Online Appendix to

Rethinking the D'Hondt Method

https://doi.org/10.1080/2474736X.2019.1625712
Replication data available at: https://doi.org/10.7910/DVN/ESLT8V

Juraj Medzihorsky

A The British 1999–2014 European Parliament Elections Dataset

The data were compiled from different sources, listed in Table A.1. Source #1 does not report the information on the numbers of cast and rejected ballots. This author was not able to find this information in other sources either. Source #2 does not report some of the last digits in some of the vote counts. These blanks were filled with zeros, which delivered the desired totals. In source #3, the votes of the Scottish Green Party from the Scotland constituency were erroneously reported as the votes of the Green Party of England and Wales. However, unlike in the case of major parties, which have a Scottish branch, the Scottish Green Party is a separate party from the Green Party of England and Wales. This was corrected accordingly.

Table A.1: Sources used in compiling the dataset of British European Parliament Elections 1999–2014

#	Years	Link
1	1999	http://web.archive.org/web/20110615093844/ http://www.europarl.org.uk/section/1999/ 1999-election-results
2	2004, 2009	http://www.europarl.org.uk/en/your-meps/ european_elections/previous_election_results/ electionresults2009/results_of_2009.html
3	2014	http://www.electoralcommission.org.uk/ __data/assets/excel_doc/0003/174828/ EPE-2014-Electoral-data.xlsx

B Supplementary Tables and Figures

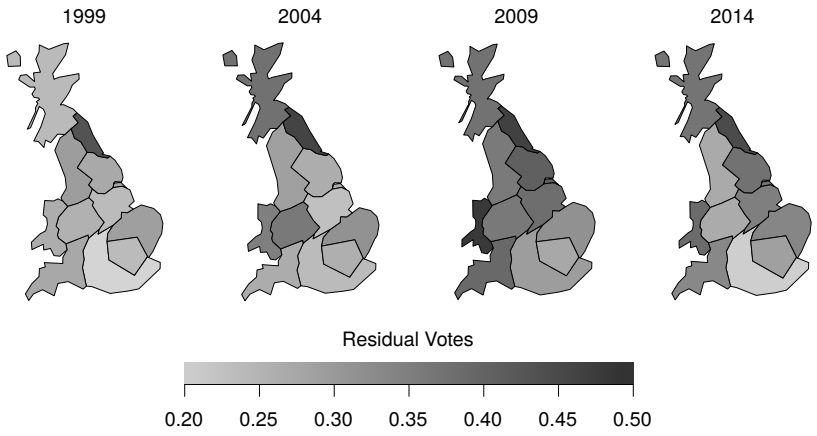

Figure B.1: Residual vote fractions by year and constituency. Constituency area proportional to valid votes. Cartograms with Eurostat (2017) and Jeworutzki (2016).

