

1 Simulation studies

In this section we present two illustrative simulation studies based on simulated data which highlight the flexibility of the Bayesian WAND for the analysis of ranked data.

1.1 Data simulation

We begin by first simulating (marginally) a cluster structure from the ANDP model using a Pólya urn scheme and then a realisation of the skill parameters, as follows.

- Simulate α from its prior distribution.
- Set $c_1 = 1$ and the (current) number of ranker clusters $N^r = 1$.
- For $i = 2, \dots, n$ simulate the allocation of ranker i to a ranker cluster according to

$$\Pr(c_i = j | c_1, \dots, c_{i-1}) = \frac{n_{ij}^r}{\alpha + i - 1}, \quad \text{for } j = 1, \dots, N^r,$$

$$\Pr(c_i = N^r + 1 | c_1, \dots, c_{i-1}) = \frac{\alpha}{\alpha + i - 1},$$

where n_{ij}^r denotes the number of rankers currently within ranker cluster j (at iteration i), and $N^r \rightarrow N^r + 1$ if $c_i = N^r + 1$.

- For each ranker cluster $s = 1, \dots, N^r$
 - Simulate γ_s from its prior distribution.
 - Set $d_{s1} = 1$ and the current number of entity clusters within ranker cluster s $N_s^e = 1$.
 - For $k = 2, \dots, K$ simulate the allocation of entity k to an entity cluster according to

$$\Pr(d_{sk} = j | d_{s1}, \dots, d_{s,k-1}) = \frac{n_{kj,s}^e}{\gamma_s + k - 1}, \quad \text{for } j = 1, \dots, N_s^e,$$

$$\Pr(d_{sk} = N_s^e + 1 | d_{s1}, \dots, d_{s,k-1}) = \frac{\gamma_s}{\gamma_s + k - 1},$$

where $n_{kj,s}^e$ denotes the number of entities currently within entity cluster j (at iteration k in ranker cluster s), and $N_s^e \rightarrow N_s^e + 1$ if $d_{sk} = N_s^e + 1$.

- Simulate $\lambda_{sj} \stackrel{\text{indep}}{\sim} G_0$ for $s = 1, \dots, N^r$, $j = 1, \dots, N_s^e$.

In the following simulation study we investigate the performance of the algorithm in two scenarios: one in which rankers were informative and another where there is a mixture of informative and uninformative rankers. Therefore, the w_i are prescribed by the scenario.

The complete ranking \mathbf{x}_i for ranker i is obtained using the well-known exponential latent variable representation of the Plackett-Luce model. We first simulate $\rho_{ij} \stackrel{\text{indep}}{\sim} \text{Exp}(\lambda_{c_i d_{c_i j}}^{w_i})$, for $j = 1, \dots, K$ and then set $\mathbf{x}_i = \text{rank}(\boldsymbol{\rho}_i)$, that is, for $j = 1, \dots, K$, $x_{ij} = q$ where $\rho_{iq} = \max(\boldsymbol{\rho}_i \setminus \{\rho_{ix_{i1}}, \dots, \rho_{ix_{ij-1}}\})$. Partial, top- M complete and top- M partial rankings are obtained from these complete rankings.

1.2 Study 1

In this study we consider two datasets with *complete* rankings of $K = 20$ entities. The first dataset (Dataset 1) contains complete rankings from $n = 40$ informative ($w_i = 1$) rankers in a single ranker group ($N^r = 1$ and $c_i = 1$). The distinct “skill” parameters λ_k were sampled from the prior distribution with $\gamma_1 = 1$ and base distribution $G_0 = \text{Ga}(1, 1)$. This simulation gave six unique entity clusters ($N_1^e = 6$). The simulated data are given in Table 11 within these supplementary materials. Note that, for ease of interpretation, we applied a permutation to the entity labels so that the λ_k were decreasing, with the most preferred entity being labelled 1 and the least preferred labelled 20.

The second dataset (Dataset 2) contains complete rankings from $n = 50$ rankers, consisting of those in Dataset 1 and an additional set from 10 uninformative rankers, with $w_{41:50} = 0$. The addition of these random rankings will allow us to investigate the extent to which the model both identifies and handles uninformative rankings. These additional (random) rankings are given in Table 12.

We investigate the effect of incomplete rankings by comparing the analysis of the complete rankings with those of top- M complete rankings for $M = 5, 10, 15$. That is, we analyse the data assuming rankers see $K_i = K = 20$ entities and report their top $n_i = 5, 10, 15, 20$ entities. These analyses will allow us to investigate the level of uncertainty introduced by only observing truncations of the rankings. Another scenario we consider is that of a so-called “restricted” analysis in which any entity not ranked by any ranker is removed from the dataset. The intuition behind this scenario is that our beliefs about the orderings of entities that appear in at least one ranking should not be affected by whether non-appearing entities are included or not. One example of this is in Dataset 1 where, under the top-5 and top-10 scenarios, entities 17, 18 and 20 do not appear in any of the rankings. Thus we also consider restricted top-5 and top-10 analyses of Dataset 1 which use $K_i = K = 17$.

We will adopt the same base distribution and prior distributions for the concentration parameters as used in the paper, that is, take $G_0 = \text{Ga}(1, 1)$, $\alpha \sim \text{Ga}(1, 1)$ and $\gamma_s \sim \text{Ga}(3, 3)$ for $s \in \mathbb{N}$, noting that these are common choices within the literature (Rodriguez et al., 2008). Here we consider two scenarios of an exchangeable prior for the ranker reliability parameters \mathbf{w} : one in which we are unsure about their ability (taking $p_i = 0.5$) and the other where we are quite confident that they are informative (taking $p_i = 0.9$).

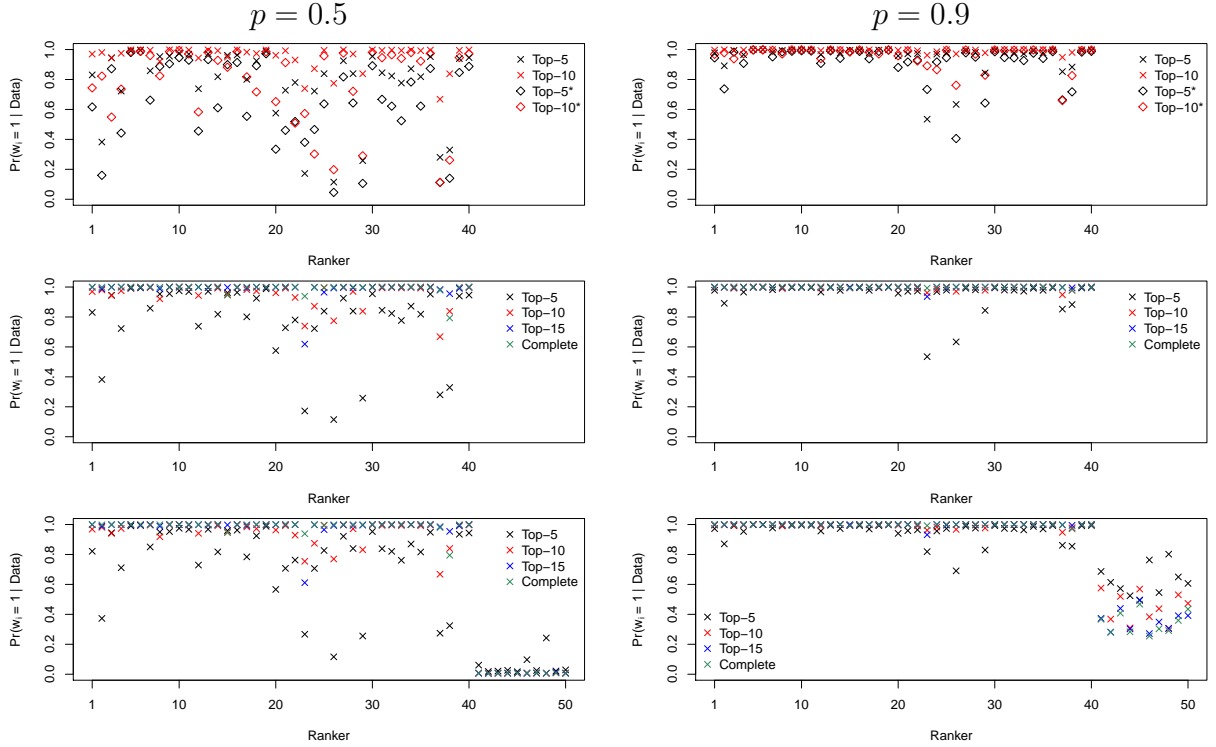


Figure 1: Plots of the posterior probability $\Pr(w_i = 1|\mathcal{D})$ that ranker i is informative for both scenarios of prior on their ability: $p_i = 0.5$ (left column) and $p_i = 0.9$ (right column). The top row of plots show the comparison between the restricted (*) and full (unrestricted) analyses for Dataset 1. Plots in the middle row are those for the full analyses using Dataset 1, with the corresponding plots using Dataset 2 in the bottom row.

Figure 1 shows the posterior probability $\Pr(w_i = 1|\mathcal{D})$ that ranker i is informative for each information scenario and for both choices of prior probabilities $p_i = \Pr(w_i = 1)$. It is not surprising that the restricted analyses, due to the loss of information, produce results which are less consistent with the “known” abilities of the rankers ($w_{1:40} = 1$ and $w_{41:50} = 0$) than the full (unrestricted) analyses which considers all entities. This finding is clear for both choices of the p_i but is more noticeable for the $p_i = 0.5$ case. The results correctly show that the majority of rankers have been identified as informative (middle row). Unsurprisingly this identification becomes clearer as more comprehensive rankings are used (going from top-5 up to complete). It is also interesting to see that the most data-poor (top-5) analysis does reasonably well.

The analyses for Dataset 2 show that the uninformative (random) rankers have been identified quite clearly, particularly for the $p_i = 0.5$ case where the posterior probability that rankers 41 to 50 are informative are very close to zero (bottom left plot). The probabilities for rankers 1 to 40 remain similar to those found when analysing Dataset 1. This (and other analyses not given here) suggests that the identification of informative rankers using this model is fairly robust to the addition of rankings from uninformative

$p = 0.5$						$p = 0.9$					
Dataset 1	1	2	3	4	≥ 5	Dataset 1	1	2	3	4	≥ 5
Top-5*	0.73	0.20	0.05	0.02	0.00	Top-5*	0.65	0.23	0.08	0.02	0.01
Top-10*	0.88	0.11	0.01	0.00	0.00	Top-10*	0.88	0.11	0.01	0.00	0.00
Top-5	0.87	0.11	0.02	0.00	0.00	Top-5	0.87	0.11	0.02	0.00	0.00
Top-10	0.98	0.02	0.00	0.00	0.00	Top-10	0.99	0.01	0.00	0.00	0.00
Top-15	0.99	0.01	0.00	0.00	0.00	Top-15	1.00	0.00	0.00	0.00	0.00
Complete	1.00	0.00	0.00	0.00	0.00	Complete	1.00	0.00	0.00	0.00	0.00

Dataset 2	1	2	3	4	≥ 5	Dataset 2	1	2	3	4	≥ 5
Top-5	0.74	0.20	0.05	0.01	0.00	Top-5	0.06	0.29	0.29	0.19	0.10
Top-10	0.88	0.10	0.02	0.00	0.00	Top-10	0.09	0.38	0.30	0.16	0.06
Top-15	0.89	0.10	0.01	0.00	0.00	Top-15	0.15	0.36	0.28	0.14	0.05
Complete	0.89	0.10	0.01	0.00	0.00	Complete	0.16	0.35	0.28	0.14	0.06

Table 1: Posterior distribution of the number of ranker clusters N^r for restricted (*) and full (unrestricted) analyses. Numbers in bold indicate modal values.

rankers. The bottom row of plots show the influence of the choice of the p_i in the prior distribution. Here we see that having a high level of confidence that uninformative rankers are informative can potentially mask their identification. It's clear here that the posterior probabilities $\Pr(w_i = 1|\mathcal{D})$ for the uninformative rankers are well separated from those for informative rankers but nevertheless they might be identified as informative if this choice were made by thresholding these posterior probabilities at say 0.5 or even higher. This suggests that the analyst should use a fairly conservative choice of the p_i and should be careful about expressing over confidence in ranker abilities *a priori*.

Table 1 gives the posterior distribution of the number of ranker clusters N^r under each analysis. For Dataset 1, we see much more posterior support for a single ranker group under the full (unrestricted) analyses compared to their restricted equivalents, with little dependence on the choice of the p_i . Also, for the analyses of Dataset 2, the posterior support for a single ranker group reduces, particularly for the $p_i = 0.9$ case. Indeed for this case, the high prior confidence that all rankers are informative changes the modal number of ranker clusters from one to two, though this comes with a higher level of posterior uncertainty. Also, as before, the probability of the correct number of ranker clusters increases as more comprehensive rankings are included within the analysis.

The posterior distribution for the number of ranker clusters in the Dataset 1 analysis gives overwhelming support for the true number $N^r = 1$ for each case, and particularly for the complete, top-15 and top-10 cases. Therefore the allocation of rankers to ranker clusters is trivial. However, the allocation is not quite as straightforward in the Dataset 2 analysis. Note that here we have retained the same prior for the upper level concentration parameter α rather than amend it to reflect the known heterogeneity in ranker beliefs in Dataset 2.

Figure 2 gives the complete linkage dendrograms for both $p_i = 0.5$ and $p_i = 0.9$ cases in the

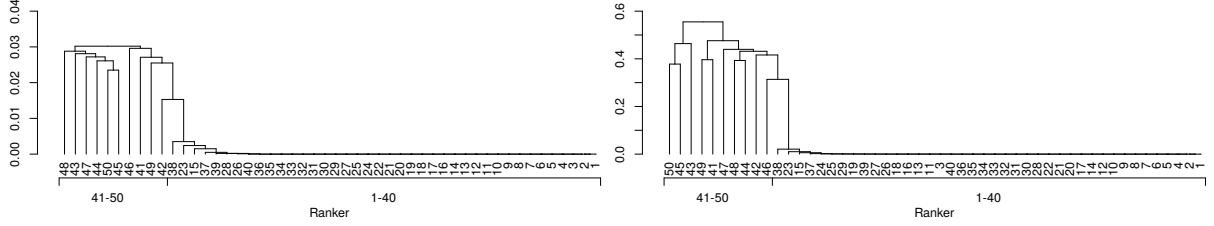


Figure 2: Dendrograms for ranker clustering within Dataset 2 under a complete analysis for $p_i = 0.5$ (left plot) and $p_i = 0.9$ (right plot).

complete analysis of Dataset 2, using the method described in section 3 in the paper. The method clearly picks out the informative and homogeneous rankers (numbered 1–40) and puts them into a single ranker group – the posterior probabilities that these rankers are co-clustered is 0.997 and 0.979 respectively. When $p_i = 0.5$ we see that the uninformative ranker most “similar” to the informative group is ranker 42. This ranker is co-clustered with each of the informative rankers at least 98.5% of the time. The other uninformative rankers also have high co-clustering probabilities and this is consistent with the very high posterior support for a single group of rankers, $\Pr(N^r = 1|\mathcal{D}, p_i = 0.5) = 0.89$. This result occurs as a consequence of the model down-weighting the influence of the uninformative rankers: $\Pr(w_i = 1|\mathcal{D}, p_i = 0.5) \ll 0.1$ for $i = 41, \dots, 50$. On the other hand when we are much more confident in the ability of the rankers (with $p_i = 0.9$) the most “similar” uninformative ranker to the informative group is co-clustered with informative rankers at least 68.6% of the time, a much smaller proportion than in the $p_i = 0.5$ case. Also the uninformative rankers do not separate themselves into a single distinct cluster, with $0.311 \leq \Delta_{ij} \leq 0.555$ for any $i \neq j \in \{41-50\}$, that is, any pair of uninformative rankers are co-clustered between 44.5%–68.9% of the time. It is perhaps not surprising that the model is not able to detect significant similarities between any pair of the ten uninformative rankers as their rankings are random permutations and they are few in number.

Table 2 gives the marginal posterior distributions of the number of entity clusters conditional on a single ranker cluster (for all analyses of each dataset). Note that the $p_i = 0.9$ analysis of Dataset 2 gives very low posterior support for a single ranker cluster, with $\Pr(N^r = 1|\mathcal{D}, p_i = 0.9) = 0.16$, and so later in Section 1.2.1 we look at results when conditioning on two ranker clusters (the posterior modal number). The table also shows that posterior support for the correct number of entity clusters ($N_1^e = 6$) increases as the information provided within each ranking increases. The cost of performing a restricted analysis is especially visible in the top–5 case. As was the case for ranker clustering, the inclusion of complete rankings in comparison to top–10 complete rankings does not have a significant effect on our posterior beliefs.

