Supplementary Materials

Lemma 2 *If there exist all different* $e_1, e_2, \cdots, e_{2k} < 1$ *and a non-zero vector* $\vec{\beta^*} = [\beta_1^*, \beta_2^*, \cdots, \beta_{2k}^*]^\top$*, s.t.*

- $\mathbf{H}^k \vec{\beta^*} = 0$*,*
- $\vec{\beta^*}$ *has* k *positive elements and* k *negative elements.*

then k*-PL for* $2k-1$ *alternatives is not identifiable.*
Proof: W.l.o.g. assume $\beta_1^*, \beta_2^*, \cdots, \beta_k^* > 0$ and $\beta_{k+1}^*, \beta_{k+2}^*, \beta_{2k}^* < 0$. $\mathbf{H}_{2k-1}^k \vec{\beta^*} = 0$ means that

$$\sum_{r=1}^{k} \beta_r^* \vec{f}_r = -\sum_{r=k+1}^{2k} \beta_r^* \vec{f}_r$$

According to the first row in $\mathbf{H}^k$, we have $\sum_r \beta_r^* = 0$. Let $S = \sum_{r=1}^k \beta_r^*$. Further let $\alpha_r^* = \beta_r^*/S$ when $r = 1, 2, \cdots, k$ and $\alpha_r^* = -\beta_r^*/S$ when $r = k+1, k+2, \cdots, 2k$. We have

$$\sum_{r=1}^{k} \alpha_r^* \vec{f}_r = \sum_{r=k+1}^{2k} \alpha_r^* \vec{f}_r$$

where $\sum_{r=1}^k \alpha_r^* = 1$ and $\sum_{r=k+1}^{2k} \alpha_r^* = 1$. This means that the model is not identifiable. ■

Lemma 4 $\sum_s \frac{1}{\prod_{t \neq s}(e_s - e_t)} = 0$ *where* $\forall s \neq t, e_s \neq e_t$*.*

Proof: The partial fraction decomposition of the first term is

$$\frac{1}{\prod_{q \neq 1}(e_1 - e_q)} = \sum_{q \neq 1} \left(\frac{B_q}{e_1 - e_q}\right)$$

where $B_q = \frac{1}{\prod_{p \neq q, p \neq 1}(e_q - e_p)}$.

Namely,

$$\frac{1}{\prod_{q \neq 1}(e_1 - e_q)} = -\sum_{q \neq 1} \left(\frac{1}{\prod_{p \neq q}(e_q - e_p)}\right)$$

We have

$$\sum_s \frac{1}{\prod_{t \neq s}(e_s - e_t)} = \frac{1}{\prod_{q \neq 1}(e_1 - e_q)} + \sum_{q \neq 1} \left(\frac{1}{\prod_{p \neq q}(e_q - e_p)}\right) = 0$$

■

Lemma 5 *For all* $\mu \leq \nu - 2$*, we have* $\sum_{s=1}^{\nu} \frac{(e_s)^\mu}{\prod_{t \neq s}(e_s - e_t)} = 0$.

Proof: Base case: When $\nu = 2, \mu = 0$, obviously

$$\frac{1}{e_1 - e_2} + \frac{1}{e_2 - e_1} = 0$$

Assume the lemma holds for $\nu = p$ and all $\mu \leq \nu - 2$, that is $\sum_{s=1}^{\nu} \frac{e_s^{\mu}}{\prod_{t \neq s}(e_s - e_t)} = 0$. When $\nu = p + 1, \mu = 0$, by Lemma 4 we have

$$\sum_{s=1}^{p+1} \frac{1}{\prod_{t \neq s}(e_s - e_t)} = 0$$

Assume $\sum_{s=1}^{p+1} \frac{e_s^q}{\prod_{t \neq s}(e_s - e_t)} = 0$ for all $\mu = q, q \leq p - 2$. For $\mu = q + 1$,

$$\begin{aligned}\sum_{s=1}^{p+1} \frac{e_s^{q+1}}{\prod_{t \neq s}(e_s - e_t)} =& \sum_{s=1}^{p+1} \frac{e_s^q e_{p+1}}{\prod_{t \neq s}(e_s - e_t)} + \sum_{s=1}^{p+1} \frac{e_s^q (e_s - e_{p+1})}{\prod_{t \neq s}(e_s - e_t)} \\ =& e_{p+1} \sum_{s=1}^{p+1} \frac{e_s^q}{\prod_{t \neq s}(e_s - e_t)} + \sum_{s=1}^{p} \frac{e_s^q}{\prod_{t \neq s}(e_s - e_t)} = 0\end{aligned}$$

The last equality is obtained from the induction hypotheses. ■

Lemma 6 *Let $f(x)$ be any polynomial of degree $\nu - 2$, then* $\sum_{s=1}^{\nu} \frac{f(e_s)}{\prod_{t \neq s}(e_s - e_t)} = 0$.

This can be easily derived from Lemma 5.

Remaining proof for Theorem 1

Now we are ready to prove that $\mathbf{H}^k \vec{\beta^*} = 0$. Note that the degree of the numerator of β_r^* is $2k - 3$ (see Equation (3)). Let $[\mathbf{H}^k]_i$ denote the i-th row of $\mathbf{H}^k$. We have the following calculations.

$$\begin{aligned}[][\mathbf{H}^k]_1 \vec{\beta^*} =& \sum_{r=1}^{2k} \frac{\prod_{p=1}^{2k-3}(pe_r + 2k - 2 - p)}{\prod_{q \neq r}(e_r - e_q)} = 0 \\ [\mathbf{H}^k]_2 \vec{\beta^*} =& \sum_{r=1}^{2k} \frac{\prod_{p=1}^{2k-3} e_r (pe_r + 2k - 2 - p)}{\prod_{q \neq r}(e_r - e_q)} = 0\end{aligned}$$

For any $2 < i \leq 2k - 1$, we have

$$\begin{aligned}&[\mathbf{H}^k]_i \vec{\beta^*} \\ =& \sum_{r=1}^{2k} \frac{e_r (1 - e_r)^{i-2}}{\prod_{p=1}^{i-2}(pe_r + 2k - 2 - p)} \frac{\prod_{p=1}^{2k-3}(pe_r + 2k - 2 - p)}{\prod_{q \neq r}(e_r - e_q)} \\ =& \sum_{r=1}^{2k} \frac{e_r (1 - e_r)^{i-2} \prod_{p=i-1}^{2k-3}(pe_r + 2k - 2 - p)}{\prod_{q \neq r}(e_r - e_q)} = 0\end{aligned}$$

The last equality is obtained by letting $v = 2k - 2$ in Lemma 6. Therefore, $\mathbf{H}^k \vec{\beta^*} = 0$. Note that $\vec{\beta^*}$ is also the solution for less than $2k - 1$ alternatives. The theorem follows after applying Lemma 2.

Theorem 2 *For $k = 2$, and any $m \geq 4$, the 2-PL is identifiable.*
Proof: We will apply Lemma 1 to prove the theorem. That is, we will show that for all non-degenerate $\vec{\theta}^{(1)}, \vec{\theta}^{(2)}, \vec{\theta}^{(3)}, \vec{\theta}^{(4)}$ such that $\text{rank}(\mathbf{F}_4^2) = 4$. We recall that $\mathbf{F}_4^2$ is a 24×4 matrix. Instead of proving $\text{rank}(\mathbf{F}_4^2) = 4$ directly, we will first obtain a 4×4 matrix $\mathbf{F}^* = T \times \mathbf{F}_4^2$ by linearly combining some row vectors of $\mathbf{F}_4^2$ via a 4×24 matrix T. Then, we show that $\text{rank}(\mathbf{F}^*) = 4$, which implies that $\text{rank}(\mathbf{F}_4^2) = 4$.

