

---

# A Latent Space Model for Rank Data.

---

**Isobel Claire Gormley**

GORMLEYI@TCD.IE

Department of Statistics, School of Computer Science and Statistics, Trinity College Dublin, Ireland.

**Thomas Brendan Murphy**

MURPHYBT@TCD.IE

Department of Statistics, School of Computer Science and Statistics, Trinity College Dublin, Ireland.

## Abstract

Proportional representation by means of a single transferable vote (PR-STV) is the electoral system employed in Irish elections. In this system, voters rank some or all of the candidates in order of preference. A latent space model is proposed for these election data where both candidates and voters are located in the same  $D$ -dimensional space. The locations are determined by the ranked preferences which are modeled using the Plackett-Luce model for rank data. Voter positions reflect their preferences while the candidate locations represent the global view of the candidates by the electorate.

## 1. Introduction

Proportional representation by means of a single transferable vote (PR-STV) is the electoral system employed in Ireland in both general (governmental) and presidential elections. In this electoral system, voters are required to rank some or all of the proposed candidates in order of preference. Votes are then transferred between candidates during counting, according to the voter preferences, as candidates are elected or eliminated from the race.

A wealth of rank data is available within the context of Irish elections. The introduction of electronic voting in several constituencies means actual voting data are now publicly available. The work presented here focuses on the Irish general election of 2002, where the current government was elected; specifically the votes from the general election in the constituency of Meath are examined. Details of this election are outlined in Section 2.

---

Appearing in *Proceedings of the 23<sup>rd</sup> International Conference on Machine Learning*, Pittsburgh, PA, 2006. Copyright 2006 by the author(s)/owner(s).

A latent space model (Section 3.1) similar to that of Hoff et al. (2002) is proposed where both voters and candidates are located simultaneously in a  $D$ -dimensional latent space. The location of each candidate is inferred from the votes cast by the electorate — the Plackett-Luce model for rank data (Section 3.2) is employed to exploit the information incorporated in the ranked preferences contained in the votes. In turn, voter locations are determined by their votes, which demonstrate their support for each of the candidates. This model is fitted within the Bayesian paradigm; the Metropolis-Hastings algorithm is the primary model fitting tool. When fitting latent space models issues such as invariant configurations and choice of dimensionality arise; these are dealt with in Sections 4.1 and 4.3 respectively.

The relative spatial locations of the candidates allow inferences to be made about the type of relationships that may exist between the candidates, as viewed by the electorate. As coalition governments often occur in countries that use proportional representation election systems, interest lies in examining if candidates from different political parties are deemed alike. Which political parties are viewed as similar by the electorate? What characteristics do closely located candidates share? What mechanisms drive Irish general elections? Such questions will be answered by examining the relative locations of the candidates.

Configurations of the candidates and electorate from the 2002 general election in the Meath constituency indicate that party politics drive voter opinions.

We conclude, in Section 6, by proposing possible modifications and extensions to the model fitted in this study.

## 2. Irish Elections

Irish elections employ a voting system called proportional representation by means of a single transfer-

able vote (PR-STV). In the PR-STV system, voters rank some or all of the candidates in order of preference. During the counting process first preferences are totalled and candidates are elected or eliminated from the race depending on a constituency specific quota. Excess votes above the quota or those belonging to an eliminated candidate are transferred to other candidates according to the ranked preferences. This process continues until all seats are filled or until a sufficient number of candidates are left in the race. A precise description of the electoral system, including the method of counting votes is given in Sinnott (1999). The transfer of votes during the 2002 general election in the Meath constituency can be viewed at <http://www.oireachtas.ie>. Good introductions to the Irish political system are given in Coakley and Gallagher (1999) and Sinnott (1995).

### 2.1. The 2002 General Election

Dáil Éireann (the Irish House of Parliament) consists of one hundred and sixty six members who are elected in a general election held at least once every five years. The members of the Dáil represent forty two constituencies. The most recent general election was held on May 17th, 2002. This election saw the introduction of electronic voting, for the first time, in three constituencies (Dublin North, Dublin West, and Meath). The remaining thirty nine constituencies used paper ballots.