Figure 3 shows dendrograms of the entity grouping structure for each of the complete analyses, conditional on a single ranker cluster. Note that the entity clusters are similar under both $p_i = 0.5$ and $p_i = 0.9$ analyses, particularly for entities 1–6 and 15–20. The other entities (on the right hand side of each dendrogram) also have a similar grouping

$p = 0.5$										
Dataset 1	1	2	3	4	5	6	7	8	9	≥ 10
Top-5*	0.00	0.11	0.23	0.27	0.18	0.12	0.06	0.02	0.01	0.00
Top-10*	0.00	0.00	0.05	0.19	0.28	0.23	0.15	0.06	0.02	0.02
Top-5	0.00	0.00	0.15	0.27	0.24	0.17	0.09	0.04	0.02	0.02
Top-10	0.00	0.00	0.03	0.13	0.22	0.25	0.19	0.11	0.05	0.03
Top-15	0.00	0.00	0.00	0.07	0.21	0.29	0.22	0.13	0.05	0.03
Complete	0.00	0.00	0.00	0.09	0.24	0.29	0.21	0.11	0.04	0.02
Dataset 2	1	2	3	4	5	6	7	8	9	≥ 10
Top-5	0.00	0.00	0.14	0.26	0.25	0.17	0.10	0.05	0.02	0.01
Top-10	0.00	0.00	0.03	0.12	0.22	0.24	0.19	0.12	0.05	0.03
Top-15	0.00	0.00	0.00	0.08	0.22	0.29	0.23	0.12	0.05	0.01
Complete	0.00	0.00	0.00	0.10	0.24	0.28	0.21	0.11	0.04	0.02
$p = 0.9$										
Dataset 1	1	2	3	4	5	6	7	8	9	≥ 10
Top-5*	0.00	0.10	0.28	0.29	0.20	0.08	0.03	0.02	0.00	0.00
Top-10*	0.00	0.00	0.04	0.18	0.28	0.25	0.15	0.07	0.02	0.01
Top-5	0.00	0.00	0.21	0.31	0.24	0.14	0.06	0.02	0.01	0.01
Top-10	0.00	0.00	0.03	0.11	0.22	0.24	0.20	0.12	0.05	0.03
Top-15	0.00	0.00	0.00	0.08	0.22	0.28	0.23	0.13	0.05	0.01
Complete	0.00	0.00	0.00	0.08	0.22	0.30	0.22	0.11	0.05	0.02
Dataset 2	1	2	3	4	5	6	7	8	9	≥ 10
Top-5	0.00	0.00	0.17	0.24	0.25	0.18	0.08	0.05	0.02	0.01
Top-10	0.00	0.00	0.02	0.10	0.24	0.24	0.21	0.11	0.05	0.03
Top-15	0.00	0.00	0.00	0.06	0.22	0.28	0.25	0.13	0.05	0.01
Complete	0.00	0.00	0.01	0.07	0.23	0.28	0.23	0.12	0.05	0.01

Table 2: Posterior distribution of the number of entity clusters, conditional on a single ranker cluster, for restricted (*) and full (unrestricted) analyses. Numbers in bold indicate modal values.

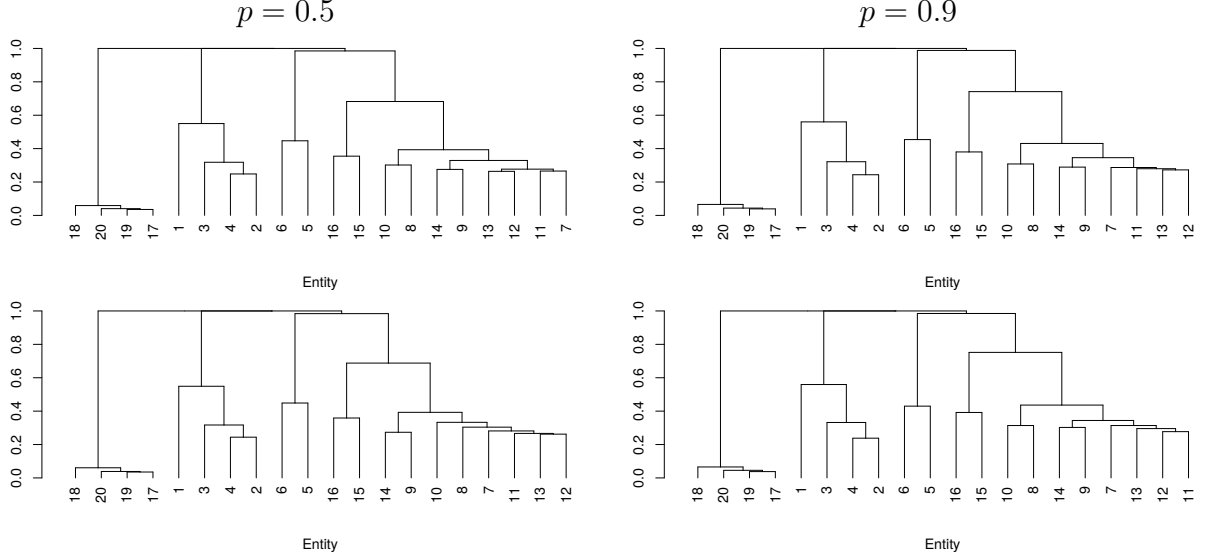


Figure 3: Dendrograms for entity clustering for Dataset 1 (top) and Dataset 2 (bottom) conditional on a single ranker cluster under both prior specifications for the Complete analyses.

structure with only minor discrepancies in the order that entity pairs cluster together (and clustering taking place at similar levels of the dissimilarity measure Δ_{ij}). Thus the dendrograms are fairly robust to the addition of the uninformative rankings. This can partly be explained by the dendrograms being conditional on a single ranker cluster and the Dataset 2 analysis correctly identifying the uninformative rankers (with $w_{41:50} = 0$).

We can explore our posterior distribution further by investigating where specific entities are likely to be ranked. Consider the posterior probability that a specific entity is ranked at least i th, that is, $P_i = \Pr(\text{entity ranked} \geq i | \mathcal{D})$. Figure 4 displays P_5 for all analyses, P_{10} for all except the top-5 case and P_{15} for the top-15 and complete analyses. Interestingly the posterior probabilities under the restricted top-5 and top-10 analyses are very similar to those under the unrestricted analysis (especially when $p_i = 0.5$). This suggests that, for these data, this aspect of the posterior distribution is robust to whether or not the unobserved entities (in the restricted analysis) are included in the analysis. The two left hand plots within Figure 4 show considerable similarity between the full (unrestricted) analyses of the two datasets (when $p_i = 0.5$). Here the WAND model has been able to identify the so-called spam rankers within Dataset 2 (see Figure 1) and so these rankers have little effect on the analysis. However, this is not the case when we take $p_i = 0.9$. In this case, the high level of confidence that the rankers are informative results in the WAND model being reluctant to classify any rankers as uninformative; again see Figure 1. This leads to the uninformative rankers contaminating the posterior distribution of the entity λ -parameters.

Earlier we questioned whether a collection of complete rankings is required if we only wish to infer, say, the top-5 entities. Looking at the results for Dataset 1 (in the top plots in Figure 4), if we consider P_5 we can see how the top-5 analysis is able to detect that entities

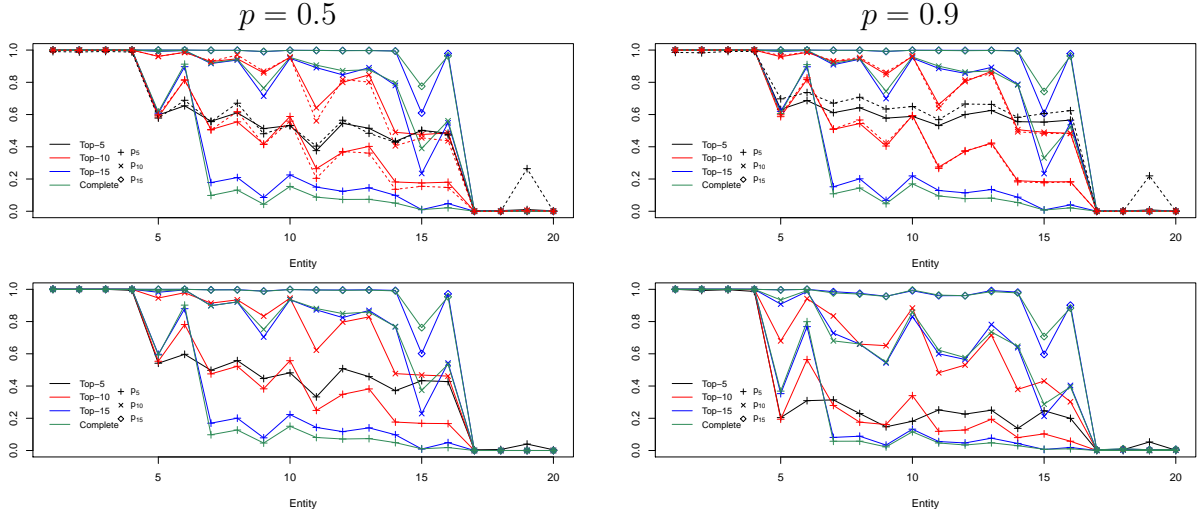


Figure 4: Posterior probabilities for each analysis plotted as frequency polygons for Dataset 1 (upper row) and Dataset 2 (lower row) under both prior specifications.

1–4 are highly likely to be within the top–5, however there is some doubt about which is the 5th strongest entity. In contrast, the other analyses (top–10, top–15 and complete) indicate that entity 6 is much more likely to be within the top–5 in comparison to the remaining entities. Furthermore notice how P_5 decreases significantly for entities 7–20 under the top–15 and complete analyses in comparison to the top–10 scenario. Similar results are obtained for P_{10} and P_{15} , that is, as we increase the information contained within the rankings, we become more certain about those entities which are the most preferred. In conclusion, although it is possible to identify the most preferred entities without complete rankings, it has been advantageous to incorporate as much information as possible. Also if interest lies in inferring the most preferred i entities, we have found that little information is lost if rankers only give top– M complete rankings where $M > i$, with more information lost as M approaches i .

1.2.1 Dataset 2 analysis conditional on two ranker clusters

Here we revisit the analysis of Dataset 2 under the choice of $p_i = 0.9$ *a priori*. In the previous section we looked at the posterior entity clustering structure conditional on a single ranker cluster. However there was little posterior support for a single ranker cluster, with $\Pr(N^r = 1 | \mathcal{D}, p_i = 0.9) \leq 0.16$ under all analyses considered, and the modal posterior number of ranker clusters is two. Therefore we now look at the (posterior) clustering structure conditional on this modal number of ranker groups (as would be done if we had no knowledge of the generating mechanism for these data).

Table 3 gives the marginal posterior distributions of the number of entity clusters within each ranker group (conditional on two ranker clusters) for all analyses. Note that, within ranker cluster 1, the posterior support for the correct number of entity clusters ($N_1^e = 6$) increases as the information provided within each ranking increases – this was also

Dataset 2	Cluster	1	2	3	4	5	6	7	8	9	≥ 10
Top-5	1	0.01	0.01	0.16	0.26	0.25	0.17	0.09	0.03	0.02	0.00
	2	0.14	0.23	0.22	0.18	0.12	0.06	0.03	0.01	0.01	0.00
Top-10	1	0.00	0.00	0.02	0.11	0.21	0.23	0.21	0.12	0.06	0.04
	2	0.11	0.21	0.24	0.19	0.13	0.07	0.03	0.01	0.01	0.00
Top-15	1	0.00	0.00	0.00	0.08	0.21	0.28	0.23	0.12	0.06	0.02
	2	0.18	0.25	0.21	0.16	0.10	0.05	0.03	0.01	0.01	0.00
Complete	1	0.00	0.00	0.00	0.08	0.25	0.28	0.20	0.12	0.05	0.02
	2	0.20	0.24	0.23	0.15	0.10	0.05	0.02	0.01	0.00	0.00

Table 3: Posterior distribution of the number of entity clusters, conditional on two ranker clusters, for each analysis of Dataset 2 with $p_i = 0.9$. Numbers in bold indicate modal values.

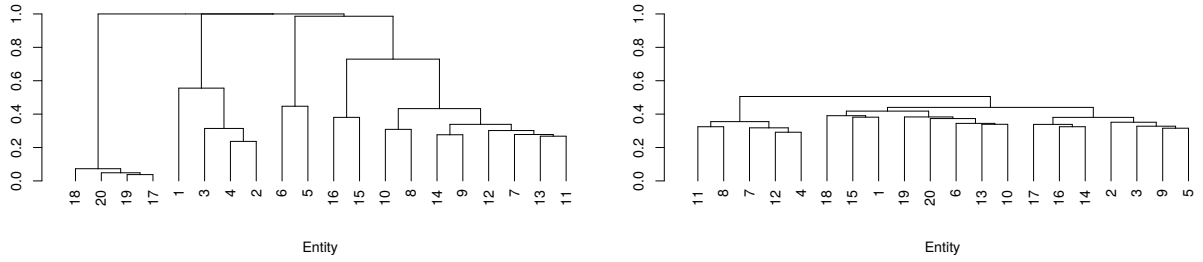


Figure 5: Dendrograms of entity clustering (conditional on 2 ranker clusters) in ranker cluster 1 (left) and ranker cluster 2 (right) for the analysis of Dataset 2 with $p_i = 0.9$.

observed when conditioning on a single ranker group. For ranker cluster 2 we see increased uncertainty within the marginal posteriors (in comparison to ranker cluster 1) with only two or three entity clusters being most probable under all analyses. Clearly the rankers in ranker cluster 2 are less able to distinguish between the entities – and is perhaps not surprising as this cluster typically houses the uninformative rankers.

Figure 5 shows dendrograms of the entity grouping structure within ranker clusters 1 and 2 (left and right respectively) for the complete analysis of Dataset 2 with $p_i = 0.9$, conditional on two ranker clusters. Note that the entity clusters within ranker cluster 1 are very similar to those when conditioning on a single ranker cluster; see Figure 3. This is probably due to ranker cluster 1 containing all informative rankers in both cases (of one or two ranker groups). For ranker cluster 2 we see that any two entities are clustered together at least 49% of the time ($\Delta_{ij} < 0.51$). Also entity clusters are formed at similar levels of dissimilarity, again highlighting the uncertainty on the entity clustering within this ranker group.

1.3 Study 2

In study 2 we look at a single dataset with $n = 40$ *complete* rankings of $K = 20$ entities from informative ($w_i = 1$) rankers. We also simulate the cluster allocations (for both rankers and entities) marginally from the prior. The distinct skill parameters and cluster allocations were simulated using $\alpha = \gamma_s = 1$ for $s \in \mathbb{N}$, and $a = 1$ so that our base distribution is $G_0 = \text{Ga}(1, 1)$. The simulation gave three ranker clusters ($N^r = 3$) containing 24, 12 and 4 rankers, which we label as rankers 1–24, 25–36 and 37–40. Also the ranker clusters contained 8, 6 and 3 entity clusters ($N_1^e = 8$, $N_2^e = 6$, $N_3^e = 3$).

Table 4 shows the entity clustering (within each ranker cluster) along with the associated true values of the skill parameters. For ease of interpretation, the entities are labelled according to the size of their “true aggregate” skill parameter, largest to smallest, so that they are labelled with the most preferred entity overall first, down to the least preferred entity overall last. Here the true aggregate values are an average of the true parameter values within each ranker cluster, weighted by the size of the ranker clusters. The complete (simulated) rankings analysed within this study can be found in Table 13.

The purpose of this study is to investigate the ability of our WAND model to (correctly) identify different ranker groups and the associated preferences therein. The analysis given here uses the same base distribution and prior distribution for the entity concentration parameters as in Section 1.2, that is, $G_0 = \text{Ga}(1, 1)$ and $\gamma_s \sim \text{Ga}(3, 3)$ for $s \in \mathbb{N}$. To reflect the known ranker heterogeneity within these data now take $a_\alpha = b_\alpha = 3$, that is, $\alpha \sim \text{Ga}(3, 3)$. We also consider the case where we have only moderate confidence in our rankers being informative by taking the $p_i = 0.5$.

The left plot in Figure 6 shows the posterior probabilities, $\Pr(w_i = 1 | \mathcal{D})$, that ranker i is informative. The plot shows that, in general, the rankers in (true) ranker clusters 1 and 2 (rankers 1–36) are well identified to be informative. However the rankers in (true) ranker cluster 3 are identified as uninformative. The reason for this misidentification is perhaps due to the (true) entity clustering structure present within ranker cluster 3. Table 4 shows that this ranker cluster contains only 3 entity clusters, with one of these containing 16 out of the 20 entities, and so it is very likely that rankings in this cluster resemble a random permutation of the K entities.

