Online Appendix

The Focal Luce Model

Matthew Kovach* Gerelt Tserenjigmid †

September 18, 2020

Appendix B: Proof of Proposition 7

Suppose we observe $c^k_{a^l_i} \in \{0,1\}$ for each $l \leq L$, $k \leq N^l$ such that $\sum_i c^k_{a^l_i} = 1$. Here $c^k_{a^l_i} = 1$ means that consumer k (or a consumer visits the menu the k-th time) chooses a^l_i from menu $A^l = \{a^l_0, a^1_l, \ldots, a^l_{m_l}\}$. Here, N^l is the number of observations from menu A^l. By our model,

$$c^k_{a^l_i} = \mathbb{1}\left\{a_i \in \arg\max_{a^l_j \in A^l} \left\{\bar{u}(a^l_j) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_j \in F(A^l)\} + \epsilon^k_{a^l_j}\right\}\right\},$$

where $\epsilon^k_{a^l_j}$ is i.i.d and follows the standard type I extreme distribution. Since

$$\bar{c}_{a^l_i} \equiv \mathbb{E}[c^k_{a^l_i}] = \frac{\exp\left(\bar{u}(a^l_i) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_i \in F(A^l)\}\right)}{\sum_{a^l_j \in A^l} \exp\left(\bar{u}(a^l_j) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_j \in F(A^l)\}\right)},$$

we can write that

$$c^k_{a^l_i} = \frac{\exp\left(\bar{u}(a^l_i) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_i \in F(A^l)\}\right)}{\sum_{a^l_j \in A^l} \exp\left(\bar{u}(a^l_j) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_j \in F(A^l)\}\right)} + \bar{\epsilon}^k_{a^l_i} = \bar{c}_{a^l_i} + \bar{\epsilon}^k_{a^l_i},$$

where $\bar{\epsilon}^k_{a^l_i}$ is i.i.d across k, $\mathbb{E}[\bar{\epsilon}^k_{a^l_i}] = 0$, and $\sum_{i=0}^{m_l} \bar{\epsilon}^k_{a^l_i} = 0$. We usually aggregate N^l separate choices to obtain choice frequencies as follows:

$$p^l_i = p(a^l_i, A^l) = \frac{\sum_{k=1}^{N^l} c^k_{a^l_i}}{N^l} = \frac{\exp\left(\bar{u}(a^l_i) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_i \in F(A^l)\}\right)}{\sum_{a^l_j \in A^l} \exp\left(\bar{u}(a^l_j) + \bar{\delta}(A^l)\,\mathbb{1}\{a^l_j \in F(A^l)\}\right)} + \frac{\sum_{k=1}^{N^l} \bar{\epsilon}^k_{a^l_i}}{N^l} = \bar{c}_{a^l_i} + \bar{\epsilon}_{a^l_i}.$$

Now let us explain the algebra behind LMP we have discussed in Appendix A.A4. Notice that

*Department of Economics, Virginia Tech. E-mail: mkovach@vt.edu

†Department of Economics, Virginia Tech. E-mail: gerelt@vt.edu

$$y_i^l \equiv \log(\frac{p_i^l}{p_0^l}) = \log\left[\frac{\exp\big(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\big) + \bar{\epsilon}_{a_i^l}\Big(\sum_{a_j^l \in A^l} \exp\big(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\big)\Big)}{\exp\big(\bar{u}(a_0^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_0^l \in F(A^l)\}\big) + \bar{\epsilon}_{a_0^l}\Big(\sum_{a_j^l \in A^l} \exp\big(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\big)\Big)}\right]$$

$$= \log\left[\frac{\exp\big(v(a_i^l) + \delta_l\, d_i^l\big) + \bar{\epsilon}_{a_i^l}\Big(\sum_{a_j^l \in A^l} \exp\big(v(a_j^l) + \delta_l\, d_j^l\big)\Big)}{1 + \bar{\epsilon}_{a_0^l}\Big(\sum_{a_j^l \in A^l} \exp\big(v(a_j^l) + \delta_l\, d_j^l\big)\Big)}\right]$$