Five seats in Dáil Éireann were allocated to the constituency of Meath and fourteen candidates ran for election within the constituency. The fourteen candidates represented seven political parties, with the major parties of Fianna Fáil and Fine Gael each having three candidates. Table 1 provides details of all the candidates and their political affiliations. The five seats in the Meath constituency were won by Dempsey, Bruton, Wallace, English and Brady. The electorate in Meath consisted of 108,717 individuals and there were a total of 64,081 valid votes cast. The actual votes cast in the Meath constituency are analyzed in this work.

The voting data from the Meath constituency is available from the Meath local authority web page (<http://www.meath.ie/election.html>). Voting data from the other constituencies where electronic voting was implemented is available from the constituency returning officer's web page (<http://www.dublincountyreturningofficer.com>).

Table 1. The fourteen candidates who ran for election in the Meath constituency in 2002. The political affiliation of each candidate is given as well as an abbreviation of their surname and party. Independent candidates are not affiliated to any party.

Candidate	Party
Brady, J. (By)	Fianna Fáil (FF)
Bruton, J. (Bt)	Fine Gael (FG)
Colwell, J. (Cl)	Independent (Ind)
Dempsey, N (Dp)	Fianna Fáil (FF)
English, D. (Eg)	Fine Gael (FG)
Farrelly, J. (Fr)	Fine Gael (FG)
Fitzgerald, B. (Fz)	Independent (Ind)
Kelly, T. (Kl)	Independent (Ind)
O'Brien, P. (Ob)	Independent (Ind)
O'Byrne, F. (Oby)	Green Party (GP)
Redmond, M. (Rd)	Christian Solidarity (CSP)
Reilly, J. (Rl)	Sinn Féin (SF)
Wallace, M. (Wl)	Fianna Fáil (FF)
Ward, P. (Wd)	Labour (Lab)

## 3. Modeling

A latent space model is combined with a model for rank data to provide a suitable tool for the simultaneous modeling of PR-STV data and the candidates and electorate that give rise to it.

### 3.1. The Latent Space Model

Hoff et al. (2002) proposed a model for social networks where the network actors are located in a latent space and the probability of a connection between two actors is determined by their proximity. In a similar vein to this work, a model is proposed for rank data where voters and candidates are located in the same  $D$  dimensional latent space  $\mathbf{Z} \subseteq \mathbb{R}^D$ . It is assumed that each of  $M$  voters has latent location  $\underline{z}_i \in \mathbf{Z}$  and each candidate  $j$  ( $j = 1, \dots, N$ ) has latent location  $\underline{\zeta}_j \in \mathbf{Z}$ . Let  $d(\underline{z}_i, \underline{\zeta}_j)$  be the squared Euclidean distance between voter  $i$  and candidate  $j$  in the latent space  $\mathbf{Z}$ , that is

$$d(\underline{z}_i, \underline{\zeta}_j) = \frac{1}{D} \sum_{d=1}^D (z_{id} - \zeta_{jd})^2$$

for  $i = 1, \dots, M$  and  $j = 1, \dots, N$ . The squared Euclidean distance is invariant to rotations and translations. Many other distance measures are available as detailed by Mardia et al. (1979) and possible alternatives are discussed in Section 6.

The distance  $d(\underline{z}_i, \underline{\zeta}_j)$  (for  $j = 1, \dots, N$ ) between voter  $i$  and the candidates describes the voter's electoral

opinions. In a similar way the proximity of two candidates in the latent space quantitatively describes their relationship as deemed by the electorate.

By exploiting the information contained in the ranked preferences the latent locations of each voter and candidate can be inferred. Thus a latent space model is incorporated with a standard rank data model to spatially model Irish voting data.

### 3.2. The Plackett-Luce Model

Many rank data models in the literature emerge from the context of predicting the final ordering of horses in a race (see Plackett (1975) and Benter (1994)). The ranking of candidates on a ballot form may be thought of in a similar manner. Marden (1995) details many of the currently available rank data models. The Plackett-Luce model (Plackett, 1975) is one such model which can be used to model PR-STV data. The model is a member of a larger group of models known as *multi stage models* (Fligner & Verducci, 1988). Mixtures of Plackett-Luce models have recently been used by Gormley and Murphy (2006) to model Irish college applications data.