The right plot in Figure 6 shows the complete linkage dendrogram determined using dissimilarities Δ_{ij} . The dendrogram suggests there are two ranker clusters (taking height $\Delta \in$

C^r	Rankers	Entity cluster							
		1	2	3	4	5	6	7	8
1	1–24	1 2.47	6 0.65	10,13 0.52	3,4,7,9,12,15 0.40	2 0.34	5,11,14,17–20 0.24	16 0.02	8 0.01
2	25–36	2–5,8 1.72	1,7,9,11 0.76	14,16 0.68	12 0.42	6,10,15,17 0.21	13,18–20 0.15		
3	37–40	2,6,10 1.54	7 1.16	1,3–5,8,9,11–20 0.64					

Table 4: True allocation of entities in to clusters, along with the corresponding true parameter value for each of the ranker clusters.

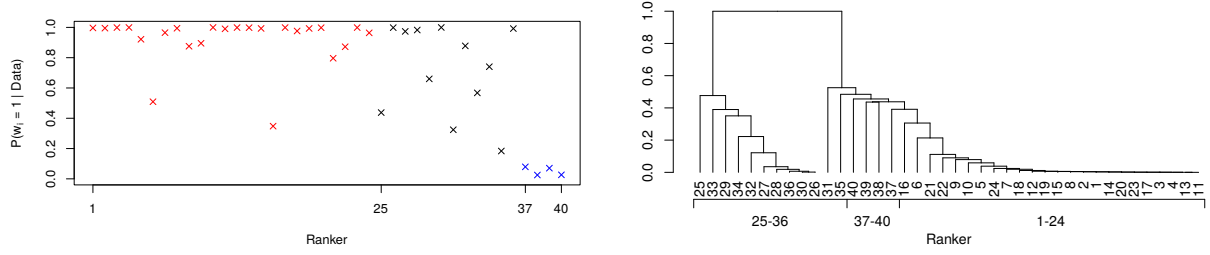


Figure 6: Plot of the posterior probability $\Pr(w_i = 1|\mathcal{D})$ that ranker i is informative (left). Dendrogram (complete linkage) computed using the dissimilarity Δ_{ij} between rankers i and j .

(0.53, 1)) which separates those rankers numbered $\{25-30, 32, 33, 34, 36\}$ from the remaining rankers. That there are two ranker clusters is supported further by the marginal posterior distribution of the number of ranker clusters: $\Pr(N^r = i|\mathcal{D}) = 0.68, 0.25, 0.06, 0.01$ for $i = 2, 3, 4, 5$. It is not surprising that the analysis has not identified the third ranker cluster as this cluster struggles to discriminate between entities and the model prefers to deem such rankers as uninformative and place them within clusters of informative rankers.

Table 5 gives the marginal posterior distribution for the number of entity clusters within each ranker cluster, conditional on the posterior modal number of ranker clusters. The modal number of entity clusters within ranker clusters 1 and 2 is six and four respectively (the corresponding true values are eight and six). Here the analysis has correctly identified that ranker cluster 1 is the stronger cluster, in that these rankers are more able to distinguish between entities. The dendrograms in Figure 7 suggest that there are five entity clusters within ranker cluster 1 (taking $\Delta_1 \in (0.58, 0.83)$) and three entity clusters in ranker cluster 2 (taking $\Delta_2 \in (0.50, 0.69)$). Notice that in ranker cluster 1, the most preferred entity in this cluster (entity 1) has its own cluster, and entities 16 and 8 (in true entity clusters 7 and 8) also form a single cluster; perhaps these are not surprising given the “true” values of the skill parameters for these entities within this ranker cluster (see Table 4). True entity cluster 6 is fairly well identified with only entity 14 not being included and entity 3 (from true cluster 4) joining the cluster. The remaining two entity clusters identified by the dendrogram house the other entities from true entity clusters 2–5. Within ranker cluster 2 the “true” entity clustering structure from which the data were simulated is largely preserved but the inferred clusters are groups of the “true” clusters, with all entities in “true” cluster 1 being clearly identified in one cluster and those in clusters 2, 3 and 4 in another cluster and those in clusters 5 and 6 in another cluster. That these entity clusters have merged is perhaps not too surprising given the true values (see Table 4) and the limited number of rankings observed.

Cluster	1	2	3	4	5	6	7	8	9	≥ 10
1	0.00	0.00	0.00	0.07	0.20	0.25	0.21	0.14	0.08	0.05
2	0.00	0.11	0.25	0.26	0.18	0.11	0.05	0.02	0.01	0.01

Table 5: Posterior distribution of the number of entity clusters, conditional on two ranker clusters.

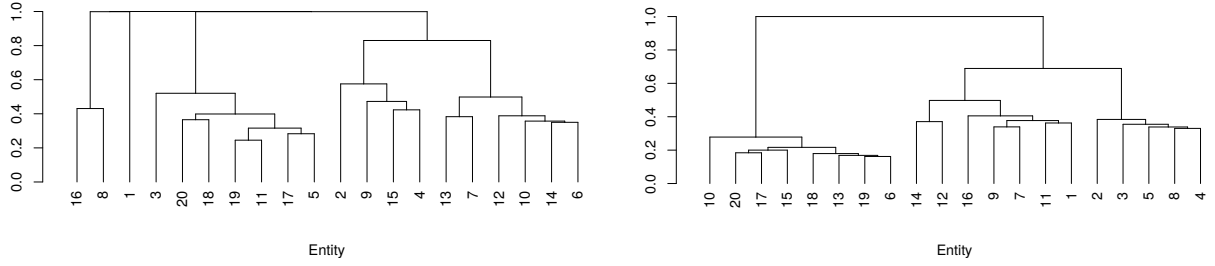


Figure 7: Dendrograms showing the dissimilarity between entities within ranker clusters 1 (left) and 2 (right), conditional on two ranker clusters ($N^r = 2$).

2 Further details on aggregate rankings

2.1 Determining the modal predictive ranking

Recall from Section 5.2 in the paper that we propose that the aggregate ranking should be the permutation of the entities which gives rise to the largest posterior predictive probability. More specifically we define the aggregate as

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{S}_K} \Pr(\mathbf{X} = \mathbf{x} | \mathcal{D}),$$

where \mathcal{S}_K is the set of permutations of K entities and $\Pr(\mathbf{X} = \mathbf{x} | \mathcal{D})$ is the posterior (predictive) probability of ranking \mathbf{x} under our model. Thus a simple way to determine the modal ranking within the posterior predictive distribution is to evaluate the posterior predictive probabilities (estimated using the MCMC output) at all possible rankings. However, when there are many entities being ranked, the number of possible rankings is prohibitively large. In such circumstances it makes sense to employ an optimisation algorithm. One algorithm that we have found to work well is based on the cyclic coordinate ascent (CCA) algorithm; see, for example, Lange (2013). In the following sections we describe the CCA algorithm and give details on how it can be used to obtain local maxima of functions defined over permutations. In Section 2.2 we outline a method for obtaining approximate aggregate rankings via a delta approximation and then conclude with a discussion of the appropriateness of aggregate rankings.

2.1.1 Cyclic coordinate ascent (CCA) algorithm

The CCA algorithm is a deterministic optimisation algorithm for finding the maximum of a function of several continuous variables. It works by first determining the optimal choice of the first coordinate (the first variable) with the others fixed at their current values. Next the first coordinate is changed to this optimal choice and the optimal choice of the second coordinate is determined (whilst all other coordinates are fixed at their current values). The algorithm then cycles through the remaining coordinates and in each case a coordinate value is replaced by its optimal value (which maximises the function) with all other coordinates remaining fixed. This process is repeated until no further improvement in the objective function is found. The resulting solution is a local maximum of the objective function. Multiple runs of the algorithm from different starting values can be used to explore possible local modes of the objective function and give confidence that the global mode has been located.

2.1.2 CCA over permutations via latent space representation

Now consider a function defined over the set of all permutations of K entities, such as the posterior predictive distribution over complete rankings. We can use the CCA algorithm to determine the modal permutation by first considering a ranking of K entities as resulting from an ordering (largest to smallest) of the coordinate values of a point

in (continuous) K -dimensional space. Consider a toy example, with $K = 3$; the point $(0.7, 0.9, 0.8)$ would produce the permutation (ranking) $(2, 3, 1)$. Thus we can view the objective function as being one over this K -dimensional latent space and employ the cyclic coordinate ascent algorithm to determine the modal ranking. Of course, the objective function (that is, the posterior predictive probability function) only changes value when one of its latent dimension values changes the permutation at which the function is evaluated. For example, in the toy example, if the first coordinate is less than 0.8 then the permutation stays as $(2, 3, 1)$, if it takes a value in the interval $(0.8, 0.9)$ the corresponding permutation is $(2, 1, 3)$ and if it is greater than 0.9 then the corresponding permutation is $(1, 2, 3)$. Thus, in terms of permutations, the search for an optimal choice in latent coordinate 1 is equivalent to moving entity 1 to all other positions whilst keeping all other entities in the same ordering. Such a move could be thought of as an “insertion” of entity 1 into the other positions. The optimal choice for latent coordinate 1 is the insertion which results in the largest value of the objective function (the posterior predictive probability). This permutation becomes the current optimal. Next attention turns to latent coordinate 2; as the latent coordinate varies it is equivalent to entity 2 being inserted in the other positions within the permutation; the optimal choice of coordinate 2 is found from these insertions as that which maximises the objective function (the posterior predictive probability). Attention then turns to position 3 (entity 3), then position 4 (entity 4), and so on up to position K (entity K). This defines a full cycle. The cycle is repeated until there is no change in the optimal permutation. This gives a local maximum of the objective function (with respect to the *Ulam distance*, as defined in Marden (1995)). As mentioned above, multiple runs of the algorithm from different starting values can be used to explore possible local modes and give confidence that the global mode has been located.

We have examined this algorithm by looking at its performance when the objective function over $K!$ permutations is a q -component mixture of Plackett-Luce distributions. Over a range of values for $q \in \{1, 2, 3, 4, 5\}$ and $K \leq 30$ the algorithm appears to perform well, with a high probability of obtaining the global mode from a relatively small number of starting points. We note that the complexity of the algorithm is roughly quadratic in K , and so may not scale well to very large K .

2.2 Approximations to aggregate rankings

Determining the modal ranking from the posterior predictive distribution may be computationally demanding and so approximations may be useful, particularly if the number of entities K is very large. First note that, when the skill parameters are known (or have been estimated), the ranking that maximises the standard Plackett-Luce probability is that in which the entities are ordered by their skill parameter values, from largest to smallest. In a Bayesian analysis, with posterior uncertainty on the skill parameters, the relevant probability to maximise is the (standard) Plackett-Luce probability marginalised over the posterior distribution of the skill parameters. More specifically the modal ranking \mathbf{x}^* is

given by

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{S}_K} E_{\lambda'|\mathcal{D}} \{\Pr(\mathbf{X} = \mathbf{x}|\lambda')\}$$

where \mathcal{S}_K is the set of permutations of K entities and $\Pr(\mathbf{X} = \mathbf{x}|\lambda')$ is the standard Plackett-Luce probability. The expectation can be approximated via the MCMC samples of the skill parameters. When the number of entities is large it is computationally infeasible to compute all the predictive probabilities. In this scenario we can take a delta approximation to the predictive probabilities and let the approximate aggregate ranking \mathbf{x}_Δ^* be

$$\mathbf{x}_\Delta^* = \arg \max_{\mathbf{x} \in \mathcal{S}_K} \Pr(\mathbf{X} = \mathbf{x} | E_{\lambda'|\mathcal{D}}\{\lambda'\}).$$

Thus this aggregate ranking can be obtained by simply ordering the entities by the posterior means of their skill parameters, from largest to smallest.

Moving now to the weighted Plackett-Luce model, the aggregate ranking should again be the modal ranking \mathbf{x}^* in the posterior predictive distribution, that is

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{S}_K} E_{\lambda', W|\mathcal{D}} \{\Pr(\mathbf{X} = \mathbf{x}|\lambda', W)\}$$

where $\Pr(\mathbf{X} = \mathbf{x}|\lambda', W)$ is the weighted Plackett-Luce probability (3). Therefore, enumerating over the posterior distribution of the weight W and noting that $\Pr(\mathbf{X} = \mathbf{x}|\lambda', W) = \Pr(\mathbf{X} = \mathbf{x}|\lambda'^W)$, a standard Plackett-Luce probability (2), and also that $\Pr(\mathbf{X} = \mathbf{x}|\lambda', W = 0) = 1/K!$ gives

$$\begin{aligned} \mathbf{x}^* &= \arg \max_{\mathbf{x} \in \mathcal{S}_K} E_{\lambda', W=1|\mathcal{D}} \{\Pr(\mathbf{X} = \mathbf{x}|\lambda')\} \Pr(W = 1|\mathcal{D}) + \Pr(W = 0|\mathcal{D})/K! \\ &= \arg \max_{\mathbf{x} \in \mathcal{S}_K} E_{\lambda', W=1|\mathcal{D}} \{\Pr(\mathbf{X} = \mathbf{x}|\lambda')\}. \end{aligned}$$

Thus the aggregate ranking for the weighted Plackett-Luce model can be obtained via the same mechanisms as the aggregate ranking for the standard Plackett-Luce model. That is, we can either enumerate the predictive probabilities for all rankings in \mathcal{S}_K or alternatively approximate the aggregate ranking by ordering the entities by the posterior means of their skill parameters, from largest to smallest.

Determining the aggregate ranking in the WAND model follows a similar pattern, that is, the aggregate ranking should be that with largest posterior predictive probability. It follows that this aggregate ranking correctly takes into account the uncertainty on the number of ranker groups, allocation of rankers to ranker groups, entity skill parameters and the entity groupings in ranker groups. Recall from Section 4.2.1 of the paper that the skill parameter vector for (ranker) cluster c is given by $\lambda'_c = g(c, D, \Lambda)$ where $g(c, D, \Lambda)_j = \lambda_{cd_{c_j}}$ for $j = 1, \dots, K$. Then, using the same arguments as above, we can write the aggregate ranking as

$$\begin{aligned} \mathbf{x}^* &= \arg \max_{\mathbf{x} \in \mathcal{S}_K} E_{N^r, c, \mathbf{N}^e, D, \Lambda, W|\mathcal{D}} \{\Pr(\mathbf{X} = \mathbf{x}|\lambda', W)\} \\ &= \arg \max_{\mathbf{x} \in \mathcal{S}_K} E_{N^r, c, \mathbf{N}^e, D, \Lambda, W=1|\mathcal{D}} \{\Pr(\mathbf{X} = \mathbf{x}|\lambda')\}. \end{aligned}$$

Thus, by again making use of a delta approximation, the aggregate ranking is approximated as that in which the entities are ordered by their (marginal) posterior mean skill parameter values $E_{N^r, c, \mathbf{N}^e, D, \Lambda, W=1|\mathcal{D}}\{\boldsymbol{\lambda}' = g(c, D, \Lambda)\}$. Note that here we must be careful to exclude any contributions from rankers with $W_i = 0$ when averaging over the MCMC samples of ranker specific parameter vectors; this was not an issue when considering the more simple models previously as there was only a single parameter vector being considered.

Within-cluster (group specific) aggregate rankings can also be obtained via the same mechanism but here the aggregate ranking is the mode of the predictive distribution obtained having first conditioned on a fixed number of ranker groups N^r . For example, conditional on two ranker groups, the (ranker) group 1 aggregate ranking \mathbf{x}_1^* is approximated by ordering the (marginal) posterior mean skill parameter values $E_{N^r=2, c=1, \mathbf{N}^e, D, \Lambda, W=1|\mathcal{D}}(\boldsymbol{\lambda}' = g(1, D, \Lambda))$.

Taking a delta approximation to the predictive distribution provides a very straightforward way of obtaining an approximate aggregate ranking. However, it is unclear as to when this delta approximation strategy gives rise to an accurate approximation to the mode of the posterior predictive distribution. Therefore, when computationally feasible, it is preferable to determine the aggregate ranking by using the CCA algorithm or by enumerating the posterior predictive distribution over all possible rankings.

2.3 Appropriateness of aggregate rankings

Although the strategies above allow us to obtain the overall aggregate ranking, it is still unclear as to when the overall aggregate ranking is meaningful in the sense that it is representative of all the rankers' views. For example, the posterior predictive distribution could be highly multi-modal and therefore a single (aggregate) ranking summary of this distribution may not be sensible. Of course, it is possible to simply examine the local modes of the posterior predictive distribution and to take a subjective view as to how different these are and whether the overall aggregate ranking is representative. However, we favour a more objective procedure in which the predictive probabilities of the local modes (computed under the full predictive distribution) are compared with that of the overall aggregate ranking. Here we would conclude that the overall aggregate ranking is a useful summary if it has substantially larger predictive probability than any of the other (local mode or within-cluster aggregate) rankings. In contrast, if numerous rankings have reasonable predictive probability then a single ranking summary may not be useful and the subgroup structure of the rankers must be accounted for when summarising the ranker views.