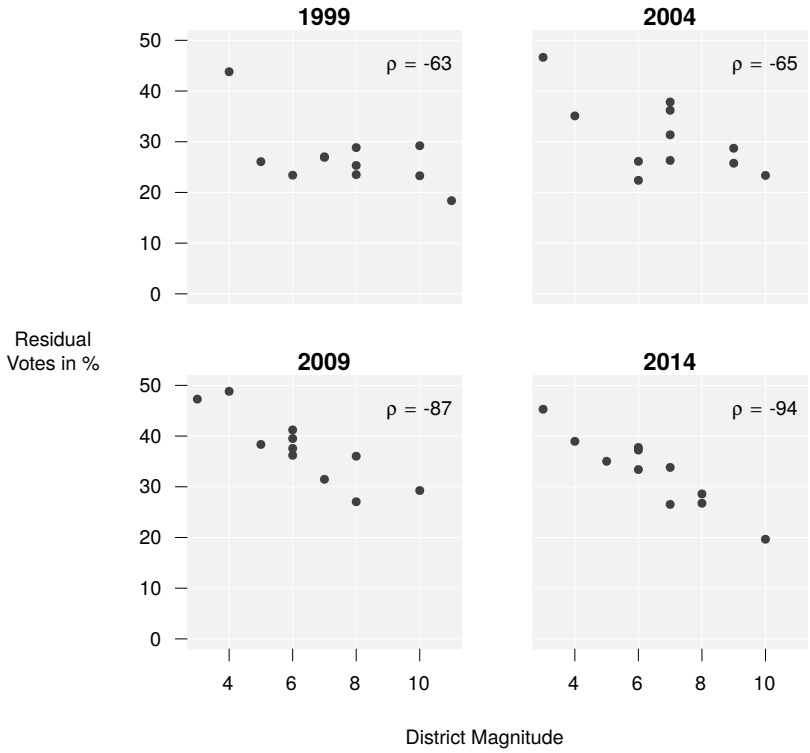

Figure B.2: Residual vote fractions and constituency magnitude by year.

Table B.2: Local residual vote fractions in rounded per cents by party, year, and region.

Year	Region	Con	Lab	LD	SNP	UKIP	BNP	Green	PC
1999	East Midlands	3	11	0	0	100	100	100	0
	East of England	17	29	26	0	0	100	100	0
	London	6	12	34	0	100	100	0	0
	North East	49	0	100	0	100	100	100	0
	North West	0	18	40	0	100	100	100	0
	Scotland	3	0	3	30	100	100	0	0
	South East	16	24	3	0	23	100	0	0
	South West	0	42	37	0	2	100	100	0
	Wales	35	7	100	0	100	0	100	0
	West Midlands	1	0	17	0	100	100	100	0
	Yorkshire and the Humber	15	0	28	0	100	100	100	0
2004	East Midlands	2	38	0	0	1	100	100	0
	East of England	5	40	30	0	0	100	100	0
	London	8	0	46	0	33	100	2	0
	North East	4	48	0	0	100	100	100	0
	North West	1	13	0	0	35	100	100	0
	Scotland	0	33	32	10	100	100	0	0
	South East	13	44	0	0	22	100	2	0
	South West	0	27	43	0	7	100	100	0
	Wales	16	0	100	0	100	100	100	7
	West Midlands	0	22	34	0	48	100	100	0
	Yorkshire and the Humber	0	6	21	0	15	100	100	0
2009	East Midlands	18	27	0	0	25	100	100	0
	East of England	6	6	29	0	0	100	100	0
	London	0	14	34	0	15	100	16	0
	North East	11	30	0	0	100	100	100	0
	North West	6	22	44	0	50	0	100	0
	Scotland	38	0	10	28	100	100	0	0
	South East	19	14	0	0	25	100	39	0
	South West	0	100	41	0	9	100	100	0
	Wales	40	37	100	0	0	100	100	31
	West Midlands	24	37	12	0	0	100	100	0
	Yorkshire and the Humber	20	48	26	0	44	0	100	0
2014	East Midlands	0	48	100	0	21	100	100	0
	East of England	0	45	100	0	18	100	100	0
	London	21	3	100	0	47	100	0	0
	North East	100	0	100	0	38	100	100	0
	North West	9	19	100	0	0	100	100	0
	Scotland	39	19	100	28	0	100	0	0
	South East	22	45	0	0	0	100	11	0
	South West	23	19	100	0	31	100	0	0
	Wales	12	46	100	0	45	100	100	0
	West Midlands	14	21	100	0	0	100	100	0
	Yorkshire and the Humber	46	29	100	0	0	100	100	0

