National Electoral Thresholds and Disproportionality

Tasos Kalandrakis Miguel R. Rueda
University of Rochester

April 14 2012

Studying electoral systems at the national level

How easy is it for small parties to gain representation?

How much does the system favor larger parties among those that are represented in the legislature?

Studying electoral systems at the national level

How easy is it for small parties to gain representation?

How much does the system favor larger parties among those that are represented in the legislature?

- We statistically estimate these two quantities from observed electoral returns.

Studying electoral systems at the national level

- Estimated quantities are interesting on their own right: natural focus on thresholds of representation in voting theory.
- Our estimates of national thresholds and disproportionality can be used (after accounting for estimation contamination) as independent variables in the study of variation in party systems.
- They can also be used to properly evaluate the impact of commonly used institutional variables: E.g., district magnitude, upper tiers, mixed systems, etc.

Estimating thresholds at the national level

The literature has provided theoretical formulas for thresholds measures at the district level (e.g., Rokkan 1968, Lijphart and Gibberd 1977).

Attempts to create formulas that apply to the national level are hampered by the (increasing) complexity electoral systems.

Estimating thresholds at the national level

The literature has provided theoretical formulas for thresholds measures at the district level (e.g., Rokkan 1968, Lijphart and Gibberd 1977).

Attempts to create formulas that apply to the national level are hampered by the (increasing) complexity electoral systems.

- Overestimation of thresholds based on district magnitude when applied to countries with large urban districts.
- Overestimation of thresholds based on district magnitude when upper tier provisions qualify lower level allocation.
- Overestimation of thresholds in SMD systems caused by a failure to account for the probability that even small parties may achieve representation in one district.

Estimating disproportionality for parties above threshold

Besides thresholds, our estimates of disproportionality have added advantages:

- We avoid overstating disproportionality for systems that have a positive threshold.
- Our estimates capture the “Political character of disproportionality” (Cox and Shugart 1991).
- We have a summary of the uncertainty that can be placed in these estimates.

A statistical model of seat allocation

- N parties $i = 1, 2, ...N$
- T elections $t = 1, 2, ... T$
- $v_{t,i}$ denotes the vote share of party i in election t
- $s_{t,i}$ denotes the seat share of party i in election t
- In each election there is an unobserved threshold $\theta_t \sim N(\theta, \sigma^2)$

A statistical model of seat allocation

- N parties $i = 1, 2, \ldots N$
- T elections $t = 1, 2, \ldots T$
- $v_{t,i}$ denotes the vote share of party i in election t
- $s_{t,i}$ denotes the seat share of party i in election t
- In each election there is an unobserved threshold $\theta_t \sim N(\theta, \sigma^2)$

Our primary interest is to estimate the expected national threshold θ, with σ a nuisance parameter.

A statistical model of seat allocation

We assume the allocation of seats for parties above realized threshold follows a multinomial distribution

The expected seat share in election t of party i is given by

$$q_i(v_t, \beta, \theta_t) = \begin{cases} \dfrac{v_{t,i}^{\beta}}{\sum\limits_{j: v_{t,j} \geq \theta_t} v_{t,j}^{\beta}} & \textit{if} \quad v_{t,i} \geq \theta_t \\ 0 & \textit{otherwise.} \end{cases}$$

$$p(s_t \mid v_t, \beta, \theta_t) = \textit{Multinomial}\left[q_1(v_t, \beta, \theta_t), ..., q_N(v_t, \beta, \theta_t); \sum_{i=1}^{N} s_{t,i}\right]$$

A statistical model of seat allocation

We assume the allocation of seats for parties above realized threshold follows a multinomial distribution

The expected seat share in election t of party i is given by

$$q_i(v_t, \beta, \theta_t) = \begin{cases} \dfrac{v_{t,i}^{\beta}}{\sum\limits_{j: v_{t,j} \geq \theta_t} v_{t,j}^{\beta}} & \text{if} \quad v_{t,i} \geq \theta_t \\ 0 & \text{otherwise.} \end{cases}$$

$$p(s_t \mid v_t, \beta, \theta_t) = \mathit{Multinomial}\left[q_1(v_t, \beta, \theta_t), \ldots, q_N(v_t, \beta, \theta_t); \sum_{i=1}^{N} s_{t,i}\right]$$

Estimation of disproportionately parameter for parties above the threshold β also of primary interest

Writing a likelihood for the model

The data gives us information that can be used to refine the location of θ_t

Writing a likelihood for the model

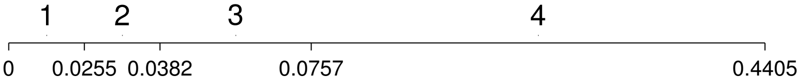
The data gives us information that can be used to refine the location of θ_t

