# Estimating Preferences from Coarse Rankings: Partially Rank-Ordered Logit

Pinjas Albagli*

May 16, 2026

**Abstract**

This paper shows that the multichoice logit model for coarse-ranking data of van Ophem et al. (1999) is algebraically equivalent to a simpler approach: treating the coarse ranking as an incomplete ranking with ties and following the procedure of Allison and Christakis (1994) for handling tied ranks in the well-known rank-ordered logit model. This equivalence has immediate practical value: researchers can estimate discrete choice models from coarse-ranking data using readily available software without the need of implementing specialized routines.

## 1 Introduction

Empirical researchers frequently encounter discrete choice data where decision makers reveal preferences over alternatives incompletely. For instance, survey respondents may be asked to identify their top 3 choices from a longer list rather than to provide a complete ranking. In such settings, we observe what can be called a "coarse ranking"—a partial order partitioning the choice set into preferred and non-preferred items—rather than a complete ordering. Despite their prevalence in applied work, coarse rankings present a methodological challenge. The standard rank-ordered logit model (Beggs et al., 1981; Hausman and Ruud, 1987) requires observing complete ordinal rankings, while the multinomial logit focuses only on the best alternative. This middle-ground setting corresponds to what we call the "partially rank-ordered logit" model

*London School of Economics and Centre for Economic Performance. Email: p.m.albagli@lse.ac.uk. I thank Alan Manning and Thomas Monk for insightful comments and discussion.

Van Ophem et al. (1999) developed the "multichoice logit" specifically for this type of data, but this model has seen limited application, possibly due to its apparent complexity and lack of implementation in standard statistical packages. This paper shows that the multichoice logit of van Ophem et al. (1999) is algebraically equivalent to a simpler approach: treating the coarse partition as an incomplete ranking with ties and following the procedure of Allison and Christakis (1994) for handling tied ranks in the rank-ordered logit model. This equivalence has immediate practical value: researchers can estimate discrete choice models from coarse-ranking data using readily available software (e.g., Stata's '`cmrologit`' command) without the need of implementing specialized routines.[1]

For concreteness, consider a discrete choice setting where individual $i \in \mathcal{I}$ faces $|\mathcal{J}| = J \geq 2$ alternatives in finite set $\mathcal{J}$.[2] In an additive random utility model (ARUM hereafter), the decision maker's induced preferences over $\mathcal{J}$ are represented by an indirect utility function of the form

$$u_{ij} = \delta_{ij} + \varepsilon_{ij}, \tag{1}$$

where alternatives are indexed by $j \in \mathcal{J}$, $\delta_{ij} = \delta\left(x_{ij}\right)$ is the systematic or deterministic component of utility, $x_{ij} \in \mathcal{X} \subseteq \mathbb{R}^K$ is a vector measuring alternative $j$'s $K$ utility-relevant characteristics, and the random taste shock $\varepsilon_{ij}$ represents the idiosyncratic component of utility.[3]

Different assumptions on the joint probability distribution of the taste shocks $\{\varepsilon_{ij}\}_{j\in\mathcal{J},i\in\mathcal{I}}$ lead to different ARUMs. For instance, the assumption that the random taste shocks are independently and identically distributed (*iid* hereafter) across $i$ and $j$ following a normal distribution leads to the probit ARUM, while assuming an *iid* type-I extreme value distribution leads to the logit ARUM with its well-known independence of irrelevant alternatives (IIA henceforth) property (see, e.g., McFadden, 1974).

While standard specifications assume the decision maker chooses only one alternative—i.e., $y_{i1} \in \arg\max_{j\in\mathcal{J}} u_{ij}$—, the ARUM class of models is, in fact, more general in the sense it can accommodate different choice situations or data structures under alternative behavioral assumptions specifying what is being chosen and how. A prominent example is the case of rank-ordered data, where an ordinal ranking of the alternatives in the choice set is observed. Such a setting may arise, for example, with stated preference survey data, where consumers are asked to rank a set of alternatives, typically after being presented with their corresponding attributes (see, e.g., Beggs et al., 1981; Hausman and Ruud, 1987; Hajivassiliou and Ruud, 1994).

This paper is concerned with estimation of the preference parameters $\beta \in \mathbb{R}^K$ in the logit

[1]See, e.g., StataCorp (2023).

[2]Set $\mathcal{I}$ can be specified either as a continuum with measure $I$ or as a finite set with size $I$. See Berry (1994) for details.

[3]Preferences are defined over bundles of attributes $x \in \mathcal{X}$. Each alternative $j \in \mathcal{J}$ is mapped into an attribute vector $x_{ij} \in \mathcal{X}$. Therefore, preferences over $\mathcal{X}$ induce preferences over $\mathcal{J}$.

ARUM under some parameterization $\delta_{ij} = \delta(x_{ij}; \beta)$ when only a coarse ranking of the alternatives in $\mathcal{J}$ is observed. In particular, for each decision maker $i$ we observe a partition of $\mathcal{J}$, $\mathcal{A}_i = \{A_i, \overline{A}_i\}$, such that each alternative in $A_i$ is preferred to any alternative in $\overline{A}_i$, while the rankings of the $|A_i| = n_i \geq 1$ "chosen" (or most preferred) alternatives in $A_i$ and the $|\overline{A}_i| = \overline{n}_i \leq J-1$ "unchosen" (or least preferred) alternatives in $\overline{A}_i$ are unobserved. That is, all the information we have is that $u_{ij} > u_{i\ell} \; \forall \, j \in A_i, \forall \, \ell \in \overline{A}_i$. Since the partition $\mathcal{A}_i$ induces a partial preference order over the choice set $\mathcal{J}$, we call this framework the "partially rank-ordered logit" model.

This situation arises naturally in applied research using stated preference data. The survey setting where respondents are asked what are the $n$ most important issues among a list of $J$ prompts is a case in point (e.g., Hatton, 2021). Another typical example are vignette studies where participants are asked to choose their $n$ most preferred items—or, alternatively, those they would consider when making a choice—from the $J$ options hypothetically available (see, for example, van Ophem et al., 1999; Yang et al., 2002). The framework, however, is not limited to the analysis of stated preference data, as it also accommodates—under certain circumstances—revealed preference data from application processes in contexts such as job search or school choice. For instance, application portfolio decisions in the Chade and Smith (2006) simultaneous search framework can be shown to reduce to a stopping rule that determines portfolio size, $n$, and a choice of the $n$ most preferred alternatives among the $J$ available—e.g., job vacancies in the case of job search or colleges in the case of school choice—under uniform admission probabilities (see Albagli, 2025).[4]

While an econometric model for this particular setting was developed by van Ophem et al. (1999), it has not been widely applied in the literature, possibly because of a lack of availability in standard statistical packages and its apparent complexity deriving from an unnecessarily intricate notation. In this paper, we show that an equivalent econometric model results from applying the method proposed by Allison and Christakis (1994) to deal with ties in the Beggs et al. (1981) rank-ordered logit model after constructing an instrumental two-tier ranking with all the chosen alternatives $j \in A_i$ tied in the first rank and the remaining unchosen alternatives $j \in \overline{A}_i$ tied in the second. This device—which, to the best of our knowledge, has not been proposed before in the literature—has the advantage that it can be readily implemented using standard maximum likelihood routines that estimate the rank-ordered logit with ties and are already available in statistical packages such as Stata (StataCorp, 2023).

The remainder of the paper is organized as follows. We discuss the multichoice and rank-

[4]The job search platform setting where firms bid on candidates studied by Roussille and Scuderi (2025) is another example where a partially rank-ordered ARUM is appropriate for revealed preference data. They propose an alternative estimation method based on indirect maximization of an integral representation of the likelihood function. Their minorize-maximize algorithm provides a computationally tractable alternative for cases with large numbers of alternatives.

ordered logit models in Section 2 and Sections 3.1 to 3.3, respectively. In Section 3.4, we show that coarse-ranking data can be seen as a special case of incomplete rankings with ties. Finally, we show in Section 4 that the resulting rank-ordered logit model with two distinct ranks is algebraically equivalent to the multichoice logit model. Section 5 concludes.

## 2 The multichoice logit model

In this section, we derive van Ophem et al.'s (1999) multichoice logit model using a more intuitive combination-set notation that avoids the need to sum over multiple indices by summing over sets instead. A central piece of this derivation is the set of all combinations of size $s$ from a given set $S$, $\mathcal{R}_s(S)$, defined in Definition 1 below. Note that an arbitrary element of $\mathcal{R}_s(S)$ is simply a finite subset of $S$ with size $s$.

