Online Supplement

What Is the Impact of Non-Randomness on Random Choice Models?

Ruxian Wang
Johns Hopkins University Carey Business School, Baltimore, MD 21202, ruxian.wang@jhu.edu

In this Online Supplement, we provide other technical proofs that are not included in the main text of the paper titled "What Is the Impact of Non-Randomness on Random Choice Models".

Proof of Proposition 1. Let $u(S)=\mu\log\big(\sum_{k\in S}\exp(u_k/\mu)\big)$. Then, we have $q_{i'}(S;u_{i'})=\Pr\big(\max_{k\in S}U_k\le u_{i'}\big)=\exp\big(-\exp\big(-((u_{i'}-u(S))/\mu+\gamma)\big)\big)$. We consider its derivatives w.r.t. $u_{i'}$ as follows:

$$\frac{\partial q_{i'}(S;u_{i'})}{\partial u_{i'}}=1/\mu\cdot\exp\big(-\exp(-((u_{i'}-u(S))/\mu+\gamma))\big)\cdot\exp(-((u_{i'}-u(S))/\mu+\gamma))\ge 0,$$
$$\frac{\partial^2 q_{i'}(S;u_{i'})}{\partial u_{i'}^2}=1/\mu^2\cdot\exp\big(-\exp(-((u_{i'}-u(S))/\mu+\gamma))\big)\cdot\exp(-((u_{i'}-u(S))/\mu+\gamma))\cdot\left(\frac{\sum_{j\in S}a_j}{a_{i'}\exp(\gamma)}-1\right).$$

Therefore, $q_{i'}(S;u_{i'})$ is concave in $u_{i'}$ if $a_{i'}\exp(\gamma)\ge\sum_{j\in S}a_j$ or equivalently $a_{i'}\ge\exp(-\gamma)\cdot\sum_{j\in S}a_j$. Note that $\gamma\approx 0.5772$ is the Euler's constant, so $\exp(-\gamma)\approx 0.5615$. By Jensen's inequality, immediately we have: if $a_{i'}\ge 0.5615\cdot\sum_{j\in S}a_j$,

$$q_{i'}(S;u_{i'})=q_{i'}(S;E[U_{i'}])\ge E[q_{i'}(S;U_{i'})]=\frac{a_{i'}}{a_{i'}+\sum_{j\in S}a_j}.$$

The last equality holds because of how the MNL model is derived; see, e.g., Theorem 2.2 in Anderson et al. (1992). If $a_{i'}\le 0.5615\cdot\sum_{j\in S}a_j$, the probability $q_{i'}(S;u_{i'})$ is convex in $u_{i'}$, then we have $q_{i'}(S;u_{i'})\le a_{i'}/(a_{i'}+\sum_{j\in S}a_j)$ also by Jensen's inequality.

For any other alternative $i\in S$, we have the choice probability $q_i(S;u_{i'})=a_i/\sum_{j\in S}a_j\cdot\big[1-\exp\big(-\exp(-(u_{i'}/\mu+\gamma))\big)\big]$. Its derivatives w.r.t. $u_{i'}$ can be derived immediately as follows:

$$\frac{\partial q_i(S;u_{i'})}{\partial u_{i'}}=\frac{-1/\mu\cdot a_i}{\sum_{j\in S}a_j}\cdot\exp\big(-\exp(-((u_{i'}-u(S))/\mu+\gamma))\big)\cdot\exp(-((u_{i'}-u(S))/\mu+\gamma))\le 0,$$
$$\frac{\partial^2 q_i(S;u_{i'})}{\partial u_{i'}^2}=\frac{-1/\mu^2\cdot a_i}{\sum_{j\in S}a_j}\cdot\exp\big(-\exp(-((u_{i'}-u(S))/\mu+\gamma))\big)\cdot\exp(-((u_{i'}-u(S))/\mu+\gamma))\cdot\left(\frac{\sum_{j\in S}a_j}{a_{i'}\exp(\gamma)}-1\right).$$

By similar argument, we can establish the comparison for the choice probabilities regarding any product i in the offer set S. □

Proof of Proposition 2. (a) Note that $\max(x,y)$ is a convex function in x for any given y. Denote $U(S):=\max_{i\in S}U_i$ and then $G(S;u_{i'})=E\big[\max\{U_i,u_{i'}:i\in S\}\big]=E[\max\{U(S),u_{i'}\}]$. By Jensen's inequality, given any $U(S)$, we have

$$E[\max\{U(S),U_{i'}\}|U(S)]\ge\max\{U(S),E[U_{i'}]\},$$

where the first expectation is taken w.r.t. $U_{i'}$. Immediately, we have

$$G(S;U_{i'})=E\big[E[\max\{U(S),U_{i'}\}|U(S)]\big]\ge E[\max\{U(S),E[U_{i'}]\}]=E[\max\{U(S),u_{i'}\}]=G(S;u_{i'}),$$

where the first expectation is a conditional expectation.

(b) The density function for $\max_{i\in S}U_i$ is also a Gumbel random variable and its density is

$$f_{\max}(x)=1/\mu\cdot\exp\big(-((x-u(S))/\mu+\gamma)\big)\cdot\exp\big(-\exp\big(-((x-u(S))/\mu+\gamma)\big)\big),$$

where $u(S) = \mu \log\big(\sum_{k\in S} \exp(u_k/\mu)\big)$. Then, the perceived consumer surplus with a deterministic utility $u_{i'}$ can be further expressed by:

$$\begin{aligned} G(S;u_{i'}) = \;& E\big[\max\{\max_{i\in S} U_i, u_{i'}\}\big] = u_{i'} \cdot \exp\big(-\exp\big(-((u_{i'} - u(S))/\mu + \gamma)\big)\big) \\ & + \int_{u_{i'}}^{\infty} x/\mu \cdot \exp\big(-((x-u(S))/\mu+\gamma)\big) \cdot \exp\big(-\exp\big(-((x-u(S))/\mu+\gamma)\big)\big) dx. \end{aligned}$$