3 Real data analyses

We now analyse three real datasets which together contain a mixture of complete, top- M and partial rankings. These datasets have been analysed in the literature and we compare their conclusions with those obtained from a WAND analysis.

3.1 Roskam’s dataset

We begin by analysing a dataset originally collected in 1968 by Roskam, more recently studied by de Leeuw (2006), and available in the R package *homals* (de Leeuw and Mair, 2009). The data consist of rankings obtained from $n = 39$ psychologists within the Psychology Department at the University of Nijmegen (Netherlands). Each ranker gives a complete ranking of $K = 9$ sub-areas (entities), listed according to how appropriate the sub-areas are to their work. The sub-areas are: SOC - Social Psychology, EDU - Educational and Developmental Psychology, CLI - Clinical Psychology, MAT - Mathematical Psychology and Psychological Statistics, EXP - Experimental Psychology, CUL - Cultural Psychology and Psychology of Religion, IND - Industrial Psychology, TST - Test Construction and Validation, and PHY - Physiological and Animal Psychology.

The heterogeneity within these data has been analysed by de Leeuw (2006) using a non-linear principal component analysis to detect groupings within the rankings. Their analysis supported the idea that there are two groups of rankings: one group which favours the qualitative fields and the other favouring the quantitative fields of psychology. A homogeneity analysis was later performed by de Leeuw and Mair (2009) which exposed groupings of entities within the rankings. More recently Choulakian (2016) performed a Taxicab correspondence analysis to look at structure both between the rankings and the entities within ranker groups. Their results support the conclusions of de Leeuw (2006) and suggest that the psychologists comprise two homogeneous groups with 23 and 16 members respectively. Within the larger ranker group they obtain the entity clustering $\{\text{MAT}, \text{EXP}\} \succ \{\text{IND}, \text{TST}\} \succ \{\text{PHY}, \text{SOC}, \text{EDU}\} \succ \text{CLI} \succ \text{CUL}$, where \succ means “is preferred to”, and quantitative areas of psychology appear to be preferred. The corresponding clustering of entities for the other ranker group is $\{\text{EDU}, \text{CLI}, \text{SOC}\} \succ \{\text{CUL}, \text{MAT}, \text{EXP}\} \succ \{\text{TST}, \text{IND}\} \succ \text{PHY}$, and here qualitative areas of psychology appear to be preferred. They also conclude that the larger ranker group is somewhat more homogeneous than the smaller group.

We now use our WAND model to investigate subgroup structure in these data and take our prior specification for the base distribution and concentration parameters to be $a = 1$ and $a_\alpha = b_\alpha = 1$, $a_\gamma = b_\gamma = 3$. These data contain orderings of individual preferences which we believe to be informative and so take $p_i = 0.75$. The posterior distribution is fairly robust to this choice; a sensitivity analysis is considered later in Section 3.1.2. We report the results from a typical run of our MCMC scheme initialised from the prior, with a burn-in of 10K iterations and then run for a further 1M iterations and thinned by 100 to obtain 10K (almost) un-autocorrelated realisations from the posterior distribution. Convergence was assessed by using multiple starting values, inspection of traceplots of

	1	2	3	4	5	6	7	≥ 8
Posterior	0.00	0.43	0.33	0.16	0.06	0.02	0.00	0.00
Prior	0.20	0.18	0.16	0.13	0.10	0.08	0.05	0.10

Table 6: Prior and posterior distribution of the number of ranker clusters (to 2 d.p.).

parameters and the logarithm of the complete data likelihood, and standard statistics available in the R package *coda* (Plummer et al., 2006). The MCMC scheme runs fairly quickly, with C code on a single thread of an Intel Core i7-4790S CPU (3.20GHz clock speed) taking around 5 minutes.

Table 6 shows both the prior and posterior distribution for the number of ranker clusters. The data clearly have been informative and suggest that it is likely that there are between two and four ranker groups, with two groups being most plausible. Note that there is almost no posterior support to suggest there is a single (homogeneous) ranker group. The posterior distribution of the allocation of rankers to ranker groups is, of course, quite complex. Rather than attempting to summarise the ranker allocation through a single summary allocation to each ranker group, such as the maximum *a posteriori* (MAP) allocation or the alternatives to the MAP allocation proposed by Dahl (2006) and Lau and Green (2007), we prefer to use a graphical summary of the posterior clustering structure. Following Medvedovic and Sivaganesan (2002), we summarise ranker heterogeneity using a dendrogram constructed from an agglomerative clustering of the rankers based on a dissimilarity matrix $\Delta = (\Delta_{ij})$. We take $\Delta_{ij} = \Pr(c_i \neq c_j | \mathcal{D})$, the posterior probability that two rankers (i and j) are *not* allocated to the same cluster. We use the complete linkage method, also known as furthest neighbour clustering, to cluster the rankers as this tends to produce balanced cluster sizes and does not suffer from “chaining”. The complete linkage method is fairly straightforward: starting with each ranker in their own singleton cluster, at each stage of the clustering process the two most similar clusters are merged to form a new cluster; the dissimilarity between this merged cluster and the remaining clusters is defined as the maximum of the pairwise dissimilarities between the members of the merged cluster and the remaining clusters. Clusters are merged one at a time until all rankers are in a single cluster. A dendrogram provides a visual representation of this agglomerative clustering procedure, with “branches” (horizontal lines) drawn between clusters that are merged; the heights of the horizontal lines represent the dissimilarity at which the clusters are merged. Further details on hierarchical clustering and dendrograms can be found, for example, in Everitt et al. (2011).

Figure 8 shows the dendrogram of rankers along with the posterior probability that each ranker is informative. The dendrogram shows that rankers 22 and 24 are the first to be merged (at height 0.01) and can therefore be considered the most similar pair of rankers; the posterior probability that these rankers are clustered together is around $1 - 0.01 = 0.99$. The dendrogram also provides useful information about the pairwise clustering probabilities between groups of rankers. For example, the horizontal line connecting rankers 3 and 29, to rankers 2, 4 and 30 indicates that no pair of rankers from the set $\{2, 3, 4, 29, 30\}$ have a posterior probability of being clustered together of less than $1 - 0.15 = 0.85$. Also note that all rankers merge at height 0.97, and so the

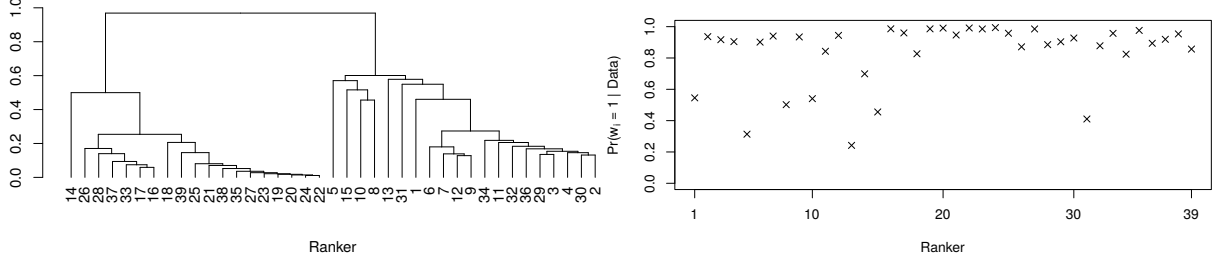


Figure 8: Roskam’s dataset: Dendrogram (left) showing the ranker cluster structure along with the posterior probability, $\Pr(w_i = 1|\mathcal{D})$, for each ranker i (right).

most dissimilar pair of rankers are clustered together with fairly low posterior probability of 0.03. Overall this is indicative of fairly strong heterogeneity in ranker preferences and is consistent with the posterior distribution in Table 6. We note that the data are consistent with most rankers being informative as $\Pr(w_i = 1|\mathcal{D}) \geq 0.8$, an increase from their prior probabilities ($p_i = 0.75$). Also the rankers whose probabilities have decreased (rankers 1, 5, 8, 10, 13, 14, 15, 31) are those with (slightly) different preferences and hence late to join the right-hand cluster in the dendrogram.

One of the novel features of a WAND analysis is its ability to reveal subgroup structure of entities within ranker clusters, that is, highlight entities that are thought to be similar within ranker groups. We now examine the entity clustering by conditioning on there being two ranker clusters. Figure 9 gives the (marginal) posterior distribution for the number of entity clusters within each ranker cluster, together with the prior distribution. The dendrograms in Figure 10 show the entity clustering structure in each ranker cluster, based on dissimilarity probabilities $\Pr(d_{ij} \neq d_{ij'}|\mathcal{D})$ of entities j and j' within ranker cluster i . We can determine the membership of the entity clusters by cutting the dendrogram at heights $\Delta_1 \in (0.45, 0.95)$ and $\Delta_2 \in (0.63, 0.89)$ for rankers groups 1 and 2 respectively and form a preference ordering of these entity clusters by ordering the marginal posterior means for the skill parameters $\lambda_{c_i d_{c_i j}}$ within each ranker group c_i . Conditioning on the ranker and entity allocations, we obtain the ordering $\{\text{EXP}, \text{MAT}\} \succ \{\text{TST}, \text{PHY}, \text{IND}\} \succ \{\text{EDU}, \text{SOC}, \text{CLI}\} \succ \{\text{CUL}\}$ (with entity cluster means 3.02, 0.72, 0.22, 0.06) in ranker cluster 1 and $\{\text{SOC}, \text{EDU}, \text{CLI}, \text{MAT}\} \succ \{\text{CUL}, \text{IND}, \text{EXP}, \text{TST}\} \succ \{\text{PHY}\}$ (with entity cluster means 1.96, 0.82, 0.12) in ranker cluster two. These entity clusters (within ranker groups) are similar to those given by Choulakian (2016). Also if we use the average value of $\Pr(w_i = 1|\mathcal{D})$ as a measure of homogeneity within a ranker cluster then we obtain 0.68 and 0.56 for clusters 1 and 2 respectively, which again agrees with the Choulakian (2016) conclusion that ranker cluster 1 is more homogeneous than ranker cluster 2.

It is also of interest to examine both the overall and within-cluster aggregate rankings. Recall from Section 5.2 of the paper that we define the aggregate ranking to be the mode of the (appropriate) posterior predictive distribution. For these data the predictive distribution is defined over the $9! = 362880$ possible permutations and, although computationally expensive, it is possible to enumerate the full predictive distribution in this case. Table 7 shows the within-cluster aggregates for ranker clusters 1 and 2 (conditional on two ranker

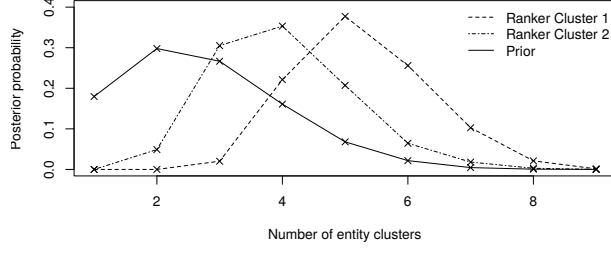


Figure 9: Prior and marginal posterior densities for the number of entity clusters within each ranker cluster (conditional on two ranker clusters).

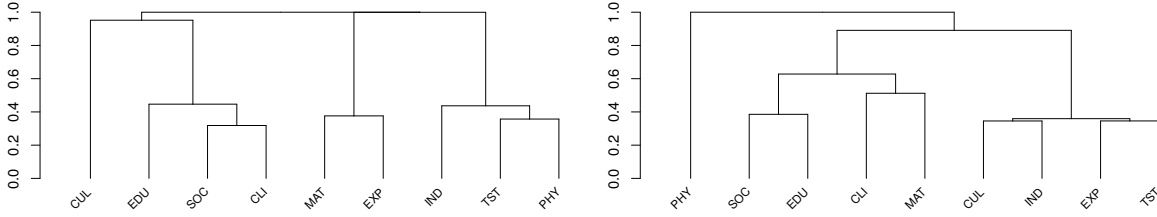


Figure 10: Roskam’s dataset: Dendrograms showing the entity clustering structure within ranker cluster 1 and 2 (left and right respectively) conditional on two ranker clusters.

clusters) along with the overall aggregate ranking. Note that, for the within-cluster aggregates, the true mode of the predictive distribution is also recovered via both the cyclic coordinate ascent and the delta approximation approaches. The within-cluster aggregates show a very similar entity ordering as to when we conditioned on specific ranker and entity clustering structure previously. It is also interesting to see that the table suggests that the ranker groups have almost opposite (reverse) preferences to each other. Table 7 also gives the probability of the within-cluster aggregate rankings under the full predictive distribution; 3.47×10^{-4} and 3.31×10^{-5} for ranker clusters 1 and 2 respectively. Although it is much more likely that a ranker shares the views expressed by (ranker) group 1 there is still non-negligible (predictive) support for the group 2 aggregate. Note that a ranking would have probability $1/9! = 2.76 \times 10^{-6}$ under a uniform distribution. Interestingly the cluster 1 aggregate ranking is also the modal ranking of the full posterior predictive distribution, that is, this is also the overall aggregate ranking when accounting for the uncertainty on the number of ranker groups. However, as alluded to above, the predictive distribution is clearly multi-modal. This is further highlighted by the cyclic coordinate ascent algorithm revealing several (local) modes with relatively large predictive probability; see Table 7. We therefore conclude that the aggregate ranking may not be a sensible summary for these data on its own as it does not effectively capture the conflicting views of the rankers. Also note that, in contrast to the within-cluster aggregates, the delta approximation to the mode of the full predictive distribution does not perform well; with this approximate aggregate ranking having fairly low predictive probability in comparison to the actual mode. This is likely an artefact of the ranking from the delta approximation

Ranker cluster	Method	Rank									Pred prob
		1	2	3	4	5	6	7	8	9	
1	Enumeration	EXP	MAT	TST	PHY	IND	EDU	SOC	CLI	CUL	3.47×10^{-4}
	CCA	as above									
	Delta	as above									
	$E(\lambda' \mathcal{D})$	3.13	2.68	0.76	0.70	0.63	0.27	0.22	0.20	0.07	
2	Enumeration	SOC	EDU	CLI	MAT	CUL	IND	EXP	TST	PHY	3.31×10^{-5}
	CCA	as above									
	Delta	as above									
	$E(\lambda' \mathcal{D})$	1.95	1.75	1.49	1.32	0.94	0.90	0.87	0.87	0.10	
Overall	Enumeration	EXP	MAT	TST	PHY	IND	EDU	SOC	CLI	CUL	3.47×10^{-4}
	CCA	as above									
	CCA	SOC	EDU	CLI	MAT	CUL	IND	EXP	TST	PHY	3.31×10^{-5}
	CCA	EDU	CLI	SOC	MAT	EXP	TST	CUL	IND	PHY	3.13×10^{-5}
	Delta	EXP	MAT	SOC	EDU	TST	CLI	IND	CUL	PHY	1.16×10^{-5}
	$E(\lambda' \mathcal{D})$	2.06	2.00	0.98	0.94	0.81	0.79	0.76	0.47	0.44	

Table 7: Roskam’s dataset: within-cluster aggregate rankings (conditional on two ranker clusters) and the overall aggregate ranking along with their corresponding probabilities under the full predictive distribution. Note that the expectations $E(\lambda'|\mathcal{D})$ are taken with respect to the appropriate posterior distribution.

being obtained via an amalgamation of the λ' -values from contrasting (ranker) clusters as opposed to from the correct mixture probabilities.

We looked at the sensitivity of the posterior distribution (and inferences) to modest changes to the prior distribution; full details are given in Section 3.1.2. Briefly, the posterior distribution was fairly insensitive to changes in the index (a) of the gamma base distribution and to changes in the parameters ($a_\alpha, b_\alpha, a_\gamma, b_\gamma$) of the gamma prior distributions for the concentration parameters. The posterior distribution was most sensitive to changes in the prior probabilities (p_i) of rankers being informative. Not surprisingly most affected by such changes were their posterior equivalents $\Pr(w_i = 1|\mathcal{D})$ though the conclusion of two ranker groups and the membership of these groups was robust. The allocation of entities to groups (within each ranker cluster) was also fairly robust, with only a minor change in the allocation when the p_i were increased from 0.65 to 0.85.