For simplicity we use $[e_r, b_r, c_r, d_r]^\top$ to denote the parameter of rth Plackett-Luce component for a_1, a_2, a_3, a_4 respectively. Namely,

$$\begin{bmatrix} \vec{\theta}^{(1)} & \vec{\theta}^{(2)} & \vec{\theta}^{(3)} & \vec{\theta}^{(4)} \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 & e_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix}$$

where for each $r \leq 4$, $\vec{\omega}^{(r)}$ is a row vector. We further let $\vec{1} = [1, 1, 1, 1]$. For proof convenience we define 5 row vectors.

$$\begin{aligned} \vec{1} &= [1, 1, 1, 1] \\ \vec{\omega}^{(1)} &= [e_1, e_2, e_3, e_4] \\ \vec{\omega}^{(2)} &= [b_1, b_2, b_3, d_3] \\ \vec{\omega}^{(3)} &= [c_1, c_2, c_3, c_4] \\ \vec{\omega}^{(4)} &= [d_1, d_2, d_3, d_4] \end{aligned}$$

Clearly we have $\sum_{i=1}^4 \vec{\omega}^{(i)} = \vec{1}$. Therefore, if there exist three $\vec{\omega}$'s, for example $\{\vec{\omega}^{(1)}, \vec{\omega}^{(2)}, \vec{\omega}^{(3)}\}$, such that $\{\vec{\omega}^{(1)}, \vec{\omega}^{(2)}, \vec{\omega}^{(3)}\}$ and $\vec{1}$ are linearly independent, then $\text{rank}(\mathbf{F}_4^2) = 4$ because each $\vec{\omega}^{(i)}$ corresponds to the probability of a_i being ranked at the top, which means that $\vec{\omega}^{(i)}$ is a linear combination of rows in $\mathbf{F}_4^2$. Because $\vec{\theta}^{(1)}, \vec{\theta}^{(2)}, \vec{\theta}^{(3)}, \vec{\theta}^{(4)}$ is non-degenerate, at least one of $\{\vec{\omega}^{(1)}, \vec{\omega}^{(2)}, \vec{\omega}^{(3)}, \vec{\omega}^{(4)}\}$ is linearly independent of $\vec{1}$. W.l.o.g. suppose $\vec{\omega}^{(1)}$ is linearly independent of $\vec{1}$. This means that not all of e_1, e_2, e_3, e_4 are equal. The theorem will be proved in the following two cases.
Case 1. $\vec{\omega}^{(2)}$, $\vec{\omega}^{(3)}$, and $\vec{\omega}^{(4)}$ are all linear combinations of $\vec{1}$ and $\vec{\omega}^{(1)}$.
Case 2. There exists a $\vec{\omega}^{(i)}$ (where $i \in \{2, 3, 4\}$) that is linearly independent of $\vec{1}$ and $\vec{\omega}^{(1)}$.
Case 1. For all $i = 2, 3, 4$ we can rewrite $\vec{\omega}^{(i)} = p_i \vec{\omega}^{(1)} + q_i$ for some constants p_i, q_i. More precisely, for all $r = 1, 2, 3, 4$ we have:

$$b_r = p_2 e_r + q_2 \tag{5}$$
$$c_r = p_3 e_r + q_3 \tag{6}$$
$$d_r = p_4 e_r + q_4 \tag{7}$$

Because $\vec{\omega}^{(1)} + \vec{\omega}^{(2)} + \vec{\omega}^{(3)} + \vec{\omega}^{(4)} = \vec{1}$, we have

$$p_2 + p_3 + p_4 = -1 \tag{8}$$
$$q_2 + q_3 + q_4 = 1 \tag{9}$$

In this case for each $r \leq 4$, the r-th column of $\mathbf{F}_4^2$, which is $f_4(\vec{\theta}^{(r)})$, is a function of e_r. Because the $\vec{\theta}$'s are non-degenerate, e_1, e_2, e_3, e_4 must be pairwise different.

We assume $p_2 \neq 0$ and $q_2 \neq 1$ for all subcases of **Case 1** (This will be denoted as **Case 1 Assumption**). The following claim shows that there exists p_i, q_i where $i \in \{2, 3, 4\}$ satisfying this condition. If $i \neq 2$ we can switch the row of alternatives a_2 and a_i. Then the assumption holds.

Claim 2 *There exists* $i \in 2, 3, 4$ *which satisfy the following conditions:*

- $q_i \neq 1$
- $p_i \neq 0$

Proof: Suppose for all $i = 2, 3, 4$, $q_i = 1$ or $p_i = 0$.

If $p_i = 0$, q_i must be positive because b_r, c_r, d_r are all positive. If $p_i \neq 0$, Then $q_i = 1$ due to the assumption above. So $q_i > 0$ for all $i = 2, 3, 4$. If there exists i s.t. $q_i = 1$, then (9) does not hold. So for all i, $q_i \neq 1$. Then $p_i = 0$ holds for all $i \in \{2, 3, 4\}$, which violates (8). ■

Case 1.1. $p_2 + q_2 \neq 0$ and $p_2 + q_2 \neq 1$.

For this case we first define a 4×4 matrix $\hat{\mathbf{F}}$ as follows.

$\hat{\mathbf{F}}$	Moments
$\begin{bmatrix} 1 & 1 & 1 & 1 \\ e_1 & e_2 & e_3 & e_4 \\ \frac{e_1 b_1}{1-b_1} & \frac{e_2 b_2}{1-b_2} & \frac{e_3 b_3}{1-b_3} & \frac{e_4 b_4}{1-b_4} \\ \frac{e_1 b_1}{1-e_1} & \frac{e_2 b_2}{1-e_2} & \frac{e_3 b_3}{1-e_3} & \frac{e_4 b_4}{1-e_4} \end{bmatrix}$	$\vec{1}$ $a_1 \succ$ others $a_2 \succ a_1 \succ$ others $a_1 \succ a_2 \succ$ others

We use $\vec{1}$ and $\vec{\omega}^{(1)}$ as the first two rows. $\vec{\omega}^{(1)}$ corresponds to the probability that a_1 is ranked in the top. We call such a probability a *moment*. Each moment is the sum of probabilities of some rankings. For example, the "$a_1 \succ$ others" moment is the total probability for $\{V \in \mathcal{L}(\mathcal{A}) : a_1 \text{ is ranked at the top of } V\}$. It follows that there exists a 4×24 matrix $\hat{T}$ such that $\hat{\mathbf{F}} = \hat{T} \times \mathbf{F}_4^2$.