The consumer surplus can be expressed by

$$G(S;u_{i'}) = E\big[\max\{U(S), u_{i'}\}\big] = u_{i'} \cdot F_{\max}(u_{i'}) + \int_{u_{i'}}^{\infty} x \cdot f_{\max}(x) dx.$$

Then, we have

$$\frac{\partial G(S;u_{i'})}{\partial u_{i'}} = F_{\max}(u_{i'}) = \Pr\big(\max_{i\in S} U_i \le u_{i'}\big) = \exp\big(-\Lambda \sum_{j\in S} a_j\big) = q_{i'}(S;u_{i'}).$$

To indicate the dependence of consumer surplus on the nominal utility of any product $i \in S$, we include the argument $u(S)$ in the density and distribution function:

$$G(S;u_{i'}) = u_{i'} \cdot \int_{-\infty}^{u_{i'}} f_{\max}(x|u(S)) dx + \int_{u_{i'}}^{\infty} x \cdot f_{\max}(x|u(S)) dx,$$

where $f_{\max}(x) = 1/\mu \cdot \exp\big(-((x-u(S))/\mu+\gamma)\big) \cdot \exp\big(-\exp\big(-((x-u(S))/\mu+\gamma)\big)\big)$. We replace $x - u(S)$ with y and derive

$$G(S;u_{i'}) = u_{i'} \cdot \int_{-\infty}^{u_{i'}-u(S)} g(y) dy + \int_{u_{i'}-u(S)}^{\infty} (y+u(S)) \cdot g(y) dy,$$

where $g(y) = 1/\mu \cdot \exp\big(-(y/\mu+\gamma)\big) \cdot \exp\big(-\exp\big(-(y/\mu+\gamma)\big)\big)$.

Then, by the chain rule for any product $i \in S$,

$$\begin{aligned} \frac{\partial G(S;u_{i'})}{\partial u_i} &= \frac{\partial}{\partial u(S)} \left(u_{i'} \cdot \int_{-\infty}^{u_{i'}-u(S)} g(y) dy + \int_{u_{i'}-u(S)}^{\infty} (y+u(S)) \cdot g(y) dy \right) \cdot \frac{\partial u(S)}{\partial u_i} \\ &= \int_{u_{i'}-u(S)}^{\infty} g(y) dy \cdot \frac{\partial u(S)}{\partial u_i} = \int_{u_{i'}}^{\infty} f_{\max}(x|u(S)) dx \cdot \frac{\partial u(S)}{\partial u_i} \\ &= \Pr\big(\max_{i\in S} U_i \ge u_{i'}\big) \cdot \frac{\partial u(S)}{\partial u_i} = \Big[1 - \exp\Big(-\Lambda \sum_{j\in S} a_j\Big)\Big] \cdot \frac{a_i}{\sum_{j\in S} a_j} = q_i(S;u_{i'}). \end{aligned}$$

We have shown $\partial G(S;u_{i'})/\partial u_i = q_i(S;u_{i'})$ for any $i \in S'$.

(c) From part (b), it is obvious to observe that $G(S;u_{i'})$ is increasing in each u_i for all $i \in S'$. Moreover, we have for any $i \in S'$

$$\frac{\partial^2 G(S;u_{i'})}{\partial u_i^2} = \frac{\partial q_i(S;u_{i'})}{\partial u_i} \ge 0.$$

The inequality holds because of Proposition 1. Therefore, the consumer surplus is increasing convex in u_i for all $i \in S'$. □

Proof of Theorem 4. By the likelihood functions $\mathcal{L}(\mathbf{n}, \mathbf{n}_0|\lambda, \mathbf{v})$ defined in equation (3) in the complete data case and $\mathcal{L}(\mathbf{n}|\lambda, \mathbf{v})$ defined in equation (7) in the incomplete data case, we have

$$\frac{\mathcal{L}(\mathbf{n}, \mathbf{n}_0|\lambda, \mathbf{v})}{\mathcal{L}(\mathbf{n}|\lambda, \mathbf{v})} = \frac{\prod_{k=1}^{K} \frac{\lambda^{n_k+n_{ki'}+n_{k0}} \cdot \exp(-\lambda)}{(n_k+n_{ki'}+n_{k0})!} \cdot \frac{(n_k+n_{ki'}+n_{k0})!}{\prod_{i\in S'_k+} n_{ki}!} \cdot \prod_{i\in S'_k+} (q_{ki}(\mathbf{v}))^{n_{ki}}}{\prod_{k=1}^{K} \sum_{n_{k0}=0}^{\infty} \frac{\lambda^{n_k+n_{ki'}+n_{k0}} \cdot \exp(-\lambda)}{(n_k+n_{ki'}+n_{k0})!} \cdot \frac{(n_k+n_{ki'}+n_{k0})!}{n_{k0}! \prod_{i\in S'_k} n_{ki}!} \cdot \prod_{i\in S'_k} (q_{ki}(\mathbf{v}))^{n_{ki}} \cdot (q_{k0}(\mathbf{v}))^{n_{k0}}}$$

$$
\begin{aligned}
&= \prod_{k=1}^{K} \frac{\frac{\lambda^{n_{k0}}}{n_{k0}!} \cdot (q_{k0}(\mathbf{v}))^{n_{k0}}}{\sum_{n_{k0}=0}^{\infty} \frac{\lambda^{n_{k0}}}{n_{k0}!} \cdot (q_{k0}(\mathbf{v}))^{n_{k0}}} = \prod_{k=1}^{K} \frac{\big(\lambda \cdot q_{k0}(\mathbf{v})\big)^{n_{k0}} \cdot \exp(-\lambda \cdot q_{k0}(\mathbf{v}))}{n_{k0}!} \\
&:= \prod_{k=1}^{K} h\big(n_{k0}; \lambda \cdot q_{k0}(\mathbf{v})\big),
\end{aligned}
$$

where $h\big(n_{k0}; \lambda \cdot q_{k0}(\mathbf{v})\big) = (\lambda \cdot q_{k0}(\mathbf{v}))^{n_{k0}} \cdot \exp(-\lambda \cdot q_{k0}(\mathbf{v}))/n_{k0}!$ is the probability mass function for the Poisson random variable with mean $\lambda \cdot q_{k0}(\mathbf{v})$. The third equality holds because $\sum_{n_{k0}=0}^{\infty} \lambda^{n_{k0}} \cdot (q_{k0}(\mathbf{v}))^{n_{k0}}/n_{k0}! = \exp(\lambda \cdot q_{k0}(\mathbf{v}))$.