The Plackett-Luce model models a ranking as a sequential process in which each voter selects the next most preferred candidate. The model is parameterized by a ‘support’ parameter

$$\underline{p}_i = (p_{i1}, p_{i2}, \dots, p_{iN})$$

for each of  $i = 1, \dots, M$  where  $\sum_{j=1}^N p_{ij} = 1$ . The parameter  $p_{ij}$  can be interpreted as the probability of voter  $i$  selecting candidate  $j$  in first place on their ballot. The probability of candidate  $j$  being given a lower preference by voter  $i$  is proportional to the support parameter  $p_{ij}$ . At each choice level the  $p_{ij}$  values are appropriately normalized subject to the constraint that previously ranked candidates are excluded from the probability vector  $\underline{p}_i$ .

The probability  $p_{ij}$  of a voter selecting a particular candidate first is a decreasing function of the distance between the voter and the candidate in the latent space. It is assumed that these probabilities take the form

$$p_{ij} = \frac{\exp\{-d(\underline{z}_i, \underline{\zeta}_j)\}}{\sum_{j'=1}^N \exp\{-d(\underline{z}_i, \underline{\zeta}_{j'})\}}$$

for  $i = 1, \dots, M$  and  $j = 1, \dots, N$ . Thus the position taken by each voter and candidate in the latent space is determined by the preferences expressed on the ballot forms.

Within the context of voting data,  $\underline{x}_i = (c(i, 1), c(i, 2), \dots, c(i, n_i))$  denotes voter  $i$ ’s ballot, where  $n_i$  is the number of preferences expressed by voter  $i$  and  $c(i, t)$  denotes the candidate chosen in  $t$ th position by voter  $i$ . Under the PR-STV voting system the number of preferences voter  $i$  expresses may vary; that is  $1 \leq n_i \leq N$ . The entire electorate’s set of votes is denoted  $\mathbf{X} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_M)$ .

Under a Plackett-Luce model with support parameters  $\mathbf{p} = (\underline{p}_1, \dots, \underline{p}_M)$  the probability of all votes  $\mathbf{X}$  is

$$\mathbf{P}\{\mathbf{X}|\mathbf{p}\} = \prod_{i=1}^M \prod_{t=1}^{n_i} \frac{p_{ic(i,t)}}{\sum_{s=t}^N p_{ic(i,s)}}$$

where for  $s \geq n_i$  the sequence of  $c(i, s)$  is any arbitrary ordering of the candidates not selected by voter  $i$ ; this ordering does not affect future calculations.

## 4. Model Fitting

The Plackett-Luce model combined with a latent space model allows for the ranked nature of the PR-STV data to be spatially modeled. The parameters of this model and their related uncertainty are estimated within a Bayesian framework.

Prior densities for voter locations,  $p_v(\underline{z}_i)$ , and for candidate locations,  $p_c(\underline{\zeta}_j)$  are assumed to be Normal and independent where  $z_{id} \sim N(\mu_v, \sigma_v^2) = N(0, 3^2)$  and  $\zeta_{jd} \sim N(\mu_c, \sigma_c^2) = N(0, 3^2)$  for  $d = 1, \dots, D$ . Thus the joint density  $P\{\mathbf{X}, \mathbf{z}, \zeta\}$  of the votes cast, the voter locations and the candidate locations is

$$\prod_{i=1}^M p_v(\underline{z}_i) \left[ \prod_{t=1}^{n_i} \frac{\exp\{-d(\underline{z}_i, \underline{\zeta}_{c(i,t)})\}}{\sum_{s=t}^N \exp\{-d(\underline{z}_i, \underline{\zeta}_{c(i,s)})\}} \right] \left[ \prod_{j=1}^N p_c(\underline{\zeta}_j) \right].$$

The locations of each voter and each candidate in the latent space are to be estimated — samples from the posterior distribution  $\mathbf{P}\{\mathbf{z}, \zeta|\mathbf{X}\}$  are generated using a Metropolis-Hastings algorithm. A random walk proposal density where each location was perturbed using normally distributed noise was employed — good acceptance rates (detailed in Section 5) were achieved in the estimation of both voter and candidate locations using this proposal.

### 4.1. Invariant Configurations

The measure of distance between voter and candidate locations in the latent space is quantified by the squared Euclidean distance. This distance is invariant to rotations and translations. As a result the model is not fully identifiable because the locations are only

identified up to rotation and translation. Procrustean methods are used to eradicate this problem.