Define

$$\begin{aligned}\vec{\theta}^{(b)} &= [\frac{1}{1-b_1}, \frac{1}{1-b_2}, \frac{1}{1-b_3}, \frac{1}{1-b_4}] \\ &= [\frac{1}{1-p_2 e_1 - q_2}, \frac{1}{1-p_2 e_2 - q_2}, \frac{1}{1-p_2 e_3 - q_2}, \frac{1}{1-p_2 e_4 - q_2}]\end{aligned}$$

and

$$\vec{\theta}^{(e)} = [\frac{1}{1-e_1}, \frac{1}{1-e_2}, \frac{1}{1-e_3}, \frac{1}{1-e_4}] \tag{10}$$

And define $\mathbf{F}^* = \begin{bmatrix} \vec{1} \\ \vec{\omega}^{(1)} \\ \vec{\theta}^{(b)} \\ \vec{\theta}^{(e)} \end{bmatrix}$. It can be verified that $\hat{\mathbf{F}} = T^* \times \mathbf{F}^*$, where

$$T^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{p_2} & -1 & \frac{1-q_2}{p_2} & 0 \\ -(p_2+q_2) & -p_2 & 0 & p_2+q_2 \end{bmatrix}$$

Because **Case 1.1** assumes that $p_2 + q_2 \neq 0$ and by Case 1 Assumption $p_2 \neq 0$, $q_2 \neq 1$, we have that T^* is invertible. Therefore, $\mathbf{F}^* = (T^*)^{-1} \times \hat{\mathbf{F}}$, which means that $\mathbf{F}^* = T \times \mathbf{F}_4^2$ for some 4×24 matrix T.

We now prove that rank($\mathbf{F}^*$) $= 4$. For the sake of contradiction, suppose that rank($\mathbf{F}^*$) < 4. It follows that there exist a nonzero row vector $\vec{t} = [t_1, t_2, t_3, t_4]$, such that $\vec{t}\mathbf{F}^* = 0$. This means that for all $r \leq 4$,

$$t_1 + t_2 e_r + \frac{t_3}{1 - p_2 e_r - q_2} + \frac{t_4}{1 - e_r} = 0$$

Let

$$f(x) = t_1 + t_2 x + \frac{t_3}{1 - p_2 x - q_2} + \frac{t_4}{1 - x}$$

Let $g(x) = (1 - p_2 x - q_2)(1 - x) f(x)$. We recall that e_1, e_2, e_3, e_4 are four roots of $f(x)$, which means that they are also the four roots of $g(x)$. Now we will verify that not all coefficients of $f(x)$ are zero. Suppose all coefficients of x in $f(x)$ are zero, then $g(x) = 0$ holds for all x. By assigning x to different values, we have

$$g(1) = t_4(1 - p_2 - q_2) = 0$$
$$g(\frac{1 - q_2}{p_2}) = \frac{t_3(p_2 + q_2 - 1)}{p_2} = 0$$

By Case 1.1 assumption $p_2 + q_2 \neq 1$, we have $t_3 = t_4 = 0$. Then from $f(x) = t_1 + t_2 x = 0$ holds for all x, we have $t_1 = t_2 = 0$, which is a contradiction.

We note that the degree of $g(x)$ is 3. Therefore, due to the Fundamental Theorem of Algebra, $g(x)$ has at most three different roots. This means that e_1, e_2, e_3, e_4 are not pairwise different, which is a contradiction. Therefore, rank($\mathbf{F}^*$) $= 4$, which means that rank($\mathbf{F}_4^2$) $= 4$.

Case 1.2. $p_2 + q_2 = 1$.

If we can find an alternative a_i, such that p_i and q_i satisfy the following conditions:

- $p_i \neq 0$
- $q_i \neq 1$
- $p_i + q_i \neq 0$
- $p_i + q_i \neq 1$

Then we can use a_i as a_2, which belongs to **Case 1.1**. Otherwise we have the following claim.

Claim 3 *If for $i \in \{3,4\}$, p_i and q_i satisfy one of the following conditions*

1. $p_i = 0$
2. $p_i \neq 0$, $q_i = 1$
3. $p_i + q_i = 0$
4. $p_i + q_i = 1$

We claim that there exists $i \in \{3,4\}$ s.t. p_i, q_i satisfy condition 2, namely $p_i \neq 0$, $q_i = 1$.

Proof: Suppose $p_i = 0$, then $q_i > 0$ because $p_i e_1 + q_i$ is a parameter in a Plackett-Luce component. If for $i = 3,4$, p_i and q_i satisfy any of conditions 1, 3 or 4, then $q_i \geq -p_i$ ($q_i > 0$ for condition 1, $q_i = -p_i$ for condition 3, $q_i = 1 - p_i > -p_1$ for condition 4). As $\sum_{i=2}^4 p_i = -1$, $\sum_{i=2}^4 q_i \geq 1 - \sum_{i=2}^4 p_i = 2$, which contradicts that $\sum_{i=2}^4 q_i = 1$. ■

Without loss of generality we let $p_3 \neq 0$ and $q_3 = 1$. We now construct $\hat{\mathbf{F}}$ as is shown in the following table.

$\hat{\mathbf{F}}$	Moments
$\begin{bmatrix} 1 & 1 & 1 & 1 \\ e_1 & e_2 & e_3 & e_4 \\ \frac{e_1 b_1}{1-e_1} & \frac{e_2 b_2}{1-e_2} & \frac{e_3 b_3}{1-e_3} & \frac{e_4 b_4}{1-e_4} \\ \frac{c_1 b_1}{1-c_1} & \frac{c_2 b_2}{1-c_2} & \frac{c_3 b_3}{1-c_3} & \frac{c_4 b_4}{1-c_4} \end{bmatrix}$	$\vec{1}$ $a_1 \succ$ others $a_1 \succ a_2 \succ$ others $a_3 \succ a_2 \succ$ others

We define $\vec{\theta}^{(b)}$ the same way as in **Case 1.1**, and define

$$\vec{\theta}^{(c)} = [\frac{1}{e_1}, \frac{1}{e_2}, \frac{1}{e_3}, \frac{1}{e_4}]$$

Define

$$\mathbf{F}^* = \begin{bmatrix} \vec{1} \\ \vec{\omega}^{(1)} \\ \vec{\theta}^{(e)} \\ \vec{\theta}^{(c)} \end{bmatrix}$$

We will show that $\hat{\mathbf{F}} = T^* \times \mathbf{F}^*$ where T^* has full rank.

For all $r = 1,2,3,4$

$$\frac{c_r b_r}{1-c_r} = \frac{(p_3 e_r + q_3)(p_2 e_r + q_2)}{1 - p_3 e_r - q_3} = \frac{(p_3 e_r + 1)(p_2 e_r + 1 - p_2)}{-p_3 e_r} = -p_2 e_r + (p_2 - 1 - \frac{p_2}{p_3}) - \frac{1-p_2}{p_3 e_r}$$

So

$$\hat{\mathbf{F}} = \begin{bmatrix} \vec{1} \\ \vec{\omega}^{(1)} \\ -\vec{1} - p_2\vec{\omega}^{(1)} + \vec{\theta}^{(e)} \\ (p_2 - 1 - \frac{p_2}{p_3})\vec{1} - p_2\vec{\omega}^{(1)} - \frac{1-p_2}{p_3}\vec{\theta}^{(c)} \end{bmatrix}$$

Suppose $p_2 \neq 1$, we have $\hat{\mathbf{F}} = T^* \times \mathbf{F}^*$ where

$$T^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & -p_2 & 1 & 0 \\ p_2 - 1 - \frac{p_2}{p_3} & -p_2 & 0 & -\frac{1-p_2}{p_3} \end{bmatrix}$$

which is full rank. So $\mathrm{rank}(\mathbf{F}^*) = \mathrm{rank}(\hat{\mathbf{F}})$.