Immediately, the log-likelihood function $\mathcal{LL}(\mathbf{n}|\lambda, \mathbf{v}) := \log(\mathcal{LL}(\mathbf{n}|\lambda, \mathbf{v}))$ can be expressed as follows

$$
\begin{aligned}
\mathcal{LL}(\mathbf{n}|\lambda, \mathbf{v}) &= \log(\mathcal{L}(\mathbf{n}, \mathbf{n}_0|\lambda, \mathbf{v})) - \log\Big(\prod_{k=1}^{K} h\big(n_{k0}; \lambda \cdot q_{k0}(\mathbf{v})\big)\Big) \\
&= \mathcal{LL}\big(\mathbf{n}, \mathbf{n}_0|\lambda\big) + \mathcal{LL}\big(\mathbf{n}, \mathbf{n}_0|\mathbf{v}\big) - \sum_{k=1}^{K} \Big[n_{k0} \cdot \log\big(\lambda \cdot q_{k0}(\mathbf{v})\big) - \lambda \cdot q_{k0}(\mathbf{v})\big)\Big] + C', \qquad \text{(EC.1)}
\end{aligned}
$$

where $C' = -\sum_{k=1}^{K} \sum_{j \in S'_k} \log(n_{kj}!)$. Notice that $\mathcal{LL}\big(\mathbf{n}, \mathbf{n}_0|\lambda\big)$ and $\mathcal{LL}\big(\mathbf{n}, \mathbf{n}_0|\mathbf{v}\big)$ defined in equations (5) are linear in $\mathbf{n}_0$ respectively, so $\mathcal{LL}(\mathbf{n}|\mathbf{v}, \lambda)$ is also linear in $\mathbf{n}_0$.

Then, we take expectations for both sides of equation (EC.1) with respect to $\mathbf{n}_0$ using the probability mass function $h\big(n_{k0}; \widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)})\big)$, where $\widehat{\lambda}^{(t-1)}$ represents the estimated market size in step $t-1$, and the choice probability $q_{k0}(\widehat{\mathbf{v}}^{(t-1)})$ corresponds to the estimated parameter $\widehat{\mathbf{v}}^{(t-1)}$ in step $t-1$. Therefore, taking expectation for both sides of equation (EC.1) yields

$$
\mathcal{LL}\big(\mathbf{n}|\mathbf{v}, \lambda\big) = \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\lambda\big) + \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\mathbf{v}\big) - \sum_{k=1}^{K} \Big[\widehat{n}_{k0}^{(t)} \cdot \log\big(\lambda \cdot q_{k0}(\mathbf{v})\big) - \lambda \cdot q_{k0}(\mathbf{v}\big]\Big] + C', \qquad \text{(EC.2)}
$$

where $\widehat{n}_{k0}^{(t)}$ is the expectation of a Poisson distributed random variable as shown in the E-step (9), i.e.

$$
\widehat{n}_{k0}^{(t)} = \sum_{n_{k0}=0}^{\infty} h\big(n_{k0}; \widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)})\big) \cdot n_{k0} = \widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)}).
$$

In the M-step, we solve the MLE problem for the complete data case with $\mathbf{n}_0$ replaced by $\widehat{\mathbf{n}}_0^{(t)}$:

$$
\max_{\lambda} \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\lambda\big) \quad \text{and} \quad \max_{\mathbf{v}} \mathcal{LL}(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\mathbf{v}),
$$

and derive the maximizers denoted by $\widehat{\lambda}^{(t)}$ and $\widehat{\mathbf{v}}^{(t)}$. In fact, we have obtained the closed-form solution for $\widehat{\lambda}^{(t)}$: $\widehat{\lambda}^{(t)} = \Big(\sum_{k=1}^{K} n_k + n_{ki} + \widehat{n}_{k0}^{(t)}\Big)/K$. Obviously, these inequalities hold

$$
\mathcal{LL}(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\lambda}^{(t)}) \geq \mathcal{LL}(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\lambda}^{(t-1)}) \text{ and } \mathcal{LL}(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\mathbf{v}}^{(t)}) \geq \mathcal{LL}(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\mathbf{v}}^{(t-1)}). \qquad \text{(EC.3)}
$$

Moreover, at the point $(\lambda, \mathbf{v}) = \big(\widehat{\lambda}^{(t-1)}, \widehat{\mathbf{v}}^{(t-1)}\big)$, equation (EC.2) can be written by

$$
\begin{aligned}
\mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t-1)}, \widehat{\mathbf{v}}^{(t-1)}\big) &= \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\lambda}^{(t-1)}\big) + \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\mathbf{v}}^{(t-1)}\big) - \sum_{k=1}^{K} \Big[\widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)}) \cdot \log\big(\widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)})\big) \\
&\quad -\widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)})\Big];
\end{aligned}
$$

while at the point $(\lambda, \mathbf{v}) = \big(\widehat{\lambda}^{(t)}, \widehat{\mathbf{v}}^{(t)}\big)$, equation (EC.2) can be written by

$$
\mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t)}, \widehat{\mathbf{v}}^{(t)}\big) = \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\lambda}^{(t)}\big) + \mathcal{LL}\big(\mathbf{n}, \widehat{\mathbf{n}}_0^{(t)}|\widehat{\mathbf{v}}^{(t)}\big) - \sum_{k=1}^{K} \Big[\widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)}) \cdot \log\big(\widehat{\lambda}^{(t)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t)})\big) - \widehat{\lambda}^{(t)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t)})\Big].
$$

The constant term C' is omitted in the above two equations for $\mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t-1)}, \widehat{\mathbf{v}}^{(t-1)}\big)$ and $\mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t)}, \widehat{\mathbf{v}}^{(t)}\big)$.