**Definition 1.** For finite set $S$ and $s \in \{1, \ldots, |S|\}$, let $\mathcal{R}_s(S) \equiv \binom{S}{s} \equiv \{\sigma \subseteq S : |\sigma| = s\}$ represent the set of all s-combinations of set $S$.

Consider the logit ARUM, i.e., equation (1) under the assumption that the taste shocks $\varepsilon_{ij}$ are *iid* type-I extreme value (or standard Gumbel) with cumulative distribution function (*cdf* hereafter)

$$F_{\varepsilon_j}(s) = \exp\left(-\exp\left(-s\right)\right). \tag{2}$$

For simplicity, assume the linear parameterization $\delta(x_{ij};\beta) = x_{ij}'\beta$ and suppose that we only observe a coarse ranking for each individual, so all the information available to estimate the preference parameters $\beta$ is that $u_{ij} > u_{i\ell}\ \forall\, j \in A_i, \forall\, \ell \in \overline{A}_i$, where $\mathcal{A}_i = \left\{A_i, \overline{A}_i\right\}$ is decision maker $i$'s coarse-ranking partition of the choice set $\mathcal{J}$. Note that this setting corresponds to the partially rank-ordered logit framework defined in the introduction.

Let $J \equiv |\mathcal{J}| \geq 2$ and $n_i \equiv |A_i| \in \{1, \ldots, J-1\}$.[5] The likelihood contribution of the observation corresponding to decision maker $i$ in the multichoice logit model is given by

$$\begin{aligned} L_i^{\text{ML}} &\equiv \mathbb{P}\left(\bigcap_{j \in A_i, \ell \in \overline{A}_i} \{u_{ij} > u_{i\ell}\}\right) \\ &= 1 - \mathbb{P}\left(\bigcup_{j \in A_i} \left\{\max_{\ell \in \overline{A}_i} u_{i\ell} \geq u_{ij}\right\}\right) \\ &= 1 - \sum_{\ell \in \overline{A}_i} \mathbb{P}\left(\bigcup_{j \in A_i} \left\{u_{i\ell} \geq \max_{m \in \{j\} \cup \overline{A}_i} u_{im}\right\}\right) \end{aligned}$$

[5] Recall that, by definition, $\overline{n}_i \equiv \left|\overline{A}_i\right| = J - n_i$. Therefore, $n_i \in \{1, \ldots, J-1\} \implies \overline{n}_i \in \{1, \ldots, J-1\}$.

$$= 1 - \sum_{\ell \in \overline{A}_i} \sum_{s=1}^{n_i} (-1)^{s+1} \sum_{\sigma \in \mathcal{R}_s(A_i)} \mathbb{P}\left( \bigcap_{j \in \sigma} \left\{ u_{i\ell} \geq \max_{m \in \{j\} \cup \overline{A}_i} u_{im} \right\} \right)$$

$$= 1 + \sum_{s=1}^{n_i} (-1)^s \sum_{\sigma \in \mathcal{R}_s(A_i)} \frac{\sum_{\ell \in \overline{A}_i} \exp(\delta_{i\ell})}{\sum_{m \in \sigma \cup \overline{A}_i} \exp(\delta_{im})}, \tag{3}$$

where the first line defines it as the probability of the intersection of all the events such that one of the chosen alternatives $j \in A_i$ is strictly preferred to one of the unchosen alternatives $\ell \in \overline{A}_i$. The first equality (in the second line) follows from De Morgan's laws and the equivalence of the events $\bigcap_{j \in A_i, \ell \in \overline{A}_i} \{u_{ij} > u_{i\ell}\}$ and $\bigcap_{j \in A_i} \{u_{ij} > \max_{\ell \in \overline{A}_i} u_{i\ell}\}$. The second equality folows from the equivalence of the events $\bigcup_{j \in A_i} \{\max_{\ell \in \overline{A}_i} u_{i\ell} \geq u_{ij}\}$ and $\bigcup_{\ell \in \overline{A}_i} \{\{u_{i\ell} = \max_{m \in \overline{A}_i} u_{im}\} \cap \bigcup_{j \in A_i} \{u_{i\ell} \geq u_{ij}\}\}$, and the fact that the latter is the union of a collection of disjoint events, each equivalent to an event of the form $\bigcup_{j \in A_i} \{u_{i\ell} \geq \max_{m \in \{j\} \cup \overline{A}_i} u_{im}\}$. The third equality follows directly from the inclusion-exclusion principle (see, e.g., Billingsley, 2012, Ch. 1, Sec. 2). Finally, the last equality follows from the equivalence of the events $\bigcap_{j \in \sigma} \{u_{i\ell} \geq \max_{m \in \{j\} \cup \overline{A}_i} u_{im}\}$ and $\{u_{i\ell} \geq \max_{m \in \sigma \cup \overline{A}_i} u_{im}\}$ for $\sigma \subseteq A_i$, the well-known closed-form result for the (singleton-)choice probabilities in the logit ARUM, and appropriately interchanging the order of summation.

# 3 Ties in the rank-ordered logit model

In this section, we briefly discuss Beggs et al.'s (1981) rank-ordered logit model for complete, strict rankings, along with the extensions proposed by Allison and Christakis (1994) to deal with incomplete rankings and ties, and show how this model can be used to estimate preferences from coarse-ranking data by means of a suitable artifice.

Again, consider the logit ARUM as specified in the previous subsection, but now assume that we observe a ranking of some or all the alternatives in the choice set. That is, we observe $r_i \colon S \to \{1, \ldots, |S|\}$, where $S \subseteq \mathcal{J}$ and $r_{ij} \equiv r_i(j)$ is the rank assigned to alternative $j$ by individual $i$. The function $r_i(\cdot)$ need not be injective—there can be ties in the sense that two or more alternatives are given the same rank—and its domain could be a proper subset of the choice set—the reported ranking can be incomplete in the sense that some alternatives are not ranked. In contrast, the true underlying ranking that represents decision maker $i$'s preferences, $\rho_i \colon \mathcal{J} \to \{1, \ldots, J\}$, is a bijection which, without loss of generality, we define in increasing order so that $\rho_{ij} < \rho_{i\ell} \iff u_{ij} > u_{i\ell}$. We assume that any ties in the reported ranking are induced by design as opposed to true indifference—e.g., survey respondents are asked to rank only the most and least preferred few alternatives, so everything else is tied in the middle. Therefore, $r_i(\cdot)$ truthfully—albeit partially—represents preferences in the sense that $r_{ij} < r_{i\ell} \implies u_{ij} > u_{i\ell}$

for $j, \ell \in S$. Observed ties in the ranking (when $r_{ij} = r_{i\ell}$) arise from survey design rather than true indifference, so the relationship between $u_{ij}$ and $u_{i\ell}$ for tied alternatives is not revealed by the data.

Just as notation was simplified in the previous subsection by introducing combination sets, permutation sets, defined in Definition 2 below, will make the notation more compact here. We can think of an arbitrary element of the $S$-permutation set of set $S$, $\mathcal{Q}(S)$, as a vector in $\mathbb{R}^{|S|}$ whose components are sequentially sampled without replacement from $S$. Armed with this notation, we can think of decision maker $i$'s true ranking as a vector in the set of $J$-permutations of the first $J$ natural numbers, i.e., $\boldsymbol{\rho}_i \equiv (\rho_{i1}, \ldots, \rho_{iJ})' \in \mathcal{Q}(\{1, \ldots, J\})$. In contrast, the reported ranking $\boldsymbol{r}_i = (r_{i1}, \ldots, r_{i|S|})'$ need not be a permutation of $\{1, \ldots, |S|\}$ due to the possibility of ties.

**Definition 2.** For finite set $S$, let $\mathcal{Q}(S) \equiv S! \equiv \left\{p \in S^{|S|} : p_k \neq p_\ell \;\forall\, k \neq \ell \in \{1, \ldots, |S|\}\right\}$ represent the set of all $|S|$-permutations of set $S$.