If $\mathrm{rank}(\mathbf{F}_4^2) \leq 3$, then there is at least one column in $\mathbf{F}_4^2$ dependent of the other columns. As all rows in $\hat{\mathbf{F}}$ are linear combinations of rows in $\mathbf{F}_4^2$, there is also at least one column in $\hat{\mathbf{F}}$ dependent of the other columns. Therefore we have $\mathrm{rank}(\hat{\mathbf{F}}) \leq 3$. Further we have $\mathrm{rank}(\mathbf{F}^*) \leq 3$. Therefore, there exists a nonzero row vector $\vec{t} = [t_1, t_2, t_3, t_4]$, s.t.

$$\vec{t}\mathbf{F}^* = 0$$

Namely, for all $r \leq 4$,

$$t_1 + t_2 e_r + \frac{t_3}{1 - e_r} + \frac{t_4}{e_r} = 0$$

Let

$$f(x) = t_1 + t_2 x + \frac{t_3}{1 - x} + \frac{t_4}{x} = 0$$
$$g(x) = x(1 - x)f(x) = x(1 - x)(t_1 + t_2) + t_3 x + t_4(1 - x)$$

If any of the coefficients in $f(x)$ is nonzero, then $g(x)$ is a polynomial of degree at most 3. There will be a maximum of 3 different roots. Since this equation holds for e_r where $r = 1, 2, 3, 4$, there exists $s \neq t$ s.t. $e_s = e_t$. Otherwise $g(x) = f(x) = 0$ for all x. We have

$$g(0) = t_4 = 0$$
$$g(1) = t_3 = 0$$

Substitute $t_3 = t_4 = 0$ into $f(x)$, we have $f(x) = t_1 + t_2 x = 0$ for all x. So $t_1 = t_2 = 0$. This contradicts the nonzero requirement of $\vec{t}$. Therefore there exists $s \neq t$ s.t. $e_s = e_t$. From (5)(6)(7) we have $\vec{\theta}^{(s)} = \vec{\theta}^{(t)}$, which is a contradition.

If $p_2 = 1$, from the assumption of **Case 1.2** $q_2 = 0$. So $b_r = e_r$ for $r = 1, 2, 3, 4$. Then from (8) we have $p_4 = -p_3 - 2$ and from (9) we have $q_4 = 0$. Since p_4 and q_4 satisfy one of the four conditions in Claim 3, we can show it must satisfy Condition 4. ($q_4 = 0$ violates Condition 2. If it satisfies Condition 1 or 3, then $p_4 = 0$. Then $d_r = p_4 a_r + q_4 = 0$, which is impossible.) So $p_4 = 1$, and $p_3 = -3$. This is the case where $\vec{\omega}^{(1)} = \vec{\omega}^{(2)} = \vec{\omega}^{(4)}$ and $\vec{\omega}^{(3)} = 1 - 3\vec{\omega}^{(1)}$. For this case, we use a_3 as a_1. After the transformation, we have $\vec{\omega}^{(2)} = \vec{\omega}^{(3)} = \vec{\omega}^{(4)} = \frac{1-\vec{\omega}^{(1)}}{3}$. We claim that this lemma holds for a more general case where $p_i + q_i = 0$ for $i = 2, 3, 4$. It is easy to check that $p_i = -\frac{1}{3}$ and $q_i = \frac{1}{3}$ belongs to this case.

Claim 4 *For all* $r = 1, 2, 3, 4$, *if*

$$\vec{\theta}^{(r)} = \begin{bmatrix} e_r \\ b_r \\ c_r \\ d_r \end{bmatrix} = \begin{bmatrix} e_r \\ p_2 e_r - p_2 \\ p_3 e_r - p_3 \\ -(1 + p_2 + p_3)e_r + (1 + p_2 + p_3) \end{bmatrix} \tag{11}$$

The model is identifiable.

Proof: We first show a claim, which is useful to the proof.

Claim 5 *Under the settings of* (11), $-1 < p_2, p_3 < 0$, $-1 < p_2 + p_3 < 0$.

Proof: From the definition of Plackett-Luce model, $0 < e_r, b_r, c_r, d_r < 1$. From (11), we have $p_2 = \frac{b_r}{e_r - 1}$. Since $b_r > 0$ and $e_r < 1$, $p_2 < 0$. Similarly we have $p_3 < 0$ and $-(1 + p_2 + p_3) < 0$. So $-1 < p_2 + p_3 < 0$. Then we have $p_2 > -1 - p_3$. So $-1 - p_3 < p_2 < 0, p_3 > -1$. Similarly we have $p_2 > -1$.

■

In this case, we construct $\hat{\mathbf{F}}$ in the following way.

$\hat{\mathbf{F}}$	Moments
$\begin{bmatrix} 1 & 1 & 1 & 1 \\ e_1 & e_2 & e_3 & e_4 \\ \frac{e_1 b_1}{1-b_1} & \frac{e_2 b_2}{1-b_2} & \frac{e_3 b_3}{1-b_3} & \frac{e_4 b_4}{1-b_4} \\ \frac{e_1 b_1 c_1}{(1-b_1)(1-b_1-c_1)} & \frac{e_2 b_2 c_2}{(1-b_2)(1-b_2-c_2)} & \frac{e_3 b_3 c_3}{(1-b_3)(1-b_3-c_3)} & \frac{e_4 b_4 c_4}{(1-b_4)(1-b_4-c_4)} \end{bmatrix}$	$\vec{1}$ $a_1 \succ$ others $a_2 \succ a_1 \succ$ others $a_2 \succ a_3 \succ a_1 \succ a_4$

Define $\vec{\theta}^{(b)}$ the same way as in **Case 1.1**

$$\begin{aligned}\vec{\theta}^{(b)} &= [\frac{1}{1-b_1}, \frac{1}{1-b_2}, \frac{1}{1-b_3}, \frac{1}{1-b_4}] \\ &= [\frac{1}{1-p_2 e_1 + p_2}, \frac{1}{1-p_2 e_2 + p_2}, \frac{1}{1-p_2 e_3 + p_2}, \frac{1}{1-p_2 e_4 + p_2}]\end{aligned}$$

And define

$$\begin{aligned}\vec{\theta}^{(bc)} =& [\frac{1}{1-(p_2+p_3)e_1 + p_2 + p_3}, \frac{1}{1-(p_2+p_3)e_2 + p_2 + p_3}, \\ & \frac{1}{1-(p_2+p_3)e_3 + p_2 + p_3}, \frac{1}{1-(p_2+p_3)e_4 + p_2 + p_3}]\end{aligned}$$

Further define

$$\mathbf{F}^* = \begin{bmatrix} \vec{1} \\ \vec{\omega}^{(1)} \\ \vec{\theta}^{(b)} \\ \vec{\theta}^{(bc)} \end{bmatrix}$$

We will show $\hat{\mathbf{F}} = T^* \times \mathbf{F}^*$ where T^* has full rank.