Given any y, consider the function $F(x) := y \cdot \log(x) - x$. We obtain its derivatives:

$$F'(x) = \frac{y}{x} - 1 \text{ and } F''(x) = -\frac{y}{x^2} < 0.$$

One can see $F(x)$ is concave in x and $\max_x F(x) = F(y)$. Immediately, we have

$$\begin{aligned}&\widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)}) \cdot \log\left(\widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)})\right) - \widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)}) \\ &\geq \widehat{\lambda}^{(t-1)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t-1)}) \cdot \log\left(\widehat{\lambda}^{(t)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t)})\right) - \widehat{\lambda}^{(t)} \cdot q_{k0}(\widehat{\mathbf{v}}^{(t)}).\end{aligned} \tag{EC.4}$$

Combining inequalities (EC.3) and (EC.4) yields

$$\mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t)}, \widehat{\mathbf{v}}^{(t)}\big) \geq \mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t-1)}, \widehat{\mathbf{v}}^{(t-1)}\big). \tag{EC.5}$$

Therefore, the log-likelihood function $\mathcal{LL}\big(\mathbf{n}|\widehat{\lambda}^{(t)}, \widehat{\mathbf{v}}^{(t)}\big)$ increases in t and is bounded above, so it must converge as $t \to \infty$. □

Proof of Theorem 7. Set $\Lambda = 1$ for notational convenience and also without loss of generality. By $q_0 = \exp\big(-\sum_{j\in S} a_j\big)$, we have $\sum_{j\in S} a_j = -\log(q_0)$ and therefore for any $i \in S$

$$q_i(q_0) = \frac{a_i(1-q_0)}{-\log(q_0)}.$$

First, we show $q_i(q_0) \geq a_i q_0$ for any $0 < q_0 < 1$. Consider

$$\frac{\partial\big(q_i(q_0)/q_0\big)}{\partial q_0} = \frac{\partial}{\partial q_0}\frac{a_i(1-q_0)}{-q_0\log(q_0)} = a_i \cdot \frac{\log(q_0) - q_0 + 1}{\big(q_0\log(q_0)\big)^2} < 0, \ \forall\ 0 < q_0 < 1. \tag{EC.6}$$

The inequality holds because $\partial(\log(q_0) - q_0 + 1)/\partial q_0 = 1/q_0 - 1 > 0$ and $(\log(q_0) - q_0 + 1)\big|_{q_0=1} = 0$. Then, $q_i(q_0)/q_0$ is decreasing in q_0. Moreover, $\lim_{q_0\to 1} q_i(q_0)/q_0 = \lim_{q_0\to 1} a_i(1-q_0)/(-q_0\log(q_0)) = a_i$. Therefore, we have $q_i(q_0) \geq a_i q_0$ for any $0 < q_0 < 1$.

Recall $q_0^* \geq \tau$, so for any $i \in S^*$

$$q_i^* \leq q_0^* \cdot \frac{a_i(1-\tau)}{-\tau\log(\tau)}. \tag{EC.7}$$

Suppose that all the products including product i' are labelled in the revenue-decreasing order. We will show that there exist two consecutive revenue-ordered assortments, denoted by S_{k-1} and S_k, such that $\gamma R(S_{k-1}) + (1-\gamma)R(S_k) \geq \frac{-\tau\log(\tau)}{1-\tau} \cdot R(S^*)$, where $S_k := \{1, 2, \ldots, k\}$ and $0 \leq \gamma \leq 1$.

For any offer set S, let $Q(S)$ be the sum of the choice probabilities for all items in set S, excluding the no-purchase option, i.e., $Q(S) = \sum_{j\in S} q_j(S) = 1 - \exp\big(-\Lambda\sum_{j\in S} a_j\big)$. We observe that $Q(S)$ is larger if the offer set S is larger. Then, for any set S, there exists k such that $Q(S_{k-1}) < Q(S^*) \leq Q(S_k)$. For such k, let $0 \leq \gamma \leq 1$ be the coefficient such that $\gamma Q(S_{k-1}) + (1-\gamma)Q(S_k) = Q(S^*)$. Immediately, we have $\gamma q_0(S_{k-1}) + (1-\gamma)q_0(S_k) = q_0(S^*)$. Next, it holds that for each $j < k$,

$$\begin{aligned}&\gamma q_j(S_{k-1}) + (1-\gamma)q_j(S_k) = \gamma q_j\big(q_0(S_{k-1})\big) + (1-\gamma)q_j\big(q_0(S_k)\big) \\ &\geq \gamma a_j q_0(S_{k-1}) + (1-\gamma)a_j q_0(S_k) = a_j q_0(S^*) \geq \frac{-\tau\log(\tau)}{1-\tau} \cdot q_j(S^*).\end{aligned}$$

The first inequality holds because $q_j(q_0) \geq a_j q_0$; the second inequality holds because of the inequality (EC.7). Note that $q_j(S^*) = 0$ for any $j \notin S^*$. Then,

$$\frac{1-\tau}{-\tau\log(\tau)}\big(\gamma q_j(S_{k-1}) + (1-\gamma)q_j(S_k)\big) - q_j(S^*) \geq 0.$$

