


A new Maxwell paired comparison model: application to a study of the effect of nicotine levels on cigarette brand choices

S. A. Cheema ^a, I. L. Hudson ^b , T. Kifayat ^c, Kalim-u-llah ^c, M. Shafqat ^c and A. Hussain ^c

^a School of Mathematical and Physical Sciences, University of Newcastle, Australia, ^b Department of Mathematical Sciences, College of Science, Engineering and Health, Royal Melbourne Institute of Technology (RMIT), Melbourne, Australia, ^c Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.

Email: irene.hudson@rmit.edu.au

Abstract: In this article, we propose a new choice model to assess the influence of nicotine level over the choice behaviors of smokers in selecting different brands of cigarette. The objectives of the study are met by considering three most commonly used cigarette brands in Pakistan, which are Benson & Hedges (BH), Marlboro (Ma) and Gold Leaf (GL). The data are gathered through a balance paired comparison (PC) experiment where 150 smokers, aged 25- 35 years, are asked about their preferences to the aforementioned brands.

An initial exploratory analysis of the data reveals that, in general, smokers prefer cigarette brands with a higher level of nicotine. We observed that almost 65% of the participants reported they prefer Gold Leaf brand (nicotine level 16.92 ± 0.82 mg/stick) over Marlboro (nicotine level 12.95 ± 0.82 mg/stick). Moreover, when choosing between Gold Leaf and Benson & Hedges (nicotine level 15.93 ± 0.69 mg/stick), around 59% recorded their preference for Gold Leaf over Benson & Hedges. Lastly, the comparative choices of the participants of our study favor Benson & Hedges over Marlboro, where almost 55% of the respondents reported that they prefer Benson & Hedges. Thus the overall preference ordering of choice behaviors based on this study can be written as $GL \rightarrow BH \rightarrow Ma$. The comparative information is then modeled by introducing a new choice model based on the Maxwell distribution. Estimation of the worth parameters and associated preference probabilities is performed by means of *Gibbs Sampling*, where two non-informative priors, the Uniform and Jefferys priors are used.

We observed that both priors are capable of retaining true preference order, however the Jefferys prior performs better as compared to the Uniform prior by closely estimating the observed extent of preferences. The estimated values of the worth parameters reveals that cigarette brand preferences are associated with higher levels of nicotine and thus we verify the findings of our initial analysis via the Maxwell distribution based choice model. For example, regardless of the prior distribution, highest utility is attached with the Gold Leaf brand (highest nicotine level), followed by Benson & Hedges and then Marlboro. Using the Jefferys prior, we observed an estimated value of the worth parameter of 0.3765 highlighting preference of Gold Leaf brand, associated with a 0.6492 value for the preference probability. Further, in the case of the second most preferred brand, namely. Benson & Hedges, the estimated worth parameter has a value of 0.3247 which is associated with a preference probability equal to 0.5927. Similarly, the estimated worth parameter underlying the choice of Marlboro brand is 0.2981 which approximates with a preference probability of 0.5599 - therefore Marlboro remained the least preferred brand. We observed similar trends in the choice behaviors when the Uniform distribution was used as a prior.

In conclusion regardless of the prior distributions, we observed common trends in the choice behaviors. For example, with respect to the both priors, the highest value of the worth parameter is associated with the GL brand, which also has the highest reported nicotine level. The GL brand is then followed by BH brand, with which the second highest value of worth parameter and nicotine level is attributed. The minimal value of the worth parameter is associated with the Ma brand which has the lowest level of nicotine among the three studied brands. The posterior standard deviations revealed that the Jefferys prior provides more stable estimates of the worth parameters as compared to Uniform prior

This paper contributes a new PC model, named the Maxwell paired comparison model to the literature. The applicability of the newly proposed comparative model is demonstrated. We have also established that nicotine level plays an instrumental role in driving the preference ordering of cigarette brands. We conclude that smokers in general tend to choose a cigarette brand with higher levels of nicotine. Public health interventions could potentially address this via public health warnings and intervention/education programs.

