# Electoral Engineering through Simulation

Michael Fowlie and Mark C. Wilson

# Electoral System Properties

- Fairness/Proportionality
- Decisiveness/Governability

- Trade-offs are not linear.

# Measures

- Proportionality
  - Loosemorehanby Index
  - Gallagher Index
- Governability
  - Effective number of parties
  - Sum of squares of Shapley Shubik Power Index
  - Entropy of Shapley Shubik Power Index

- The Electoral Sweet Spot paper claims district magnitude of 2 or 3 is optimal.
- Uses data from 609 real elections in 81 different countries.
- We aim to validate this theoretically by simulation of artificial societies.

# Models for Artificial Societies

- Polya Eggenberger (Urn model)
- Spatial Model

- We assume districts are independent.

# Urn Process (Intuition)

- Urn starts with one ball of each colour.
- We select a ball randomly from the Urn, record it, and place the ball and $\alpha$ copies of the ball back into the Urn.

- $\alpha$ represents the homogeneity of preferences.
- Each colour represents a preference order.

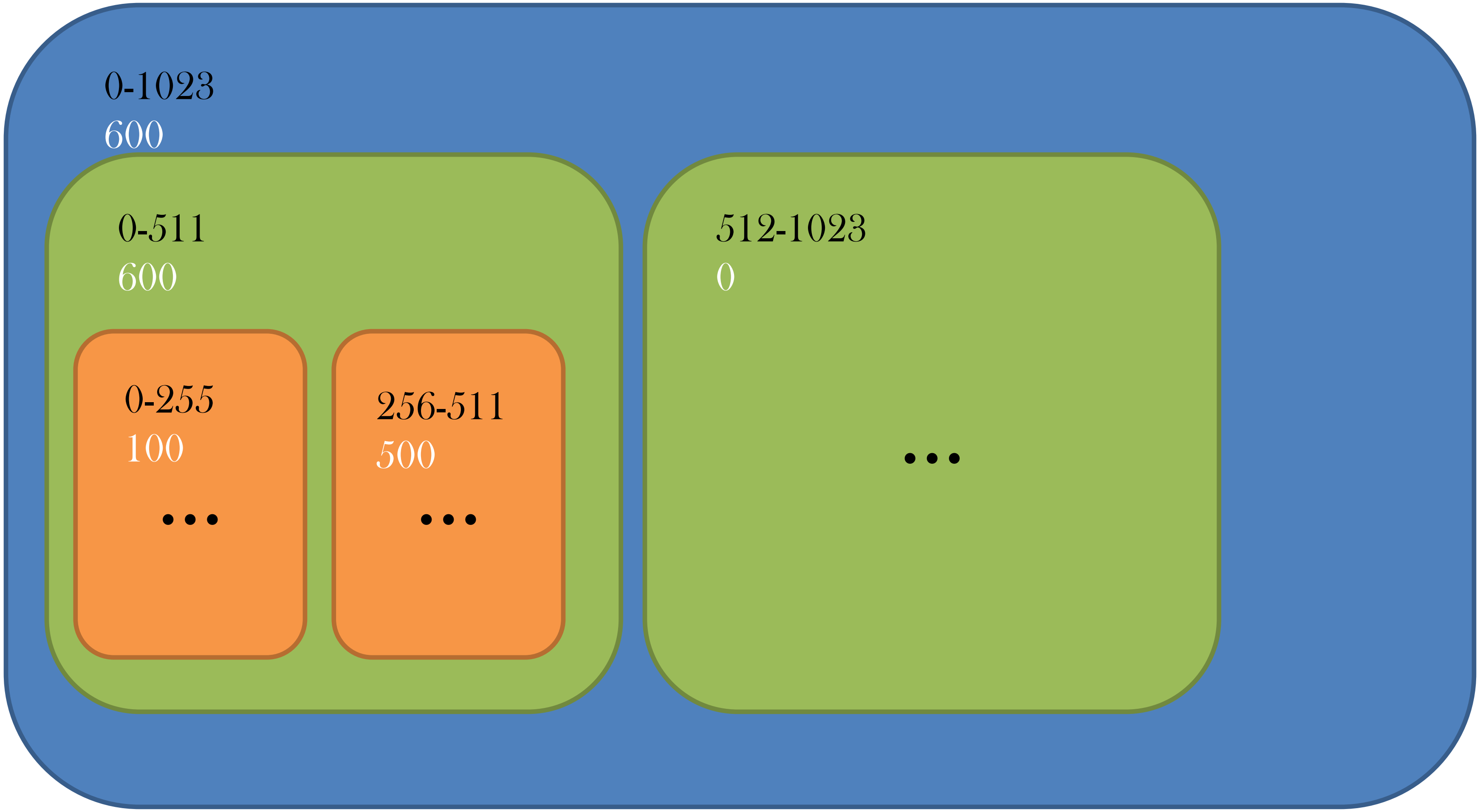
# Sampling from the Urn process

- $m =$ number of parties.
- $n =$ number of voters.
- $M =$ number of preference orders.
- Naive approach would be to simply store the balls in a list and sample from the list.
- $M = m!$, very large for even relatively small $m$. $M \approx 4{,}000{,}000$ for $m = 10$
- Typically $n < M$
- $n \approx 30{,}000$ in the case of New Zealand.

# Efficiently Sampling from the Urn

- We take advantage of the fact that almost all preference orders have a weight of 1.
- Time complexity of order $n \cdot \log(m!)$
- We select $x$ uniformly at random between 0 and 1.
- We select the smallest node with the cumulative probability distribution < x.

# Tree Structure



- Each node in the tree represents a range. Starting from the root, we recursively select a child range that contains $x$.
- Each node has a parameter $w$ representing the total weight of the range.
- We can compute whether a node contains $x$ in constant time.
- We never actually use most of the nodes. Nodes are lazily created.

# Spatial Model

- We create parties with coordinates in an arbitrary dimensional space (e.g. 2D) called the space of issues.
- Each voter has random coordinates and selects a preference order by distance from itself to each party. The lower the distance, the more a voter prefers a party.
- We vary the mean for each coordinate on the district level.

# Spatial Model

- Selecting the number of dimensions and variance in each.
- Placement of parties.
  - Normal distribution
  - Normal distribution + traditional major "left" and "right" parties extremely close to the origin.
- We may be able to obtain typical values for these from real world data. (e.g. PCA on votes for each party, by district)

# Spatial Model

- Variance of distribution used for selecting means for larger district should be lower than that of a smaller district.
- We create a larger district by two methods
  - by combining multiple independent districts
  - the case where the variance is unchanged for the larger district.
- In reality the variance of the distribution for larger districts should be between these two figures. (e.g. it’s not very plausible that a far left district and far right district would be combined)