The last two rows of $\hat{\mathbf{F}}$

$$\begin{aligned}
\frac{e_r b_r}{1-b_r} &= -e_r - \frac{1}{p_2} + \frac{1+p_2}{p_2(1-p_2e_r+p_2)} \\
\frac{e_r b_r c_r}{(1-b_r)(1-b_r-c_r)} &= \frac{e_r(p_2e_r-p_2)(p_3e_r-p_3)}{(1-p_2e_r+p_2)(1-(p_2+p_3)e_r+p_2+p_3)} \\
&= \frac{p_2p_3e_r(e_r-1)^2}{(1-p_2e_r+p_2)(1-(p_2+p_3)e_r+p_2+p_3)} \\
&= \frac{p_3(2p_2+p_3)}{p_2(p_2+p_3)^2} + \frac{p_3}{p_2+p_3}e_r - \frac{(1+p_2)}{p_2(1-p_2e_r+p_2)} \\
&+ \frac{p_2(1+p_2+p_3)}{(1-(p_2+p_3)e_r+p_2+p_3)(p_2+p_3)^2}
\end{aligned}$$

So

$$\hat{\mathbf{F}} = \begin{bmatrix} \vec{1} \\ \vec{\omega}^{(1)} \\ -\frac{1}{p_2}\vec{1} - \vec{\omega}^{(1)} + \frac{1+p_2}{p_2}\vec{\theta}^{(b)} \\ \frac{p_3(2p_2+p_3)}{p_2(p_2+p_3)^2}\vec{1} + \frac{p_3}{p_2+p_3}\vec{\omega}^{(1)} - \frac{1+p_2}{p_2}\vec{\theta}^{(b)} + \frac{p_2(1+p_2+p_3)}{(p_2+p_3)^2}\vec{\theta}^{(bc)} \end{bmatrix}$$

Then we have $\hat{\mathbf{F}} = T^* \times \mathbf{F}^*$ where

$$T^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{1}{p_2} & -1 & \frac{1+p_2}{p_2} & 0 \\ \frac{p_3(2p_2+p_3)}{p_2(p_2+p_3)^2} & \frac{p_3}{p_2+p_3} & -\frac{1+p_2}{p_2} & \frac{p_2(1+p_2+p_3)}{(p_2+p_3)^2} \end{bmatrix}$$

From Claim 5, we have $-1 < p_2 < 0$ and $-1 < p_2 + p_3 < 0$, so $\frac{1+p_2}{p_2} \neq 0$ and $\frac{p_2(1+p_2+p_3)}{(p_2+p_3)^2} \neq 0$. So T has full rank. Then $\mathrm{rank}(\mathbf{F}^*) = \mathrm{rank}(\hat{\mathbf{F}})$.

If $\mathrm{rank}(\mathbf{F}_4^2) \leq 3$, then there is at least one column in $\mathbf{F}_4^2$ dependent of other columns. As all rows in $\hat{\mathbf{F}}$ are linear combinations of rows in $\mathbf{F}_4^2$, $\mathrm{rank}(\hat{\mathbf{F}}) \leq 3$. Since $\mathrm{rank}(\mathbf{F}^*) = \mathrm{rank}(\hat{\mathbf{F}})$, we have $\mathrm{rank}(\mathbf{F}^*) \leq 3$. Therefore, there exists a nonzero row vector $\vec{t} = [t_1, t_2, t_3, t_4]$, s.t.

$$\vec{t}\mathbf{F}^* = 0$$

Namely, for all $r \leq 4$,

$$t_1 + t_2e_r + \frac{t_3}{1-p_2a_r+p_2} + \frac{t_4}{1-(p_2+p_3)e_r+p_2+p_3} = 0$$

Let

$$\begin{aligned}
f(x) &= t_1 + t_2x + \frac{t_3}{1-p_2x+p_2} + \frac{t_4}{1-(p_2+p_3)x+p_2+p_3} \\
g(x) &= (1-p_2x+p_2)(1-(p_2+p_3)x+p_2+p_3)(t_1+t_2x) \\
&\quad + t_3(1-(p_2+p_3)x+p_2+p_3) + t_4(1-p_2x+p_2)
\end{aligned}$$

If any of the coefficients of $g(x)$ is nonzero, then $g(x)$ is a polynomial of degree at most 3. There will be a maximum of 3 different roots. As the equation holds for all e_r where $r = 1, 2, 3, 4$. There exists $s \neq t$ s.t. $e_s = e_t$. Otherwise $g(x) = f(x) = 0$ for all x. We have

$$g(\frac{1+p_2}{p_2}) = \frac{-t_3 p_3}{p_2} = 0$$
$$g(\frac{1+p_2+p_3}{p_2+p_3}) = \frac{t_4 p_3}{p_2+p_3} = 0$$

From Claim 5 we know $p_2, p_3 < 0$ and $p_2 + p_3 < 0$. So $t_3 = t_4 = 0$. Substitute it into $f(x)$ we have $f(x) = t_1 + t_2 x = 0$ for all x. So $t_1 = t_2 = 0$. This contradicts the nonzero requirement of $\vec{t}$. Therefore there exists $s \neq t$ s.t. $e_s = e_t$. According to (5)(6)(7) we have $\vec{\theta}^{(s)} = \vec{\theta}^{(t)}$, which is a contradition.

■

Case 1.3. $p_2 + q_2 = 0$.

If there exists i such that $p_i + q_i = 1$, then we can use a_i as a_2 and the proof is done in **Case 1.2**. It may still be possible to find another i such that p_i, q_i satisfy the following two conditions:

1. $p_i \neq 0$ and $q_i \neq 1$;
2. $p_i + q_i \neq 0$.

If we can find another i to satisfy the two conditions, then the proof is done in **Case 1.1**. Then we can proceed by assuming that the two conditions are not satisfied by any i. We will prove that the only possibility is $p_i + q_i = 0$ for $i = 2, 3, 4$.

Suppose for $i = 3, 4$, p_i and q_i violate Condition 1. If $p_i = 0$, then $q_i > 0$. If at least one of them has $q_i = 1$, then $e_r + b_r + c_r + d_r > 1$, which is impossible. If both alternatives violates Condition 1 and $p_3 = p_4 = 0$, then $0 < q_3, q_4 < 1$. According to (8) $p_2 = -1$. As $p_2 + q_2 = 0$, we have $q_2 = 1$. From (9), $q_3 + q_4 = 2$, which is impossible. So there exists $i \in \{3, 4\}$ such that $p_i + q_i = 0$. Then from $\sum_i \theta_i^r = 1$ we obtain the only case we left out, which is

$$\begin{aligned}
&e_r \\
&b_r = p_2 e_r - p_2 \\
&c_r = p_3 e_r - p_3 \\
&d_r = -(1 + p_2 + p_3) e_r + (1 + p_2 + p_3)
\end{aligned}$$

This case has been proved in Claim 4.

Case 2: There exists $\vec{\omega}^{(i)}$ that is linearly independent of $\vec{1}$ and $\vec{\omega}^{(1)}$. W.l.o.g. let it be $\vec{\omega}^{(2)}$. Define matrix

$$\mathbf{G} = \begin{bmatrix} \vec{1} \\ \vec{\omega}^{(1)} \\ \vec{\omega}^{(2)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ e_1 & e_2 & e_3 & e_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix}$$

The rank of $\mathbf{G}$ is 3. Since $\mathbf{G}$ is constructed using linear combinations of rows in $\mathbf{F}_4^2$, the rank of $\mathbf{F}_4^2$ is at least 3.

If $\vec{\omega}^{(3)}$ or $\vec{\omega}^{(4)}$ is independent of rows in $\mathbf{G}$, then we can append it to $\mathbf{G}$ as the fourth row so that the rank of the new matrix is 4. Then $\mathbf{F}_4^2$ is full rank. So we only need to consider the case where $\vec{\omega}^{(3)}$ and $\vec{\omega}^{(4)}$ are linearly dependent of $\vec{1}$, $\vec{\omega}^{(1)}$, and $\vec{\omega}^{(2)}$. Let

$$\vec{\omega}^{(3)} = x_3\vec{\omega}^{(1)} + y_3\vec{\omega}^{(2)} + z_3\vec{1} \tag{12}$$

$$\vec{\omega}^{(4)} = x_4\vec{\omega}^{(1)} + y_4\vec{\omega}^{(2)} + z_4\vec{1} \tag{13}$$

where $x_3 + x_4 = -1$, $y_3 + y_4 = -1$, $z_3 + z_4 = 1$.