$$= \log\left[\exp\big(v(a_i^l) + \delta_l\, d_i^l\big) + \frac{\big(\bar{\epsilon}_{a_i^l} - \exp\big(v(a_i^l) + \delta_l\, d_i^l\big)\bar{\epsilon}_{a_0^l}\big)\Big(\sum_{a_j^l \in A^l} \exp\big(v(a_j^l) + \delta_l\, d_j^l\big)\Big)}{1 + \bar{\epsilon}_{a_0^l}\Big(\sum_{a_j^l \in A^l} \exp\big(v(a_j^l) + \delta_l\, d_j^l\big)\Big)}\right]$$

$$= \log\Big[\exp\big(v(a_i^l) + \delta_l\, d_i^l\big) + \zeta_i^l\Big] = v(a_i^l) + \delta_l\, d_i^l + \bar{\zeta}_i^l,$$

where $v(a_i^l) = u(a_i^l) - u(a_0^l)$, $\delta_l = \bar{\delta}(A^l)\mathrm{sign}\big(\mathbb{1}\{a_i^l \in F(A^l)\} - \mathbb{1}\{a_0^l \in F(A^l)\}\big)$ and $d_i^l = \big|\mathbb{1}\{a_i^l \in F(A^l)\} - \mathbb{1}\{a_0^l \in F(A^l)\}\big|$.

We assume that $\delta_l \neq 0$ for each l. We shall prove Proposition 7 in three steps.

Step 1: for each $l \leq L$, $\bar{\zeta}_i^l \to^P 0$ as $N^l \to \infty$.

Proof of Step 1. By direct calculations,

$$\bar{\zeta}_i^l = \log\big(1 + \frac{\zeta_i^l}{\exp(v(a_i^l) + \delta_l\, d_i^l)}\big) = \frac{\zeta_i^l}{\exp(v(a_i^l) + \delta_l\, d_i^l)} + O((\zeta_i^l)^2).$$

Therefore, it is enough to prove that $\zeta_i^l \to^P 0$ as $N^l \to \infty$. By the Law of Large Number, for each l, $\bar{\epsilon}_{a_i^l} = \frac{\sum_{k=1}^{N^l} \bar{\epsilon}^k_{a_i^l}}{N^l} \to^P 0$ as $N^l \to \infty$. Therefore,

$$\zeta_i^l = \frac{\big(\bar{\epsilon}_{a_i^l} - \exp\big(v(a_i^l) + \delta_l\, d_i^l\big)\bar{\epsilon}_{a_0^l}\big)\Big(\sum_{a_j^l \in A^l} \exp\big(v(a_j^l) + \delta_l\, d_j^l\big)\Big)}{1 + \bar{\epsilon}_{a_0^l}\Big(\sum_{a_j^l \in A^l} \exp\big(v(a_j^l) + \delta_l\, d_j^l\big)\Big)} \to^P 0 \text{ as } N^l \to \infty.$$

Before we proceed to Step 2, let $x_i^l = a_i^l - a_0^l$. Then we solve the following least squares problem:

$$\min_{\alpha \in \mathbb{R}^n,\, \delta_l \in \mathbb{R},\, d_i^l \in \{0,1\}} \sum_{1 \leq i \leq m^l,\, l \leq L} (y_i^l - \alpha\, x_i^l - \delta_l\, d_i^l)^2,$$

where $y_i^l = \alpha\, x_i^l + \delta_l\, d_i^l + \bar{\zeta}_i^l$. Let $(\hat{\alpha}, \hat{\delta} \equiv (\hat{\delta}_l)_l, \hat{d} \equiv (\hat{d}_i^l)_{i,l})$ be the solution to the above LMP.

Step 2: $\lim_{N^l \to \infty \text{ for all } l} Pr(\hat{d} = d) = 1$.