Procrustean methods (Krzyszowski, 1988) match one configuration of points to another as well as possible in a least squares sense. Transformations such as dilation, rotation and translation are used to create the match. In this context only translations and rotations are applicable without altering the likelihood of the data due to the definition of the probabilities  $p_{ij}$ .

Assume  $C^R = (\mathbf{z}^R, \zeta^R)$  is a reference configuration of the voter and candidate locations which is centered around the origin. To match the estimated configuration  $\hat{C}$  to the reference configuration  $C^R$ ,  $\hat{C}$  is first translated so that is also centered around the origin.  $\hat{C}$  is then rotated to provide the best match with  $C^R$  in a least squares sense.

To obtain  $Q$ , the optimal orthogonal rotation matrix, the sum

$$\begin{aligned} S &= \sum_{i=1}^{M+N} \sum_{d=1}^D (c_{id}^R - \hat{c}_{id})^2 \\ &= \text{trace} \left\{ C^R C^{R'} + \hat{C} \hat{C}' - 2C^R \hat{C}' \right\} \quad (1) \end{aligned}$$

is minimized. The newly rotated configuration is denoted  $\hat{C}Q$ . Thus (1) becomes

$$S = \text{trace} \left\{ C^R C^{R'} + \hat{C} \hat{C}' - 2C^R Q' \hat{C}' \right\}$$

and the minimization problem becomes the constrained maximization of  $2C^R Q' \hat{C}'$ . It follows that  $Q = VU'$  where  $USV'$  is the singular value decomposition of  $C^{R'} \hat{C}$ . Thus by centering each estimated configuration around the origin and rotating the configuration using the rotation matrix  $Q$  the estimated configuration  $\hat{C}$  is best matched with the reference configuration  $C^R$ .

Samples of the configuration  $(\mathbf{z}, \zeta)$  are generated using the Metropolis-Hastings algorithm. Initial iterations of the algorithm are constrained to only accept uphill moves to achieve an estimate of the *maximum a posteriori* (MAP) configuration of candidate and voter locations. This MAP configuration is henceforth employed as  $C^R$ , the reference configuration, to which each subsequently estimated configuration  $\hat{C}$  is matched.  $C^R$  is not assumed to be the correct configuration but is merely used as a standard to which others are matched.

## 4.2. Estimation Of Voter And Candidate Latent Locations

The location of each voter  $\underline{z}_i$  and each candidate  $\underline{\zeta}_j$  within a  $D$  dimensional latent space is to be estimated. A random walk Metropolis-Hastings algorithm is used to sample from the joint density  $\mathbf{P}\{\mathbf{z}, \zeta | \mathbf{X}\}$ . Once the MAP reference configuration  $C^R$  has been obtained, the algorithm proceeds as usual accepting or rejecting proposed voter and candidate locations. At the end of each iteration of the Metropolis-Hastings algorithm (subsequent to obtaining the MAP configuration and to an additional burn-in period) the newly estimated configuration  $\hat{C}$  is translated and rotated to best match the reference configuration  $C^R$ .

To determine whether or not voter location  $\underline{z}_i$  should be updated to location  $\underline{z}_i^*$  an acceptance probability,  $\alpha(\underline{z}_i^*, \underline{z}_i)$ , is required:

$$\alpha(\underline{z}_i^*, \underline{z}_i) = \min \left\{ \frac{P(\mathbf{z}^*, \zeta | \mathbf{X}) q(\underline{z}_i | \underline{z}_i^*)}{P(\mathbf{z}, \zeta | \mathbf{X}) q(\underline{z}_i^* | \underline{z}_i)}, 1 \right\} \quad (2)$$

where independence of voter locations and a symmetric random walk proposal distribution  $q(\cdot)$  are assumed. Estimates of  $\hat{\underline{z}}_i$  are generated from the posterior distribution via the following algorithm:

1. Generate a value  $\epsilon$  from the symmetric proposal density  $q(\epsilon) = N(\mu_{vprop}, \sigma_{vprop}^2)$  and form the proposal point  $z_{id}^* = z_{id} + \epsilon$  for  $d = 1, \dots, D$ .
2. Compute the acceptance probability  $\alpha(\underline{z}_i^*, \underline{z}_i)$  as detailed in (2).
3. Generate a value  $u \sim \text{Uniform}(0, 1)$ .
4. If  $u \leq \alpha(\underline{z}_i^*, \underline{z}_i)$  then define  $\underline{z}_i = \underline{z}_i^*$ , otherwise define  $\underline{z}_i = \underline{z}_i$ .