Claim 6 *There exists $i \in \{3, 4\}$ such that $x_i + z_i \neq 0$.*

Proof: If in the current setting $\exists i \in \{3, 4\}$ s.t. $x_i + z_i \neq 0$, then the proof is done. If in the current setting $x_3 + z_3 = x_4 + z_4 = 0$, but $\exists i \in \{3, 4\}$ s.t. $y_i + z_i = 0$, then we can switch the role of e_r and b_r, namely

$$\vec{\omega}^{(3)} = y_3\vec{\omega}^{(1)} + x_3\vec{\omega}^{(2)} + z_3\vec{1}$$

$$\vec{\omega}^{(4)} = y_4\vec{\omega}^{(1)} + x_4\vec{\omega}^{(2)} + z_4\vec{1}$$

Then the proof is done. If for all $i \in \{3, 4\}$ we have $x_i + z_i = 0$ and $y_i + z_i = 0$, then we switch the role of e_r and c_r and get

$$\vec{\omega}^{(3)} = \frac{1}{x_3}(\vec{\omega}^{(1)} - y_3\vec{\omega}^{(2)} - z_3\vec{1})$$

$$\vec{\omega}^{(4)} = \frac{1}{x_4}(\vec{\omega}^{(1)} - y_4\vec{\omega}^{(2)} - z_4\vec{1})$$

If $\frac{1-z_3}{x_3} \neq 0$, namely $z_3 \neq 1$, the proof is done. Suppose $z_3 = 1$, then $x_3 = y_3 = -1$. We have $\vec{\omega}^{(3)} = 1 - \vec{\omega}^{(1)} - \vec{\omega}^{(2)}$. Then $\vec{\omega}^{(4)} = \vec{0}$, which is impossible. ■

Without loss of generality let $x_3 + z_3 \neq 0$. Similar to the previous proofs, we want to construct a matrix $\mathbf{G}'$ using linear combinations of rows from $\mathbf{F}_4^2$. Let the first 3 rows for $\mathbf{G}'$ to be $\mathbf{G}$. Then rank$(\mathbf{G}') \geq 3$. Since rank$(\mathbf{F}_4^2) \leq 3$ and all rows in $\mathbf{G}'$ are linear combinations of rows in $\mathbf{F}_4^2$, we have rank$(\mathbf{G}') \leq 3$. So rank$(\mathbf{G}') = 3$. This means that any linear combinations of rows in $\mathbf{F}_4^2$ is linearly dependent of rows in $\mathbf{G}$.

Consider the moment where a_1 is ranked at the top and a_2 is ranked at the second position. Then $[\frac{e_1 b_1}{1-e_1}, \frac{e_2 b_2}{1-e_2}, \frac{e_3 b_3}{1-e_3}, \frac{e_4 b_4}{1-e_4}]$ is linearly dependent of $\mathbf{G}$. Adding $\vec{\omega}^{(2)}$ to it, we have

$$\vec{\theta}^{(eb)} = [\frac{b_1}{1-e_1}, \frac{b_2}{1-e_2}, \frac{b_3}{1-e_3}, \frac{b_4}{1-e_4}]$$

which is linearly dependent of $\mathbf{G}$.

Similarly consider the moment that a_1 is ranked at the top and a_3 is ranked at the second position. We obtain $[\frac{e_1 c_1}{1-e_1}, \frac{e_2 c_2}{1-e_2}, \frac{e_3 c_3}{1-e_3}, \frac{e_4 c_4}{1-e_4}]$. Add $\vec{\omega}^{(3)}$ to it, we get

$$\vec{\theta}^{(ec)} = [\frac{c_1}{1-e_1}, \frac{c_2}{1-e_2}, \frac{c_3}{1-e_3}, \frac{c_4}{1-e_4}]$$

which is linearly dependent of $\mathbf{G}$.

Recall from (10)

$$\vec{\theta}^{(e)} = [\frac{1}{1-e_1}, \frac{1}{1-e_2}, \frac{1}{1-e_3}, \frac{1}{1-e_4}]$$

Then

$$\begin{aligned}\vec{\theta}^{(ec)} &= [\frac{x_3e_1+y_3b_1+z_3}{1-e_1}, \frac{x_3e_2+y_3b_2+z_3}{1-e_2}, \frac{x_3e_3+y_3b_3+z_3}{1-e_3}, \frac{x_3e_4+y_3b_4+z_3}{1-e_4}] \\ &= (x_3+z_3)\vec{\theta}^{(e)} + y_3\vec{\theta}^{(eb)} - x_3\vec{1}\end{aligned}$$

Because both $\vec{\theta}^{(eb)}$ and $\vec{\theta}^{(ec)}$ are linearly dependent of $\mathbf{G}$, $\vec{\theta}^{(e)}$ is also linearly dependent of $\mathbf{G}$. Make it the 4th row of $\mathbf{G}'$. Suppose the rank of $\mathbf{G}'$ is still 3. We will first prove this lemma under the assumption below, and then discuss the case where the assumption does not hold.

Assumption 1: Suppose $\vec{1}, \vec{\omega}^{(1)}, \vec{\theta}^{(e)}$ are linearly independent.

Then $\vec{\omega}^{(2)}$ is a linear combination of $\vec{1}, \vec{\omega}^{(1)}$ and $\vec{\theta}^{(e)}$. We write $\vec{\omega}^{(2)} = s_1 + s_2\vec{\omega}^{(1)} + s_3\vec{\theta}^{(e)}$ for some constants s_1, s_2, s_3. We have $s_3 \neq 0$ because $\vec{\omega}^{(2)}$ is linearly independent of $\vec{1}$ and $\vec{\omega}^{(1)}$. Elementwise, for $r = 1, 2, 3, 4$ we have

$$b_r = s_1 + s_2e_r + \frac{s_3}{1-e_r} \tag{14}$$

Let

$$\mathbf{G}'' = \begin{bmatrix}\mathbf{G} \\ \vec{\theta}^{(eb)}\end{bmatrix}$$

$\vec{\theta}^{(eb)}$ is linearly dependent of $\mathbf{G}$. There exists a non-zero vector $\vec{h} = [h_1, h_2, h_3, h_4]$ such that $\vec{h} \cdot \mathbf{G}'' = 0$. Namely $h_1\vec{1} + h_2\vec{\omega}^{(1)} + h_3\vec{\omega}^{(2)} + h_4\vec{\theta}^{(eb)} = 0$. Elementwise, for all $r = 1, 2, 3, 4$

$$h_1 + h_2e_r + h_3b_r + h_4\frac{b_r}{1-e_r} = 0 \tag{15}$$

where $h_4 \neq 0$ because otherwise $\text{rank}(\mathbf{G}) = 2$. Substitute (14) into (15), and multiply both sides of it by $(1-e_r)^2$, we get

$$(h_1 + h_2e_r + h_3b_r)(1-e_r)^2 + h_4(s_1 + s_2e_r)(1-e_r) + h_4s_3 = 0$$

Let

$$f(x) = (h_1 + h_2e_r + h_3b_r)(1-e_r)^2 + h_4(s_1 + s_2e_r)(1-e_r) + h_4s_3$$

We claim that not all coefficients of x are zero, because $f(1) = h_4s_3 \neq 0$ ($s_3 \neq 0$ and $h_4 \neq 0$ by assumption). Then there are a maximum of 3 different roots, each of which uniquely determines b_r by (14). This means that there are at least two identical components. Namely $\exists s \neq t$ s.t. $\vec{\theta}^{(s)} = \vec{\theta}^{(t)}$.