Proof of Step 2. For any $\tilde{d} \in \{0,1\}^{\sum_{l=1}^L m_l}$, let

$$X(\tilde{d}) = \begin{bmatrix} x_1^1 & \tilde{d}_1^1 & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & \ddots & \vdots & \vdots \\ x_{m_1}^1 & \tilde{d}_{m_1}^1 & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & \ddots & \vdots & \vdots \\ x_1^l & 0 & 0 & \dots & \tilde{d}_1^l & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ x_{m_l}^l & 0 & 0 & \dots & \tilde{d}_{m_l}^l & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ x_1^L & 0 & 0 & \dots & 0 & \dots & 0 & \tilde{d}_1^L \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ x_{m_L}^L & 0 & 0 & \dots & 0 & \dots & 0 & \tilde{d}_{m_L}^L \end{bmatrix}, \; \beta = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ \delta_1 \\ \vdots \\ \delta_L \end{bmatrix} \text{ and } \zeta(\tilde{d}) = \begin{bmatrix} \bar{\zeta}_1^1 + (d_1^1 - \tilde{d}_1^1)\delta_1 \\ \vdots \\ \bar{\zeta}_{m_1}^1 + (d_{m_1}^1 - \tilde{d}_{m_1}^1)\delta_1 \\ \vdots \\ \bar{\zeta}_1^l + (d_1^l - \tilde{d}_1^l)\delta_l \\ \vdots \\ \bar{\zeta}_{m_l}^l + (d_{m_l}^l - \tilde{d}_{m_l}^l)\delta_l \\ \vdots \\ \bar{\zeta}_1^L + (d_1^L - \tilde{d}_1^L)\delta_L \\ \vdots \\ \bar{\zeta}_{m_L}^L + (d_{m_L}^L - \tilde{d}_{m_L}^L)\delta_L \end{bmatrix}$$

For given $\tilde{d}$, let the vector $(\alpha(\tilde{d}), \delta(\tilde{d}))$ be the solution to the minimization problem:

$$(B1) \qquad \min_{\alpha \in \mathbb{R}^n,\, \delta_l \in \mathbb{R}} \sum_{1 \le i \le m^l,\, l \le L} (y_i^l - \alpha\, x_i^l - \delta_l\, \tilde{d}_i^l)^2.$$

Moreover, let

$$S(\tilde{d}) = \sum_{1 \le i \le m^l,\, l \le L} (y_i^l - \alpha(\tilde{d})\, x_i^l - \delta(\tilde{d})_l\, \tilde{d}_i^l)^2.$$

Indeed, by the definition of $(\hat{\alpha}, \hat{\delta} \equiv (\hat{\delta}_l)_l, \hat{d} \equiv (\hat{d}_i^l)_{i,l})$, we shall have $(\hat{\alpha}, \hat{\delta}) = (\alpha(\hat{d}), \delta(\hat{d}))$ and $S(\hat{d}) \le S(\tilde{d})$ for any $\tilde{d}$.

Note that Equation (B1) is an OLS regression for $y = X(\tilde{d})\beta + \zeta(\tilde{d})$. Therefore, following standard calculations, we have

$$S(\tilde{d}) = \sum_{1 \le i \le m^l,\, l \le L} (y_i^l - \alpha(\tilde{d})\, x_i^l - \delta(\tilde{d})_l\, \tilde{d}_i^l)^2 = \zeta(\tilde{d})^T \Big(I - X(\tilde{d})^T \big(X(\tilde{d})^T X(\tilde{d}) \big)^{-1} X(\tilde{d}) \Big) \zeta(\tilde{d}).$$