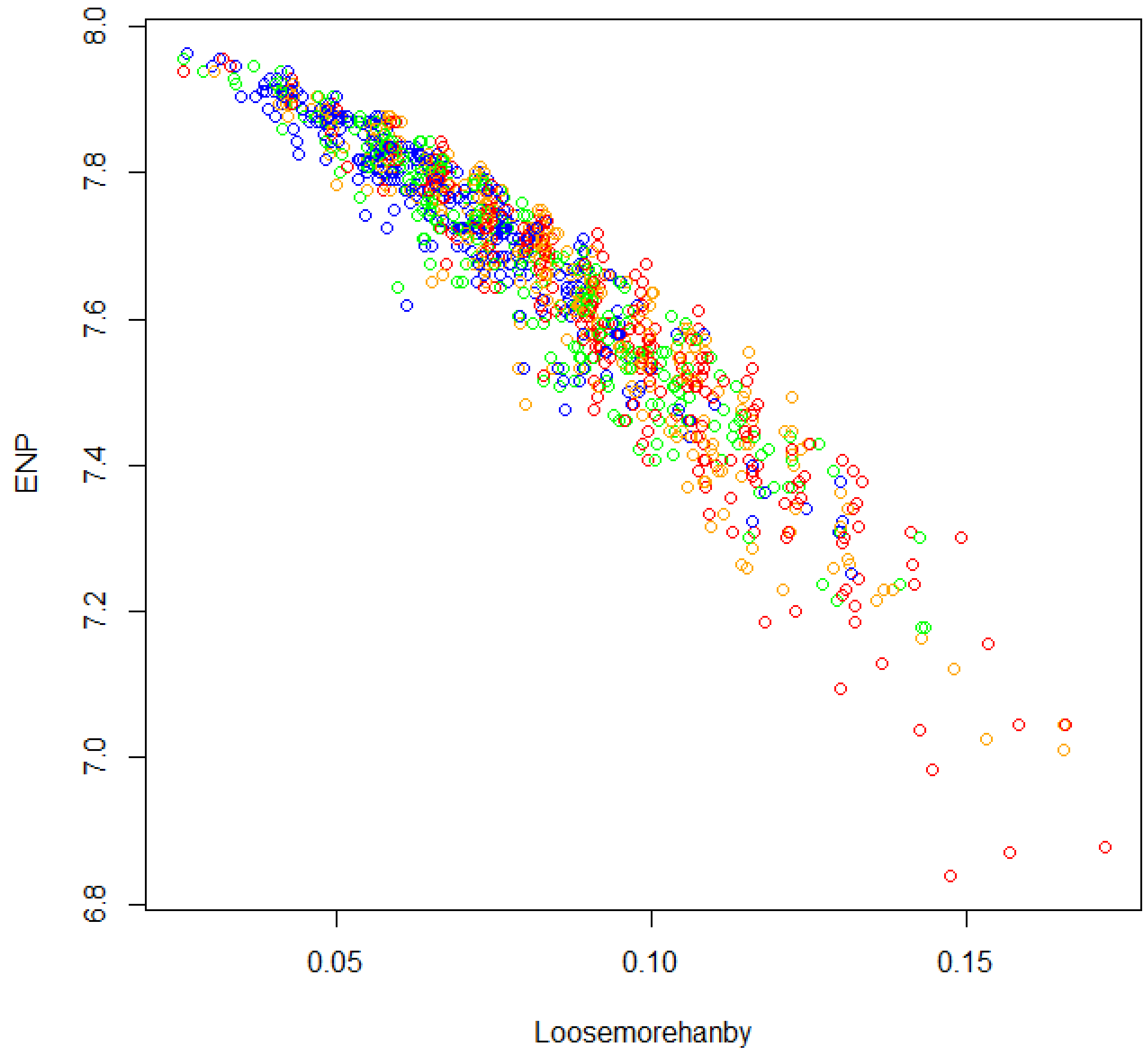
# Preliminary Results

- These results are for STV with different values of district magnitude.