If Assumption 1 does not hold, namely $\vec{\theta}^{(e)}$ is a linear combination of $\vec{1}$ and $\vec{\omega}^{(1)}$, let

$$\frac{1}{1-e_r} = p_5e_r + q_5 \tag{16}$$

Define

$$f(x) = \frac{1}{1-x} - p_5 x - q_5$$

If $f(x)$ has only 1 root or two identical roots between 0 and 1, then all columns of $\mathbf{G}$ have identical e_r-s. This means $\vec{\omega}^{(1)}$ is dependent of $\vec{1}$, which is a contradiction. So we only consider the situation where $f(x)$ has two different roots between 0 and 1, denoted by u_1 and u_2 ($u_1 \neq u_2$). Because e_1, e_2, e_3, e_4 are roots of $f(x)$, there must be at least two identical e_r's, with the value u_1 or u_2.

Substitute (16) into $\vec{\theta}^{(eb)}$, we have $\vec{\theta}^{(eb)} = [b_1(p_5 e_1 + q_5), b_2(p_5 e_2 + q_5), b_3(p_5 e_3 + q_5), b_4(p_5 e_4 + q_5)]$, which is linearly dependent of $\mathbf{G}$. So there exists nonzero vector $\vec{\gamma_1} = [\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14}]$ such that

$$\gamma_{11} + \gamma_{12} e_r + \gamma_{13} b_r + \gamma_{14} b_r (p_5 e_r + q_5) = 0$$

From which we get

$$(\gamma_{13} + \gamma_{14} p_5 e_r + \gamma_{14} q_5) b_r = -(\gamma_{11} + \gamma_{12} e_r) \tag{17}$$

We recall that $e_r = u_1$ or $e_r = u_2$ for $r = 1, 2, 3, 4$. Since $u_1 \neq u_2$, there exists $i \in \{1, 2\}$ s.t. $\gamma_{13} + \gamma_{14} p_5 u_i + \gamma_{14} q_5 \neq 0$. W.l.o.g. let it be u_1. If at least two of the e_r's are u_1, without loss of generality let $e_1 = e_2 = u_1$. Then using (17) we know $b_1 = b_2 = \frac{-(\gamma_{11}+\gamma_{12}u_1)}{(\gamma_{13}+\gamma_{14}p_5 u_1+\gamma_{14}q_5)}$. From (12)(13) we can further obtain $c_1 = c_2$ and $d_1 = d_2$. So $\vec{\theta}^{(1)} = \vec{\theta}^{(2)}$, which is a contradiction.

If there is only one of the e_r's, which is u_1, w.l.o.g. let $e_1 = u_1$ and $e_2 = e_3 = e_4 = u_2$. We consider the moment where a_2 is ranked at the top and a_1 the second, which is $[\frac{e_1 b_1}{1-b_1}, \frac{e_2 b_2}{1-b_2}, \frac{e_3 b_3}{1-b_3}, \frac{e_4 b_4}{1-b_4}]$. Add $\vec{\omega}^{(1)}$ to it and we have $\vec{\theta}^{(be)} = [\frac{e_1}{1-b_1}, \frac{e_2}{1-b_2}, \frac{e_3}{1-b_3}, \frac{e_4}{1-b_4}]$, which is linearly dependent of $\mathbf{G}$. So there exists nonzero vector $\vec{\gamma_2} = [\gamma_{21}, \gamma_{22}, \gamma_{23}, \gamma_{24}]$ such that

$$\gamma_{21} + \gamma_{22} e_r + \gamma_{23} b_r + \gamma_{24} \frac{e_r}{1-b_r} = 0 \tag{18}$$

Let

$$f(x) = \gamma_{21} + \gamma_{22} u_2 + \gamma_{23} x + \gamma_{24} \frac{u_2}{1-x}$$
$$g(x) = (1-x) f(x) = (1-x)(\gamma_{21} + \gamma_{22} u_2 + \gamma_{23} x) + \gamma_{24} u_2$$

If any coefficient of $g(x)$ is nonzero, then $g(x)$ has at most 2 different roots. As $g(x) = 0$ holds for b_2, b_3, b_4, $\exists s \neq t$ s.t. $b_s = b_t$. Since $e_s = e_t = u_2$, from (12)(13) we know $c_s = c_t$ and $d_s = d_t$. So $\vec{\theta}^{(s)} = \vec{\theta}^{(t)}$. Otherwise we have $g(x) = f(x) = 0$ for all x. So

$$g(1) = \gamma_{24} u_2 = 0$$

Since $0 < u_2 < 1$, we have $\gamma_{24} = 0$. Substitute it into $f(x)$ we have $f(x) = \gamma_{21} + \gamma_{22} u_2 + \gamma_{23} x = 0$ holds for all x. So we have $\gamma_{21} + \gamma_{22} u_2 = 0$ and $\gamma_{23} = 0$. Substitute $\gamma_{23} = \gamma_{24} = 0$ into (18) we get $\gamma_{21} + \gamma_{22} e_r = 0$, which holds for both $e_r = u_1$ and $e_r = u_2$. As $u_1 \neq u_2$, we have $\gamma_{22} = 0$. Then we have $\gamma_{21} = 0$. This contradicts the nonzero requirement of $\vec{\gamma_2}$. So there exists $s \neq t$ s.t. $\vec{\theta}^{(s)} = \vec{\theta}^{(t)}$, which is a contradiction.

■

Lemma 3 *Given a random utility model $\mathcal{M}(\vec{\theta})$ over a set of m alternatives $\mathcal{A}$, let $\mathcal{A}_1, \mathcal{A}_2$ be two non-overlapping subsets of $\mathcal{A}$, namely $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{A}$ and $\mathcal{A}_1 \cap \mathcal{A}_2 = \emptyset$. Let V_1, V_2 be rankings over $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively, then we have* $\Pr(V_1, V_2|\vec{\theta}) = \Pr(V_1|\vec{\theta})\Pr(V_2|\vec{\theta})$.

Proof: In an RUM, given a ground truth utility $\vec{\theta} = [\theta_1, \theta_2, \ldots, \theta_m]$ and a distribution $\mu_i(\cdot|\theta_i)$ for each alternative, an agent samples a random utility X_i for each alternative independently with probability density function $\mu_i(\cdot|\theta_i)$. The probability of the ranking $a_{i_1} \succ a_{i_2} \succ \cdots \succ a_{i_m}$ is

$$\begin{aligned}\Pr(a_{i_1} \succ \cdots \succ a_{i_m}|\vec{\theta}) &= \Pr(X_{i_1} > X_{i_2} > \cdots > X_{i_m}) \\ &= \int_{-\infty}^{\infty}\int_{x_{i_m}}^{\infty}\cdots\int_{x_{i_2}}^{\infty}\mu_{i_m}(x_{i_m})\mu_{i_{m-1}}(x_{i_{m-1}})\ldots\mu_{i_1}(x_{i_1})dx_{i_1}dx_{i_2}\ldots dx_{i_m}\end{aligned}$$

W.l.o.g. we let $i_1 = 1, \ldots, i_m = m$. Let $\mathcal{S}_{X_1>X_2>\cdots>X_m}$ denote the subspace of $\mathbb{R}^m$ where $X_1 > X_2 > \cdots > X_m$ and let $\mu(\vec{x}|\vec{\theta})$ denote $\mu_m(x_m)\mu_{m-1}(x_{m-1})\ldots\mu_1(x_1)$. Thus we have

$$\Pr(a_1 \succ \cdots \succ a_m|\vec{\theta}) = \int_{\mathcal{S}_{X_1>X_2>\cdots>X_m}}\mu(\vec{x}|\vec{\theta})d\vec{x}$$

We first prove the following claim.