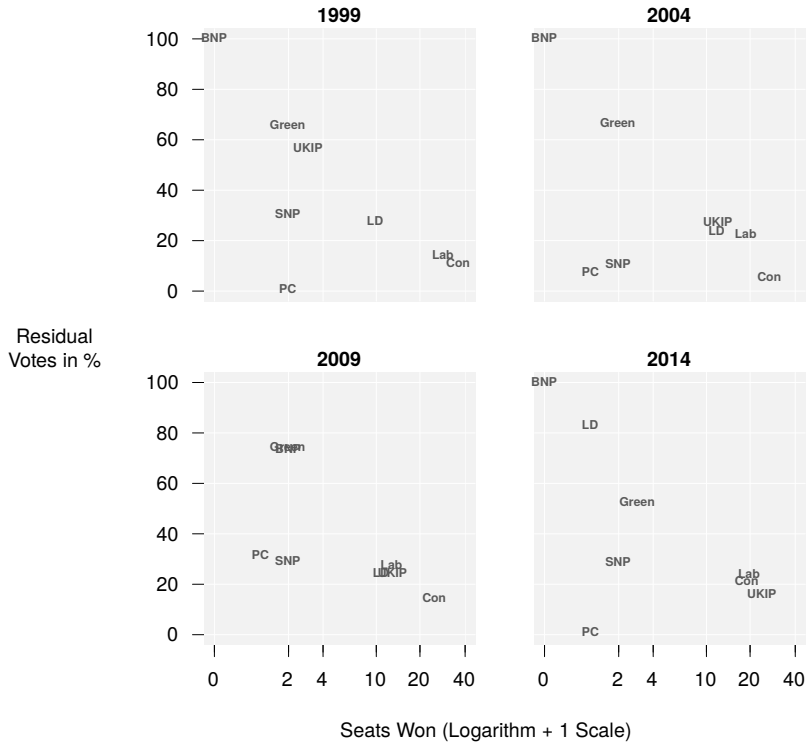

Figure B.3: Logarithm of seats won plus one and residual vote fractions by party and year. The x-axis rescales s_p, the number of seats of party p, with $\ln(s_p + 1)$

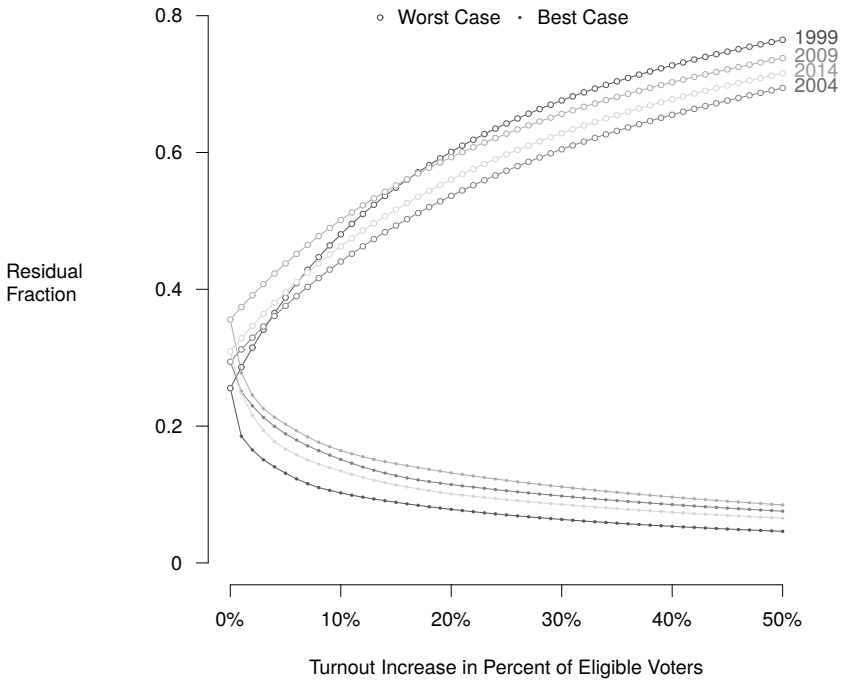

Figure B.4: Best and worst case scenarios for the aggregate fractions of unrepresented votes in British EP elections.