## 3.1 Complete, strict rankings

As a starting point, suppose that the observed ranking is strict and complete, truthfully representing decision maker $i$'s preferences, so that $r_i(\cdot) = \rho_i(\cdot)$ is a bijection. Notice that the marginal distribution of the idiosyncratic shocks given in equation (2) induces a marginal distribution on utilities with *cdf*

$$F_{u_{ij}}(s) \equiv \mathbb{P}\left(u_{ij} \leq s\right) = F_{\varepsilon_j}(s - \delta_{ij}). \tag{4}$$

Beggs et al. (1981) exploit the implication of the IIA property that the conditional distribution of the utility of a given alternative is independent of the ranking of the other alternatives, and show that the likelihood contribution of observation $i$ with ranking $\boldsymbol{\rho}_i$ is given by

$$L_i^{\text{ROL}} = \prod_{j \in \mathcal{J}} \frac{\exp\left(\delta_{ij}\right)}{\sum\limits_{k \in \Omega_{ij}} \exp\left(\delta_{ik}\right)}, \tag{5}$$

where $\Omega_{ij} \equiv \{\ell \in \mathcal{J} : \rho_{i\ell} \geq \rho_{ij}\}$ is the set of alternatives that $i$ does not strictly prefer to alternative $j$—i.e., it includes $j$ and any strictly worse alternatives.

Thinking in terms of the inverse mapping $y_i : \{1, \ldots, J\} \to \mathcal{J}$, where $y_{ik} \equiv y_i(k)$ represents (the identity of) the alternative that individual $i$ ranks in position $k \in \{1, \ldots, J\}$, will prove useful in extending the model to accommodate incomplete rankings and ties.[6] Note that the observed inverse ranking $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iJ})' \in \mathcal{Q}(\mathcal{J})$ is simply a $J$-permutation of the choice set $\mathcal{J}$.

[6]Notice these are equivalent representations of the same underlying data structure since $\rho_{ij} = k \iff y_{ik} = j$.

### 3.2 Incomplete, strict rankings

Now, following Allison and Christakis (1994), suppose that the observed ranking is strict but incomplete, with individuals ranking only their $|S| < |\mathcal{J}|$ most-preferred alternatives. Then, the domain of $r_i : S \to \{1, \ldots, |S|\}$ is a strict subset of the choice set, and $\boldsymbol{y}_i \in \mathcal{Q}(S)$. Since the observed ranking truthfully reveals preferences over alternatives $j \in S \subset \mathcal{J}$, the likelihood contribution of observation $i$ has the same form as the one in equation (5) but restricted to alternatives in $S$, i.e.,

$$L_i^{\text{ROL}} = \prod_{j \in S} \frac{\exp(\delta_{ij})}{\sum\limits_{k \in \Omega_{ij} \cup \overline{S}} \exp(\delta_{ik})} = \prod_{j=1}^{|S|} \frac{\exp(\delta_{iy_{ij}})}{\sum\limits_{k \in \Lambda_{ij} \cup \overline{S}} \exp(\delta_{ik})}, \tag{6}$$

where $\Omega_{ij} \equiv \{\ell \in S : r_{i\ell} \geq r_{ij}\} = \{y_{ij}, \ldots, y_{i|S|}\} \equiv \Lambda_{ij}$ and $\overline{S} \equiv \mathcal{J} \setminus S$. The last term in this individual likelihood is simply the probability of choosing the last-ranked alternative among all the remaining alternatives.

### 3.3 Weak rankings (ties)

As a final generalization, consider the case where observed rankings include ties, causing $r_i : S \to \{1, \ldots, |S|\}$ to be non-injective. For concreteness, suppose that individual $i$'s observed ranking of the $|S|$ alternatives in subset $S \subseteq J$ comprises $N_i \leq |S|$ distinct ranks. Let the correspondence $Y_i : \{1, \ldots, N_i\} \rightrightarrows \mathcal{P}(S)$—where $\mathcal{P}(S)$ is the power set of set $S$—identify the subset of alternatives tied in rank $j \in \{1, \ldots, N\}$, with shorthand notation $Y_{ij} \equiv Y_i(j)$. Let $d_{ij} \equiv |Y_{ij}|$ represent the number of alternatives tied in rank $j$.

Noting that the likelihood contribution for the no-ties case in equation (5) is isomorphic to the partial likelihood for survival data in the Cox (1972) relative risk model, Allison and Christakis (1994) propose emulating the treatment of tied failure times in that model to generalize (5) to the tied-ranks setting.[7] This treatment consists of considering every possible way of breaking the ties. To this effect, let $\Lambda_i(j, p, \ell) \equiv \bigcup_{s=j}^{N_i} Y_{is} \setminus \{p_1, \ldots, p_{\ell-1}\}$ for $j \in \{1, \ldots, N_i\}$, $p \in \mathcal{Q}(Y_{ij})$, and $\ell \in \{1, \ldots, d_{ij}\}$, with shorthand notation $\Lambda_{ijp\ell}$, represent the set of all ranked alternatives—including those with a rank higher than $j$—ranked no better than the $\ell^{\text{th}}$ least preferred alternative among those originally tied in position $j$ when the tie is broken according to permutation $p$.[8] Then, the likelihood contribution of individual $i$ is given by

$$L_i^{\text{ROL}} = \prod_{j=1}^{N_i} \sum_{p \in \mathcal{Q}(Y_{ij})} \prod_{\ell=1}^{d_{ij}} \frac{\exp(\delta_{ip_\ell})}{\sum\limits_{k \in \Lambda_{ijp\ell} \cup \overline{S}} \exp(\delta_{ik})}, \tag{7}$$

[7] See Appendix A for details.

[8] For notational consistency, set $\{p_1, \ldots, p_0\} \equiv \emptyset$. Also, note that $s \neq j \implies Y_{is} \setminus \{p_1, \ldots, p_{\ell-1}\} = Y_{is}$, $\forall p \in \mathcal{Q}(Y_{ij})$, $\forall \ell \in \{1, \ldots, d_{ij}\}$ since $Y_{is}$ and $Y_{ij}$ are disjoint sets.

which is simply the likelihood of all possible true rankings $(\rho_{i1}, \ldots, \rho_{i|S|})$ that are consistent with the observed ranking $(r_{i1}, \ldots, r_{i|S|})$. An example is presented in Appendix B to make matters less abstract.

### 3.4 Coarse rankings

Consider again the coarse-ranking setting from Section 2 where, for each individual, we observe only a coarse partition of the choice set, $\mathcal{A}_i = \{A_i, \overline{A}_i\}$ with $n_i \equiv |A_i|$ and $\overline{n}_i \equiv |\overline{A}_i| = J - n_i$. For concreteness, and abstracting away from the uncertainty induced by employer selection on the other side of the market and the dynamic aspects of job applications, suppose that a job seeker chooses an exogenously determined number of job vacancies to apply to, $A_i$, from a set of available vacancies, $\mathcal{J}$, according to the logit ARUM.

Job seeker $i$ applies to the $n_i$ highest-utility vacancies, so that $j \in A_i \iff \rho_{ij} \leq n_i$ for $j \in \mathcal{J}$. We observe a sample of application data for $I$ job seekers, including a job seeker identifier $i$, a job vacancy identifier $j$, and a dummy $a_{ij} \in \{0, 1\}$ indicating whether individual $i$ applied to vacancy $j$. The data structure may seem completely different from the one in Section 3.3, but it is, in fact, observationally equivalent to a special case. To see this, note that

$$a_{ij} = \begin{cases} 1 & \text{if } \rho_{ij} \leq n_i \\ 0 & \text{if } \rho_{ij} > n_i \end{cases}$$

contains exactly the same information as pseudo-ranking

$$r_{ij} = \begin{cases} 1 & \text{if } j \in A_i \\ 2 & \text{if } j \in \overline{A}_i, \end{cases}$$

which gives rise to a special case of the data structure in Section 3.3 with $N_i = 2$, $d_{i1} = n_i$, $d_{i2} = \overline{n}_i$, $Y_{i1} = A_i$, $Y_{i2} = \overline{A}_i$, $S = \mathcal{J}$, and $\overline{S} = \emptyset$. Our coarse-ranking setting, however, is also observationally equivalent to an even simpler data structure with $N_i = 1$, $d_{i1} = n_i$, $Y_{i1} = S = A_i$, and $\overline{S} = \overline{A}_i$, where only the vacancies in $A_i$ are ranked in position 1 and the discarded vacancies in $\overline{A}_i$ are not ranked at all. The reason is all the information we have boils down to $u_{ij} > u_{i\ell}, \forall j \in A_i, \forall \ell \in \overline{A}_i$, which can be equivalently represented by any of these data structures.