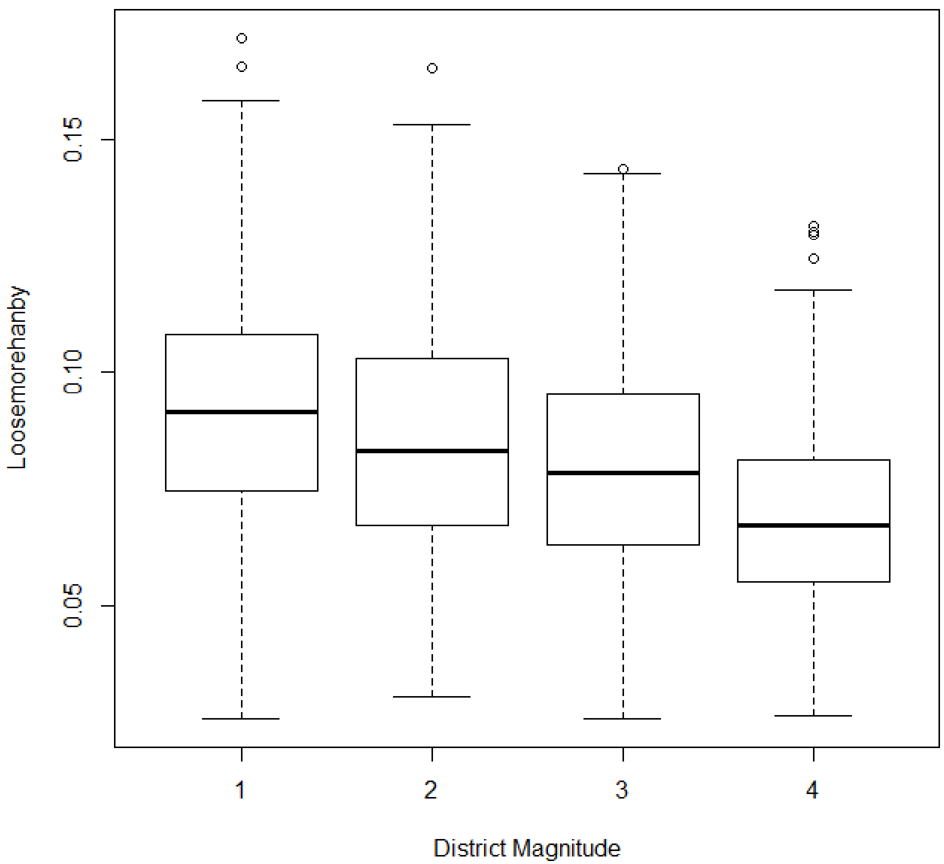
| Colour | DM |
| --- | --- |
| Red | 1 |
| Orange | 2 |
| Green | 3 |
| Blue | 4 |