$S_{t,i}$	$v_{t,i}$
295	48.7781
321	47.9714
6	2.5547
3	0.4405
0	0.0757
0	0.0382
0	0.0255

Writing a likelihood for the model

The data gives us information that can be used to refine the location of θ_t

$S_{t,i}$	$v_{t,i}$
295	48.7781
321	47.9714
6	2.5547
3	0.4405
0	0.0757
0	0.0382
0	0.0255



Writing a likelihood for the model

We can see z_t as the realization of a random variable Z_t (with support in $(\{1, \ldots, n_t\})$ that indicates in which interval is θ_t located,

The log-likelihood is then

$$L(\theta, \sigma, \beta \mid v_t, s_t) = \sum_{t=1}^{T} \log \left(\sum_{z_t=1}^{n_t} p(z_t \mid v_t, \theta, \sigma) p(s_t \mid v_t, \beta, u_{t,z_t}) \right)$$

Writing a likelihood for the model

We can see z_t as the realization of a random variable Z_t (with support in $(\{1, \ldots, n_t\})$ that indicates in which interval is θ_t located,

The log-likelihood is then

$$L(\theta, \sigma, \beta \mid v_t, s_t) = \sum_{t=1}^{T} \log \left(\sum_{z_t=1}^{n_t} p(z_t \mid v_t, \theta, \sigma) p(s_t \mid v_t, \beta, u_{t,z_t}) \right)$$

The likelihood is not amenable to regular ML estimation

The MAP-EM estimator

We first write the ‘complete data likelihood’ a likelihood that is conditional on the observed data $X = \{(v_t, s_t)\}_{t=1}^{T}$ and all the unobserved data $Z = \{z_t\}_{t=1}^{T}$ and $\Theta = \{\theta_t\}_{t=1}^{T}$

$$L(\theta, \sigma, \beta \mid X, Z, \Theta)$$

Then we iteratively apply the following steps:

1. Expectation:

$$Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) = \mathbb{E}_{Z,\Theta}\left[L(\theta, \sigma, \beta \mid X, Z, \Theta) \mid X, \theta_m, \sigma_m, \beta_m\right]$$

2. Maximization:

$$(\theta_{m+1}, \sigma_{m+1}, \beta_{m+1}) = \arg\max_{(\theta,\sigma,\beta)} \{Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m)\}$$

The MAP-EM estimator

We first write the 'complete data likelihood' a likelihood that is conditional on the observed data $X = \{(v_t, s_t)\}_{t=1}^{T}$ and all the unobserved data $Z = \{z_t\}_{t=1}^{T}$ and $\Theta = \{\theta_t\}_{t=1}^{T}$

$$L(\theta, \sigma, \beta \mid X, Z, \Theta)$$

Then we iteratively apply the following steps:

1. Expectation:

$$Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) = \mathbb{E}_{Z,\Theta}\left[L(\theta, \sigma, \beta \mid X, Z, \Theta) \mid X, \theta_m, \sigma_m, \beta_m\right]$$

2. Maximization:

$$(\theta_{m+1}, \sigma_{m+1}, \beta_{m+1}) = \arg\max_{(\theta,\sigma,\beta)} \{Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) + p(\theta, \sigma, \beta)\}$$

Data

- We implement the MAP-EM estimator on results data of 415 elections held after WWII from 30 European countries
- We identify a total of 101 electoral systems or partitions within a systems

General patterns

	$\widehat{\beta} > 1$	$\widehat{\beta} = 1$	Total
$\widehat{\theta} > 0$	27 (26.7)	29 (28.7)	56 (55.4)
$\widehat{\theta} = 0$	29 (28.7)	16 (15.8)	45 (44.6)
Total	56 (55.4)	45 (44.6)	101 (100)

General patterns

	$\widehat{\beta} > 1$	$\widehat{\beta} = 1$	Total
$\widehat{\theta} > 0$	27 (26.7)	29 (28.7)	56 (55.4)
$\widehat{\theta} = 0$	29 (28.7)	16 (15.8)	45 (44.6)
Total	56 (55.4)	45 (44.6)	101 (100)

- About one quarter of the systems exhibit positive thresholds and disproportional allocations for parties above the threshold
- There are no cases of significant reverse disproportionality

Observations on the estimated threshold $\widehat{\theta}$

Country	System	$\widehat{\theta}$	Legal Threshold
Netherlands	1946-1952	0.92	1
		(0.52)	
	1956-2010	0.63	0.67
		(0.13)	
Greece	1993-2004	3.18	3
		(0.33)	
	2007-2009	3.14	3
		(0.68)	

The estimator recovers the national legal thresholds when these are instituted without qualifying provisions