Adhering to the first representation with $N_i = 2$, the likelihood contribution in equation (7) simplifies to

$$L_i^{\text{ROL}} = \left[ \sum_{p \in \mathcal{Q}(A_i)} \prod_{\ell=1}^{n_i} \frac{\exp(\delta_{ip_\ell})}{\sum_{k \in \mathcal{J} \setminus \{p_1, \ldots, p_{\ell-1}\}} \exp(\delta_{ik})} \right] \left[ \sum_{p \in \mathcal{Q}(\overline{A}_i)} \prod_{\ell=1}^{\overline{n}_i} \frac{\exp(\delta_{ip_\ell})}{\sum_{k \in \overline{A}_i \setminus \{p_1, \ldots, p_{\ell-1}\}} \exp(\delta_{ik})} \right]$$

$$= \sum_{p \in \mathcal{Q}(A_i)} \prod_{\ell=1}^{n_i} \frac{\exp\left(\delta_{ip_\ell}\right)}{\sum\limits_{k \in \mathcal{J} \setminus \{p_1,\ldots,p_{\ell-1}\}} \exp\left(\delta_{ik}\right)}, \tag{8}$$

which is consistent with the second, simpler representation with $N_i = 1$. To see this, notice that the first equality in equation (8) factors the likelihood contribution as a product of the form $L_i^{\text{ROL}} = \mathcal{L}_{1i} \cdot \mathcal{L}_{2i}$, and the second term simplifies to

$$\begin{aligned}
\mathcal{L}_{2i} &= \sum_{p \in \mathcal{Q}\left(\overline{A}_i\right)} \prod_{\ell=1}^{\overline{n}_i} \frac{\exp\left(\delta_{ip_\ell}\right)}{\sum\limits_{k \in \overline{A}_i \setminus \{p_1,\ldots,p_{\ell-1}\}} \exp\left(\delta_{ik}\right)} \\
&= \sum_{q_1 \in \overline{A}_i} \frac{\exp\left(\delta_{iq_1}\right)}{\sum\limits_{k \in \overline{A}_i} \exp\left(\delta_{ik}\right)} \sum_{p \in \mathcal{Q}\left(\overline{A}_i \setminus \{q_1\}\right)} \prod_{\ell=1}^{\overline{n}_i - 1} \frac{\exp\left(\delta_{ip_\ell}\right)}{\sum\limits_{k \in \overline{A}_i \setminus (\{q_1\} \cup \{p_1,\ldots,p_{\ell-1}\})} \exp\left(\delta_{ik}\right)} \\
&= \sum_{q_1 \in \overline{A}_i} \frac{\exp\left(\delta_{iq_1}\right)}{\sum\limits_{k \in \overline{A}_i} \exp\left(\delta_{ik}\right)} \sum_{q_2 \in \overline{A}_i \setminus \{q_1\}} \frac{\exp\left(\delta_{iq_2}\right)}{\sum\limits_{k \in \overline{A}_i \setminus \{q_1\}} \exp\left(\delta_{ik}\right)} \sum_{p \in \mathcal{Q}\left(\overline{A}_i \setminus \{q_1,q_2\}\right)} \prod_{\ell=1}^{\overline{n}_i - 2} \frac{\exp\left(\delta_{ip_\ell}\right)}{\sum\limits_{k \in \overline{A}_i \setminus (\{q_1,q_2\} \cup \{p_1,\ldots,p_{\ell-1}\})} \exp\left(\delta_{ik}\right)} \\
&\vdots \\
&= \sum_{q_1 \in \overline{A}_i} \frac{\exp\left(\delta_{iq_1}\right)}{\sum\limits_{k \in \overline{A}_i} \exp\left(\delta_{ik}\right)} \cdots \sum_{q_{\overline{n}_i-1} \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-2}\right\}} \frac{\exp\left(\delta_{iq_{\overline{n}_i-1}}\right)}{\sum\limits_{k \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-2}\right\}} \exp\left(\delta_{ik}\right)} \sum_{p \in \mathcal{Q}\left(\overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-1}\right\}\right)} \prod_{\ell=1}^{1} \frac{\exp\left(\delta_{ip_\ell}\right)}{\sum\limits_{k \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-1}\right\}} \exp\left(\delta_{ik}\right)} \\
&= \sum_{q_1 \in \overline{A}_i} \frac{\exp\left(\delta_{iq_1}\right)}{\sum\limits_{k \in \overline{A}_i} \exp\left(\delta_{ik}\right)} \cdots \sum_{q_{\overline{n}_i-1} \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-2}\right\}} \frac{\exp\left(\delta_{iq_{\overline{n}_i-1}}\right)}{\sum\limits_{k \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-2}\right\}} \exp\left(\delta_{ik}\right)} \frac{\exp\left(\delta_{iq_{\overline{n}}}\right)}{\exp\left(\delta_{iq_{\overline{n}}}\right)} \\
&= \frac{\sum\limits_{q_1 \in \overline{A}_i} \exp\left(\delta_{iq_1}\right)}{\sum\limits_{k \in \overline{A}_i} \exp\left(\delta_{ik}\right)} \cdots \frac{\sum\limits_{q_{\overline{n}_i-1} \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-2}\right\}} \exp\left(\delta_{iq_{\overline{n}_i-1}}\right)}{\sum\limits_{k \in \overline{A}_i \setminus \left\{q_1,\ldots,q_{\overline{n}_i-2}\right\}} \exp\left(\delta_{k}\right)} \\
&= 1.
\end{aligned}$$

The second to fourth equalities follow from recursively factoring out the terms in each permutation. The fifth equality follows from the fact that $\mathcal{Q}\left(\overline{A}_i \setminus \{q_1,\ldots,q_{\overline{n}_i-1}\}\right) = \overline{A}_i \setminus \{q_1,\ldots,q_{\overline{n}_i-1}\} = \{q_{\overline{n}_i}\}$ is a singleton, containing the last term in permutation $q = (q_1,\ldots,q_{\overline{n}_i}) \in \mathcal{Q}\left(\overline{A}_i\right)$.

Thus far, we have shown that the treatment of ties in the rank-ordered logit model proposed by Allison and Christakis (1994) provides a valid approach to estimating the preference parameters from coarse-ranking data by simply treating the coarse partition of the choice set as a pseudo-ranking with two (or one) distinct ranks. This approach has the advantage that maximum likelihood estimation routines are readily available in software such as Stata (see, e.g., the entry for command `cmrologit` in StataCorp, 2023). Moreover, while we have used a revealed-preference setting to motivate the approach, the scope for its applications is wider, including many stated-preference settings typically found in applied research. The natural question arises of how this approach compares to the van Ophem et al. (1999) multichoice logit model. It turns out, as we show in the following section, they are equivalent.

# 4 Equivalence Result

*Remark.* Setting $f(\emptyset, S) = f(S, \emptyset) = f(\emptyset, \emptyset) = 0 \; \forall \, S \in \mathcal{P}(\mathcal{J})$ in Lemma 1 below is without loss of generality for the purpose of this paper since the proof of Proposition 1—where this lemma is invoked—does not deal with empty sets.

**Lemma 1.** *Let $\mathcal{J}$ be a finite set and denote its power set by $\mathcal{P}(\mathcal{J})$. The set function $f : \mathcal{P}(\mathcal{J}) \times \mathcal{P}(\mathcal{J}) \to \mathbb{R}$ defined by*

$$f\left(S_1, S_2\right) = \begin{cases} \dfrac{\sum\limits_{k \in S_1} \exp\left(\delta_k\right)}{\sum\limits_{\ell \in S_2} \exp\left(\delta_\ell\right)} & \text{if } S_1, S_2 \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

*satisfies the following properties.*

1. *For all nonempty $S_1, S_2 \subseteq \mathcal{J}$,*

$$\sum_{j \in S_1} f(\{j\}, S_2) = f(S_1, S_2).$$

2. *For all nonempty $S_1 \subseteq \mathcal{J}$, $S_2 \subseteq \mathcal{J} \setminus S_1$, and $S_3 \subseteq S_1$,*

$$f(S_3, S_1 \cup S_2) = 1 - f\big((S_1 \setminus S_3) \cup S_2, S_1 \cup S_2\big) \text{ and } f(S_1, S_1 \cup S_2) = 1 - f(S_2, S_1 \cup S_2).$$