Claim 7 *Given a random utility model $\mathcal{M}(\vec{\theta})$, for any parameter $\vec{\theta}$ and any $\mathcal{A}_s \subseteq \mathcal{A}$, we let $\vec{\theta}_s$ denote the components of $\vec{\theta}$ for alternatives in $\mathcal{A}_s$, and let V_s be a full ranking over A_s (which is a partial ranking over $\mathcal{A}$). Then we have* $\Pr(V_s|\vec{\theta}) = \Pr(V_s|\vec{\theta}_s)$.

Proof: Let m_s be the number of alternatives in $\mathcal{A}_s$. Let $\mathcal{S}_{X_1>X_2>\cdots>X_{m_s}}$ denote the subspace of $\mathbb{R}^{m_s}$ where $X_1 > X_2 > \cdots > X_{m_s}$. W.l.o.g. let V_s be $a_1 \succ a_2 \cdots \succ a_{m_s}$. Then we have

$$\begin{aligned}\Pr(V_s|\vec{\theta}) &= \int_{\mathcal{S}_{X_1>X_2>\cdots>X_{m_s}}\times\mathbb{R}^{m-m_s}}\mu(\vec{x}|\vec{\theta})d\vec{x} \\ &= \int_{-\infty}^{\infty}\int_{x_{m_s}}^{\infty}\cdots\int_{x_2}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mu_m(x_m)\ldots\mu_1(x_1)dx_{m_s+1}\cdots dx_m dx_1\ldots dx_{m_s} \\ &= \int_{-\infty}^{\infty}\int_{x_{m_s}}^{\infty}\cdots\int_{x_2}^{\infty}\mu_{m_s}(x_{m_s})\mu_{m_s-1}(x_{m_s-1})\ldots\mu_1(x_1)dx_1dx_2\ldots dx_{m_s} \\ &= \int_{\mathcal{S}_{X_1>X_2>\cdots>X_{m_s}}}\mu(\vec{x}_s|\vec{\theta}_s)d\vec{x} = \Pr(V_s|\vec{\theta}_s)\end{aligned}$$

■

Let $\mathcal{A}_1 = \{a_{11}, a_{12}, \ldots, a_{1m_1}\}$ and $\mathcal{A}_2 = \{a_{21}, a_{22}, \ldots, a_{2m_2}\}$. Without loss of generality we let V_1 and V_2 be $a_{11} \succ a_{12} \succ \cdots \succ a_{1m_1}$ and $a_{21} \succ a_{22} \succ \cdots \succ a_{2m_2}$ respectively. For any $\vec{\theta}$, let $\vec{\theta}_1$ denote the subvector of $\vec{\theta}$ on $\mathcal{A}_1$. Let $\mathcal{S}_1$ denote $\mathcal{S}_{X_{11}>X_{12}>\cdots>X_{1m_1}}$. $\vec{\theta}_2$ and $\mathcal{S}_2$ are defined similarly. According to Claim 7, we have $\Pr(V_1|\vec{\theta}) = \Pr(V_1|\vec{\theta}_1) = \int_{\mathcal{S}_1} \mu(\vec{x}_1|\vec{\theta}_1)d\vec{x}_1$ and $\Pr(V_2|\vec{\theta}) = \Pr(V_2|\vec{\theta}_2) = \int_{\mathcal{S}_2} \mu(\vec{x}_2|\vec{\theta}_2)d\vec{x}_2$. Then we have

$$\begin{aligned}
\Pr(V_1, V_2|\vec{\theta}) &= \int_{\mathcal{S}_1 \times \mathcal{S}_2 \times \mathbb{R}^{m-m_1-m_2}} \mu(\vec{x}|\vec{\theta})d\vec{x} \\
&= \int_{\mathcal{S}_1 \times \mathcal{S}_2} \mu(\vec{x}_1, \vec{x}_2|\vec{\theta}_1, \vec{\theta}_2)d\vec{x} && \text{(Claim 7)} \\
&= \int_{\mathcal{S}_1} \int_{\mathcal{S}_2} \mu(\vec{x}_1|\vec{\theta}_1)\mu(\vec{x}_2|\vec{\theta}_2)d\vec{x}_1 d\vec{x}_2 && \text{(Fubini's Theorem)} \\
&= \int_{\mathcal{S}_1} \mu(\vec{x}_1|\vec{\theta}_1)d\vec{x}_1 \int_{\mathcal{S}_2} \mu(\vec{x}_2|\vec{\theta}_2)d\vec{x}_2 \\
&= \Pr(V_1|\vec{\theta}_1)\Pr(V_2|\vec{\theta}_2)
\end{aligned}$$

■

Theorem 4 *Algorithm 1 is consistent w.r.t.* 2*-PL, where there exists* $\epsilon > 0$ *such that each parameter is in* $[\epsilon, 1]$.

Proof: We will check all assumptions in Theorem 3.1 in **?**.

Assumption 3.1: Strict Stationarity: the $(n\times 1)$ random vectors $\{v_t; -\infty < t < \infty\}$ form a strictly stationary process with sample space $\mathcal{S} \subseteq \mathbb{R}^n$.

As the data are generated i.i.d., the process is strict stationary.

Assumption 3.2: Regularity Conditions for $g(\cdot,\cdot)$: the function $g : \mathcal{S} \times \Theta \to \mathbb{R}^q$ where $q < \infty$, satisfies: (i) it is continuous on Θ for each $P \in \mathcal{S}$; (ii) $E[g(P, \vec{\theta})]$ exists and is finite for every $\theta \in \Theta$; (iii) $E[g(P, \vec{\theta})]$ is continuous on Θ.

Our moment conditions satisfy all the regularity conditions since $g(P, \vec{\theta})$ is continuous on Θ and bounded in $[-1, 1]^9$.

Assumption 3.3: Population Moment Condition. The random vector v_t and the parameter vector θ_0 satisfy the $(q \times 1)$ population moment condition: $E[g(P, \theta_0)] = 0$.

This assumption holds by the definition of our GMM.

Assumption 3.4 Global Identification. $E[g(P, \vec{\theta'})] \neq 0$ for all $\vec{\theta'} \in \Theta$ such that $\vec{\theta'} \neq \theta_0$.

This is proved in Theorem 2.

Assumption 3.7 Properties of the Weighting Matrix. W_t is a positive semi-definite matrix which converges in probability to the positive definite matrix of constants W.

This holds because $W = I$.

Assumption 3.8 Ergodicity. The random process $\{v_t; -\infty < t < \infty\}$ is ergodic.

Since the data are generated i.i.d., the process is ergodic.

Assumption 3.9 Compactness of Θ. Θ is a compact set.

$\Theta = [\epsilon, 1]^9$ is compact.

Assumption 3.10 Domination of $g(P, \vec{\theta})$. $E[\sup_{\theta\in\Theta} ||g(P, \vec{\theta})||] < \infty$.

This assumption holds because all moment conditions are finite.

Theorem 3.1 Consistency of the Parameter Estimator. If Assumptions 3.1-3.4 and 3.7-3.10 hold then $\hat{\theta}_T \xrightarrow{p} \theta_0$

■