Keywords: Choice behaviors, comparative models, Maxwell distribution, worth parameters, priors

1. INTRODUCTION

Cigarettes are a thin paper wrapped, cylindrical shaped, and tobacco based product containing nicotine, carbon monoxide and tar. Nicotine is a vital element of immune-suppressance and is the paramount root for addiction, as observed by Bergen and Caporaso (1999). In this investigation, we studied smokers' choice behaviors of cigarette brands in reference to their nicotine level, via an application of the Maxwell distribution. The data were collected through balance paired comparison (PC) design from 150 adults. The utility of PC methods in analyzing choice data is well documented in various fields of research, such as medicine, sensory analysis, sports ranking and voting behaviors. For example, Amlani and Schafer (2009) have applied PC models in the ranking of hearing aids and electroacoustic devices, whereas, Sung and Wu (2018) demonstrated the PC approach as an alternative to the Likert scale for psychological measurements and assessment. Further, Dittrich et al. (2002) examined shifts in post-materialistic values through the PC approach, and Mazzuchi, et al. (2008) used PC models for an application of choice models to resolve a reliability problem. Green-Armytage (2004) studied voters' preference behaviors through cardinal weighted PC studies and Schauburger and Tutz (2015) applied PC models to predict the outcomes of sports competitions.

In the simplest form, PC methods typically involve circumstances where choices are required to be documented while conducting a pairwise comparison between two objects or treatments. For demonstration purposes, let us say, k items or objects are to be compared by n judges through a simple question "Do you prefer item i over item j ?" It is then trivial to calculate that, for k items, there are actually $\binom{k}{2}$ paired comparisons for every judge and that every response is indeed a Bernoulli outcome, see also Cattelan et al. (2013) and Veghes (2014).

Interest then lies in the estimation of θ_i and θ_j , known as the worth parameters, denoting the underlying utility of item i and item j which influences the preference ordering of judges, respectively. Thus in the case of k items, θ^j will be a $1 \times k$ dimensional vector consisting of worth parameters associated with a single judge. Furthermore, specification of pairwise comparison implies the constraint $\sum_{i=1}^k \theta_i = 1$, to resolve the problem of parameter estimability (various other constraints are also possible, see Rayner and Best (2001)).

In their pioneering and seminal work Bradley and Terry (1952) proposed a logit function for the worth parameters to quantify the probability of preferring item i over item j , as follows

$$p_{ij} = P_r(i > j) = \frac{\theta_i}{\theta_i + \theta_j}, \quad (1)$$

where the preference probability of item j over item i can be estimated trivially by $p_{ji} = P_r(j > i) = 1 - p_{ij}$.

Since then, many researchers have proposed various strategies to model comparative information. For example Latta (1979) used the Cauchy distribution as the distribution and proposed the Cauchy paired comparison method, while Stern (1990) employed a Poisson process to the Gamma paired comparison model for players' scoring. More recently Cattelan et al. (2013) introduced dynamic paired comparison models which include a time varying parameter in contrast to the static PC schemes mentioned above, and of late, Schauburger and Tutz (2017) generalized PC models to structural equation modelling.

In this paper, we present an application of the Maxwell distribution, a commonly used distribution in speed theory; see also Brush (2003) and Uffink (2007). The newly proposed model is performed using a Bayesian framework, under two non-informative priors, the Jeffrey prior and the Uniform prior. In section 2, we present our model along with the derivation of the marginal posterior distributions of the worth parameters. Section 3 demonstrates the applicability and comparative performance of the proposed model under the aforementioned priors using a data set which comprises 150 smokers, aged 25-35 years, who were asked about their preferences to the three most popular cigarette brands in Pakistan. Finally, section 4 summarizes with concluding remarks and future research ideas.

2. PROPOSED MODEL

Let X be a Maxwell random variable, then the probability distribution function (pdf) of the Maxwell distribution is given as,

$$f(x_i) = \sqrt{\frac{2}{\pi}} x_i^2 \frac{e^{-\frac{x_i^2}{2\theta_i^2}}}{\theta_i^3}, \quad x_i > 0, \theta_i > 0 \quad (2)$$

where, θ_i is a scale parameter measuring the speed in a Euclidian space. In our proposed strategy, we exploit the scale parameter of the Maxwell distribution to quantify the worth of a given cigarette brand in reference to varying nicotine levels across brands.