Moreover, by similar argument for (EC.6), we can show $\frac{1-\tau}{-\tau\log(\tau)} > 1$ for any $0<\tau<1$. Therefore, we have the following comparison:

$$\begin{aligned}
&\frac{1-\tau}{-\tau\log(\tau)}\cdot\Big[\gamma R(S_{k-1})+(1-\gamma)R(S_k)\Big]-R(S^*)\\
&=\sum_{j=1}^{k-1}p_j\cdot\left[\frac{1-\tau}{-\tau\log(\tau)}\big(\gamma q_j(S_{k-1})+(1-\gamma)q_j(S_k)\big)-q_j(S^*)\right]+\frac{1-\tau}{-\tau\log(\tau)}\cdot(1-\gamma)p_k\cdot q_k(S_k)-\sum_{j\geq k, j\in S^*}p_j\cdot q_j(S^*)\\
&\geq p_k\cdot\sum_{j=1}^{k-1}\left[\frac{1-\tau}{-\tau\log(\tau)}\big(\gamma q_j(S_{k-1})+(1-\gamma)q_j(S_k)\big)-q_j(S^*)\right]+\frac{1-\tau}{-\tau\log(\tau)}\cdot(1-\gamma)p_k\cdot q_k(S_k)-\sum_{j\geq k, j\in S^*}p_j\cdot q_j(S^*)\\
&\geq p_k\cdot\Big(\sum_{j=1}^{k-1}\gamma q_j(S_{k-1})+(1-\gamma)q_j(S_k)-q_j(S^*)\Big)+(1-\gamma)p_k\cdot q_k(S_k)-p_k\cdot\sum_{j\geq k, j\in S^*}q_j(S^*)\\
&=p_k\cdot\Big(\gamma\sum_{j\in S_{k-1}}q_j(S_{k-1})+(1-\gamma)\sum_{j\in S_k}q_j(S_k)-\sum_{j\in S^*}q_j(S^*)\Big)=p_k\cdot\Big(\gamma Q(S_{k-1})+(1-\gamma)Q(S_k)-Q(S^*)\Big)=0.
\end{aligned}$$

Thus, for the optimal assortment S^*, there exist two revenue-ordered assortments S_{k-1} and S_k such that

$$\text{either}\quad \frac{1-\tau}{-\tau\log(\tau)}\cdot R(S_{k-1})\geq R(S^*)\quad\text{or}\quad\frac{1-\tau}{-\tau\log(\tau)}\cdot R(S_k)\geq R(S^*).$$

Finally, we have established the performance guarantee for the revenue-ordered assortments. □

Proof of Proposition 3. Without loss of generality, assume $\alpha_0=0$ for the no-purchase option and then $a_0=\exp(\alpha_0)=1$ for notational brevity in this proof. We consider the first-order condition for the aggregate revenue w.r.t. p_k:

$$\begin{aligned}
&\frac{\beta_k\cdot\exp(\alpha_k-\beta_kp_k)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)}\cdot\left(\sum_{i\in S}\frac{p_i\cdot\exp(\alpha_i-\beta_ip_i)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)}+1/\beta_k-p_k\right)\cdot\Big[1-\exp\Big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)\Big)\Big]\\
&+\left(p_{i'}-\sum_{i\in S}\frac{p_i\cdot\exp(\alpha_i-\beta_ip_i)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)}\right)\cdot\exp\Big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)\Big)\cdot\beta_k\Lambda\exp(\alpha_k-\beta_kp_k)=0.
\end{aligned}$$

It is straightforward to see that $p_k<\infty$ at optimality, so we have

$$\begin{aligned}
&\left(\sum_{i\in S}\frac{p_i\cdot\exp(\alpha_i-\beta_ip_i)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)}+\frac{1}{\beta_k}-p_k\right)\cdot\frac{1-\exp\Big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)\Big)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)}\\
&+\left(p_{i'}-\sum_{i\in S}\frac{p_i\cdot\exp(\alpha_i-\beta_ip_i)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)}\right)\cdot\exp\Big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_jp_j)\Big)\cdot\Lambda=0.
\end{aligned}$$

We observe that the above equality holds at optimality for any $k\in S$, so (p_k-1/β_k) is the same at optimality for any $k\in S$, denoted it by p.

Then, $p_k=p+1/\beta_k$ for any $k\in S$, and the problem (15) can be rewritten by

$$\begin{aligned}
\max_{p,p_{i'}}R(p,p_{i'})\ :=&\sum_{i\in S}\frac{(p+1/\beta_i)\cdot\exp(\alpha_i-1-\beta_ip)}{1+\sum_{j\in S}\exp(\alpha_j-1-\beta_jp)}\cdot\Big[1-\exp\Big(-\Big(1+\sum_{j\in S}\exp(\alpha_j-1-\beta_jp)\Big)\\
&\cdot\exp(-\alpha_{i'}+\beta_{i'}p_{i'}-\gamma)\Big)\Big]+p_{i'}\cdot\exp\Big(-\Big(1+\sum_{j\in S}\exp(\alpha_j-1-\beta_jp)\Big)\cdot\exp(-\alpha_{i'}+\beta_{i'}p_{i'}-\gamma)\Big).
\end{aligned}$$

We consider the first-order condition for $R(p,p_{i'})$ w.r.t. $p_{i'}$:

$$\frac{\partial R(p,p_{i'})}{\partial p_{i'}}=\ \exp\Big(-\Big(1+\sum_{j\in S}\exp(\alpha_j-1-\beta_jp)\Big)\cdot\exp(-\alpha_{i'}+\beta_{i'}p_{i'}-\gamma)\Big)\cdot\Big[1-\beta_{i'}\Big(p_{i'}$$

$$-\sum_{i\in S}\frac{(p+1/\beta_i)\cdot\exp(\alpha_i-1-\beta_i p)}{1+\sum_{j\in S}\exp(\alpha_j-1-\beta_j p)}\Big)\cdot\Big(1+\sum_{j\in S}\exp(\alpha_j-1-\beta_j p)\Big)\cdot\exp(-\alpha_{i'}+\beta_{i'}p_{i'}-\gamma)\Big]=0.$$

Then, we have

$$\Big(p_{i'}-\sum_{i\in S}\frac{(p+1/\beta_i)\cdot\exp(\alpha_i-1-\beta_i p)}{1+\sum_{j\in S}\exp(\alpha_j-1-\beta_j p)}\Big)\cdot\Big(1+\sum_{j\in S}\exp(\alpha_j-1-\beta_j p)\Big)\cdot\exp(-\alpha_{i'}+\beta_{i'}p_{i'}-\gamma)=\frac{1}{\beta_{i'}}.$$