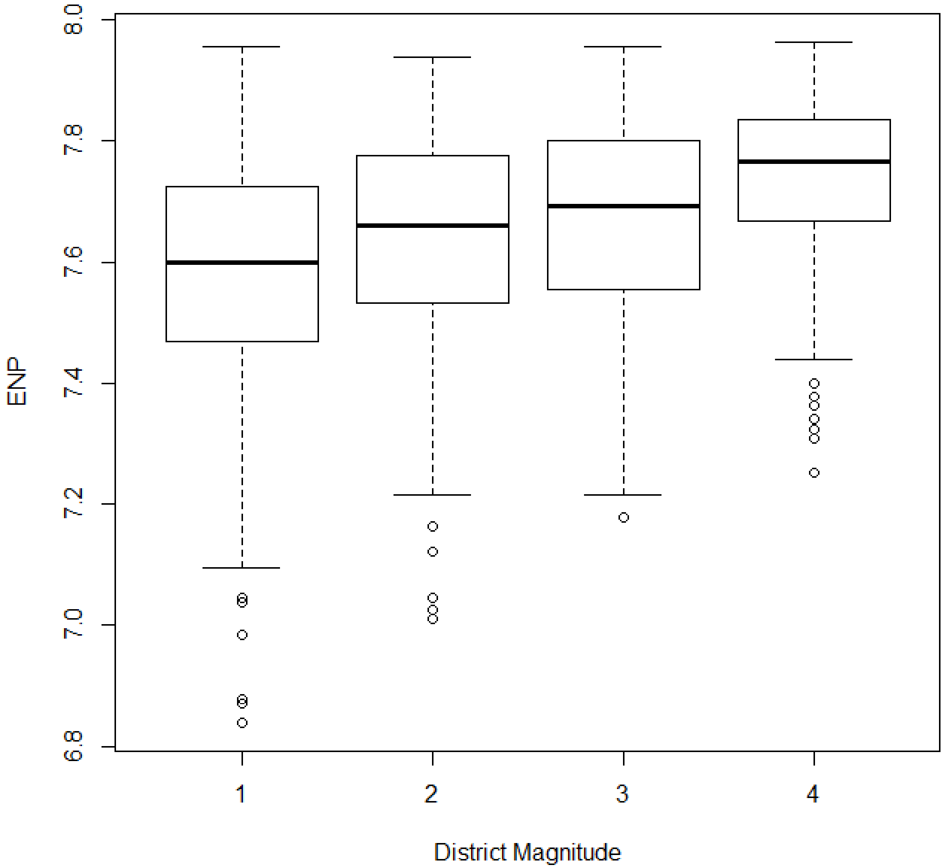
- Points near the bottom left corner are better.