3.1.1 Model assessment via posterior predictive checks

For Roskam’s data, each ranker rates $n_i = K = 9$ entities and so the (posterior) predictive distribution for each ranker contains probabilities for each of the $9! = 362880$ possible permutations. Figure 11 (left) shows the predictive distribution for ranker 1, calculated as described in Section 5.1 of the paper. The cross shows the predictive probability of ranking \mathbf{x}_1 provided by ranker 1, and suggests that this ranking is not unusual, that is, ranking \mathbf{x}_1 looks like a plausible realisation from its predictive distribution. We can

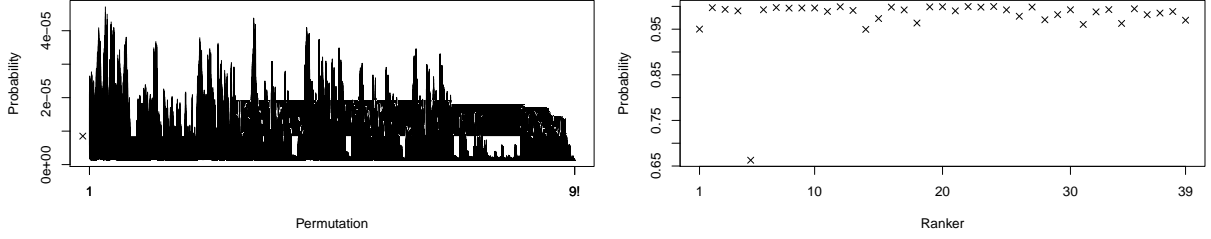


Figure 11: Roskam’s dataset: Posterior predictive distribution (left) for ranker 1 (left), the cross shows the posterior predictive probability of the observed data \mathbf{x}_1 . Plot (right) showing the diagnostic probability for each ranking (high values indicate that an observed ranking is close to the mode of its posterior predictive distribution).

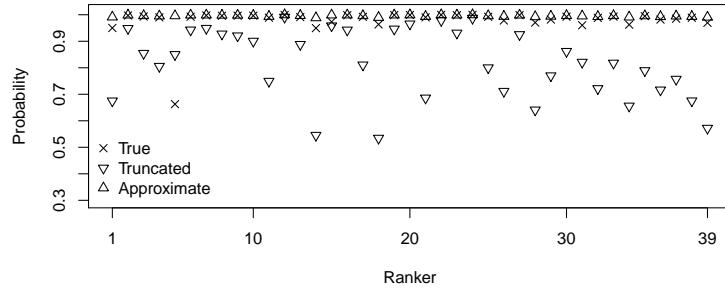


Figure 12: Roskam’s dataset: diagnostic probabilities for each observed ranking \mathbf{x}_i calculated using their full, truncated and approximate predictive distributions.

summarise this distribution (and those for other rankers) in a diagnostic plot (right hand plot in Figure 11) that shows the proportion of the $9!$ permutations that have a predictive probability less than or equal to that of the observed ranking. The plot shows that all observed rankings are consistent with their predictive distributions and this suggests that the WAND model provides a reasonable description of the data. The most outlying ranking is that from ranker 5 and, even so, the ranking has a larger predictive probability than around 65% of all $9!$ possible rankings. In Section 5.1 of the paper we noted that calculating the full predictive distribution is only computationally feasible when n_i is fairly small. Here $n_i = 9$ is on the cusp of being feasible and so we show, in Figure 12, the diagnostic probabilities determined from the truncated and approximate predictive distributions (for each ranker) using $L = 1$, as described in Section 5.1 of the paper. The figure confirms that, for these data, the alternative predictive distributions lead to a similar conclusion, namely that the observed rankings are consistent with the WAND model.

3.1.2 Prior sensitivity analysis for Roskam’s data

Here we look at the sensitivity of the posterior distribution to changes in the prior probability that a ranker is informative. We consider two alternative choices to the one used

in the paper ($p_i = 0.75$), namely $p_i = 0.65$ and $p_i = 0.85$. For ease of reference we also include the results for the $p_i = 0.75$ case.

Overall, we found that the posterior distribution was fairly robust to the choice of p_i *a priori*. Perhaps unsurprisingly the aspect of our posterior distribution most sensitive to changes in the p_i was their posterior equivalents $\Pr(w_i = 1|\mathcal{D})$; see Figure 13 (right column). However we note that the rankers whose informative probability decreases (prior \rightarrow posterior) remain the same in each case: these are rankers $\{1, 5, 8, 10, 13, 14, 15, 31\}$. The dendrograms of the ranker clustering structure are similar for each prior choice and clearly indicate that there are two groups of rankers. Also the allocation of rankers to clusters is similar in each case; see Figure 13 (left column). Interestingly we observe increasing posterior support for two ranker clusters as the p_i decrease – this is also the posterior mode in each case; see Figure 14. In addition, conditional on there being two ranker clusters, the marginal posterior of the number of entity clusters N_s^e (within each ranker cluster $s = 1, 2$) remains fairly robust to the prior choice; see Figure 14. Also the dendrograms of the entity clustering structure are similar in each case; see Figure 15.

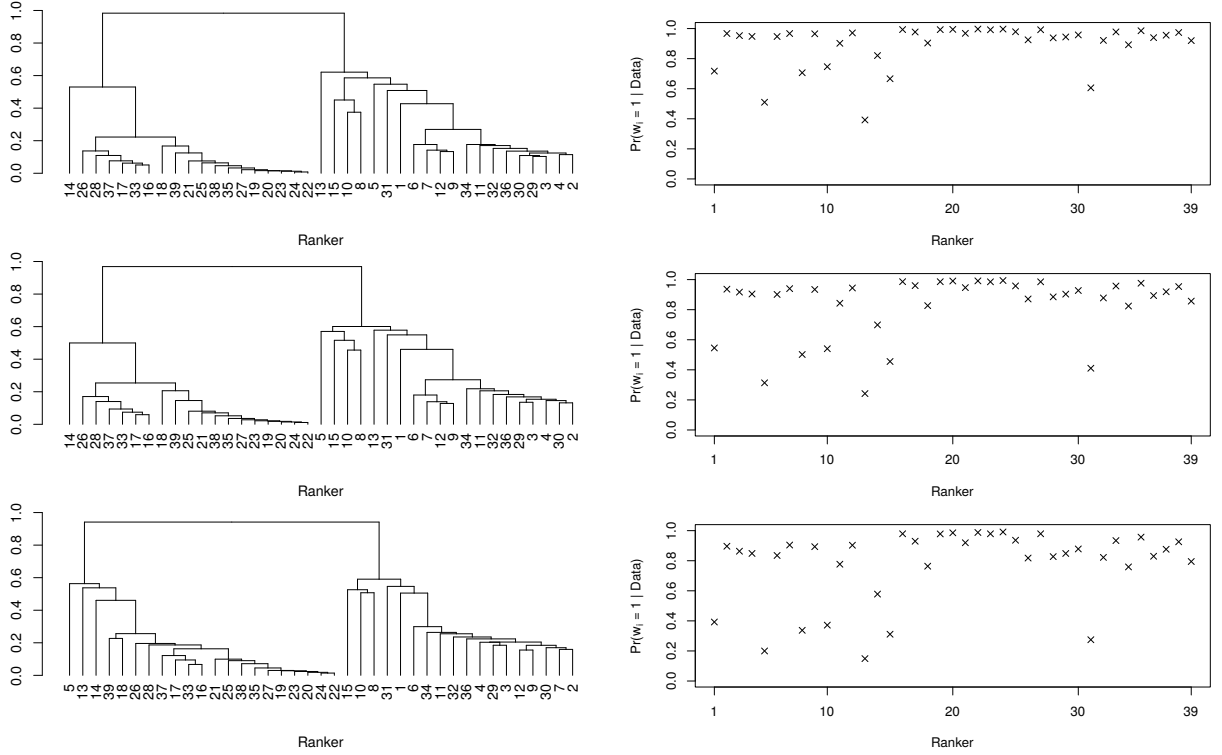


Figure 13: Roskam's dataset: Dendrogram (left) showing the cluster structure of the rankers along with the posterior probability $\Pr(w_i = 1|\mathcal{D})$ for each ranking i (right) for $p_i = 0.85, 0.75, 0.65$ (from top to bottom respectively).

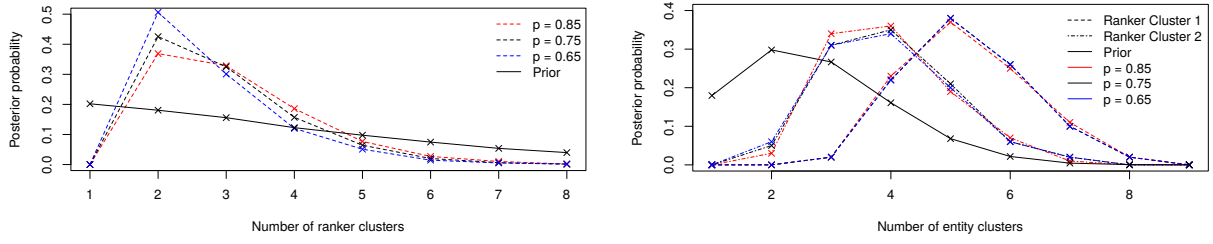


Figure 14: Roskam's dataset: Prior and marginal posterior densities for the number of rankers clusters (left plot) and the number of entity clusters within each ranker cluster, conditional on two ranker clusters, (right plot) for $p_i = 0.85, 0.75, 0.65$.

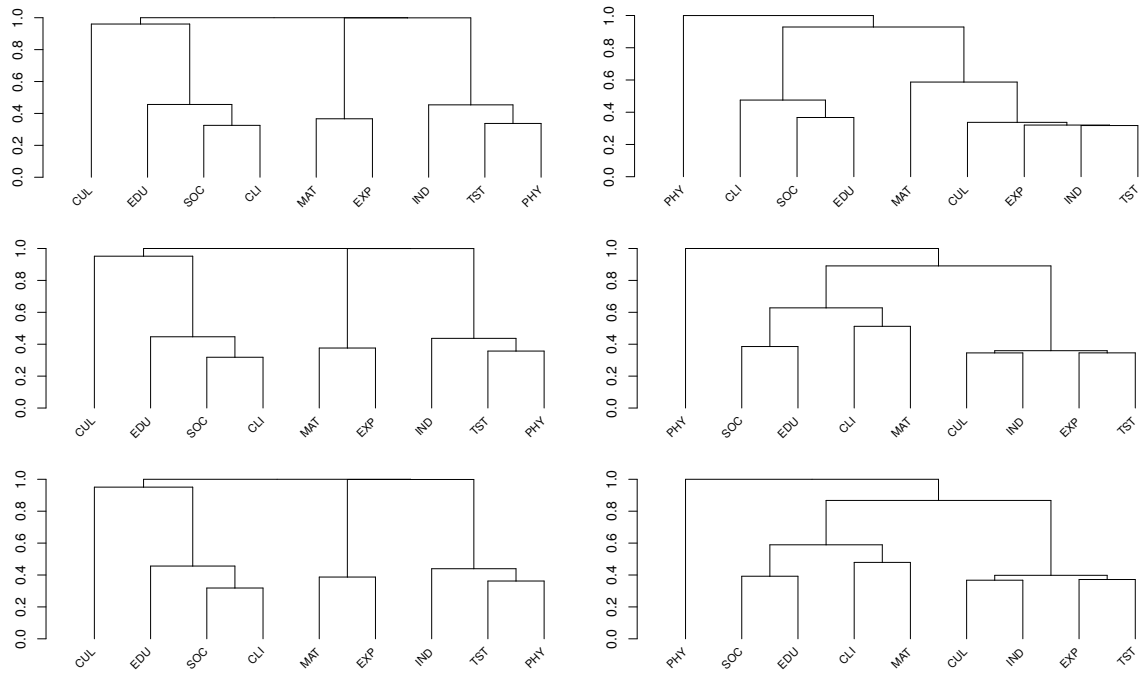


Figure 15: Roskam's dataset: Dendrograms of entity clustering structure within ranker cluster 1 (left) and ranker cluster 2 (right). These are shown for each prior specification, $p_i = 0.85, 0.75, 0.65$, from top to bottom respectively.

3.2 Rankings of NBA teams

We now consider another dataset of ranks, studied by Deng et al. (2014) and involving rankings of NBA (National Basketball Association) teams. In their paper, Deng et al. propose a model (named “Bayesian Aggregation of Ranked Data”, BARD) which aims to aggregate rankings and identify *relevant entities*; essentially the top- M strongest entities for some M . With rank aggregation in mind, the BARD model only permits a single ranker group and does not cater for entity clustering. However, as for our weighted Plackett–Luce model, BARD does allow for the possibility that rankers may not be equally reliable and can therefore down-weight rankings that are in some sense unusual. It will be interesting to see how our inferences compare with the conclusions under BARD, particularly whether we are able to identify similar groups of uninformative rankers and also to see if we detect a strong signal for a single ranker group.

In 2011/12 the NBA league contained $K = 30$ teams (entities) and the dataset we consider has a ranking of these teams from each of $n = 34$ rankers. The first six complete rankings were obtained from “professional” news agencies and the other top-8 complete rankings obtained from amateurs. Further, each amateur was asked to classify themselves into one of the following groups: “Avid fans” (never missed an NBA game), “Fans” (watched NBA games frequently), “Infrequent watchers” (occasionally watched NBA games) and “Not interested” (never watched NBA games). Each ranker considered all teams and so we have $K_i = K$ for $i = 1, \dots, n$. The rankers are numbered as follows: Professionals (1–6), Avid fans (7–12), Fans (13–18), Infrequent watchers (19–25) and Not interested (26–34). Therefore we have $n_i = K = 30$ for $i = 1, \dots, 6$ and $n_i = 8$ for $i = 7, \dots, n$. The data are reproduced in Table 15. Further details on how these data were collected can be found in Deng et al. (2014).

We now analyse these data using our WAND model to investigate any (potential) subgroup structure, in particular it will be interesting to see whether the self-assessed groups behave differently and also to determine how our aggregate ranking compared with that formed under BARD. We take the same prior for the base distribution ($a = 1$) as in the previous example. However, to reflect weak prior beliefs that there are several ranker groups, we take $a_\alpha = b_\alpha = 3$ in addition to the previous choice for entities, $a_\gamma = b_\gamma = 3$. The prior we choose for each ranker’s ability is based on how much attention they reportedly pay to the NBA, with professional rankers likely to be most informative, followed by the Avid fans, then Fans and so on. We do this by giving the same p_i -value for each ranker in the same “ability” group, with $p_i = 0.9$ for professionals, $p_i = 0.7$ for Avid fans, $p_i = 0.5$ for Fans, $p_i = 0.3$ for Infrequent watchers and $p_i = 0.1$ for Not interested. Of course, in general, information on rankers might well not be available and so, for comparison purposes we include an analysis with $p_i = 0.5$ in Section 3.3. We found that, as seen in the simulation studies and (other) real data analyses, although the $\Pr(w_i = 1|\mathcal{D})$ were fairly sensitive to changes in the p_i , many other aspects of the posterior distribution (such as the allocation distributions) were not.

As in the previous analysis, we report the results from a typical run of our MCMC scheme initialised from the prior, with a burn-in of 10K iterations and then run for a further 1M iterations and thinned by 100 to obtain 10K (almost) un-autocorrelated realisations from

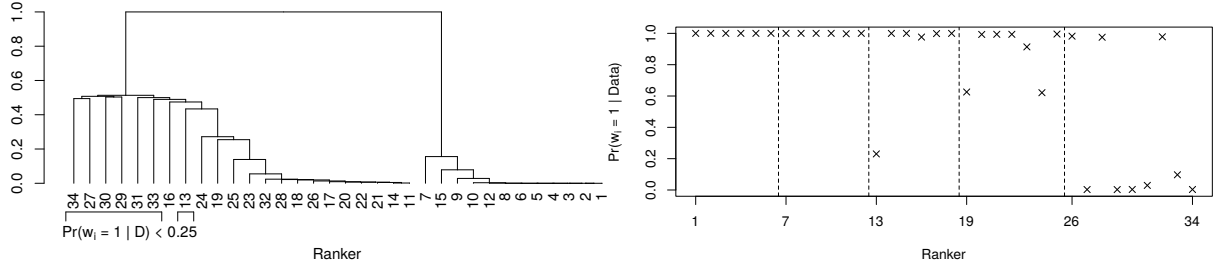


Figure 16: NBA dataset: Dendrogram (left) showing the clustering structure of rankers and highlighting those rankers with $\Pr(w_i = 1|\mathcal{D}) < 0.25$. Plot (right) of the posterior probabilities $\Pr(w_i = 1|\mathcal{D})$ for each ranker, with vertical lines separating the self-certified groups.

the posterior distribution. As in the previous analysis, convergence was assessed by using multiple starting values, inspection of traceplots of parameters and the logarithm of the complete data likelihood, together with standard diagnostics available in the R package *coda*. Again the MCMC scheme runs reasonably quickly, with C code on a single thread of an Intel Core i7-4790S CPU (3.20GHz clock speed) taking just under 18 minutes.