Note that for any given p, it must hold $p_{i'}-\sum_{i\in S}\frac{(p+1/\beta_i)\cdot\exp(\alpha_i-1-\beta_i p)}{1+\sum_{j\in S}\exp(\alpha_j-1-\beta_j p)}>0$. The LHS of the above equation is increasing in $p_{i'}$, so for any given p, there exists a unique $p_{i'}$ corresponding to the above equation. □

Proof of Theorem 8. We consider its log-derivatives as follows: for any $i\in S$,

$$\frac{\partial\log\big(R_i(\mathbf{p})\big)}{\partial p_i}=\frac{1}{p_i}-\beta_i+\frac{\beta_i\exp(\alpha_i-\beta_i p_i)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)}-\frac{\exp\big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)\big)\cdot\beta_i\Lambda\exp(\alpha_i-\beta_i p_i)}{1-\exp\big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)\big)}.$$

For any $i,k\in S, i\neq k$, we have the cross-derivative:

$$\begin{aligned}\frac{\partial^2\log\big(R_i(\mathbf{p})\big)}{\partial p_i\partial p_k}&=\frac{\beta_i\exp(\alpha_i-p_i)}{\sum_{j\in S^+}\exp(\alpha_j-p_j)}\cdot\frac{\beta_k\exp(\alpha_k-p_k)}{\sum_{j\in S^+}\exp(\alpha_j-p_j)}\\&\quad-\frac{\exp\big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)\big)}{\big(1-\exp\big(-\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)\big)\big)^2}\cdot\Lambda^2\beta_i\exp(\alpha_i-\beta_i p_i)\cdot\beta_k\exp(\alpha_k-\beta_k p_k)\\&=\Lambda^2\beta_i\beta_k\cdot\exp(\alpha_i-p_i)\exp(\alpha_k-p_k)\cdot\Big(\frac{1}{x^2}-\frac{1}{\exp(x)+\exp(-x)-2}\Big)\geq 0,\end{aligned}$$

where $x=\Lambda\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)$. The inequality holds because for any $x\geq 0$ we have $x^2\leq\exp(x)+\exp(-x)-2$ as follows:

$$\begin{aligned}\frac{\partial\big(\exp(x)+\exp(-x)-2-x^2\big)}{\partial x}\Big|_{x=0}&=\big(\exp(x)-\exp(-x)-2x\big)\big|_{x=0}=0,\\\frac{\partial^2\big(\exp(x)+\exp(-x)-2-x^2\big)}{\partial x^2}&=\exp(x)+\exp(-x)-2\geq 0,\ \forall x.\end{aligned}$$

Then, $\big(\exp(x)+\exp(-x)-2-x^2\big)$ is convex in x, and therefore $\partial\big(\exp(x)+\exp(-x)-2-x^2\big)/\partial x$ is increasing in x. Thus, $\partial\big(\exp(x)+\exp(-x)-2-x^2\big)/\partial x\geq 0$ for any $x\geq 0$, i.e., $\big(\exp(x)+\exp(-x)-2-x^2\big)$ is increasing in x. Moreover, $\big(\exp(x)+\exp(-x)-2-x^2\big)\big|_{x=0}=0$, so $x^2\leq\exp(x)+\exp(-x)-2$ for any $x\geq 0$.

Recall $\Lambda=\exp(-\alpha_{i'}+\beta_{i'}p_{i'}-\gamma)$. Then, $\partial\log\big(R_i(\mathbf{p})\big)/\partial p_i$ can be rewritten explicitly in $p_{i'}$:

$$\frac{\partial\log\big(R_i(\mathbf{p})\big)}{\partial p_i}=\frac{1}{p_i}-\beta_i+\frac{\beta_i\exp(\alpha_i-\beta_i p_i)}{\sum_{j\in S^+}\exp(\alpha_j-\beta_j p_j)}-\frac{\exp\big(-\sum_{j\in S^+}\exp(\hat\alpha_j-\beta_j p_j+\beta_{i'}p_{i'})\big)\cdot\beta_i\exp(\hat\alpha_i-\beta_{i'}p_i+\beta_{i'}p_{i'})}{1-\exp\big(-\sum_{j\in S^+}\exp(\hat\alpha_j-\beta_j p_j+\beta_{i'}p_{i'})\big)}.$$

where $\hat\alpha_j=\alpha_j-\alpha_{j'}-\gamma$ for any $j\in S^+$. For any $i\in S$ and i', we have the cross-derivative:

$$\begin{aligned}\frac{\partial^2\log\big(R_i(\mathbf{p})\big)}{\partial p_i\partial p_{i'}}&=-\frac{\exp(-y)\cdot\beta_i\beta_{i'}\exp(\hat\alpha_i-\beta_i p_i+\beta_{i'}p_{i'})}{1-\exp(-y)}+\frac{\exp(-y)\beta_i\beta_{i'}\exp(\hat\alpha_i-\beta_i p_i+p_{i'})\cdot y}{(1-\exp(-y))^2}\\&=\frac{\exp(-y)\cdot\beta_i\beta_{i'}\exp(\hat\alpha_i-\beta_i p_i+\beta_{i'}p_{i'})}{(1-\exp(-y))^2}\cdot\big(-1+\exp(-y)+y\big)\geq 0.\end{aligned}$$

where $y=\sum_{j\in S^+}\exp(\hat\alpha_j-\beta_j p_j+\beta_{i'}p_{i'})$. The inequality holds because $-1+\exp(-y)+y\geq 0$ for any $y\geq 0$. To see this inequality, consider $\partial(-1+\exp(-y)+y)/\partial y=1-\exp(-y)\geq 0$ and $(-1+\exp(-y)+y)\big|_{y=0}=0$. Therefore, for any $i\in S$, the payoff function $R_i(\mathbf{p})$ is log-supermodular.