3. *For all nonempty $S_1, S_2, S_3 \subseteq \mathcal{J}$,*

$$f(S_1, S_2) f(S_2, S_3) = f(S_1, S_3) \text{ and } f(S_1, S_2) f(S_3, S_1) = f(S_3, S_2).$$

4. *For all nonempty $S_1 \subseteq \mathcal{J}$ and $S_2 \subseteq \mathcal{J} \setminus S_1$, and for all $s \in \{1, \ldots, |S_1| - 1\}$,*

$$\sum_{\ell \in S_1} f(\{\ell\}, S_1 \cup S_2) \sum_{\sigma \in \mathcal{R}_s(S_1 \setminus \{\ell\})} f(S_2, \sigma \cup S_2) = \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, \sigma \cup S_2) - \binom{|S_1|}{s} f(S_2, S_1 \cup S_2),$$

*where $\mathcal{R}_s\left(S\right)$ denotes the set of $s$-combinations of set $S$.*

*Proof.* To establish property 1, let $S_1$ and $S_2$ be two nonempty subsets of $\mathcal{J}$ and notice that

$$\sum_{j \in S_1} f(\{j\}, S_2) = \sum_{j \in S_1} \frac{\exp(\delta_j)}{\sum\limits_{k \in S_2} \exp(\delta_k)}$$

$$= \frac{\sum\limits_{j \in S_1} \exp(\delta_j)}{\sum\limits_{k \in S_2} \exp(\delta_k)}$$

$$= f(S_1, S_2),$$

where the first and third equalities follow from the definition of $f(\cdot, \cdot)$.

Next, to establish the first part of property 2, let $S_1 \subseteq \mathcal{J}$ and $S_2 \subseteq \mathcal{J} \setminus S_1$ be two nonempty, disjoint subsets of $\mathcal{J}$. Let $S_3 \subseteq S_1$. Then,

$$f(S_3, S_1 \cup S_2) = \frac{\sum\limits_{j \in S_3} \exp(\delta_j)}{\sum\limits_{k \in S_1 \cup S_2} \exp(\delta_k)}$$

$$= 1 - \frac{\sum\limits_{j \in (S_1 \setminus S_3) \cup S_2} \exp(\delta_j)}{\sum\limits_{k \in S_1 \cup S_2} \exp(\delta_k)}$$

$$= 1 - f\big((S_1 \setminus S_3) \cup S_2, S_1 \cup S_2\big),$$

where, again, the first and third equalities follow from the definition of $f(\cdot, \cdot)$. The second part of property 2 directly follows as a special case with $S_3 = S_1$.

Now, let $S_1, S_2, S_3 \subseteq \mathcal{J}$ be three nonempty subsets of $\mathcal{J}$, and note that

$$f(S_1, S_2) f(S_2, S_3) = \frac{\sum\limits_{\ell \in S_1} \exp(\delta_\ell)}{\sum\limits_{\ell \in S_2} \exp(\delta_\ell)} \frac{\sum\limits_{\ell \in S_2} \exp(\delta_\ell)}{\sum\limits_{\ell \in S_3} \exp(\delta_\ell)}$$

$$= \frac{\sum\limits_{\ell \in S_1} \exp(\delta_\ell)}{\sum\limits_{\ell \in S_3} \exp(\delta_\ell)}$$

$$= f(S_1, S_3),$$

where the first and third equalities follow, once again, from the definition of $f(\cdot, \cdot)$. Moreover, notice that

$$f(S_1, S_2) f(S_3, S_1) = f(S_3, S_1) f(S_1, S_2)$$

$$= f(S_3, S_2)$$

by the first part of property 3 proved above, establishing the second part of property 3.

Finally, to prove property 4, let $S_1 \subseteq \mathcal{J}$ and $S_2 \subseteq \mathcal{J} \setminus S_1$ be two nonempty, disjoint subsets of $\mathcal{J}$. Fix $s \in \{1, \ldots, |S_1| - 1\}$ and note that

$$\begin{aligned}
\sum_{\ell \in S_1} f(\{\ell\}, S_1 \cup S_2) \sum_{\sigma \in \mathcal{R}_s(S_1 \setminus \{\ell\})} f(S_2, \sigma \cup S_2) &= \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, \sigma \cup S_2) \sum_{\ell \in S_1 \setminus \sigma} f(\{\ell\}, S_1 \cup S_2) \\
&= \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, \sigma \cup S_2)\, f(S_1 \setminus \sigma, S_1 \cup S_2) \\
&= \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, \sigma \cup S_2) \left[1 - f(\sigma \cup S_2, S_1 \cup S_2)\right] \\
&= \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, \sigma \cup S_2) - \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, S_1 \cup S_2) \\
&= \sum_{\sigma \in \mathcal{R}_s(S_1)} f(S_2, \sigma \cup S_2) - \binom{|S_1|}{s} f(S_2, S_1 \cup S_2)
\end{aligned}$$

where the first equality follows from the fact that, for all $\ell \in S_1$ and $\sigma \in \mathcal{R}_s(S_1)$, $\sigma \in \mathcal{R}_s(S_1 \setminus \{\ell\}) \iff \ell \in S_1 \setminus \sigma$;[9] the second equality follows from property 1; the third equality follows from property 2; the fourth equality follows from property 3; and the last equality follows from the definition of $\mathcal{R}_s(\cdot)$. ☐

**Proposition 1.** *Let $A$ be a nonempty, finite set of size $n \equiv |A| \in \mathbb{N}$. Let $\mathcal{S}_n$ represent the class of all finite sets of size $J \geq n+1$. Then,*

$\forall\, n \in \mathbb{N},\ \forall\, \mathcal{J} \in \mathcal{S}_n : \mathcal{J} \supset A,\ \forall\ \{\delta_j\}_{j \in \mathcal{J}} : \delta_j \in \mathbb{R}\ \forall j \in \mathcal{J},$

$$\sum_{p \in \mathcal{Q}(A)} \prod_{\ell=1}^{n} \frac{\exp(\delta_{p_\ell})}{\sum_{k \in \mathcal{J} \setminus \{p_1, \ldots, p_{\ell-1}\}} \exp(\delta_k)} = 1 + \sum_{s=1}^{n} (-1)^s \sum_{\sigma \in \mathcal{R}_s(A)} \frac{\sum_{\ell \in \overline{A}} \exp(\delta_\ell)}{\sum_{k \in \sigma \cup \overline{A}} \exp(\delta_k)}, \tag{9}$$

*where $\mathcal{Q}(S)$ and $\mathcal{R}_s(S)$ denote the sets of $|S|$-permutations and $s$-combinations of set $S$, respectively, $\overline{A} \equiv \mathcal{J} \setminus A$, and $\{p_1, \ldots, p_0\} \equiv \emptyset$.*

*Proof.* Consider the set function $f : \mathcal{P}(\mathcal{J}) \times \mathcal{P}(\mathcal{J}) \to \mathbb{R}$ defined in Lemma 1 and rewrite the statement as

[9] To see this, consider $\ell \in S_1$ and nonempty $\sigma \subseteq S_1$, and notice that

$$\sigma \in \mathcal{R}_{|\sigma|}(S_1 \setminus \{\ell\}) \iff \sigma \subseteq S_1 \setminus \{\ell\} \iff \sigma \not\supset \{\ell\} \iff \{\ell\} \subseteq S_1 \setminus \sigma \iff \ell \in S_1 \setminus \sigma.$$

$$\forall n\in\mathbb{N},\,\forall\,\mathcal{J}\in\mathcal{S}_n:\mathcal{J}\supset A,\,\forall\,\{\delta_j\}_{j\in\mathcal{J}}:\delta_j\in\mathbb{R}\,\forall j\in\mathcal{J},$$

$$\sum_{p\in\mathcal{Q}(A)}\prod_{\ell=1}^{n} f\left(\{p_\ell\},\mathcal{J}\setminus\{p_1,\ldots,p_{\ell-1}\}\right)=1+\sum_{s=1}^{n}(-1)^s\sum_{\sigma\in\mathcal{R}_s(A)} f\left(\overline{A},\sigma\cup\overline{A}\right),$$

which can be proved by mathematical induction and invocation of the properties of $f(\cdot,\cdot)$ established in Lemma 1 as follows.