Our analysis of the posterior realisations reveals very little posterior support for a single homogeneous group of rankers, with most support for two ranker groups ($\Pr(N^r = 1|\mathcal{D}) = 0.00$, $\Pr(N^r = 2|\mathcal{D}) = 0.80$ and $\Pr(N^r = 3|\mathcal{D}) = 0.17$). Figure 16 (left) shows a dendrogram of the posterior clustering structure of rankers and confirms the conclusion that there are two distinct groups of rankers: one with rankers 1–10, 12, 15 and the other with rankers 11, 14, 17–26, 28 and 32. Nearly all the other rankers are classed as uninformative, with $\Pr(w_i = 1|\mathcal{D}) < 0.25$, except informative ranker 16 who is (roughly) equally likely to be allocated to each cluster; see Figure 16 (right). Unsurprisingly, uninformative rankers are typically those who pay less attention to the NBA, with average values of $\Pr(w_i = 1|\mathcal{D})$ for rankers in the self-certified groups (from professionals down to the not interested individuals) of 1, 1, 0.87, 0.88, 0.34 respectively. A similar conclusion was found under BARD through its ranking quality parameters; see Figure 8 in Deng et al. (2014).

Figure 17 shows the marginal posterior distribution for the number of entity clusters within each ranker cluster (conditional on there being two ranker clusters) together with the prior distribution. The posterior mean number of entity clusters for ranker clusters 1 and 2 is 8.88 and 4.58 respectively, with corresponding standard deviations 1.55 and 1.29. These distributions suggest that rankers within cluster 1 are able to distinguish between many more entities than those in cluster 2. Again this should come as no surprise as ranker cluster 2 mainly consists of rankers who typically pay little attention to the NBA. The dendrograms in Figure 18 show the entity clustering in each ranker cluster, and suggest that there are six distinct entity clusters within ranker cluster 1 (taking $\Delta_1 \in (0.81, 1)$) and three entity clusters in ranker cluster 2 (taking $\Delta_2 \in (0.61, 0.95)$). We note that the MCMC-based estimate of the MAP clustering gives six and two entity clusters respectively, though there are relatively few MCMC iterations contributing to this MAP

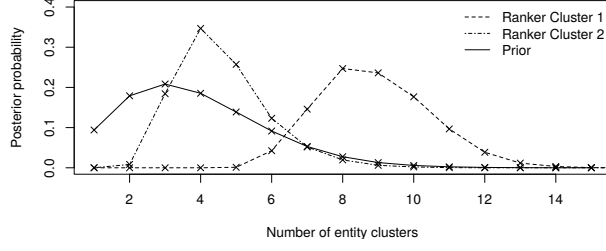


Figure 17: Prior and marginal posterior densities for the number of entity clusters within each ranker cluster (conditional on two ranker clusters).

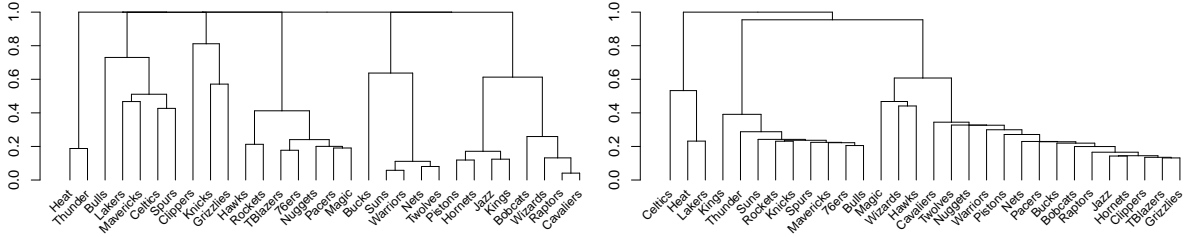


Figure 18: NBA dataset: Dendrograms showing the entity cluster structure within ranker clusters 1 and 2 (left and right respectively) conditional on two ranker clusters.

allocation for cluster 2.

It is also of interest to look at the differences in preferences between the two ranker clusters by examining the within-cluster aggregate rankings; see Table 8. As before, these are determined via both the cyclic coordinate ascent algorithm (initialised at numerous random permutations) and the delta approximation in which the entities are ordered by their marginal posterior mean for each ranker cluster. Note that complete enumeration is not possible in this case as the predictive distribution is defined over $30! = \mathcal{O}(10^{32})$ rankings. For this analysis, there are several discrepancies between the delta approximation to the group 2 aggregate and the “true” mode of the predictive distribution (obtained via the CCA algorithm). However the delta approximation method still produces a within-cluster aggregate that has predictive probability of the same order of magnitude. This is likely caused by the large amount of uncertainty on the entity positions within this cluster; particularly for those entities not in the top-5. The delta approximation to the group 1 aggregate was fairly robust with only slight discrepancies in the form of swaps between the entities in positions 5 and 6 along with those in positions 28 and 29. The horizontal lines in this table show the MCMC-based estimate of the MAP entity clustering described above and the (quite small) number of occurrences of this clustering in the MCMC sample is also given. The individuals in ranker cluster 1 strongly favour the Heat (entity 1) and Thunder (2), with the Bulls (10) as the 3rd most preferred team. Those in ranker cluster 2 also favour the Heat but differ in their preferences for second and third positions – here being the Lakers (6) and Celtics (4). There are many differences in preference orderings between the ranker clusters, for example, the Thunder and Bulls appear in positions 12

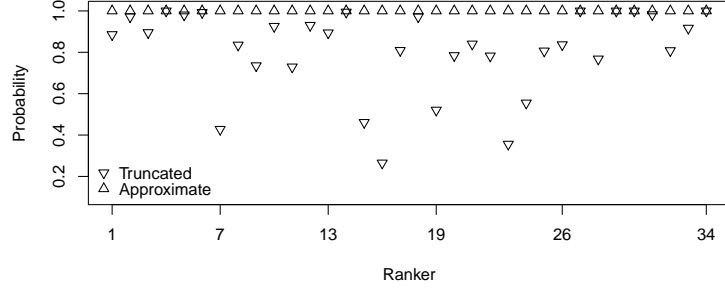


Figure 19: NBA dataset: diagnostic probabilities for each observed ranking \mathbf{x}_i calculated using their truncated and approximate predictive distributions.

and 5 in ranker cluster 2.

Of course, it is also possible to obtain the overall aggregate ranking as a summary of the (full) predictive distribution. Table 8 shows the overall aggregate rankings determined via the cyclic coordinate ascent and delta approximation approaches. As before, the delta approximation performs rather poorly in this case however the cyclic coordinate ascent approach reveals that the overall aggregate is that of ranker cluster 1 (conditional on 2 groups). This is perhaps not surprising given that group one typically houses the professional and avid fans whereas those rankers in group two struggle to distinguish between many of the entities. Also note that this ranking is $\mathcal{O}(10^8)$ times more likely than the group two aggregate under the full predictive distribution. We therefore conclude that, for this analysis, the overall aggregate is a sensible summary of these data. Note that the analysis given in Deng et al. (2014) looks at “relevant entities”, defined to be those entities within the top-16, and concludes that these are $\{1, 2, \dots, 15, 18\}$. The overall aggregate ranking under our WAND model gives the top-16 as $\{1, 2, \dots, 15, 17\}$, with entity 18 appearing in 17th position; see Table 8. For these data it is perhaps not surprising that there is considerable overlap between the aggregates formed from the WAND and BARD analyses given there is only a single ranker group that are capable of effectively discriminating between the entities.

3.2.1 Model assessment via posterior predictive checks

Figure 19 shows the diagnostic probabilities for each ranking, calculated from their truncated and approximate posterior predictive distributions using $L = 1$, as described in Section 5.1 of the paper. Note that the probabilities from the truncated predictive are likely to be fairly conservative as there are $30! = \mathcal{O}(10^{32})$ possible complete rankings (for rankers 1–6) and $30!/22! = \mathcal{O}(10^{11})$ possible top-8 rankings (for rankers 7–34), whereas the diagnostic probabilities are determined by ranking within the (possibly) most likely $\mathcal{O}(10^4)$ permutations. Overall the figure suggests that the observed rankings are consistent with the WAND model.

Rank	Ranker cluster 1			Ranker cluster 2			Overall			
	CCA	Delta	$E(\lambda' \mathcal{D})$	CCA	Delta	$E(\lambda' \mathcal{D})$	CCA	Entity	Delta	$E(\lambda' \mathcal{D})$
1	1	1	5.63	1	1	3.18	1	Heat	1	4.35
2	2	2	5.22	6	6	3.03	2	Thunder	2	2.61
3	10	10	1.48	4	4	2.23	10	Bulls	6	1.99
4	6	6	0.92	8	8	0.20	6	Lakers	4	1.52
5	9	9	0.86	10	10	0.20	9	Mavericks	10	0.81
6	3	4	0.75	26	9	0.19	3	Spurs	9	0.52
7	4	3	0.74	27	3	0.19	4	Celtics	3	0.46
8	5	5	0.53	3	18	0.19	5	Clippers	5	0.28
9	11	11	0.32	18	11	0.18	11	Knicks	11	0.26
10	12	12	0.20	9	20	0.18	12	Grizzlies	8	0.13
11	7	7	0.05	11	2	0.17	7	Pacers	18	0.12
12	13	13	0.05	2	26	0.15	13	Nuggets	12	0.12
13	14	14	0.05	20	14	0.12	14	Magic	20	0.10
14	17	17	0.05	14	27	0.10	17	TBlazers	14	0.09
15	8	8	0.04	15	15	0.09	8	76ers	26	0.08
16	15	15	0.04	23	29	0.07	15	Hawks	15	0.07
17	18	18	0.03	13	23	0.07	18	Rockets	13	0.07
18	19	19	0.02	25	13	0.07	19	Bucks	7	0.06
19	20	20	0.00	28	22	0.06	20	Suns	27	0.05
20	22	22	0.00	19	25	0.06	22	Warriors	17	0.05
21	21	21	0.00	29	21	0.05	21	Nets	23	0.04
22	23	23	0.00	22	7	0.05	23	Twolves	29	0.04
23	25	25	0.00	21	19	0.05	25	Pistons	19	0.04
24	24	24	0.00	9	30	0.05	24	Hornets	22	0.04
25	16	16	0.00	17	28	0.05	16	Jazz	25	0.04
26	26	26	0.00	12	16	0.04	26	Kings	21	0.03
27	27	27	0.00	30	24	0.04	27	Wizards	30	0.03
28	29	28	0.00	16	5	0.04	29	Cavaliers	28	0.03
29	28	29	0.00	5	17	0.04	28	Raptors	24	0.03
30	30	30	0.00	24	12	0.04	30	Bobcats	16	0.03
Pred prob	2.11×10^{-17}	2.10×10^{-17}		4.29×10^{-25}	1.10×10^{-25}		2.11×10^{-17}		6.55×10^{-24}	

Table 8: NBA analysis: within-cluster aggregate rankings (conditional on two ranker clusters) and the overall aggregate ranking along with their corresponding probabilities under the full predictive distribution. The horizontal lines indicate the MAP entity clustering. Note that for ranker groups 1 and 2 this clustering only occurs on 24 and 67 (out of 8038) iterations from this MCMC sample.

3.3 Prior sensitivity analysis

In the paper we used a relatively “informative” prior on the ranker weights w_i based on the (self declared) ranker abilities. In general this type of information is unlikely to be available and so in this section we look at the sensitivity of the posterior distribution to changes in the prior probability that a ranker is informative. Our experience in using the WAND model suggests that, when there is little information about the ranker abilities, it is best to make conservative choices of the p_i *a priori*, that is, use less extreme values nearer 0.5 than 0 or 1. We now report the main differences between the posterior given in the paper and one obtained from assuming that each ranker is equally likely to be informative/uninformative ($p_i = 0.5$).

Overall we found that the posterior distribution was fairly robust to the *a priori* choice of p_i . This is perhaps not surprising as it is consistent with the simulation studies and (other) real data analyses. As in other analyses, the aspect of the posterior distribution most sensitive to changes in the p_i was their posterior equivalents $\Pr(w_i = 1|\mathcal{D})$; see Figure 20 (right column). Unsurprisingly the largest changes were for those rankers whose p_i had increased the most (rankers 19–25 and 26–34). However both analyses give similar rankers i with $\Pr(w_i = 1|\mathcal{D}) < 0.25$; only rankers 31 and 33 are no longer under this threshold for the $p_i = 0.5$ analysis, although ranker 31 is close, with $\Pr(w_{31} = 1|\mathcal{D}) = 0.26$. The dendrograms of the ranker clustering structure are similar for both priors and clearly indicate that there are two groups of rankers. Also the allocation of rankers to clusters is similar in both cases; see Figure 20 (left column). This is supported further by the marginal posterior distribution for the number of ranker groups: both analyses show that these data have been informative and suggest two ranker groups; see Figure 21 (left). Interestingly, conditional on there being two ranker clusters, the marginal posterior for the number of entity clusters N_s^e (within each ranker cluster $s = 1, 2$) suggests there are slightly more entity clusters within ranker group 2 under the $p_i = 0.5$ analysis; see Figure 21 (right). This is perhaps due to increases in $\Pr(w_i = 1|\mathcal{D})$ for some of the rankers within this group leading to their rankings being more influential in the entity clustering. The dendrograms of the entity clustering structure in ranker group 1 are similar in both cases; see Figure 22 (left column). Although the entity dendrograms for ranker group 2 may appear to be slightly different for each analysis, on closer inspection it becomes clear that only Magic and Wizards have changed association from the right hand group to the central group for the $p_i = 0.5$ analysis.

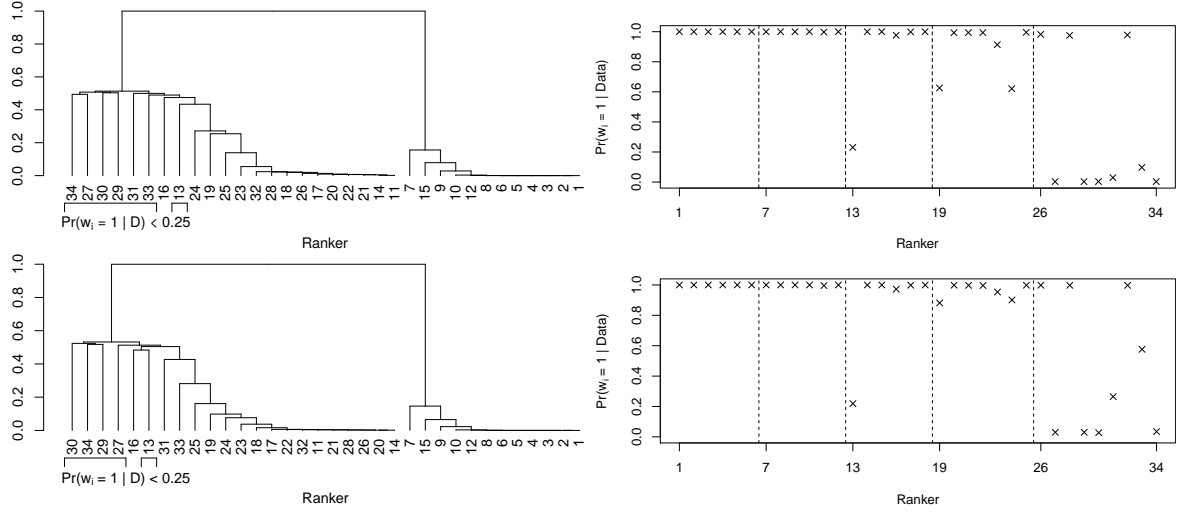


Figure 20: NBA dataset: Dendrogram (left) showing the clustering structure of rankers and highlighting those rankers with $\Pr(w_i = 1 | \mathcal{D}) < 0.25$. Plot (right) of the posterior probabilities $\Pr(w_i = 1 | \mathcal{D})$ for each ranker, with vertical lines separating the self-certified groups. The top row shows the results for the “staggered” choice of p_i (from the paper) and the bottom row shows the corresponding results when $p_i = 0.5$ for all rankers.

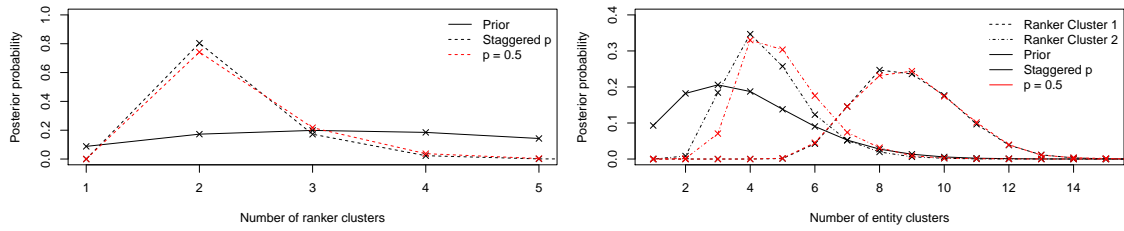


Figure 21: NBA dataset: Prior and marginal posterior densities for the number of rankers clusters (left plot) and the number of entity clusters within each ranker cluster, conditional on two ranker clusters, (right plot) for “staggered” choice of p_i (from the paper) and $p_i = 0.5$.