2.1. The Model

In the case of binary PC models, that is where permissible choice categories remain “yes” or “no”, the outcome of every pairwise comparison is considered as an independent Bernoulli trial. Therefore, using the pdf given in equation (2), the probability that the h' th smoker will prefer brand i , b_i , over brand j , b_j , is quantifiable as,

$$p_{ij} = \frac{2}{\pi} \int_0^\infty \int_{b_j}^\infty \frac{b^2_i b^2_j}{\theta^3_i \theta^3_j} e^{-\frac{1}{2} \left(\frac{b^2_i}{2\theta^2_i} - \frac{b^2_j}{2\theta^2_j} \right)} db_i db_j,$$

which upon simplification reduces to the proposed Maxwell model to quantify the preference probabilities based on choice data, as follows,

$$p_{ij} = 1 + \frac{2}{\pi} \left[\frac{\theta_i \theta_j (\theta^2_i - \theta^2_j)}{(\theta^2_i + \theta^2_j)^2} - \arctan \left(\frac{\theta_i}{\theta_j} \right) \right]. \tag{3}$$

The quantification of the preference probability of item j over item i is able to be calculated from equation (3) because $p_{ji} = 1 - p_{ij}$, and is written as,

$$p_{ji} = 1 + \frac{2}{\pi} \left[\frac{\theta_i \theta_j (\theta^2_j - \theta^2_i)}{(\theta^2_i + \theta^2_j)^2} + \arctan \left(\frac{\theta_j}{\theta_i} \right) \right].$$

The likelihood function where k items are compared by n judges, is then given as,

$$L(\underline{\theta}; \underline{x}) = \prod_{i < j}^k \frac{n!}{r_{ij}!(n-r_{ij})!} p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}}, \tag{4}$$

where \underline{X} is the data matrix and r_{ij} is the number of preferences of item i over item j .

2.2. Priors

For demonstration purposes, we consider two non-informative priors for the worth parameters.

Uniform Prior:

Under the Uniform prior, when k items are compared by n judges, $P(\theta_1, \theta_2, \dots, \theta_k) \propto 1$, and the kernel density of the joint posterior distribution is given as,

$$P(\theta_1, \theta_2, \dots, \theta_k | \underline{x}) \propto \prod_{i < j}^k p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}}, \tag{5}$$

where the normalizing constant, say C_u , is calculated as,

$$C_u = \int_{\theta_1=0}^1 \int_{\theta_2=0}^{1-\theta_1} \dots \int_{\theta_k=0}^{1-\sum_{i=1}^{k-1} \theta_i} \prod_{i < j}^k p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}} d\theta_1 d\theta_2 \dots d\theta_{k-1}. \tag{6}$$

Jefferys Prior:

Under the Jefferys prior, $P(\theta_1, \theta_2, \dots, \theta_k) \propto \det[I(\underline{\theta})]^{1/2}$ where, $I(\underline{\theta})$ in the case of a k items PC comparison is a $k \times k$ Fisher information matrix and is calculable as follows,

$$I(\underline{\theta}) = -E \left[\frac{\partial^2 \ln L(\underline{\theta})}{\partial \theta_i \partial \theta_j} \right].$$

The kernel density of the joint posterior distribution under the Jefferys prior is given as,

$$P(\theta_1, \theta_2, \dots, \theta_k | \underline{x}) \propto P(\theta_1, \theta_2, \dots, \theta_k) \prod_{i < j}^k p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}}, \tag{7}$$

where the normalizing constant, say C_j , is given as,

$$C_j = \int_{\theta_1=0}^1 \int_{\theta_2=0}^{1-\theta_1} \dots \int_{\theta_k=0}^{1-\sum_{i=1}^{k-1} \theta_i} P(\theta_1, \theta_2, \dots, \theta_k) \prod_{i < j}^k p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}} d\theta_1 d\theta_2 \dots d\theta_{k-1}. \tag{8}$$

2.3. Quantification of the marginal posterior distribution – Gibbs sampling

Gibbs sampling is one of the Markov chain Monte Carlo methods used to handle complex multiple integrals such as that given in equation (5) and equation (7). We employed the Gibbs sampler to attain the marginal posterior distribution of the worth parameters by iteratively conditioning on interim values in a continuing cycle. The prominence of Gibbs sampling in Bayesian working paradigm has been observed by many researchers. For example, Gelfand et al. (1990) used the Gibbs sample for Bayesian analysis of normal data models. Carter and Kohn (1994) performed a linear state space model by Gibbs sample. For more details, one may also consult to Dellaportas and Smith (1993), Diaconis et al. (2008) and Jone and Johnson (2008).

Let $P(\underline{\theta}; \underline{x})$ be the joint posterior density, where $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$, then the conditional densities are given by, $P(\theta_1 | \theta_2, \theta_3, \dots, \theta_k), P(\theta_2 | \theta_1, \theta_3, \dots, \theta_k) \dots P(\theta_k | \theta_1, \theta_2, \dots, \theta_{k-1})$.