Start by establishing that the statement holds for the special case $n=1$. To this end, let $\mathcal{J}$ be a finite set of arbitrary size $J\equiv|\mathcal{J}|\geq 2$ and, without loss of generality, suppose that $A=\{a\}\subset\mathcal{J}$. Note that, since $A$ is a singleton, $\mathcal{Q}(A)$ is a singleton itself, i.e., $\mathcal{Q}(A)=\{a\}=A$. Therefore,

$$\begin{aligned}\sum_{p\in\mathcal{Q}(A)}\prod_{\ell=1}^{n} f\left(\{p_\ell\},\mathcal{J}\setminus\{p_1,\ldots,p_{\ell-1}\}\right)&=\sum_{p\in A} f\left(\{p\},\mathcal{J}\right)\\&=f\left(A,\mathcal{J}\right)\\&=1-f\left(\overline{A},A\cup\overline{A}\right)\\&=1+\sum_{s=1}^{n}(-1)^s\sum_{\sigma\in\mathcal{R}_s(A)} f\left(\overline{A},\sigma\cup\overline{A}\right),\end{aligned}$$

where the second equality follows from property 1 in Lemma 1, the third equality follows from property 2, and the last equality follows from the fact that $\mathcal{R}_{|S|}(S)=\{S\}\ \forall\,S$. Thus, since $J$ and the underlying $\{\delta_j\}_{j\in\mathcal{J}}$ are arbitrary, the statement holds for $n=1$.

The next step is formulating the induction hypothesis. Fix natural number $k\geq 2$ and suppose that the statement holds for $n=k-1$. That is, suppose that the statement

$$\forall\,\widetilde{\mathcal{J}}\in\mathcal{S}_{k-1}:\widetilde{\mathcal{J}}\supset B,\,\forall\,\{\delta_j\}_{j\in\widetilde{\mathcal{J}}}:\delta_j\in\mathbb{R}\,\forall j\in\widetilde{\mathcal{J}},$$

$$\sum_{r\in\mathcal{Q}(B)}\prod_{\ell=1}^{k-1} f\left(\{r_\ell\},\widetilde{\mathcal{J}}\setminus\{r_1,\ldots,r_{\ell-1}\}\right)=1+\sum_{s=1}^{k-1}(-1)^s\sum_{\sigma\in\mathcal{R}_s(B)} f\left(\overline{B},\sigma\cup\overline{B}\right)$$

holds for any finite set $B$ of size $|B|=k-1$, where $\overline{B}\equiv\widetilde{\mathcal{J}}\setminus B$. Notice that, in particular, this implies that if $A$ is a finite set of size $k$ and $B\in\mathcal{R}_{k-1}(A)$, then

$$\forall\,\mathcal{J}\in\mathcal{S}_k:\mathcal{J}\supset A,\,\forall\,\{\delta_j\}_{j\in\mathcal{J}}:\delta_j\in\mathbb{R}\,\forall j\in\mathcal{J},$$

$$\sum_{r\in\mathcal{Q}(B)}\prod_{\ell=1}^{k-1} f\left(\{r_\ell\},\left(B\cup\overline{A}\right)\setminus\{r_1,\ldots,r_{\ell-1}\}\right)=1+\sum_{s=1}^{k-1}(-1)^s\sum_{\sigma\in\mathcal{R}_s(B)} f\left(\overline{A},\sigma\cup\overline{A}\right)$$

since $\widetilde{\mathcal{J}}=B\cup\overline{A}\in\mathcal{S}_{k-1}$, where $\overline{A}\equiv\mathcal{J}\setminus A$.[10]

[10]Note that $\overline{B}=\widetilde{\mathcal{J}}\setminus B=(B\cup\overline{A})\setminus B=\overline{A}$ since $B\subset A\implies B\cap\overline{A}=\emptyset$. Also notice that $\mathcal{J}\in\mathcal{S}_k\implies\mathcal{J}\in\mathcal{S}_{k-1}$ and $\mathcal{J}\supset A\implies\mathcal{J}\supset B$.

Now, consider the case $n = k$. Let $A$ be a finite set of size $k$ and consider $\mathcal{J} \in \mathcal{S}_k$ such that $\mathcal{J} \supset A$. For $j \in A$, let $B_j \equiv A \setminus \{j\}$ and $\widetilde{\mathcal{J}}_j \equiv B_j \cup \overline{A} = \mathcal{J} \setminus \{j\}$, where $\overline{A} \equiv \mathcal{J} \setminus A$. Then,

$$\sum_{p \in \mathcal{Q}(A)} \prod_{\ell=1}^{n} f\left(\{p_\ell\}, \mathcal{J} \setminus \{p_1, \ldots, p_{\ell-1}\}\right) = \sum_{p_1 \in A} f\left(\{p_1\}, A \cup \overline{A}\right) \sum_{r \in \mathcal{Q}(B_{p_1})} \prod_{\ell=1}^{k-1} f\left(\{r_\ell\}, \widetilde{\mathcal{J}}_{p_1} \setminus \{r_1, \ldots, r_{\ell-1}\}\right)$$

$$= \sum_{p_1 \in A} f\left(\{p_1\}, A \cup \overline{A}\right) \left[1 + \sum_{s=1}^{k-1} (-1)^s \sum_{\sigma \in \mathcal{R}_s(B_{p_1})} f\left(\overline{A}, \sigma \cup \overline{A}\right)\right]$$

$$= \sum_{p_1 \in A} f\left(\{p_1\}, A \cup \overline{A}\right) + \sum_{s=1}^{k-1} (-1)^s \sum_{p_1 \in A} f\left(\{p_1\}, A \cup \overline{A}\right) \sum_{\sigma \in \mathcal{R}_s(A \setminus \{p_1\})} f\left(\overline{A}, \sigma \cup \overline{A}\right)$$

$$= f\left(A, A \cup \overline{A}\right) + \sum_{s=1}^{k-1} (-1)^s \left[\sum_{\sigma \in \mathcal{R}_s(A)} f\left(\overline{A}, \sigma \cup \overline{A}\right) - \binom{k}{s} f\left(\overline{A}, A \cup \overline{A}\right)\right]$$

$$= 1 - f\left(\overline{A}, A \cup \overline{A}\right) + \sum_{s=1}^{k-1} (-1)^s \sum_{\sigma \in \mathcal{R}_s(A)} f\left(\overline{A}, \sigma \cup \overline{A}\right) + \left[\sum_{s=1}^{k-1} \binom{k}{s} (-1)^{s+1}\right] f\left(\overline{A}, A \cup \overline{A}\right)$$

$$= 1 + \sum_{s=1}^{k-1} (-1)^s \sum_{\sigma \in \mathcal{R}_s(A)} f\left(\overline{A}, \sigma \cup \overline{A}\right) + (-1)^k \sum_{\sigma \in \mathcal{R}_k(A)} f\left(\overline{A}, \sigma \cup \overline{A}\right)$$

$$= 1 + \sum_{s=1}^{k} (-1)^s \sum_{\sigma \in \mathcal{R}_s(A)} f\left(\overline{A}, \sigma \cup \overline{A}\right).$$

The first equality follows from factoring out the first term of the product in the summand on the left-hand side, which corresponds to the first element of every possible permutation of the set $A$. The second equality follows from the induction hypothesis since $B_{p_1} \in \mathcal{R}_{k-1}(A)$ and $\widetilde{\mathcal{J}}_{p_1} \in \mathcal{S}_{k-1} \, \forall \, p_1 \in A$. The fourth equality follows from properties 1 and 4 in Lemma 1, while the fifth equality follows from property 2. The sixth equality follows from the binomial theorem and the fact that $\mathcal{R}_k(A) = \{A\}$.[11] □

[11] Recall the binomial theorem: $\forall \, k \in \{0\} \cup \mathbb{N}, \; \forall \, a, b \in \mathbb{R}, (a+b)^k = \sum_{s=0}^{k} \binom{k}{s} a^{n-s} b^s$. Now, for $(a, b) = (1, -1)$, we obtain $0 = \sum_{s=0}^{k} \binom{k}{s} (-1)^s = 1 + \sum_{s=1}^{k-1} \binom{k}{s} (-1)^s + (-1)^k \iff \sum_{s=1}^{k-1} \binom{k}{s} (-1)^{s+1} = 1 + (-1)^k$.