Similar methodology applies in the case of estimating candidate locations. A random walk proposal density is again employed where  $\epsilon \sim N(\mu_{cprop}, \sigma_{cprop}^2)$ . The acceptance probability of updating candidate location  $\underline{\zeta}_j$  to location  $\underline{\zeta}_j^*$  is given by

$$\alpha(\underline{\zeta}_j^*, \underline{\zeta}_j) = \min \left\{ \frac{P\{\mathbf{z}, \zeta^* | \mathbf{X}\} q(\underline{\zeta}_j | \underline{\zeta}_j^*)}{P\{\mathbf{z}, \zeta | \mathbf{X}\} q(\underline{\zeta}_j^* | \underline{\zeta}_j)}, 1 \right\} \quad (3)$$

where independence of candidates locations is assumed. Estimates  $\hat{\underline{\zeta}}_j$  are generated from the posterior distribution via the following algorithm:

1. Generate a value  $\epsilon$  from the symmetric proposal density  $q(\epsilon) = N(\mu_{cprop}, \sigma_{cprop}^2)$  and form the proposal point  $\zeta_{jd}^* = \zeta_{jd} + \epsilon$  for  $d = 1, \dots, D$ .
2. Compute the acceptance probability  $\alpha(\zeta_j^*, \zeta_j)$  as detailed in (3).
3. Generate a value  $u \sim \text{Uniform}(0, 1)$ .
4. If  $u \leq \alpha(\zeta_j^*, \zeta_j)$  then define  $\zeta_j = \zeta_j^*$ , otherwise define  $\zeta_j = \zeta_j$ .

The algorithm sequentially estimates the voter locations and then the candidate locations until sufficient mixing of the Markov chain is achieved. Locations estimated subsequent to both the uphill-only and burn-in periods are considered when calculating final estimates.

### 4.3. Dimensionality

The dimensionality  $D$  of the latent space is a further variable which requires estimation. Several techniques have been discussed in the literature as potential methods of selecting the optimal dimensionality of a space. Selecting the optimal  $D$  can be viewed as a model selection process between models with different dimensions.

The deviance information criterion (DIC) (Spiegelhalter et al., 2002) is a measure of model complexity defined as a Bayesian measure of accuracy (or deviance) penalized by an additional complexity term. The complexity term is labeled the ‘effective number of parameters’ and is the difference between the posterior mean of the deviance and the deviance at the posterior mean of the estimates of the model parameters. Spiegelhalter et al. (2002) tentatively suggest that combining the Bayesian deviance and the complexity term forms a criterion which may be used for comparing models.

Pritchard et al. (2000) suggest a model selection criterion based on an approximation of the posterior distribution  $\mathbf{P}\{D|\mathbf{X}\}$ . It is computationally similar to the DIC but penalizes the mean of the Bayesian deviance by a quarter of its variance as opposed to the effective number of parameters.

In practice, the DIC and Pritchard et al.’s criterion gave conflicting results (see Table 2). The multivariate analysis technique of applying principal components analysis (Mardia et al., 1979) to the resulting configurations of different dimensions provided more intuitive results. Principal components analysis (PCA) is a data reduction technique which finds linear combinations (components) of the variables with maximal

variance. The aim of PCA is to summarize the data while losing as little information as possible.

The number of principal components required to adequately represent the data is a question similar to that of choosing the optimal dimension  $D$ . The variances of the fitted principal components describe the proportion of the variance of the data explained by the components. Thus these variances provide insight into the number of components necessary to adequately summarize the data in the sense that the addition of further components is not deemed to provide any supplemental information.

PCA is applied to the configuration of candidates only as the predominant interest lies in the interpretation of the relative locations of the candidates. The variances of the resulting principal components are examined and the optimal number of dimensions  $D$  is selected to be the number of dimensions after which the addition of another dimension was not deemed to be beneficial.