Table B.3: Comparison of the mixture index $\pi^*(\mathbf{v}, \mathbf{s})$ with Monroe's (1994) O_v on all possible apportionments of 5 seats to three parties with 60, 28, and 12 votes. Minima for both indexes in bold type. Example adapted from Gallagher (1991).

A	B	C	Indexes	
60	28	12	π^*	O_v
5	0	0	0.4	0.0516
4	1	0	**0.25**	0.0258
3	2	0	0.3	**0.0227**
2	3	0	0.5333	0.0605
1	4	0	0.65	0.0983
0	5	0	0.72	0.1361
4	0	1	0.4	0.0346
3	1	1	0.4	0.0231
2	2	1	0.4	0.0324
1	3	1	0.5333	0.0647
0	4	1	0.65	0.1009
3	0	2	0.7	0.0808
2	1	2	0.7	0.0808
1	2	2	0.7	0.084
0	3	2	0.7	0.1009
2	0	3	0.8	0.1386
1	1	3	0.8	0.1386
0	2	3	0.8	0.1404
1	0	4	0.85	0.1963
0	1	4	0.85	0.1963
0	0	5	0.88	0.254

Table B.4: Gallagher's 1991 example of the D'Hondt method distributing five seats to three parties. Votes-to-divisor ratios rewarded with mandates in bold type.

	Divisor					
Party	1	2	3	4	5	Seats
A	**60**	**30**	**20**	**15**	12	*4*
B	**28**	14	$9.\bar{3}$	7	5.6	*1*
C	12	6	4	3	2.4	*0*

C R Package `seatdist`

This paper is accompanied by `seatdist`, an R package which implements a variety of apportionment algorithms and disproportionality measures. The `seatdist` package is developed from the `SciencesPo` (Marcelino 2016) package, but offers more apportionment algorithms and disproportionality measures and corrects some errors.

C.1 Seat Apportionment Methods in `seatdist`

Apportionment algorithms are accessible through a unified interface provided by `seatdist::giveseats()` which takes the following arguments

`v`	a numeric vector of vote counts;
`ns`	numeric, the number of seats to allocate;
`method`	character, name of the method, see Table C.5 for divisor methods and Table C.6 for largest remainder quotas;
`thresh`	numeric, threshold of exclusion; if in [0,1], treated as a fraction; if in (1, 100), treated as a percent; if larger than 100, treated as a vote count;
`quota`	character, quota for `method="largest remainders"`; see Table C.6, defaults to `NA`.

For the Largest Remainders method (`method="lr"` or `"largest remainders"`) the Imperiali quota (`quota="im"`) or Reinforced Imperiali (`quota="rei"`) can assign in the first round more seats than available, in which case the function terminates its execution with an error message.

The `seatdist::giveseats()` function returns a named list with items

`method`	character, name of the apportionment method used;
`seats`	numeric, vector with seats.

For illustration, 10 seats can be apportioned to parties with 60,000, 28,000, and 12,000 votes under a system with a 5% threshold with the Largest Remainders method and the Hagenbach-Bischoff quota in the following way

```
> seatdist::giveseats(v=c(A=60, B=28, C=12)*1e3, ns=1e1,
                      method="lr", quota="hb", thresh=5e-2)

thresh treated as a fraction
$method
"Largest Remainders with Hagenbach-Bischoff quota"

$seats
A B C
6 3 1
```

Table C.5: Divisor method implemented in `seatdist::giveseats()`. For background on the methods see e.g. Grilli di Cortona et al. (1999).