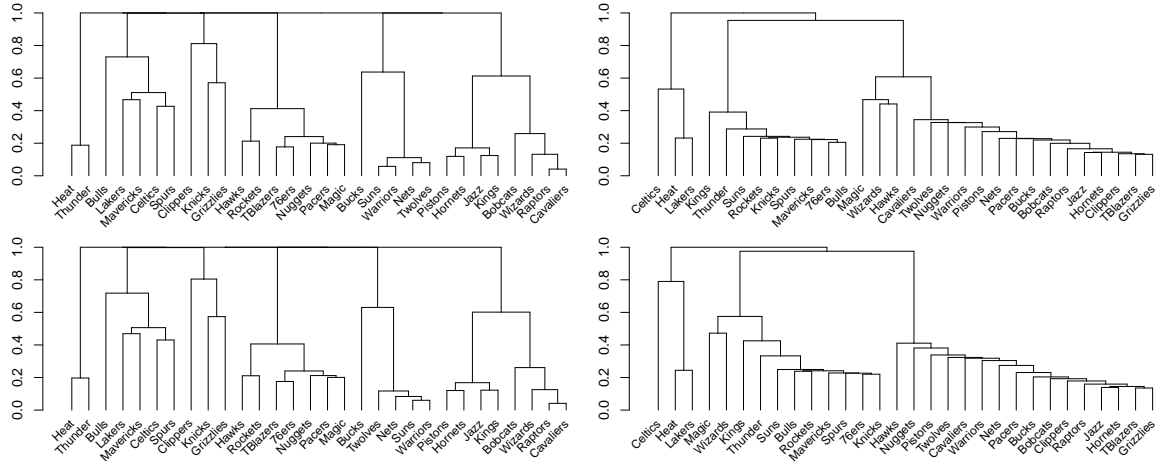


Figure 22: NBA dataset: Dendrograms showing the entity cluster structure within ranker clusters 1 and 2 (left and right respectively) conditional on two ranker clusters for “staggered” choice of p_i (from the paper) and $p_i = 0.5$, top and bottom respectively.

3.4 Partial rankings of fabric detergents

Here we analyse data that comprise *partial rankings* of fabric detergents previously studied by Best (1993) and Best and Rayner (1997). The data, originally given in Gacula (1993), consist of the partial rankings of $K = 10$ fabric detergents (labelled A to J) provided by $n = 30$ rankers. Each ranker is provided with $n_i = K_i = 4$ detergents and asked to rank them in terms of softness (detergents were allocated to rankers by using a balanced incomplete block design). The data are reproduced in Table 16.

We use the WAND model to analyse these data and take our prior specification for the base distribution and concentration parameters to be $a = 1$ and $a_\alpha = b_\alpha = 1$, $a_\gamma = b_\gamma = 3$, as in previous analyses. As we have no information about the rankers, we will assume that each ranker is equally likely to be informative or uninformative *a priori* and so take $p_i = 0.5$. We report the results from a typical run of our MCMC scheme initialised from the prior, with a burn-in of 50K iterations and then run for a further 600K iterations and thinned by 60, giving 10K (almost) un-autocorrelated realisations from the posterior distribution. Convergence was assessed by using multiple starting values, inspection of traceplots of parameters and the logarithm of the complete data likelihood, together with standard diagnostics available in the R package *coda* (Plummer et al., 2006). The MCMC scheme runs fairly quickly, with C code on a single thread of an Intel Core i7-4790S CPU (3.20GHz clock speed) taking around 40 seconds.

Table 9 shows the prior and posterior distribution for the number of ranker clusters. The data clearly have been informative and suggest it is likely that there is only a single ranker group, though there is also non-negligible posterior support for two and three ranker groups. Figure 23 displays the dendrogram of rankers along with the posterior probability that each ranker is informative. The dendrogram shows that rankers 14 and 29 are the first to be merged and are therefore the most similar (perhaps not surprising as they gave identical rankings). The height of the horizontal line connecting rankers 14 and 29 represents the dissimilarity between them, which in this case is roughly 0.14, and so the posterior probability that rankers 14 and 29 are clustered together is around $1 - 0.14 = 0.86$. The horizontal lines between groups of rankers provide useful information about the pairwise clustering probabilities. For example, the horizontal line connecting rankers 14 and 29, to rankers 19, 15, 23, and 16 indicates that no pair of rankers from the set $\{14, 15, 16, 19, 23, 29\}$ have a posterior probability of being clustered together of less than $1 - 0.16 = 0.84$. The horizontal line connecting all the rankers is at height 0.24, indicating that no pair of rankers have a posterior probability of being co-clustered of less than 0.76, so even the most dissimilar pair of rankers are clustered together with fairly

	1	2	3	4	5	6	7	≥ 8
Posterior	0.44	0.26	0.14	0.08	0.04	0.02	0.01	0.01
Prior	0.21	0.19	0.16	0.13	0.10	0.07	0.05	0.09

Table 9: Detergent dataset: Prior and posterior distribution of the number of ranker clusters (to 2 d.p.).

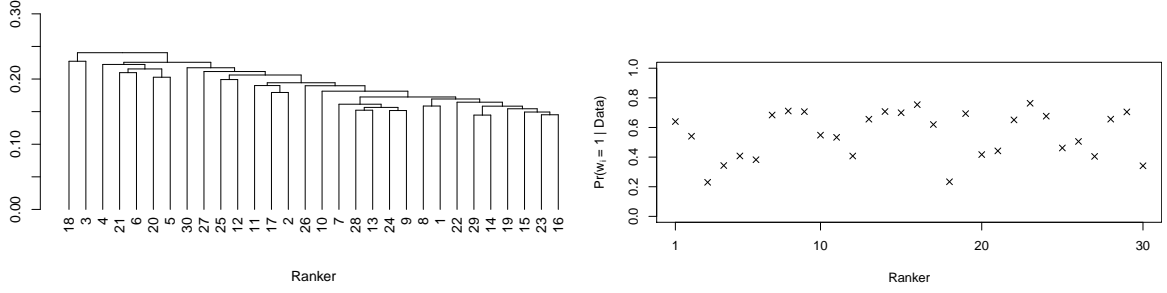


Figure 23: Detergent dataset: Dendrogram (left) showing the ranker cluster structure along with the posterior probability, $\Pr(w_i = 1 | \mathcal{D})$, for each ranker i (right).

	1	2	3	4	5	6	7	≥ 8
Posterior	0.01	0.16	0.29	0.28	0.16	0.07	0.02	0.01
Prior	0.17	0.28	0.26	0.17	0.08	0.03	0.01	0.00

Table 10: Prior and posterior distribution of the number of entity clusters conditional on a single ranker cluster (to 2 d.p.).

high posterior probability. Overall this is indicative of fairly strong homogeneity in ranker preferences and is consistent with the (marginal) posterior distribution in Table 9. It is interesting to see that this analysis reveals fairly modest probabilities for rankers being informative with $\Pr(w_i = 1 | \mathcal{D}) \in (0.2, 0.8)$ for all rankers. However, this is perhaps not too surprising given that each ranking only contains preferences for four entities: there are only $4! = 24$ possible permutations of these entities and so each ranking might not look that different from a uniformly random permutation of the entities.

One of the novel features of a WAND analysis is its ability to reveal subgroup structure of entities within ranker clusters. For illustrative purposes we now present results conditioned on a single ranker cluster; this is the posterior modal number as well as that suggested by the dendrogram in Figure 23. Table 10 gives the (marginal) posterior distribution for the number of entity clusters together with its prior distribution. Again we summarise the clustering structure of entities within this ranker cluster by examining a complete linkage dendrogram but now using dissimilarity probabilities $\Pr(d_{ij} \neq d_{ij'} | \mathcal{D})$ of entities j and j' within ranker cluster i (here with $i = 1$). The dendrogram is displayed in Figure 24. We see that no pair of fabric detergents from $\{B, C, A, G, E, I, J\}$ have a posterior probability of being co-clustered of less than about 0.4. Similarly, no pair of fabric detergents from $\{F, H, D\}$ have a posterior probability of being co-clustered of less than about 0.4. However, pairs of detergents, one from each of these two sets, are very rarely co-clustered; this is indicated by the horizontal line with a height of just less than one which represents a pairwise posterior clustering probability of roughly zero. The impression given by the complete linkage dendrogram is strongly suggestive of heterogeneous structure within the detergents, and with two clusters, $\{B, C, A, G, E, I, J\}$ and $\{F, H, D\}$, perhaps most prominently suggested. Alternatively, if a single summary clustering is sought assuming a particular number of clusters (such as the roughly joint

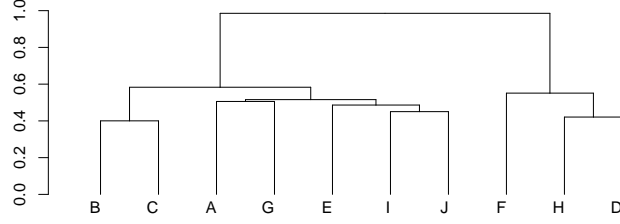


Figure 24: Detergent dataset: Dendrogram showing the entity (fabric detergent) clustering structure within ranker cluster 1 (conditional on one ranker cluster).

posterior modes of three and four clusters), then this could be done in a variety of ways. For example, we could use the MAP allocation determined from the posterior output or, much more simply, determine the clusters by “cutting” the dendrogram at a particular dissimilarity value which gives the required number of clusters — this is easily done by drawing a vertical line across the dendrogram at some height Δ , starting at the top, and moving this line downwards until the number of entity groups that remain connected beneath the line is the required number. For example, we can obtain three entity clusters by taking $\Delta \in (0.55, 0.58)$ and this gives clusters $\{B, C\}$, $\{A, G, E, I, J\}$ and $\{F, H, D\}$. If we then look at the marginal posteriors for the skill parameters $\lambda_{1,d_{1j}}$ for entities in ranker group 1 and order by posterior mean, we obtain the preference ordering $\{B, C\} \succ \{A, G, E, I, J\} \succ \{F, H, D\}$ with entity cluster means 1.70, 0.86 and 0.15 respectively. Note that, for this data analysis, the aggregate ranking (conditional on a single ranker cluster) gives a very similar entity ordering. Here the aggregate ranking is obtained via the delta approximation, that is, by using the marginal posterior means of the skill parameters (marginalised over the distribution of entity clusters). Here the ordering of entities is $(B, C, I, J, A, G, E, F, H, D)$ with (respective) posterior mean skill parameters $(1.91, 1.80, 1.50, 1.40, 1.36, 1.18, 1.17, 0.47, 0.45, 0.16)$. This preference ordering is consistent with the analysis of these data in Best and Rayner (1997) but note that the WAND analysis provides much more information than simply a single preference ordering.

3.4.1 Model assessment via posterior predictive checks

Figure 25 shows the diagnostic probabilities for each ranking, calculated from their (exact) posterior predictive distributions, as described in Section 5.1 of the paper. The figure strongly suggests that the observed rankings are consistent with the WAND model.

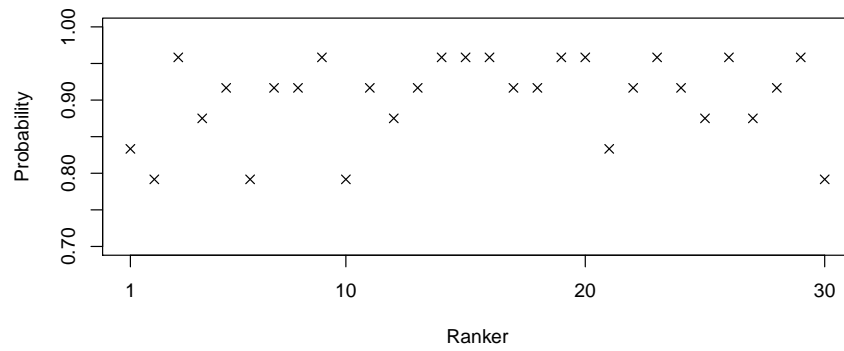
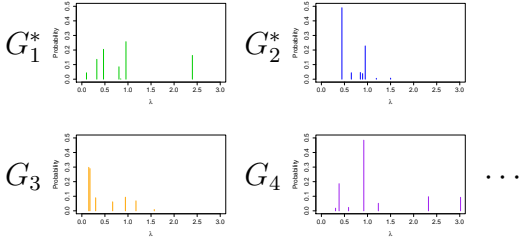
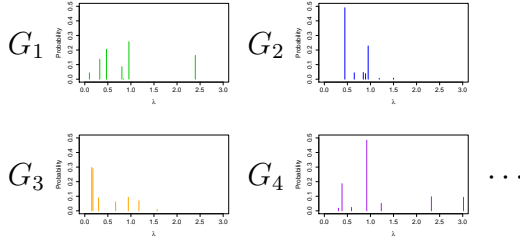
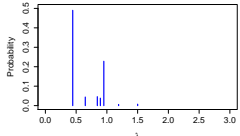
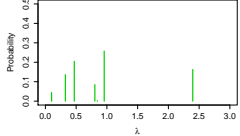
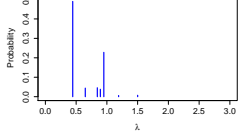
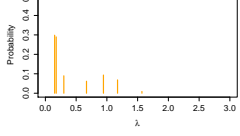


Figure 25: Detergent dataset: diagnostic probability for each ranking, calculated from the posterior predictive distribution.

4 Graphical representation of the NDP and the ANDP

The following figure gives a graphical representation of the NDP and the ANDP, based on in Figure 1 in Rodriguez et al. (2008).

NDP	ANDP
$Q \sim \text{NDP}(\alpha, \gamma, G_0)$ $G'_i \sim Q$ $\lambda'_i \sim G'_i$	$Q \alpha, \gamma, G_0 \sim \text{ANDP}(\alpha, \gamma, G_0)$ $(\lambda'_1, \dots, \lambda'_n) Q \sim Q$
	
	$\lambda_1^* = (\lambda_{11}^*, \dots, \lambda_{1K}^*) \sim G_1 \quad \lambda_2^* \sim G_2$ $\lambda_3^* \sim G_3 \quad \lambda_4^* \sim G_4 \quad \dots$
$\lambda'_1 \sim G'_1 \stackrel{d}{=} G_2$  $\lambda'_2 \sim G'_2 \stackrel{d}{=} G_1$  $\lambda'_3 \sim G'_3 \stackrel{d}{=} G_2$  $\lambda'_4 \sim G'_4 \stackrel{d}{=} G_3$  \vdots	$\lambda'_1 = \lambda_2^*$ $\lambda'_2 = \lambda_1^*$ $\lambda'_3 = \lambda_2^*$ $\lambda'_4 = \lambda_3^*$ \vdots

5 Datasets

		Ranker																																															
		1										10										20										30										40							
38	3	4	12	16	3	4	2	15	1	1	4	7	1	7	3	12	2	1	7	16	13	2	19	10	1	13	13	1	11	7	6	4	16	10	3	4	7	10	3	3									
	5	14	2	12	2	3	1	2	3	4	1	9	4	1	1	3	16	11	4	11	2	5	3	3	8	14	6	4	9	2	3	6	1	1	8	13	16	3	10	4									
	2	13	3	1	12	2	5	1	5	3	2	2	3	4	9	2	3	2	2	4	10	13	8	9	5	15	4	12	13	1	1	1	6	11	4	2	13	14	16	2									
	8	5	15	9	4	1	15	6	4	6	8	5	16	6	10	1	5	3	1	13	3	7	4	14	3	6	1	7	6	8	15	5	9	4	15	1	11	11	1	14									
	9	7	1	2	1	8	12	3	14	8	10	1	6	12	2	8	13	6	3	3	14	1	12	1	9	8	2	10	4	3	8	15	3	2	6	5	1	12	4	7									
	6	2	7	7	5	10	4	4	7	10	6	11	9	10	4	7	1	10	15	9	1	15	6	12	2	4	3	3	7	6	10	3	10	3	1	8	10	1	11	1									
	1	6	6	4	7	13	3	7	2	2	14	10	2	8	11	4	14	9	8	10	11	4	1	6	12	1	7	5	2	4	2	2	2	2	8	16	3	4	16	2	10								
	7	1	14	14	15	9	9	9	6	5	12	3	12	3	6	9	15	16	12	1	4	11	5	11	6	5	11	14	1	5	5	7	4	6	2	16	15	8	8	13									
	16	3	9	3	6	15	13	19	9	12	3	6	7	2	19	10	4	15	5	2	5	10	2	13	4	3	5	13	5	14	4	8	5	12	5	9	3	2	9	16									
	10	8	16	5	11	11	7	13	10	9	13	8	13	11	13	6	8	7	10	12	6	6	13	4	11	10	12	8	12	9	12	14	8	13	11	10	14	5	13	8									
	11	12	5	6	16	5	16	16	13	7	5	4	15	5	8	16	10	14	6	5	8	3	10	2	13	11	14	2	14	12	9	16	14	5	7	7	5	4	6	6									
	12	11	4	11	14	12	6	11	12	13	9	15	11	9	12	5	6	4	14	6	9	12	14	5	7	2	20	6	8	10	14	11	12	16	12	14	8	7	15	11									
	4	15	8	8	13	16	11	14	11	14	7	13	18	13	14	11	7	13	11	14	16	14	20	16	10	16	16	15	3	11	13	9	11	9	10	6	2	13	19	12									
	15	19	13	10	18	14	14	5	8	15	11	16	5	16	5	18	11	8	17	7	7	9	16	7	17	9	10	11	10	15	7	10	7	15	9	15	6	17	14	5									
	14	10	10	13	8	17	8	12	16	11	17	14	8	15	16	14	20	12	16	8	12	16	15	15	18	12	8	20	16	13	11	13	13	14	13	12	9	6	5	15									
	13	9	11	15	9	6	10	10	15	20	15	12	14	14	20	15	9	5	9	15	15	8	7	19	15	7	15	9	15	16	16	12	15	7	17	11	20	9	12	20									
	18	16	20	19	10	7	18	8	19	17	16	18	10	18	18	13	12	20	19	19	17	20	11	8	16	20	9	16	17	20	17	20	19	18	14	18	12	20	7	17									
	20	17	18	18	17	19	20	20	18	16	19	17	17	20	7	17	18	17	13	18	20	17	9	17	14	19	19	17	20	19	19	19	18	19	19	19	17	19	18	9									
	17	20	19	20	20	20	17	17	17	18	20	20	20	19	15	19	17	18	20	17	19	19	18	20	19	17	17	19	19	18	18	17	20	20	20	20	18	15	17	19									
	19	18	17	17	19	18	19	18	20	19	18	19	19	17	17	20	19	19	18	20	18	18	17	18	20	18	18	18	18	17	20	18	17	17	18	17	19	18	20	18									

Table 11: The 40 informative rankings within the dataset used in Simulation Study 1. The additional 10 uninformative rankings (numbered 41–50 within the paper) are given within Table 12.