## 5. Results

A latent space model for rank data was applied to the set of votes from the Meath constituency in the 2002 Irish general election.

### 5.1. The 2002 General Election: Meath Constituency

Five seats in Dáil Éireann were available for election in the Meath constituency in the 2002 general election. A latent space model, incorporating the Plackett-Luce model for rank data, was fitted to the 64,081 electronic votes over the range of dimensions  $D = 1, 2$  and 3. Proposal densities were fixed to be

$$N(\mu_{vprop}, \sigma_{vprop}^2) = N(0, 3^2)$$

$$N(\mu_{cprop}, \sigma_{cprop}^2) = N(0, 0.005^2).$$

The DIC and Pritchard et al.’s criterion were computed (Table 2) to determine the appropriate dimension for the latent space but DIC suggested  $D = 3$  whereas Pritchard et al.’s criterion suggested  $D = 1$ .

Table 3 shows the variation captured by each principal component when different dimensions of latent space model were fitted to the data. Both dimensions  $D = 1$  and  $D = 2$  appear to summarize the data well. When a three dimensional model was fitted the additional principal component only accounted for 7% of the variance of the data.

Both the one dimensional and two dimensional con-

Table 2. The DIC values and Pritchard et al.’s criterion values for latent space models of dimension  $D = 1, 2$  and  $3$  fitted to the Meath data. Small values of both criteria indicate the best fitting model. DIC indicated three dimensions as the optimal model where Pritchard et al.’s criterion suggested one. Thus PCA (see Table 3) was employed as the method of selecting the optimal  $D$ .

Dimension	DIC	Pritchard et al.
1	1216797	<b>1209812</b>
2	1118221	3080834
3	<b>1071723</b>	7113250

Table 3. The variance,  $\sigma_d^2$ , captured by each principal component fitted to configurations resulting from the Meath constituency voting data, for different dimensions. Principal components analysis was applied to the average candidate configuration only as the main interest of this study lies in the relative locations of the candidates.

Dimension	Variances		
	$\sigma_1^2$	$\sigma_2^2$	$\sigma_3^2$
1	1	-	-
2	0.81	0.19	-
3	0.70	0.23	0.07

figurations are analyzed to examine relationships between the electoral candidates.

### 5.2. One Dimensional Results

Figure 1 illustrates the one dimensional configuration of the fourteen candidates. Each candidate is represented by a two letter abbreviation of their surname and political party as detailed in Table 1. It is immediately clear that party politics plays a large role in the electorates view of the candidates. The Fianna Fáil candidates (Brady, Dempsey and Wallace) are located on the far left of the single dimension with the Fine Gael candidates (Bruton, English and Farrelly) located on the far right. Fianna Fáil and Fine Gael are the two largest (and rival) Irish political parties. The other candidates lie between the two poles created by the Fianna Fáil and Fine Gael candidates but closer to Fine Gael. Interestingly Ward, who is a Labour Party candidate, is located closest to the Fine Gael candidates — Fine Gael and Labour have a history of forming coalition governments (most recently from 1994–1997). Also of note are the narrow interval estimates for the estimated candidate positions (mean  $\pm 2$  standard deviations are shown). This suggests low uncertainty in the candidate locations in one dimension.

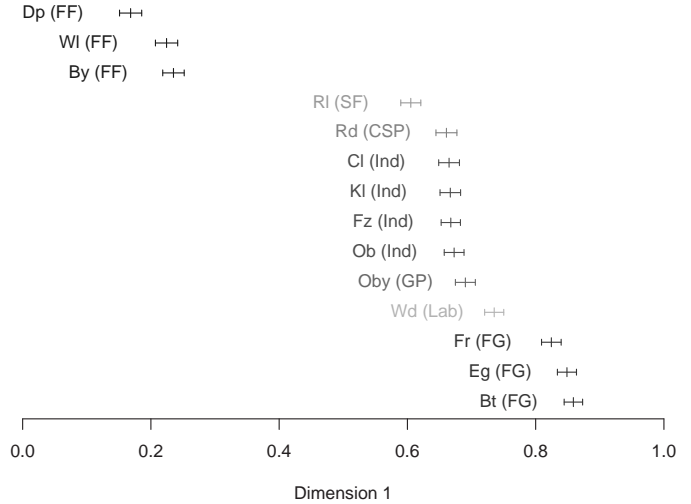


Figure 1. The one dimensional configuration of the candidate means, averaged over a Metropolis-Hastings algorithm, and their associated uncertainty (indicated by  $\pm 2$  standard deviation intervals). Each of the fourteen candidates as detailed in Table 1 are denoted by a two letter abbreviation of their surname and political party. Candidates from different parties are plotted in different levels of gray scale.