Then according to the Gibbs sampler, we assume initial values such as $(\theta_2^{(0)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$ and pursue the conditional distribution of θ_1 such that $P(\theta_1^{(1)} | \theta_2^{(0)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$. The iterative procedure will continue until it converges.

For demonstration purposes, we provide the expression for the marginal posterior distribution of k' th worth parameter θ_k , under the Uniform and Jefferys Prior,

$$P(\theta_k | \underline{X}) = \int_{\theta_1=0}^{1-\theta_k} \int_{\theta_2=0}^{1-\theta_1-\theta_k} \dots \int_{\theta_{k-1}=0}^{1-\sum_{i=1}^{k-2} \theta_i - \theta_k} \frac{1}{c_u} \prod_{i < j} p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}} d\theta_1 d\theta_2 \dots d\theta_{k-1}, \tag{9}$$

$$P(\theta_k | \underline{X}) = \int_{\theta_1=0}^{1-\theta_k} \int_{\theta_2=0}^{1-\theta_1-\theta_k} \dots \int_{\theta_{k-1}=0}^{1-\sum_{i=1}^{k-2} \theta_i - \theta_k} \frac{1}{c_j} P(\theta_1, \theta_2, \dots, \theta_k) \prod_{i < j} p_{ij}^{r_{ij}} p_{ji}^{n-r_{ij}} d\theta_1 d\theta_2 \dots d\theta_{k-1}, \tag{10}$$

given by equations (9) and (10), respectively.

3. APPLICATION AND RESULTS

In this study, we are interested in examining the choices of 150 smokers (judges) when asked about their preferred cigarette brand among three commonly used and accessible brands (item number is three). As the data is collected through a balanced PC experiment, all judges are requested to report their preferences on three possible pairs, that is $(GL, Ma), (GL, BH)$ and (BH, Ma) , by asking “Do you prefer brand i over brand j ?”

Table 1 presents the reported responses of 150 respondents when asked about their pairwise choices. When comparing GL and Ma, 98 respondents (65.3%) out of 150 smokers reported GL as their preferred brand, whereas 52 (34.6%) chose Ma over GL. Similarly, 89 participants (59.3%) preferred GL brand over BH brand, which was preferred by 61 (40.6%) respondents. In a comparison of BH and Ma, 83 contestants (55.3%) favored BH, while 67 participants (44.6%) stayed with Ma.

It is important to note that, based on the reported nicotine levels, brands can be ranked in descending order such as $GL(16.92 \pm 0.82) \rightarrow BH(15.93 \pm 0.69) \rightarrow Ma(12.95 \pm 0.82)$. Thus the data support the anecdotal evidence and understanding that nicotine level plays an important role in driving choice behaviors of a smoker in choosing a preferred brand.

Table 1. The Choice data matrix

Pairs	Yes	No	Total
$GL > Ma$	98 (65.30%)	52 (34.60%)	150
$GL > BH$	89 (59.30%)	61 (40.60%)	150
$BH > Ma$	83 (55.30%)	67 (44.60%)	150

Figure 1 and Figure 2 display the marginal posterior distributions (MPDs) of the worth parameters (equations (9) and (10)) associated with each brand under both priors, the Uniform and Jefferys prior, respectively. A similar behavior of marginal posterior distributions under both non-informative priors is observed. Note the highest peak of the marginal posterior distribution, indicating higher utility of a brand, belongs to the GL brand, then followed by the BH brand and lastly the Ma brand.

We report on the estimates of the worth parameters and preference probabilities. Using Gibbs sampling, in Table 2 we give the posterior means estimated by using MPD of equations (9 & 10) as estimates of the worth parameters under both priors, along with the posterior standard deviations (reported in parenthesis).

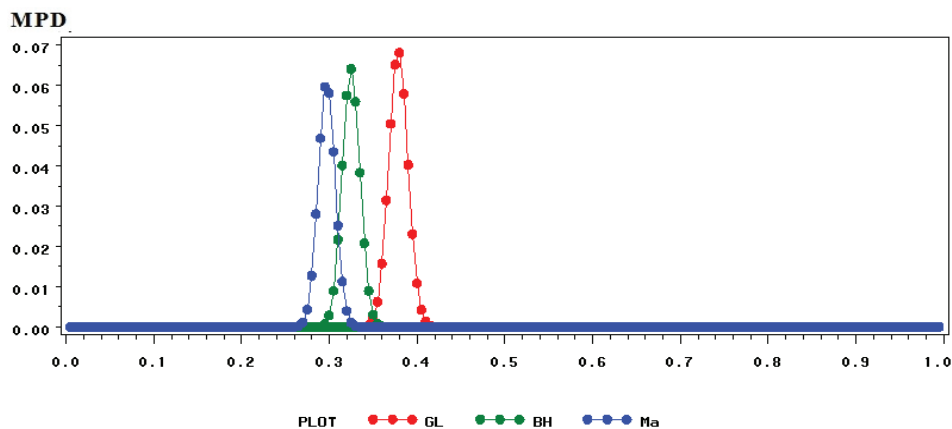


Figure 1. Marginal posterior distribution (MPD) of the worth parameters using the Uniform prior.