**Theorem 1.** *Consider a sample of decision makers $i \in \{1, \dots, I\}$, each choosing a subset comprising $n_i \in \{1, \dots, J-1\}$ alternatives from choice set $\mathcal{J} = \{1, \dots, J\}$, $A_i \subset \mathcal{J}$, according to the logit ARUM. Then, the rank-ordered logit model with two tied ranks and the multichoice logit model are equivalent in the sense that the likelihood functions of these models coincide.*

*Proof.* The statement follows directly as a corollary of Proposition 1. □

# 5 Conclusion

Despite the prevalence of partially rank-ordered discrete choice data in survey contexts, the multichoice logit model of van Ophem et al. (1999) has seen limited adoption in applied work, possibly due to the perceived need for specialized programming. In contrast, the Beggs et al. (1981) rank-ordered logit for complete ordinal rankings and the subsequent treatment of ties by Allison and Christakis (1994) have become quite popular, with statistical packages typically offering canned routines for their estimation. We show that the partially rank-ordered logit model is a special case of the rank-ordered logit with ties and is also equivalent to the multichoice logit.

By unifying seemingly different estimators under a common data structure and econometric model, we hope to provide applied researchers with a broader set of tools to estimate preferences from coarse rankings. The fact that the multichoice logit, the rank-ordered logit with two (or one) tied ranks, and the integral-representation approach to partially ordered data of Roussille and Scuderi (2025) are equivalent representations of the partially rank-ordered logit model reveals that the three approaches yield the same estimator, the maximum likelihood estimator for this model. The differences lie in how the estimator is computed.

Some representations of the same likelihood function may be easier to maximize than others, and, in the case of the tied rank-ordered logit, a maximum likelihood routine is readily available in statistical software such as Stata or R. This is an attractive alternative for moderately-sized choice sets, and the result that the two- and one-tier ranking representations are equivalent simplifies computations.[12] For large choice sets, the combinatorial nature of this closed-form representation of the likelihood function may render estimation computationally intractable, but the minorize-maximize algorithm of Roussille and Scuderi (2025) provides a convenient alternative—although, to the best of our knowledge, it requires specialized programming.

[12]For example, using the option `incomplete(#)` with Stata command `cmrologit`—where # represents the rank assigned to the unchosen alternatives—will produce the same estimates with fewer computations since the objective function will consider only the term $\mathcal{L}_{1i}$ from the decomposition $L_i^{\mathrm{ROL}} = \mathcal{L}_{1i} \cdot \mathcal{L}_{2i}$ in Equation (8). In contrast, the two-tier representation—i.e., not specifying the option `incomplete(#)`— will also compute the term $\mathcal{L}_{2i}$ which is identically 1.

# References


Albagli, P. (2025): "Estimating Labor Market Power from Job Applications," Working Paper.

Allison, P. D. and N. A. Christakis (1994): "Logit Models for Sets of Ranked Items," *Sociological Methodology*, 24, 199–228.

Beggs, S., S. Cardell, and J. Hausman (1981): "Assessing the Potential Demand for Electric Cars," *Journal of Econometrics*, 17, 1–19.

Berry, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 25, 242–262.

Billingsley, P. (2012): *Probability and Measure*, John Wiley and Sons, anniversary ed.

Chade, H. and L. Smith (2006): "Simultaneous Search," *Econometrica*, 74, 1293–1307.

Cox, D. R. (1972): "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, 34, 187–220.

Hajivassiliou, V. A. and P. A. Ruud (1994): "Chapter 40: Classical Estimation Methods for LDV Models Using Simulation," Elsevier, vol. 4 of *Handbook of Econometrics*, 2383–2441.

Hatton, T. J. (2021): "Public Opinion on Immigration in Europe: Preference and Salience," *European Journal of Political Economy*, 66, 101969.

Hausman, J. A. and P. A. Ruud (1987): "Specifying and Testing Econometric Models for Rank-Ordered Data," *Journal of Econometrics*, 34, 83–104.

McFadden, D. (1974): "Chapter Four: Conditional Logit Analysis of Qualitative Choice Behavior," Academic Press, Frontiers in Econometrics, 105–142.

Roussille, N. and B. Scuderi (2025): "Bidding for Talent: A Test of Conduct in a High-Wage Labor Market," Working Paper 33848, National Bureau of Economic Research.

StataCorp (2023): *Stata 18 Choice Models Reference Manual*, Stata Press, College Station, TX.

van Ophem, H., P. Stam, and B. van Praag (1999): "Multichoice Logit: Modeling Incomplete Preference Rankings of Classical Concerts," *Journal of Business & Economic Statistics*, 17, 117–128.

Yang, S., G. M. Allenby, and G. Fennell (2002): "Modeling Variation in Brand Preference: The Roles of Objective Environment and Motivating Conditions," *Marketing Science*, 21, 14–31.

# Appendix

## A Partial likelihood in the Cox relative risk regression model

This appendix derives the partial likelihood in the Cox (1972, 1975) relative risk regression model, closely following Chapter 4 of Kalbfleisch and Prentice (2002) while emphasizing its isomorphism to the likelihood contribution of an observation in the Beggs et al. (1981) rank-ordered logit model given in equation (5), which is at the heart of Allison and Christakis (1994)'s generalization to accommodate ties.

Consider the modeling of failure times for a sample of $J = |\mathcal{J}|$ items $j \in \mathcal{J}$. The simplest version of the Cox (1972) relative risk regression model, known as the proportional hazards model, specifies the hazard rate

$$\begin{aligned}\lambda(t \mid x) &\equiv \lim_{h \to 0^+} \frac{\mathbb{P}\left(t \leq T < t+h \mid T \geq t, x\right)}{h} \\ &= \lambda_0(t) \exp\left(x'\beta\right), \end{aligned} \tag{A.1}$$

where the absolutely continuous random variable $T$ is failure time, $x \in \mathbb{R}^K$ is a vector of fixed covariates measured at or before time $0$, and $\lambda_0(\cdot)$ is an unspecified baseline hazard function. Our objective is the estimation of parameter vector $\beta \in \mathbb{R}^K$.

Let us start with the simplest setting by assuming there are no ties in failure times. Suppose that the failure times of a subset $S \subseteq \mathcal{J}$ of the items in the sample are uncensored, with the remaining $J - |S|$ failure times arbitrarily and independently right-censored. For $j \in \{1, \ldots, |S|\}$, let $t_j$ denote the $j^{\text{th}}$ smallest uncensored failure time —so that $t_1 < \cdots < t_{|S|}$—, and let $y_j \in S$ be the identity of the item with observed failure time $t_j$. Suppose that right-censoring is such that $T_\ell > t_{|S|}\ \forall \ell \in \overline{S} \equiv \mathcal{J} \setminus S$. Then, the set of items at risk of failure just before time $t_j$ is given by $\Lambda_j \cup \overline{S}$, where $\Lambda_j \equiv \left\{y_j, \ldots, y_{|S|}\right\}$, and the partial likelihood (Cox, 1972, 1975) is given by

$$\begin{aligned} L &= \prod_{j=1}^{|S|} \frac{\lambda(t_j \mid x_{y_j})\,\mathrm{d}t_j}{\sum\limits_{k \in \Lambda_j \cup \overline{S}} \lambda(t_j \mid x_k)\,\mathrm{d}t_j} \\ &= \prod_{j=1}^{|S|} \frac{\lambda_0(t_j)\ \exp\left(x'_{y_j}\beta\right)\mathrm{d}t_j}{\sum\limits_{k \in \Lambda_j \cup \overline{S}} \lambda_0(t_j)\ \exp\left(x'_k\beta\right)\mathrm{d}t_j} \\ &= \prod_{j=1}^{|S|} \frac{\exp\left(x'_{y_j}\beta\right)}{\sum\limits_{k \in \Lambda_j \cup \overline{S}} \exp\left(x'_k\beta\right)}. \end{aligned} \tag{A.2}$$

Allison and Christakis (1994) notice that the partial likelihood in (A.2) for this survival analysis problem is identical to the likelihood contribution of an individual observation in the rank-ordered logit model in equation (5) with $\delta_k = x_k'\beta$, and propose to use the extensions of the Cox model to the case of tied failure times to accommodate tied ranks in the rank-ordered logit model.