### 5.3. Two Dimensional Results

Good acceptance rates of 35% and 33% were achieved for the voter and candidate positions respectively when a two dimensional model was fitted. Figure 2 illustrates the final average position of each of the fourteen candidates in the Meath constituency. Each candidate is denoted by the abbreviations detailed in Table 1. Party politics are again demonstrated as the mechanism which drives this election. The first principal component separates candidates by their political ideals — the estimated positions shows a clear divide between Fianna Fáil and the other parties in the  $x$ -axis direction. The second principal component illustrates the presence of an ideological cleavage (left to right wing) of the candidates. For example, the Christian Solidarity Party espouse right wing conservative values and their candidate Redmond (Rd) is located highest in the second principal component.

The plot also includes ellipses which show approximate 95% posterior set estimates of each candidate location to represent the uncertainty in the estimated locations. The uncertainty associated with all candidate locations is low. Furthermore, there is considerable overlap between candidates from the same party.

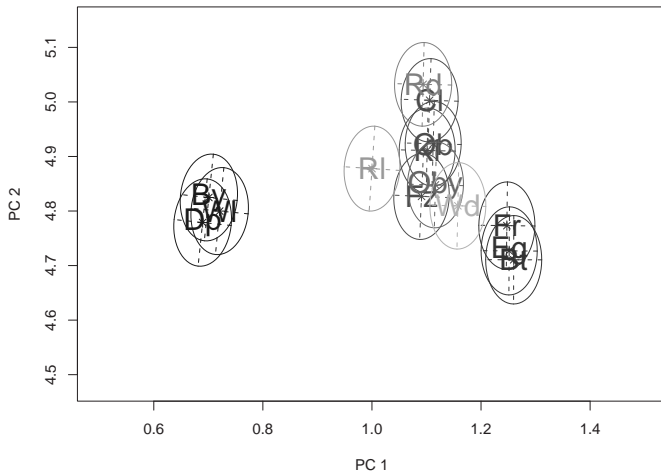


Figure 2. The two dimensional configuration of the candidate means with their associated uncertainty. The candidate initials indicate their posterior mean positions and the ellipses are approximate 95% posterior sets which indicate the uncertainty in the candidate positions. Candidates from different parties are plotted in different levels of gray scale. The position of each candidate and the ellipses are estimated by 8500 Metropolis-Hastings iterations (post burn-in), thinned after every 10th iteration.

## 6. Conclusions And Extensions

A latent space model, incorporating a Plackett-Luce model, provides good methodology for statistically modeling PR-STV rank data. The latent space aspect of the model gives an interpretable framework for the results of model fitting and the Plackett-Luce model works well in modeling PR-STV data.

The latent configurations suggest that party politics drive general elections in Ireland. Other factors such as the level of a candidate’s public profile are also influential but some of these factors are confounded with party membership.

The Plackett-Luce model is in fact a special case of Benter’s model (Benter, 1994) which is another potential rank data model. Under Benter’s model the probability of vote  $i$  is:

$$P\{x_i | p_i, \alpha\} = \prod_{t=1}^{n_i} \left( \frac{p_{ic(i,t)}^{\alpha_t}}{\sum_{s=t}^N p_{ic(i,s)}^{\alpha_t}} \right) \quad 1 \geq \alpha_t \geq 0$$

where  $\alpha_t$  accounts for changing randomness at each choice level. The parameter  $p_i$  has the same interpretation as the Plackett-Luce support parameter. The  $\alpha_t$  values ‘dampen’ the lower choice level probabilities to capture the randomness associated with lower choices. Gormley and Murphy (2005) use a mixture of Benter’s models to model data from the 1997 Irish

presidential election and the 2002 Dublin West general election data. Results of that analysis also show party politics take an important role in the 2002 general election.