Observations on the estimated threshold $\widehat{\theta}$

Country	System	$\widehat{\theta}$	Legal Threshold
Germany	1957-1983	3.88 (0.53)	5
	1994-2009	3.03 (0.78)	5
Croatia	1992 PR	0.05 (0.58)	3

It also captures cases where there are provisions that allow small parties to gain representation below the nominal legal threshold

Observations on the estimated disproportionality $\widehat{\beta}$

Country	System	$\widehat{\beta}$
Croatia	1990	3.32
		(0.69)
	1992 SMD	4.29
		(0.89)
Germany	1987	1.00
		(0.07)
Slovakia	1992-1994	1.01
		(0.08)

- Out of 12 systems/partitions that have majoritarian or plurality formulas there are only two that are consistent with the "Cube law"
- Systems for which $\widehat{\beta}$ is closest to 1 have some form of PR with compensatory upper tiers or PR allocation in nationwide districts

Conclusions

1. We develop a method to empirically estimate national representation thresholds and disproportionality.
2. We aim to incorporate institutional variables (district size, number of districts, allocation rules, tier structure) in the estimator to discern their effect on $\widehat{\theta}$ and $\widehat{\beta}$
3. Estimates can be used for the study of electoral systems and party systems

Thanks!

The EM estimator

We start by writing the 'complete data likelihood' a likelihood that is conditional on the observed data $X = \{(v_t, s_t)\}_{t=1}^{T}$ and all the unobserved data $Z = \{z_t\}_{t=1}^{T}$ and $\Theta = \{\theta_t\}_{t=1}^{T}$

$$L(\theta, \sigma, \beta \mid X, Z, \Theta) = \sum_{t=1}^{T} \log\left(p(\theta_t \mid v_t, \theta, \sigma, z_t) p(z_t \mid v_t, \theta, \sigma) p(s_t \mid v_t, \beta, u_{t,z_t})\right)$$

where $p(\theta_t \mid v_t, \theta, \sigma, z_t)$ is the density of θ_t (a truncated normal in $(\ell_{t,z_t}, u_{t,z_t}])$

$$p(\theta_t \mid v_t, \theta, \sigma, z_t) = \frac{\mathbb{I}_{(\ell_{t,z_t}, u_{t,z_t}]}(\theta_t) f(\theta_t \mid \theta, \sigma)}{p(z_t \mid v_t, \theta, \sigma)}$$

E-Step

To calculate the expected value of the complete data likelihood we need the conditional unobserved data distributions

$$\mathbb{P}(z_t \mid X, \theta_m, \sigma_m, \beta_m) = \frac{p(z_t \mid v_t, \theta_m, \sigma_m) p(s_t \mid v_t, \beta_m, u_{t,z_t})}{\sum_{z_t'=1}^{n_t} p(z_t' \mid v_t, \theta_m, \sigma_m) p(s_t \mid v_t, \beta_m, u_{t,z_t'})},$$

$$\mathbb{P}(\theta_t \mid X, Z, \theta_m, \sigma_m, \beta_m) = p(\theta_t \mid v_t, \theta_m, \sigma_m, z_t).$$

The prior $p(\theta, \sigma, \beta)$

Following Fraley and Raftery (2007) we assume

$$\begin{aligned} \sigma^2 &\sim \mathit{InverseGamma}\left(\frac{\nu}{2}, \frac{s^2}{2}\right) \\ \theta &\sim N\left(0, \frac{\sigma^2}{\kappa}\right) \end{aligned}$$

We choose flat priors with $\kappa = 10^{-5}$ and $\nu = s^2 = 2$

The prior $p(\theta, \sigma, \beta)$

Following Fraley and Raftery (2007) we assume

$$\begin{aligned}\sigma^2 &\sim \mathit{InverseGamma}\left(\frac{\nu}{2}, \frac{s^2}{2}\right)\\ \theta &\sim N\left(0, \frac{\sigma^2}{\kappa}\right)\end{aligned}$$

We choose flat priors with $\kappa = 10^{-5}$ and $\nu = s^2 = 2$

- Introduction of these priors does not have an appreciable effect on cases where the near zero σ problem does not arise
- We avoid numerical instability in the rest of cases

Standard errors

Standard errors are calculated in the usual way using the Hessian of the original likelihood

We can approximate the Hessian of the original likelihood using the following expression

$$H(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) = \ddot{Q}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}; \widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) \left(I_3 - \dot{M}(\widehat{\theta}, \widehat{\sigma}, \widehat{\beta}) \right),$$

where

$$M(\theta_m, \sigma_m, \beta_m) = \arg\max_{\theta, \sigma, \beta} \{ Q(\theta, \sigma, \beta; \theta_m, \sigma_m, \beta_m) + p(\theta, \sigma, \beta) \}.$$