Method	`method`	Formula	Sequence
D'Hondt	`"dh"`	x	$1, 2, 3, 4, 5, \ldots$
Jefferson	`"je"`	”	”
Hagenbach-Bischoff	`"hb"`	”	”
Adams	`"ad"`	$x - 1$	$0, 1, 2, 3, 4, \ldots$
Smallest Divisors	`"sd"`	”	”
Nohlen	`"no"`	$x + 1$	$2, 3, 4, 5, 6, \ldots$
Imperiali	`"im"`	$(x + 1)/2$	$1, 1.5, 2, 2.5, 3, 3.5, \ldots$
Sainte-Laguë	`"sl"`	$2x - 1$	$1, 3, 5, 7, 9, \ldots$
Webster	`"we"`	”	”
Hungarian Sainte-Laguë	`"hu"`	$2x - 1;\ x > 1$	$1.5, 3, 5, 7, 9, \ldots$
Modified Sainte-Laguë	`"msl"`	$(2x - 1)5/7;\ x > 1$	$1, 2.14, 3.57, 5, 6.43, \ldots$
Danish	`"da"`	$3x - 2$	$1, 4, 7, 10, 13, \ldots$
Huntington-Hill	`"hh"`	$\sqrt{x(x-1)}$	$0, 1.41, 2.45, 3.46, 4.47 \ldots$
Equal Proportions	`"ep"`	”	”

Table C.6: Quotas implemented for the Largest Remainders method (`method="lr"`) in `seatdist::giveseats()`. For background on the methods see e.g. Grilli di Cortona et al. (1999).

Quota	`quota`	Formula
Hare	`"ha"`	$\frac{e}{l}$
Droop	`"dr"`	$\left\lfloor 1 + \frac{e}{l+1} \right\rfloor$
Hagenbach-Bischoff	`"hb"`	$\frac{e}{l+1}$
Imperiali	`"im"`	$\frac{e}{l+2}$
Reinforced Imperiali	`"rei"`	$\frac{e}{l+3}$

C.2 Measures of Disproportionality in `seatdist`

The `seatdist` package computes 24 disproportionality measures (Table C.7) accessible through a unified interface provided by the function `seatdist::disproportionality()`.

The function takes the following arguments:

`s`	a numeric vector of seat counts or fractions;
`v`	a numeric vector of vote counts or fractions; for `measure = "ortona"` this can alternatively be a vector with seats under the highest possible proportionality
`measure`	character, see Table C.7;
`ignore_zeros`	logical: should parties with 0 votes and 0 seats be ignored?
`k`	numeric, k for the Generalized Gallagher index, defaults to 2;
`eta`	η for the Atkinson index, defaults to 2;
`alpha`	α for the Generalized Entropy index, defaults to 2;
`thresh`	numeric, threshold for the Fragnelli and the Gambarelli & Biella indexes, defaults to `"NULL"`;
`powind`	character, power index for the Fragnelli and the Gambarelli & Biella indexes, defaults to the Shapley-Shubik index, `"shapley shubik"`. no other power indexes implemented yet.

The function returns a named list with the following items

`measure`	character, the measure used;
`distance`	numeric, distance from proportionality.

For illustration, the Gallagher index can be computed for parties with 60,000, 28,000, and 12,000 votes and 6, 3, and 1 seats in the following way:

```
> seatdist::disproportionality(v=c(60,28,12)*1e3,
                               s=c(6,3,1),
                               measure="gallagher")
$measure
[1] "Gallagher"

$distance
[1] 0.02
```

Table C.7: Disproportionality measures in the `seatdist` package. Values for the `measure=` argument in `seatdist::disproportionality()` below index names. For indexes without citations in the table see also Karpov 2008 and Chessa and Fragnelli 2012.