When defining the latent space squared Euclidean distance was implemented as a measure of ‘distance’ between two members of the space. While the measure worked well it can be somewhat confusing as, for example, a voter positioned quite far away from Brady may have voted for him alone but is positioned so that he is far away from all other candidates while being as close as possible to Brady. Hoff (2005) made use of the inner product as a latent space distance measure and such a method could be implemented in this context.

Principal components analysis selected the optimal dimension of the latent space — the method worked well but is somewhat ad-hoc. Raftery et al. (2006) introduced the AICM (Akaike Information Criterion Monte Carlo) and BICM (Bayesian Information Criterion Monte Carlo) which are derived through the estimation of the harmonic mean estimator. The AICM and BICM are easily calculated from the posterior output of Monte Carlo simulations and therefore could easily be implemented as a method of estimating  $D$ . Reversible jump Metropolis-Hastings with delayed rejection is another possible but complicated method of selecting  $D$ .

In terms of the Bayesian tools used to fit the model, the random walk proposal worked well in practice but a more sophisticated proposal could be implemented. Also, a diffuse prior was used and a more structured prior on the voters could be employed — for example, a mixture of normals as was used in a social networks context by Handcock et al. (2005) may provide a more informed prior.

## Acknowledgments

Both authors would like to thank Adrian Raftery and other members of the Working Group on Model-Based Clustering at the University of Washington, Seattle for important inputs into this research. We would also like to thank Slava Mikhailov and Liam Weeks from the Department of Political Science in Trinity College Dublin for providing important insights into Irish elections.

Isobel Claire Gormley was supported by a Government of Ireland Research Scholarship in Science, Engineering and Technology provided by the Irish Research Council for Science, Engineering and Technology, funded by the National Development Plan.

## References

- Benter, W. (1994). Computer-based horse race handicapping and wagering systems: A report. In W. T. Ziemba, V. S. Lo and D. B. Haush (Eds.), *Efficiency of racetrack betting markets*, 183–198. San Diego and London: Academic Press.
- Coakley, J., & Gallagher, M. (1999). *Politics in the Republic of Ireland*. London: Routledge in association with PSAI Press. 3rd edition.
- Fligner, M. A., & Verducci, J. S. (1988). Multistage ranking models. *J. Amer. Statist. Assoc.*, *83*, 892–901.
- Gormley, I. C., & Murphy, T. B. (2005). *Exploring Irish Election Data: A Mixture Modelling Approach* (Technical Report 05/08). Department of Statistics, Trinity College Dublin, Dublin 2, Ireland.
- Gormley, I. C., & Murphy, T. B. (2006). Analysis of Irish third-level college applications data. *J. Roy. Statist. Soc. Ser. A*, *169*, 361–379.
- Handcock, M. S., Raftery, A. E., & Tantrum, J. M. (2005). *Model-based clustering for social networks* (Technical Report 482). Department of Statistics, University of Washington, Seattle, WA, USA.
- Hoff, P. D. (2005). Bilinear Mixed-Effects Models for Dyadic Data. *J. Amer. Statist. Assoc.*, *100*, 286–295.
- Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent Space Approaches to Social Network Analysis. *J. Amer. Statist. Assoc.*, *97*, 1090–1098.
- Krzanowski, W. J. (1988). *Principles of multivariate analysis: A user's perspective*. Clarendon Press.
- Marden, J. I. (1995). *Analyzing and modeling rank data*. London: Chapman & Hall.
- Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. London: Academic Press.
- Plackett, R. L. (1975). The analysis of permutations. *Applied Statistics*, *24*, 193–202.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*, 945–959.
- Raftery, A. E., Newton, M. A., Satagopan, J. M., & Krivitsky, P. N. (2006). *Estimating the Integrated Likelihood via Posterior Simulation Using the Harmonic Mean Identity*. (Technical Report 499). Department of Statistics, University of Washington, Seattle, Washington, USA.
- Sinnott, R. (1995). *Irish voters decide: Voting behaviour in elections and referendums since 1918*. Manchester: Manchester University Press.
- Sinnott, R. (1999). The electoral system. In J. Coakley and M. Gallagher (Eds.), *Politics in the Republic of Ireland*, 99–126. London: Routledge & PSAI Press. 3rd edition.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. Roy. Statist. Soc. Ser. B*, *64*, 583–639.