Ranker									
1									10
11	7	20	13	17	8	15	19	18	18
1	11	15	12	13	7	17	10	20	6
18	20	16	2	10	1	5	4	1	19
15	19	17	17	6	11	7	3	11	1
12	15	1	19	18	19	6	12	7	16
13	14	14	20	9	6	18	9	2	10
20	12	7	6	15	9	14	7	10	14
10	2	9	10	1	16	9	15	6	20
6	18	13	4	14	14	1	18	15	8
7	3	10	16	20	17	4	8	5	13
19	9	5	18	16	5	13	6	3	2
14	8	19	3	5	13	19	17	4	3
4	5	6	9	19	12	2	13	13	15
16	4	3	8	2	4	12	14	9	11
8	6	2	7	3	20	16	20	19	17
2	10	18	5	8	18	3	5	12	5
17	1	4	14	11	10	10	11	16	7
3	16	12	11	12	2	8	2	14	9
9	17	8	15	4	3	11	16	17	12
5	13	11	1	7	15	20	1	8	4

Table 12: Additional 10 uninformative rankings used within Simulation Study 1

		Ranker																																																			
1		24																				25																				36				37				40			
40	1	2	1	13	13	18	1	14	7	9	10	19	1	1	1	13	1	10	7	5	9	1	20	18	1	2	8	9	12	3	16	14	12	5	16	8	7	10	10	7													
	20	1	6	1	6	13	13	4	2	4	1	14	10	11	20	6	10	1	14	15	11	12	14	7	3	8	1	16	8	11	18	12	3	9	13	7	10	19	9	17													
	7	15	10	10	1	1	15	10	1	7	15	3	18	13	3	17	20	18	12	1	1	5	10	13	5	1	5	4	2	8	2	13	8	14	2	4	3	18	6	16													
	2	3	12	15	15	3	5	1	5	14	12	1	7	3	2	20	13	3	11	10	5	13	1	5	11	7	3	7	16	4	5	16	13	4	5	2	6	2	2	18													
	17	18	14	7	16	17	2	15	18	18	14	17	15	7	7	5	9	13	6	12	10	18	13	4	2	4	17	17	20	2	8	5	17	2	20	3	12	7	3	6													
	10	19	13	4	9	20	7	6	6	13	6	10	9	18	17	14	2	20	13	9	19	4	18	12	12	9	4	11	3	1	1	2	2	7	1	9	1	17	5	8													
	4	20	4	9	7	4	10	7	12	1	17	4	14	14	13	12	6	12	2	7	14	7	3	10	10	5	2	5	5	16	3	7	9	13	4	11	17	9	19	2													
	19	7	18	11	11	15	6	13	3	15	13	20	20	6	14	1	15	5	1	14	6	9	15	6	6	3	16	2	1	5	7	11	4	3	18	10	15	4	20	12													
	12	14	3	17	4	6	9	17	14	16	9	9	3	9	4	11	7	9	17	2	4	17	6	14	7	14	11	1	6	9	20	4	10	16	11	5	13	3	16	3													
	14	10	19	20	17	19	4	20	10	6	7	13	13	15	6	19	12	19	4	3	7	15	4	9	9	12	7	12	10	14	10	8	1	12	3	16	2	13	1	19													
	9	12	15	12	20	16	14	9	13	12	4	7	5	12	19	10	18	7	19	6	13	6	7	11	15	16	20	8	11	12	11	9	16	8	10	1	18	8	15	13													
	15	4	5	6	3	7	8	12	17	5	19	12	17	4	12	4	19	14	15	20	12	14	17	1	8	11	12	3	13	7	4	15	5	15	19	14	8	6	12	10													
	6	13	2	2	10	10	20	19	16	20	3	5	2	2	9	18	11	2	3	13	15	3	12	2	4	17	15	19	4	6	14	3	18	18	14	19	11	5	7	14													
	5	6	7	3	12	12	11	3	20	2	18	2	6	5	18	9	4	6	9	11	2	10	2	20	14	15	9	14	14	10	13	6	14	6	8	20	19	14	18	5													
	18	11	20	14	14	11	12	2	4	3	2	6	12	10	11	16	5	11	10	4	3	11	9	15	17	13	14	20	9	19	6	1	7	1	9	12	9	15	11	1													
	13	5	17	18	2	14	18	5	9	10	20	15	4	19	10	15	14	15	18	17	18	2	5	3	20	19	10	18	7	15	15	18	11	19	12	13	5	20	17	20													
	3	9	9	5	18	5	17	11	11	19	11	11	19	20	15	2	17	17	5	18	17	19	11	17	18	18	6	10	17	20	9	10	15	11	7	17	20	16	14	9													
	11	17	11	19	19	2	19	16	15	17	5	18	11	17	5	7	3	4	20	19	16	16	19	19	16	10	18	6	19	17	12	19	20	10	17	15	16	12	4	11													
	16	16	16	16	5	9	3	18	19	11	8	16	16	8	16	3	16	8	16	8	8	8	16	16	19	20	13	15	15	13	17	20	19	17	15	6	14	11	13	4													
	8	8	8	8	8	8	16	8	8	8	16	8	8	16	8	8	8	16	8	16	20	20	8	8	13	6	19	13	18	18	19	17	6	20	6	18	4	1	8	15													

Table 13: The 40 rankings which comprise the dataset for Simulation Study 2, vertical lines separate the rankings within each different ranker group.

Ranker	Rank								
	1								9
1	1	5	4	6	2	7	3	9	8
2	1	3	2	6	7	5	4	8	9
3	1	6	4	7	3	2	8	5	9
4	1	6	7	3	2	5	4	8	9
5	2	7	4	3	9	5	1	6	8
6	2	3	5	8	4	1	6	7	9
7	2	1	5	3	4	7	8	6	9
8	2	3	5	1	6	8	9	4	7
9	2	8	3	1	4	6	5	7	9
10	2	3	1	4	9	5	8	6	7
11	2	7	4	1	8	6	5	3	9
12	3	2	1	8	4	5	7	6	9
13	3	1	6	7	8	4	9	5	2
14	3	1	5	4	7	8	2	9	6
15	3	2	4	8	5	9	1	6	7
16	4	8	5	7	1	9	2	3	6
17	4	5	8	9	1	7	6	3	2
18	4	8	5	9	3	2	6	7	1
19	5	4	7	8	9	2	3	6	1
20	5	4	2	7	8	9	3	1	6
21	5	2	9	8	4	7	1	3	6
22	5	7	4	9	8	3	2	1	6
23	5	4	9	7	3	2	8	1	6
24	5	4	9	8	3	7	2	1	6
25	5	4	2	8	9	3	1	7	6
26	5	7	8	1	4	9	2	6	3
27	5	4	9	7	1	2	8	3	6
28	1	4	5	8	9	7	6	2	3
29	6	1	8	4	2	3	7	5	9
30	6	1	4	3	2	5	8	7	9
31	6	3	2	5	1	7	8	9	4
32	7	4	8	1	2	3	6	5	9
33	7	5	4	8	1	9	2	6	3
34	7	4	2	8	6	1	3	5	9
35	7	5	8	4	2	9	3	1	6
36	7	1	5	4	3	2	6	8	9
37	7	4	5	8	1	9	6	2	3
38	9	5	3	4	7	8	2	1	6
39	9	4	8	5	1	2	3	7	6

Entity No	1	2	3	4	5	6	7	8	9
Area	SOC	EDU	CLI	MAT	EXP	CUL	IND	TST	PHY

Table 14: Roskam's psychology data

		Professionals						Avid Fans						Fans				Infrequent Watchers						Not-interested Individuals											
Entity No	Team	1	...				6	7	...				12	13	...				18	19	...				25	26	...				34				
1	Heat	1	9	1	1	1	1	1	1	4	6	1	6	6	1	4	1	6	1	6	1	6	1	9	6	1	6	22	6	26	28	1	1	1	10
2	Thunder	9	1	2	9	2	10	2	2	1	2	4	2	4	6	1	2	11	6	18	6	4	2	4	1	2	1	5	4	11	2	25	4	6	1
3	Spurs	2	2	9	2	10	2	6	11	2	1	6	1	10	4	10	4	1	4	1	19	9	4	3	18	6	10	9	10	10	18	8	8	21	16
4	Celtics	10	10	10	10	9	9	4	10	11	4	10	10	26	8	2	3	4	2	10	4	1	6	1	11	4	4	14	1	23	8	9	6	5	23
5	Clippers	4	5	13	3	5	4	10	3	3	3	11	4	18	26	3	15	21	9	3	14	26	26	6	8	14	18	6	18	15	10	7	10	9	21
6	Lakers	6	11	5	6	6	5	3	5	6	10	22	3	14	3	18	9	9	8	22	18	20	3	30	3	8	29	7	8	8	19	14	14	27	11
7	Pacers	3	6	6	12	11	3	15	6	9	9	27	5	24	18	6	6	2	29	20	20	23	23	13	27	11	15	8	27	21	23	20	11	11	18
8	76ers	5	12	12	11	3	6	7	4	10	5	13	9	16	20	9	28	8	10	26	9	3	20	20	25	15	11	28	25	12	11	6	7	13	15
9	Mavericks	11	13	11	4	4	13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
10	Bulls	12	3	4	5	13	12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
11	Knicks	14	4	3	14	12	14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
12	Grizzlies	15	14	15	17	15	7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
13	Nuggets	17	7	8	13	7	11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
14	Magic	7	17	7	7	14	17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
15	Hawks	8	18	17	8	8	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
16	Jazz	19	8	18	18	17	19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
17	TrailBlazers	21	19	14	19	18	18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
18	Rockets	16	15	24	15	24	15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
19	Bucks	13	21	20	22	20	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
20	Suns	20	23	19	21	19	22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
21	Nets	18	22	26	20	26	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
22	Warriors	22	20	23	23	22	25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
23	Timberwolves	23	16	22	24	23	21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
24	Hornets	28	26	21	25	21	23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
25	Pistons	25	25	25	27	25	24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
26	Kings	29	28	16	26	29	26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
27	Wizards	24	27	29	16	27	27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
28	Raptors	27	24	27	28	16	28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
29	Cavaliers	26	29	28	29	30	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
30	Bobcats	30	30	30	30	28	30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Table 15: NBA data

Ranker	Rank				\mathcal{U}_i
	1	2	3	4	
1	1	2	3	4	{5,6,7,8,9,10}
2	1	5	2	6	{3,4,7,8,9,10}
3	8	7	1	3	{2,4,5,6,9,10}
4	9	10	4	1	{2,3,5,6,7,8}
5	7	1	5	9	{2,3,4,6,8,10}
6	8	10	1	6	{2,3,4,5,7,9}
7	3	2	9	6	{1,4,5,7,8,10}
8	10	2	7	4	{1,3,5,6,8,9}
9	2	10	5	8	{1,3,4,6,7,9}
10	9	7	2	8	{1,3,4,5,6,10}
11	3	10	9	5	{1,2,4,6,7,8}
12	3	6	10	7	{1,2,4,5,8,9}
13	3	5	4	8	{1,2,6,7,9,10}
14	7	5	6	4	{1,2,3,8,9,10}
15	9	6	8	4	{1,2,3,5,7,10}
16	2	3	1	4	{5,6,7,8,9,10}
17	1	2	5	6	{3,4,7,8,9,10}
18	8	1	7	3	{2,4,5,6,9,10}
19	9	10	1	4	{2,3,5,6,7,8}
20	1	7	5	9	{2,3,4,6,8,10}
21	10	8	1	6	{2,3,4,5,7,9}
22	2	9	3	6	{1,4,5,7,8,10}
23	2	10	7	4	{1,3,5,6,8,9}
24	2	5	10	8	{1,3,4,6,7,9}
25	2	9	8	7	{1,3,4,5,6,10}
26	3	9	5	10	{1,2,4,6,7,8}
27	3	6	10	7	{1,2,4,5,8,9}
28	3	5	4	8	{1,2,6,7,9,10}
29	7	5	6	4	{1,2,3,8,9,10}
30	6	8	9	4	{1,2,3,5,7,10}

Entity No	1	2	3	4	5	6	7	8	9	10
Detergent	A	B	C	D	E	F	G	H	I	J

Table 16: Detergent data

References

- Best, D. (1993). Extended analysis for ranked data. *Australian Journal of Statistics* 35(3), 257–262.
- Best, D. and J. Rayner (1997). Product maps for ranked preference data. *Journal of the Royal Statistical Society: Series D (The Statistician)* 46(3), 347–354.
- Choulakian, V. (2016). Globally homogenous mixture components and local heterogeneity of rank data. *arXiv:1608.05058*.
- Dahl, D. B. (2006). Model-based clustering for expression data via a Dirichlet process mixture model. In K.-A. Do, P. Müller, and M. Vannucci (Eds.), *Bayesian Inference for Gene Expression and Proteomics*, pp. 201–218. Cambridge University Press.
- de Leeuw, J. (2006). Nonlinear principal component analysis. In M. Greenacre and J. Blasius (Eds.), *Multiple correspondence analysis and related methods*, Chapter 4, pp. 107–134. CRC Press.
- de Leeuw, J. and P. Mair (2009). Gifi methods for optimal scaling in R: The package homals. *Journal of Statistical Software* 31(4), 1–20.
- Deng, K., S. Han, K. J. Li, and J. S. Liu (2014). Bayesian aggregation of order-based rank data. *Journal of the American Statistical Association* 109(507), 1023–1039.
- Everitt, B. S., S. Landau, M. Leese, and D. Stahl (2011). *Cluster analysis*. Wiley, Chichester, UK.
- Gacula, M. (1993). *Design and analysis of sensory optimization*. Food and Nutrition Press.
- Lange, K. (2013). *Optimization*. Springer Texts in Statistics. Springer New York.
- Lau, J. W. and P. J. Green (2007). Bayesian model-based clustering procedures. *Journal of Computational and Graphical Statistics* 16(3), 526–558.
- Marden, J. I. (1995). *Analysing and Modeling Rank Data*. London: Chapman and Hall.
- Medvedovic, M. and S. Sivaganesan (2002). Bayesian infinite mixture model based clustering of gene expression profiles. *Bioinformatics* 18(9), 1194–1206.
- Plummer, M., N. Best, K. Cowles, and K. Vines (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* 6(1), 7–11.
- Rodriguez, A., D. B. Dunson, and A. E. Gelfand (2008). The Nested Dirichlet process. *Journal of the American Statistical Association* 103(483), 1131–1154.