0.70676084	-2.0811305

From the results, it is clear that the own price elasticity of brand 1 (-1.05) is lower than its SKU level elasticities (-1.5 and -2.0). The own elasticity of brand 2 is -2.08 while its SKU elasticities are -2.5 and -3.0.

References

Chintagunta, Pradeep K. (1993) "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decision of Households", *Marketing Science*, Vol. 12, No. 2, 184-208