Firm i''s pay-off function (17) can be rewritten by

$$R_{i'}(\mathbf{p}):=p_{i'}\cdot\exp\Big(-\sum_{j\in S^+}\exp(\hat\alpha_j-\beta_j p_j+\beta_{i'}p_{i'})\Big).$$

Then, we derive its log-derivatives as follows:

$$\frac{\partial \log(R_{i'}(\mathbf{p}))}{\partial p_{i'}} = \frac{1}{p_{i'}} - \sum_{j \in S+} \beta_{i'} \exp(\hat{\alpha}_j - \beta_j p_j + \beta_{i'} p_{i'}) \text{ and } \frac{\partial^2 \log(R_{i'}(\mathbf{p}))}{\partial p_{i'} \partial p_i} = \beta_{i'} \beta_i \exp(\hat{\alpha}_j - \beta_j p_j + \beta_{i'} p_{i'}) \geq 0.$$

Thus, we have shown that the price competition is log-supermodular and therefore there exists a Nash equilibrium; see Vives (2001). □

Proof of Proposition 4. We set $u_0 = -\gamma$ and then have $\Lambda = \exp(-u_0 - \gamma) = 1$ without loss of generality. We consider the first-order condition for $R(\mathbf{p})$ w.r.t. p_k:

$$\frac{\partial R(\mathbf{p})}{\partial p_k} = \frac{\beta_k \exp(\alpha_k - \beta p_k)}{A} \cdot \left[(1/\beta_k - p_k) \cdot \big(1 - \exp(-A)\big) - \sum_{i \in S} p_i \cdot \exp(\alpha_i - \beta_i p_i) \cdot \frac{\exp(-A)(A+1) - 1}{A} \right] = 0,$$

where $A = \sum_{j \in S} \exp(\alpha_j - \beta_j p_j)$ for notational convenience. Then, we have

$$(1/\beta_k - p_k) \cdot \big(1 - \exp(-A)\big) = \sum_{i \in S} p_i \cdot \exp(\alpha_i - \beta_i p_i) \cdot \frac{\exp(-A)(A+1) - 1}{A}.$$

The RHS is independent of p_k, so $(p_k - 1/\beta_k)$ is the same for any product $k \in S$, denoted by p. Then, $p_k = p + 1/\beta_k$ for any $k \in S$, and the total revenue $R(\mathbf{p})$ can be rewritten by

$$R(p) := \sum_{i \in S} \frac{(p + 1/\beta_i) \cdot \exp(\alpha_i - 1 - \beta_i p)}{\sum_{j \in S} \exp(\alpha_j - 1 - \beta_j p)} \cdot \Big[1 - \exp\Big(- \sum_{j \in S} \exp(\alpha_j - 1 - \beta_j p) \Big) \Big].$$

We consider the first-order condition as follows:

$$\frac{\partial R(p)}{\partial p} = -\exp(-B) \cdot \sum_{j \in S} \beta_j \exp(\alpha_j - 1 - \beta_j p) \cdot G(p) = 0.$$

where $G(p)$ is given by

$$G(p) = \sum_{j \in S} \frac{\exp(\alpha_j - 1 - \beta_j p)}{\beta_j} \cdot \frac{1 + B - \exp(B)}{B^2} + p \quad \text{and} \quad B = \sum_{j \in S} \exp(\alpha_j - 1 - \beta_j p).$$

To show the unimodality of $R(p)$, it is sufficient to show $G(p) = 0$ has a unique solution. To do so, consider

$$\frac{\partial G(p)}{\partial p} = -\frac{1}{B} + \frac{\exp(B)}{B} + \frac{1}{B^3}\big(2 + B - 2\exp(B) + B\exp(B)\big) \cdot \sum_{j \in S} \frac{\exp(\alpha_j - 1 - \beta_j p)}{\beta_j} \cdot \sum_{j \in S} \beta_j \exp(\alpha_j - 1 - \beta_j p).$$

We observe $\partial\big(2 + B - 2\exp(B) + B\exp(B)\big)/\partial B = 1 - \exp(B) + B\exp(B) \geq 0$ for any $B \geq 0$ because $\partial\big(1 - \exp(B) + B\exp(B)\big)/\partial B = B\exp(B) \geq 0$ for any $B \geq 0$ and $(1 - \exp(B) + B\exp(B))\big|_{B=0} = 0$. Notice $\big(2 + B - 2\exp(B) + B\exp(B)\big)\big|_{B=0} = 0$, so $2 + B - 2\exp(B) + B\exp(B) \geq 0$ for any $B \geq 0$. By the Cauchy-Schwarz inequality, we have

$$\sum_{j \in S} 1/\beta_j \cdot \exp(\alpha_j - 1 - \beta_j p) \cdot \sum_{j \in S} \beta_j \cdot \exp(\alpha_j - 1 - \beta_j p) \geq B^2,$$

where again $B = \sum_{j \in S} \exp(\alpha_j - 1 - \beta_j p)$. Therefore, we have

$$\frac{\partial G(p)}{\partial p} \geq -\frac{1}{B} + \frac{\exp(B)}{B} + \frac{1}{B^3}\big(2 + B - 2\exp(B) + B\exp(B)\big) \cdot B^2 = 1 + \frac{1}{B} + \exp(B) - \frac{\exp(B)}{B}.$$

We observe that

$$\frac{\partial\big(1 + 1/B + \exp(B) - \exp(B)/B\big)}{\partial B} = \frac{1}{B^2} \cdot \big(-1 + (1 - B + B^2) \cdot \exp(B)\big) \geq 0, \ \forall B \geq 0,$$

because $\partial\big(-1 + (1 - B + B^2) \cdot \exp(B)\big)/\partial B = (B + B^2)\exp(B) \geq 0$ for any $B \geq 0$ and $\big(-1 + (1 - B + B^2) \cdot \exp(B)\big)\big|_{B=0} = 0$. Moreover,

$$\lim_{B \to 0} \Big(1 + \frac{1}{B} + \exp(B) - \frac{\exp(B)}{B}\Big) = \lim_{B \to 0} (1 + \exp(B)) + \lim_{B \to 0} \frac{1 - \exp(B)}{B} = 1.$$