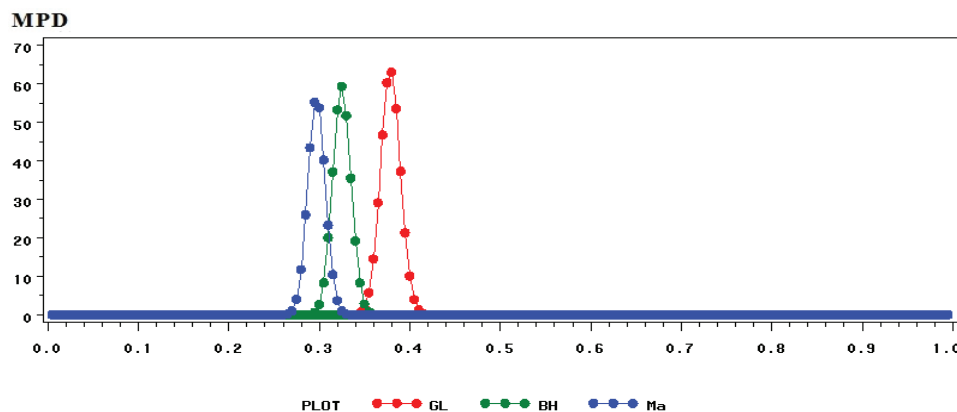


Figure 2. Marginal posterior distribution (MPD) of the worth parameters using the Jefferys prior.

Table 2. Estimated worth parameters, θ 's, and posterior standard deviations under the Uniform prior and Jefferys prior.

Worth parameters (θ 's)	Uniform prior			Jefferys prior		
	$\hat{\theta}_{GL}$	$\hat{\theta}_{BH}$	$\hat{\theta}_{Ma}$	$\hat{\theta}_{GL}$	$\hat{\theta}_{BH}$	$\hat{\theta}_{Ma}$
	0.3774 (0.01524)	0.3239 (0.01463)	0.2973 (0.00076)	0.3765 (0.00901)	0.3247 (0.00876)	0.2981 (0.00991)

Regardless of the prior distributions, we observed common trends in the choice behaviors. For example, with respect to the both priors, the highest value of the worth parameter is associated with the GL brand, which also has the highest reported nicotine level. The GL brand is then followed by BH brand, with which the second highest value of worth parameter and nicotine level is attributed. The minimal value of the worth parameter is associated with the Ma brand which has the lowest level of nicotine among the three studied brands.

Based on these estimates of worth parameters, in Table 3 we report the estimated preference probabilities of all pairwise comparisons of the three cigarette brands. By comparing the estimated preference probabilities, obtained by our proposed model (equation (3) reported in Table 3) with the data based estimates given in Table 1, the newly proposed scheme is successful in retaining relevant information in choice data. Under both priors, we observed that

under both priors, the proposed procedure not only maintained the preference ordering of brands, but also closely estimated the preference probabilities.

Table 3. Estimated preference probabilities, $\hat{p}_{i>j}$, under the Uniform prior and Jefferys prior.

Preference probabilities ($\hat{p}_{i>j}$)	Uniform prior			Jefferys prior		
	$\hat{p}_{GL>Ma}$	$\hat{p}_{GL>BH}$	$\hat{p}_{BH>Ma}$	$\hat{p}_{GL>Ma}$	$\hat{p}_{GL>BH}$	$\hat{p}_{BH>Ma}$
	0.6429	0.5987	0.5877	0.6492	0.5927	0.5599

4. CONCLUSIONS

This paper contributes a new PC model, named the Maxwell paired comparison model to the literature. The applicability of the newly proposed comparative model is demonstrated by analyzing 150 smokers' brand choice. The data is collected through a balanced PC design where all respondents are questioned about their preferred cigarette brand among the three most popular brands smoked in Pakistan. These brands are selected on the basis of their accessibility, popularity and nicotine level.