# Polya Eggenberger (Urn model)

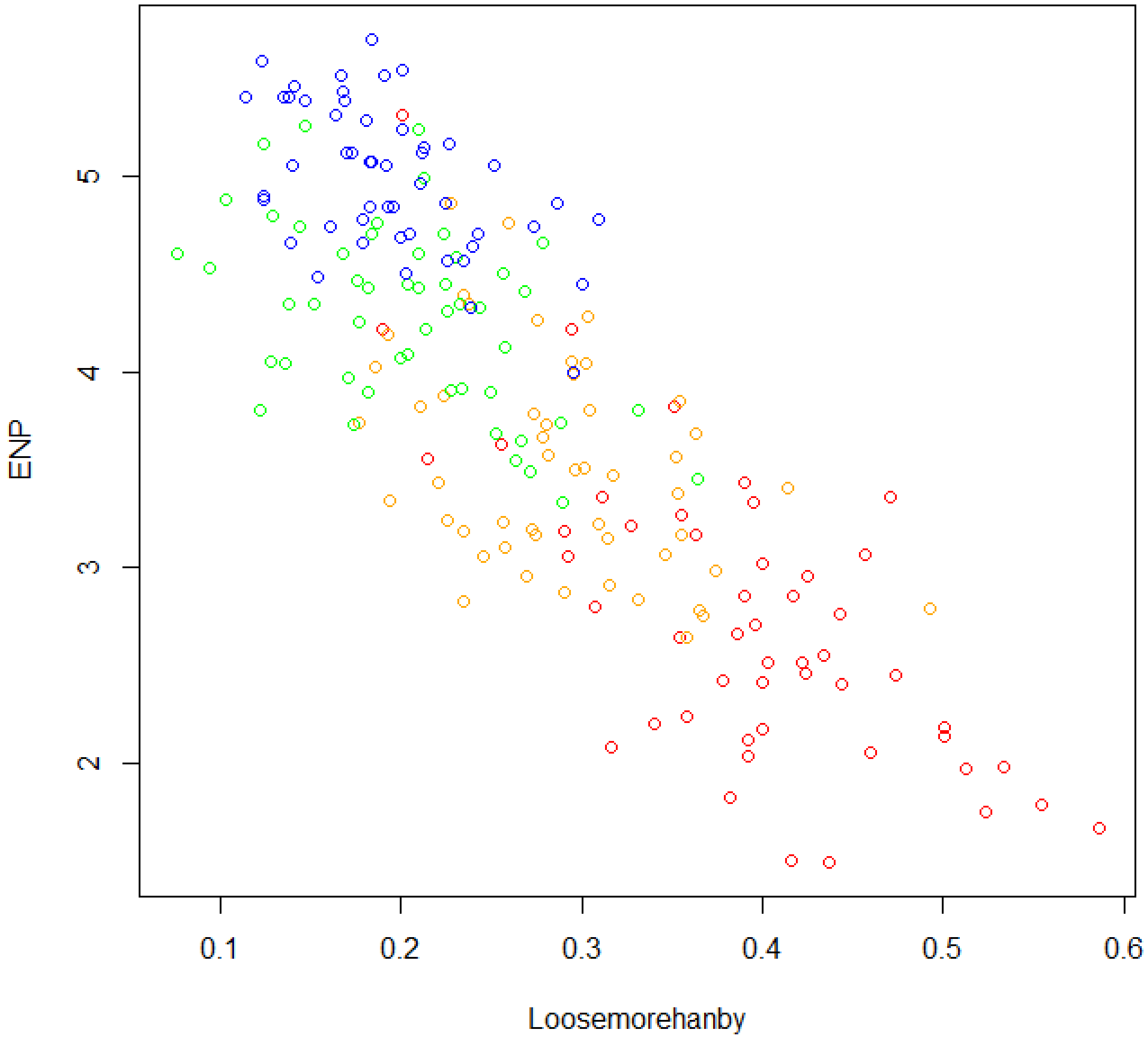


# Polya Eggenberger (Urn model)

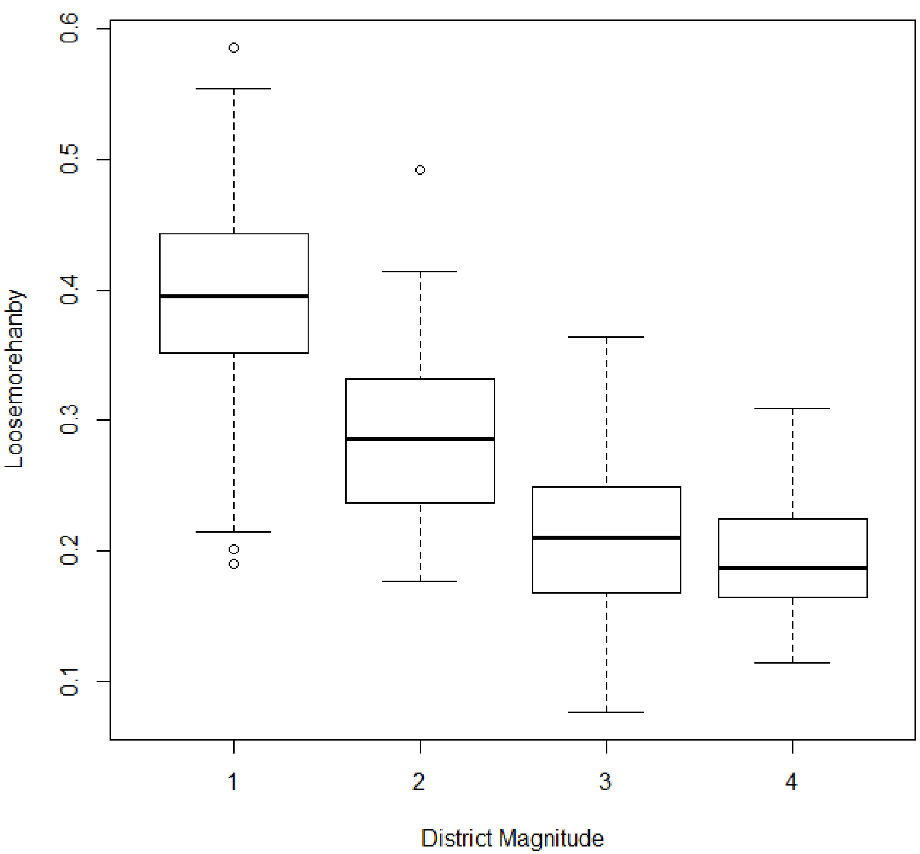


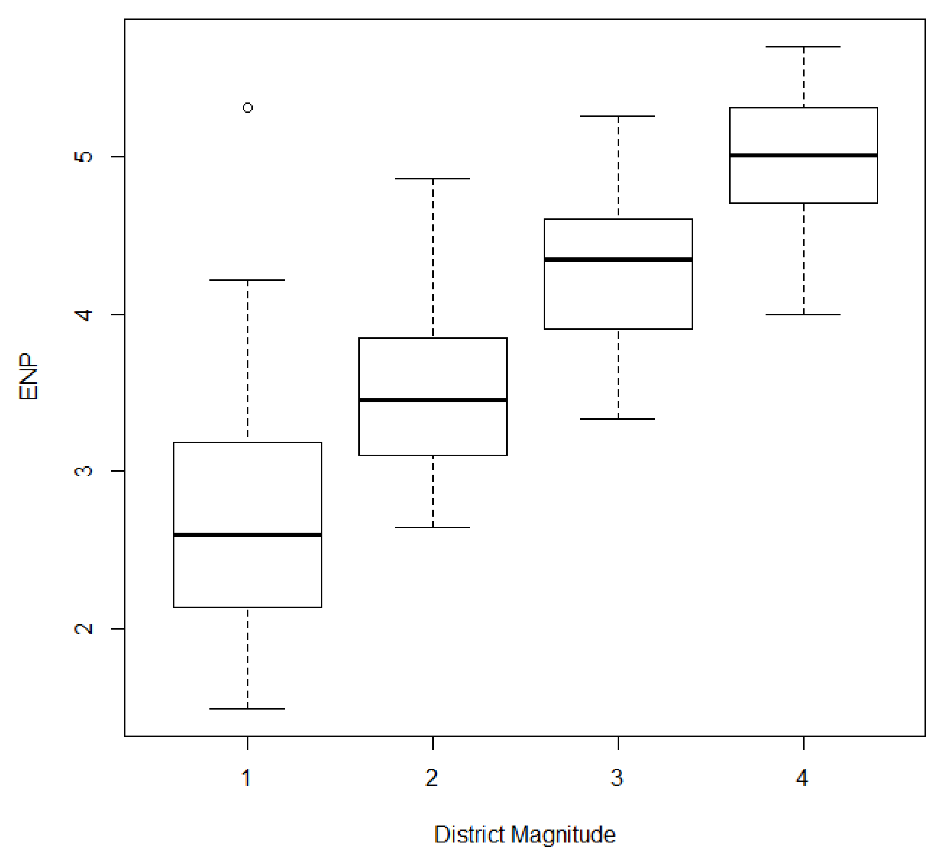



# Spatial Model



# Spatial Model



# References


- John Carey and Simon Hix (2011) 'The Electoral Sweet Spot: Low-Magnitude Proportional Electoral Systems', *American Journal of Political Science 55*(2) 383-339.
- New Zealand Election Study 2008. http://www.nzes.org/