Kalbfleisch and Prentice (2002) argue that a natural way to adjust for ties in the uncensored failure times is to consider all the possible ways of breaking these ties and using the resulting average partial likelihood. Let us modify the setting in the previous paragraph by assuming that, out of the $|S|$ uncensored failure times, we observe $N \leq |S|$ distinct failure times such that $t_1 < \cdots < t_N$. Suppose that $d_j$ items fail at time $t_j$, where $j \in \{1, \ldots, N\}$. Let $Y_j \equiv \{\ell \in S : t(\ell) = t_j\}$ be the set of items that fail at time $t_j$, where $t(\ell)$ denotes the observed (uncensored) failure time of item $\ell \in S$. The definition of the set of items at risk of failure just before time $t_j$ can be generalized to this setting by defining $\Lambda_{jp\ell} \equiv \bigcup_{s=j}^{N} Y_s \setminus \{p_1, \ldots, p_{\ell-1}\}$, where $p \in \mathcal{Q}(Y_j)$ is one of the $d_j!$ permutations of $Y_j$ and $\ell \in \{1, \ldots, d_j\}$.[13] Then, considering every possible way of breaking the ties at failure time $t_j$, the average partial likelihood contribution at $t_j$ is

$$L_j = \frac{1}{d_j!} \sum_{p \in \mathcal{Q}(Y_j)} \prod_{\ell=1}^{d_j} \frac{\exp\left(x_{p_\ell}'\beta\right)}{\sum\limits_{k \in \Lambda_{jp\ell} \cup \overline{S}} \exp\left(x_k'\beta\right)}.$$

Therefore, the average partial likelihood is proportional to

$$L = \prod_{j=1}^{N} \sum_{p \in \mathcal{Q}(Y_j)} \prod_{\ell=1}^{d_j} \frac{\exp\left(x_{p_\ell}'\beta\right)}{\sum\limits_{k \in \Lambda_{jp\ell} \cup \overline{S}} \exp\left(x_k'\beta\right)}. \tag{A.3}$$

# B Example of rank-ordered logit with ties

**Example.** Consider the following example of the setting in Section 3.3. Individual $i$ ranks only $|S| = 6$ out of the $J = 7$ available alternatives such that $\mathcal{J} = \{a, b, c, d, e, f, g\}$, $S = \{a, b, c, d, e, f\}$, $\overline{S} = \{g\}$, and the observed ranking is $\boldsymbol{r}_i = (1, 2, 2, 2, 3, 3)$. Therefore, we observe $N_i = 3$ distinct ranks with $d_{i1} = 1$, $d_{i2} = 3$, and $d_{i3} = 2$ tied alternatives, respectively. Then, the sets of tied alternatives at each of the 3 observed ranks are $Y_{i1} = \{a\}$, $Y_{i2} = \{b, c, d\}$, and $Y_{i3} = \{e, f\}$, respectively, with the corresponding permutation sets $\mathcal{Q}(Y_{i1}) = \{a\}$, $\mathcal{Q}(Y_{i2}) = \{(b, c, d), (b, d, c), (c, b, d), (c, d, b), (d, b, c), (d, c, b)\}$, and $\mathcal{Q}(Y_{i3}) = \{(e, f), (f, e)\}$. There is only one *no-better-than* set corresponding to rank 1, namely $\Lambda_i(1, a, 1) = (Y_{i1} \setminus \emptyset) \cup Y_{i2} \cup Y_{i3} = S$. For rank 2, we need to consider every way of breaking the tie, given by all the permutations in $\mathcal{Q}(Y_{i2})$. Note that the *no-better-than* set corresponding to $\ell = 1$ is the same for any permutation,

[13] Set $\{p_1, \ldots, p_0\} \equiv \emptyset$, $\forall p \in \mathcal{Q}(Y_j)$, $\forall j \in \{1, \ldots, N\}$ to make the notation consistent.

i.e., $\Lambda(2,p,1) = (Y_{i2} \setminus \emptyset) \cup Y_{i3} = \{b,c,d,e,f\}$ $\forall p \in \mathcal{Q}(Y_{i2})$. When the tie is broken according to permutation $(b,c,d)$, we have $\Lambda_i(2,(b,c,d),2) = (Y_{i2} \setminus \{b\}) \cup Y_{i3} = \{c,d,e,f\}$, $\Lambda_i(2,(b,c,d),3) = (Y_{i2} \setminus \{b,c\}) \cup Y_{i3} = \{d,e,f\}$. Table B.1 below summarizes all the relevant *no-better-than* sets $\Lambda_{ijp\ell}$ for this example, resulting in the likelihood contribution in Equation (B.1) below.

**Table B.1.** Example *no-better-than* sets $\Lambda_i(j,p,\ell)$

| | $j=1$ | $j=2$ | | | | | | $j=3$ | |
|---|---|---|---|---|---|---|---|---|---|
| | $p \in \mathcal{Q}(Y_{i1})$ | $p \in \mathcal{Q}(Y_{i2})$ | | | | | | $p \in \mathcal{Q}(Y_{i3})$ | |
| $\ell$ | $a$ | $(b,c,d)$ | $(b,d,c)$ | $(c,b,d)$ | $(c,d,b)$ | $(d,b,c)$ | $(d,c,b)$ | $(e,f)$ | $(f,e)$ |
| 1 | $\{a,b,c,d,e,f\}$ | $\{b,c,d,e,f\}$ | $\{b,c,d,e,f\}$ | $\{b,c,d,e,f\}$ | $\{b,c,d,e,f\}$ | $\{b,c,d,e,f\}$ | $\{b,c,d,e,f\}$ | $\{e,f\}$ | $\{e,f\}$ |
| 2 | | $\{c,d,e,f\}$ | $\{c,d,e,f\}$ | $\{b,d,e,f\}$ | $\{b,d,e,f\}$ | $\{b,c,e,f\}$ | $\{b,c,e,f\}$ | $\{f\}$ | $\{e\}$ |
| 3 | | $\{d,e,f\}$ | $\{c,e,f\}$ | $\{d,e,f\}$ | $\{b,e,f\}$ | $\{c,e,f\}$ | $\{b,e,f\}$ | | |

$$
\begin{aligned}
L_i^{\mathrm{ROL}} &= \left[\sum_{p\in\mathcal{Q}(Y_{i1})} \prod_{\ell=1}^{1} \frac{\exp(\delta_{ip_\ell})}{\sum\limits_{k\in\Lambda_{i1p\ell}\cup\overline{S}} \exp(\delta_{ik})}\right] \left[\sum_{p\in\mathcal{Q}(Y_{i2})} \prod_{\ell=1}^{3} \frac{\exp(\delta_{ip_\ell})}{\sum\limits_{k\in\Lambda_{i2p\ell}\cup\overline{S}} \exp(\delta_{ik})}\right] \left[\sum_{p\in\mathcal{Q}(Y_{i3})} \prod_{\ell=1}^{2} \frac{\exp(\delta_{ip_\ell})}{\sum\limits_{k\in\Lambda_{i3p\ell}\cup\overline{S}} \exp(\delta_{ik})}\right] \\
&= \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{f,g\}} \exp(\delta_{ik})} \\
&+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,g\}} \exp(\delta_{ik})} \\
&+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{f,g\}} \exp(\delta_{ik})} \\
&+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,g\}} \exp(\delta_{ik})} \\
&+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{f,g\}} \exp(\delta_{ik})} \\
&+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,g\}} \exp(\delta_{ik})}
\end{aligned}
$$

$$+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{b,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{f,g\}} \exp(\delta_{ik})}$$

$$+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{b,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,g\}} \exp(\delta_{ik})}$$

$$+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{f,g\}} \exp(\delta_{ik})}$$

$$+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,g\}} \exp(\delta_{ik})}$$

$$+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{b,c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{f,g\}} \exp(\delta_{ik})}$$

$$+ \frac{\exp(\delta_{ia})}{\sum\limits_{k\in\{a,b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{id})}{\sum\limits_{k\in\{b,c,d,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ic})}{\sum\limits_{k\in\{b,c,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ib})}{\sum\limits_{k\in\{b,e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{if})}{\sum\limits_{k\in\{e,f,g\}} \exp(\delta_{ik})} \frac{\exp(\delta_{ie})}{\sum\limits_{k\in\{e,g\}} \exp(\delta_{ik})}. \tag{B.1}$$

# Appendix references


Allison, P. D. and N. A. Christakis (1994): "Logit Models for Sets of Ranked Items," *Sociological Methodology*, 24, 199–228.

Beggs, S., S. Cardell, and J. Hausman (1981): "Assessing the Potential Demand for Electric Cars," *Journal of Econometrics*, 17, 1–19.

Cox, D. R. (1972): "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, 34, 187–220.

——— (1975): "Partial Likelihood," *Biometrika*, 62, 269–276.

Kalbfleisch, J. D. and R. L. Prentice (2002): *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, Ltd.