The data analysis is conducted under a Bayesian framework while considering two non-informative priors, the Uniform prior and Jefferys prior. The capability of our proposed model in handling choice data is clear. We observed that under both priors, the proposed procedure not only maintained the preference ordering of brands, but also closely estimated the preference probabilities. However, the Jefferys prior seems to be a more suitable prior than the Uniform prior. Furthermore, we observed that nicotine level plays an instrumental role in driving the preference ordering of cigarette brands. We conclude that smokers in general tend to choose a cigarette brand with higher levels of nicotine. Public health interventions could potentially address this via public health warnings and intervention/education programs.

Of interest for future work is to explore the performance of our newly proposed Maxwell model under informative priors, particularly the Chi-square distribution, given the Maxwell distribution has inherently similar characteristics to the Chi-square distribution. Moreover, a study to assess the accuracy of the worth parameters based on the width of MPDs is also in view. In the future we also plan to amend our model to accommodate *ties* and also adjust for ties in further models using informative priors.

REFERENCES

- Amlani, A. M., and Schafer, E. C. (2009). Application of paired-comparison methods to hearing aids. *Trends in Amplification*, 13(4), 241-259.
- Bergen, A. W., and Caporaso, N. (1999). Cigarette smoking. *Journal of the National Cancer Institute*, 91(16), 1365-1375.
- Bradley, R. A., and Terry, M. E. (1952). Rank analysis of incomplete block designs: The method of paired comparisons. *Biometrika*, 39(3/4), 324-345.
- Brush, S. G. (2003). *The Kinetic Theory of Gases: Anthology of Historical Papers with Historical Commentary*. London: Imperial College Press.
- Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3), 541-553.
- Cattelan, M., Varin, C. and Firth, D. (2013). Dynamic Bradley–Terry modelling of sports tournaments. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 62(1), 135-150.
- Dellaportas, P. and Smith, A. F. (1993). Bayesian inference for generalized linear and proportional hazards models via Gibbs sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 42(3), 443-459.
- Diaconis, P., Khare, K. and Saloff-Coste, L. (2008). Gibbs sampling, exponential families and orthogonal polynomials. *Statistical Science*, 23(2), 151-178.
- Dittrich, R. Francis, B., Hatzinger, R., and Katzenbeisser, W. (2012). Missing observations in paired comparison data. *Statistical Modelling*, 12(2), 117-143.
- Gelfand, A. E. (2000). Gibbs sampling. *Journal of the American statistical Association*, 95(452), 1300-1304.
- Green-Armytage, J. (2004). Cardinal-weighted pairwise comparison. *Voting matters*, 19, 6-13.
- Jackson, J. E. and Fleckenstein, M. (1957). An evaluation of some statistical techniques used in the analysis of paired comparison data. *Biometrics*, 13(1), 51-64.
- Jones, G. L. and Johnson, A. A. (2008). Comment: Gibbs Sampling, Exponential Families, and Orthogonal Polynomials. *Statistical science*, 23(2), 183-186.

- Latta, R. B. (1979). Composition rules for probabilities from paired comparisons. *The Annals of Statistics*, 7(2), 349-371.
- Mazzuchi, T. A., Linzey, W. G. and Bruning, A. (2008). A paired comparison experiment for gathering expert judgment for an aircraft wiring risk assessment. *Reliability Engineering & System Safety*, 93(5), 722-731.
- Rayner, J. C.W. and Best, D. J. (2001). *A contingency table approach to nonparametric testing*. Chapman and Hall/CRC, Boca Raton, FL.
- Schauberger, G. and Tutz, G. (2017). Subject-specific modelling of paired comparison data: A lasso-type penalty approach. *Statistical Modelling*, 17(3), 223-243.
- Stern, S. E. (2011). Moderated paired comparisons: A generalized Bradley Terry model for continuous data using a discontinuous penalized likelihood function. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 60(3), 397-415.
- Sung, Y. T. and Wu, J. S. (2018). The Visual Analogue Scale for Rating, Ranking and Paired-Comparison (VAS-RRP): A new technique for psychological measurement. *Behavior research methods*, 50(4), 1694-1715.
- Veghes, C. (2014). Using the paired comparison method in improving soccer rankings: The case of the Romanian First Division. *Annales Universitatis Apulensis Series Oeconomica*, 16(2), 379-388.
- Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. E. J. Butterfield (Ed.), *Philosophy of Physics, Part B Handbook of the Philosophy of Science*. The Netherlands: Elsevier.