By way of contradiction, suppose that $\hat{d} \neq d$; i.e., $\Delta \equiv \zeta(\hat{d}) - \zeta(d) = \zeta(\hat{d}) - \bar{\zeta} \neq 0$. First, note that

$$S(\hat{d}) = (\bar{\zeta} + \Delta)^T \big(I - X(\hat{d})^T \big(X(\hat{d})^T X(\hat{d}) \big)^{-1} X(\hat{d}) \big) (\bar{\zeta} + \Delta) =$$
$$= \Delta^T \big(I - X(\hat{d})^T \big(X(\hat{d})^T X(\hat{d}) \big)^{-1} X(\hat{d}) \big) \Delta$$
$$+ (\bar{\zeta} + \Delta)^T \big(I - X(\hat{d})^T \big(X(\hat{d})^T X(\hat{d}) \big)^{-1} X(\hat{d}) \big) \bar{\zeta} + \bar{\zeta}^T \big(I - X(\hat{d})^T \big(X(\hat{d})^T X(\hat{d}) \big)^{-1} X(\hat{d}) \big) \Delta.$$

Notice that for any $\tilde{d} = (\tilde{d}_i^l)_{l,i} \in \{0,1\}^{\sum_{l=1}^L m_l}$, the sum of the squares of residuals of an OLS regression for $z = X(\tilde{d})\gamma + \Delta$ is strictly positive; i.e.,

$$\Delta^T \big(I - X(\tilde{d})^T \big(X(\tilde{d})^T X(\tilde{d}) \big)^{-1} X(\tilde{d}) \big) \Delta = (z - X(\tilde{d})\hat{\gamma})^T (z - X(\tilde{d})\hat{\gamma}) > 0.$$

Moreover, since there are only finitely many different $\tilde{d}$, there is $\eta > 0$ such that

$$\Delta^T\big(I - X(\tilde{d})^T\big(X(\tilde{d})^T X(\tilde{d})\big)^{-1}X(\tilde{d})\big)\Delta > \eta \text{ for all } \tilde{d}.$$

However, in order to have $S(\hat{d}) \leq S(d)$, we shall have

$$\begin{aligned} -\eta \;\; > \;\; & (\bar{\zeta}+\Delta)^T\big(I - X(\hat{d})^T\big(X(\hat{d})^T X(\hat{d})\big)^{-1}X(\hat{d})\big)\bar{\zeta} + \bar{\zeta}^T\big(I - X(\hat{d})^T\big(X(\hat{d})^T X(\hat{d})\big)^{-1}X(\hat{d})\big)\Delta \\ & -\bar{\zeta}^T\big(I - X(d)^T\big(X(d)^T X(d)\big)^{-1}X(d)\big)\bar{\zeta}. \end{aligned}$$

However, since $\bar{\zeta} \to^P 0$ as $N^l \to \infty$ for all l by Step 1,

$$\begin{aligned} & (\bar{\zeta}+\Delta)^T\big(I - X(\hat{d})^T\big(X(\hat{d})^T X(\hat{d})\big)^{-1}X(\hat{d})\big)\bar{\zeta} + \bar{\zeta}^T\big(I - X(\hat{d})^T\big(X(\hat{d})^T X(\hat{d})\big)^{-1}X(\hat{d})\big)\Delta \\ & \qquad -\bar{\zeta}^T\big(I - X(d)^T\big(X(d)^T X(d)\big)^{-1}X(d)\big)\bar{\zeta} \to^P 0 \text{ as } N^l \to \infty \text{ for all } l. \end{aligned}$$

Therefore, the above is a contradiction to the assumption that $\Delta \neq 0$. In other words, $\hat{d} = d$ as $N^l \to \infty$ for all l; i.e., the focal sets will be correctly identified.

Step 3: $(\hat{\alpha}, \hat{\delta}) \to^P (\alpha, \delta)$ as $N^l \to \infty$ for all l.

Given Step 2, by standard results the OLS regression returns consistent estimators for α and δ.