Thus, for any p we have

$$\frac{\partial G(p)}{\partial p} \geq 1 + \frac{1}{B} + \exp(B) - \frac{\exp(B)}{B} \geq 1 > 0,$$

so $G(p)$ is increasing in p. We also observe $G(0) < 0$ because $1 + B - \exp(B) < 0$ for any $B > 0$, and $\lim_{p\to\infty} G(p) = \infty$. Therefore, there exists a unique solution to $G(p) = 0$ and then $R(p)$ is unimodal in p. □

Proof of Theorem 9. For the payoff function:

$$R_i(\mathbf{p}) = \frac{p_i \cdot \exp(\alpha_i - \beta_i p_i)}{\sum_{j\in S} \exp(\alpha_j - \beta_j p_j)} \cdot \Big[1 - \exp\Big(-\Lambda \sum_{j\in S} \exp(\alpha_j - \beta_j p_j)\Big)\Big],$$

we consider its log-derivatives as follows:

$$\begin{aligned}
\frac{\partial \log\big(R_i(\mathbf{p})\big)}{\partial p_i} &= \frac{1}{p_i} - \beta_i + \frac{\beta_i \exp(\alpha_i - \beta_i p_i)}{\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)} - \frac{\exp\big(-\Lambda\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)\big)\cdot \Lambda\beta_i \exp(\alpha_i - \beta_i p_i)}{1 - \exp\big(-\Lambda\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)\big)},\\
\frac{\partial^2 \log\big(R_i(\mathbf{p})\big)}{\partial p_i \partial p_k} &= \frac{\beta_i \exp(\alpha_i - \beta_i p_i)}{\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)} \cdot \frac{\beta_k \exp(\alpha_k - \beta_k p_k)}{\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)}\\
&\quad - \frac{\exp\big(-\Lambda\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)\big)}{\big(1 - \exp\big(-\Lambda\sum_{j\in S}\exp(\alpha_j - \beta_j p_j)\big)\big)^2} \cdot \Lambda^2 \beta_i\beta_k \exp(\alpha_i - \beta_i p_i)\exp(\alpha_k - \beta_k p_k)\\
&= \Lambda^2 \beta_i \beta_k \exp(\alpha_i - \beta_i p_i)\exp(\alpha_k - \beta_k p_k) \cdot \Big(\frac{1}{x^2} - \frac{1}{\exp(x) + \exp(-x) - 2}\Big),
\end{aligned}$$

where $x = \Lambda \sum_{j\in S} \exp(\alpha_j - \beta_j p_j)$. To show $R_i(\mathbf{p})$ is log-supermodular, it is sufficient to show $x^2 \leq \exp(x) + \exp(-x) - 2$ for any $x \geq 0$.

We consider

$$\begin{aligned}
\frac{\partial\big(\exp(x) + \exp(-x) - 2 - x^2\big)}{\partial x}\Big|_{x=0} &= \big(\exp(x) - \exp(-x) - 2x\big)\big|_{x=0} = 0,\\
\frac{\partial^2\big(\exp(x) + \exp(-x) - 2 - x^2\big)}{\partial x^2} &= \exp(x) + \exp(-x) - 2 \geq 0,\ \forall x.
\end{aligned}$$

Then, $\big(\exp(x)+\exp(-x)-2-x^2\big)$ is convex in x, and therefore $\partial\big(\exp(x)+\exp(-x)-2-x^2\big)/\partial x$ is increasing in x. Thus, $\partial\big(\exp(x)+\exp(-x)-2-x^2\big)/\partial x \geq 0$ for any $x \geq 0$, i.e., $\big(\exp(x)+\exp(-x)-2-x^2\big)$ is increasing in x. Moreover, $\big(\exp(x) + \exp(-x) - 2 - x^2\big)\big|_{x=0} = 0$, so $x^2 \leq \exp(x) + \exp(-x) - 2$ for any $x \geq 0$. Finally, we have shown the price competition is log-supermodular and therefore there exists a Nash equilibrium. Moreover, we can also establish its uniqueness by considering diagonal dominance. □

Proof of Proposition 5. By the proof of Theorem 10, $\max_{j\in S_i}(u_{ij} + \xi_{ij})$ is another Gumbel random variable with the scale parameter μ_2 and mean $u(S_i) := \mu_2 \log\big(\sum_{j\in S_i}\exp(u_{ij}/\mu_2)\big)$. We define $U(S_i) := \max_{j\in S_i} U_{ij} = \max_{j\in S_i}(u_{ij} + \xi_{ij}) + \xi_i$. Then, the $U(S_i)$'s are independent Gumbel random variables with the same scale parameter μ_1 for all $i = 1, 2, \ldots, m$, and $E[U(S_i)] = u(S_i)$.

Therefore, we have

$$\begin{aligned}
q_{ij}\big(\mathbf{S}; u_0\big) &= \Pr\big(U(S_i) \geq U(S_j), \forall j \neq i, U(S_i) \geq u_0\big) \cdot \Pr\big(U_{ij} \geq U_{ik}, \forall k \neq j\big)\\
&= \frac{\exp\big(u(S_i)/\mu_1\big)}{\sum_{i=1}^m \exp\big(u(S_i)/\mu_1\big)} \cdot \Big[1 - \exp\Big(-\Lambda\sum_{i=1}^m \exp\big(u(S_i)/\mu_1\big)\Big)\Big] \cdot \Pr(u_{ij} + \xi_{ij} \geq u_{ik} + \xi_{ik}, \forall k \neq j)\\
&= \frac{\exp\big(u(S_i)/\mu_1\big)}{\sum_{i=1}^m \exp\big(u(S_i)/\mu_1\big)} \cdot \Big[1 - \exp\Big(-\Lambda\sum_{i=1}^m \exp\big(u(S_i)/\mu_1\big)\Big)\Big] \cdot \frac{\exp(u_{ij}/\mu_2)}{\sum_{k\in S_i}\exp(u_{ik}/\mu_2)}.
\end{aligned}$$

The second equality holds because of Theorem 1; the third equality holds because of the MNL model. □