Appendix C: Maximum Likelihood Estimation

Consider a data that defined in Appendix B. Let $c_l^k = (c^k_{a_i^l})_{i=1}^{m_l}$, $n^l_{a_i^l} = \sum_{k=1}^{N^l} c^k_{a_i^l}$, and $p_i^l = p(a_i^l, A^l) = \frac{n^l_{a_i^l}}{N^l}$. Let us first derive the likelihood function. By our model,

$$c^k_{a_i^l} = \mathbb{1}\Big\{a_i \in \arg\max_{a_j^l \in A^l} \big\{\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\} + \epsilon^k_{a_j^l}\big\}\Big\},$$

where $\epsilon^k_{a_j^l}$ is i.i.d and follows the standard type I extreme distribution. Since

$$Pr(c^k_{a_i^l} = 1) = \frac{\exp\big(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\big)}{\sum_{a_j^l \in A^l} \exp\big(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\big)},$$

$$Pr(c_l = c_l^k) = \prod_{i=1}^{m_l}\left(\frac{\exp\big(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\big)}{\sum_{a_j^l \in A^l} \exp\big(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\big)}\right)^{c^k_{a_i^l}}.$$

Therefore, the likelihood function for given $\bar{u}, \bar{\delta}, F$ is

$$\mathscr{L}\Big(\{c_l^k\}_{k \leq N^l, l \leq L} | \bar{u}, \bar{\delta}, F\Big) = \prod_{l=1}^{L}\prod_{i=1}^{m_l}\left(\frac{\exp\big(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\big)}{\sum_{a_j^l \in A^l} \exp\big(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\big)}\right)^{n^l_{a_i^l}}.$$

Then to find the maximum likelihood estimators, we shall solve

$$\max_{\bar{u},\bar{\delta},F} \mathscr{L}\left(\{c_l^k\}_{k\leq N^l, l\leq L} | \bar{u}, \bar{\delta}, F\right);$$

equivalently, we solve

$$\max_{\bar{u},\bar{\delta},F} \sum_{l=1}^{L} \left[\sum_{i=1}^{m_l} n_{a_i^l}^l \left(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\right) - N^l \log\left(\sum_{j=1}^{m^l} \exp\left(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\right)\right)\right].$$

Let us derive the first order conditions.

FOC for u: for each a,

$$(C1) \qquad 0 = \sum_{l:a=a_i^l} N^l \left(p_i^l - \frac{\exp\left(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\right)}{\sum_{a_j^l \in A^l} \exp\left(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\right)}\right).$$

FOC for $\bar{\delta}$: for each $l \leq L$,

$$(C2) \qquad 0 = \sum_{i=1}^{m^l} \mathbb{1}\{a_i^l \in F(A^l)\} \left(p_i^l - \frac{\exp\left(\bar{u}(a_i^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\right)}{\sum_{a_j^l \in A^l} \exp\left(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\right)}\right).$$

FOC for F: for l and $a_i^l \in A^l$, $a_i^l \in F(A^l)$ if and only if

$$(C3) \qquad \frac{\exp\left(p_i^l\,\bar{\delta}(A^l)\right) - 1}{\exp\left(\bar{\delta}(A^l)\right) - 1} > \frac{\exp\left(\bar{u}(a_i^l)\right)}{\exp\left(\bar{u}(a_i^l)\right) + \sum_{j\neq i} \exp\left(\bar{u}(a_j^l) + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\right)}.$$

For linear models, $u(x) = \alpha \cdot x = \sum_{k=1}^{n} \alpha_i x_i$, Equation (C1) should be replaced with FOCs for α: for any $k \leq n$,

$$0 = \sum_{l=1}^{l} N^l \sum_{i=1}^{m^l} a_{i,k}^l \left(p_i^l - \frac{\exp\left(\alpha \cdot a_i^l + \bar{\delta}(A^l)\,\mathbb{1}\{a_i^l \in F(A^l)\}\right)}{\sum_{a_j^l \in A^l} \exp\left(\alpha \cdot a_j^l + \bar{\delta}(A^l)\,\mathbb{1}\{a_j^l \in F(A^l)\}\right)}\right).$$