Index	Formula
D'Hondt (Gallagher 1991) `"dhondt"`	$\delta = \max_i \frac{s_i}{v_i}$
Monroe (1994) `"monroe"`	$I_M = \sqrt{\frac{\sum_i (s_i - v_i)^2}{1 + \sum_i v_i^2}}$
Max. Abs. Dev. `"maxdev"`	$I_{MAD} = \max_i \{\lvert s_i - v_i \rvert\}$
Rae (1967) `"rae"`	$I_{Rae} = \frac{1}{p} \sum_i \lvert s_i - v_i \rvert$
Loosemore & Hanby (1971) `"loosemore hanby"`	$I_{LH} = \frac{1}{2} \sum_i \lvert s_i - v_i \rvert$
Grofman `"grofman"`	$I_{Grof} = \frac{1}{e} \sum_i \lvert s_i - v_i \rvert;\ e = \frac{1}{\sum_i v_i^2}$
Lijphart `"lijphart"`	$I_L = \frac{\lvert s_a - v_a \rvert + \lvert s_b - v_b \rvert}{2};\ v_a > v_b > \ldots$
Gallagher (1991) `"gallagher"`	$I_{Gal} = \sqrt{\frac{1}{2} \sum_i (s_i - v_i)^2}$

Table C.8: Table C.7 continued from the previous page

Index	Formula
Generalized Gallagher `"kindex"`	$I_K = \sqrt[k]{\frac{1}{k}\sum_i (s_i - v_i)^k}$
Gatev `"gatev"`	$I_{Gat} = \sqrt{\frac{\sum_i (s_i - v_i)^2}{\sum_i (s_i^2 + v_i^2)}}$
Ryabtsev `"ryabtsev"`	$I_{Ryb} = \sqrt{\frac{\sum_i (s_i - v_i)^2}{\sum_i (s_i + v_i)^2}}$
Szalai (Stewart 2006) `"szalai"`	$I_{Sz} = \sqrt{\frac{1}{p}\sum_i \left(\frac{s_i - v_i}{s_i + v_i}\right)^2}$
Weighted Szalai (Stewart 2006) `"weighted szalai"`	$I_{WSz} = \sqrt{\frac{1}{2}\sum_i \frac{(s_i - v_i)^2}{s_i + v_i}}$
Aleskerov & Platonov `"aleskerov"`	$I_{AP} = \frac{\sum_i k_i \frac{s_i}{v_i}}{\sum_i k_i};\ k_i = \mathbb{1}\left(\frac{s_i}{v_i} > 1\right)$
Gini `"gini"`	The Gini coefficient of inequality
Atkinson `"atkinson"`	$I_A = 1 - \left[\sum_i v_i \left(\frac{s_i}{v_i}\right)^{(1-\eta)}\right]^{\frac{1}{1-\eta}}$

Table C.9: Table C.7 continued from the previous page

Index	Formula
Generalized Entropy `"gen entropy"`	$I_{GE} = \frac{1}{\alpha^2 - \alpha}\left[\sum_i v_i \left(\frac{s_i}{v_i}\right)^{\alpha} - 1\right]$
Sainte-Laguë (1910) `"sainte lague"`	$I_{SL} = \sum_i \frac{(s_i - v_i)^2}{v_i}$
Cox & Shugart `"cox shugart"`	$I_{CS} = \frac{\sum_i (s_i - \bar{s})(v_i - \bar{v})}{\sum_i (v_i - \bar{v})^2}$
Farina (Kestelman 2005) `"farina"`	$I_{Far} = \arccos\left[\frac{\sum_i s_i v_i}{\sqrt{\sum_i s_i^2 \sum_i v_i^2}}\right]\frac{10}{9}$
Ortona `"ortona"`	$I_O = \frac{\sum_i \lvert s_i - v_i\rvert}{\sum_i \lvert u_i - v_i\rvert};\ u_i = \mathbb{1}(s_i = \max_i s_i)$
Fragnelli `"fragnelli"`	$I_{Frag} = \frac{1}{2}\sum_i \lvert\varphi_i(s) - \varphi_i(v)\rvert;$
Gambarelli & Biella `"gambarelli biella"`	$I_{GB} = \max_i\{\lvert s_i - v_i\rvert,\ \lvert\varphi_i(s) - \varphi_i(v)\rvert\}$
Cosine Dissimilarity `"cosine"`	$I_{CD} = 1 - \frac{\sum_i s_i v_i}{\sqrt{\sum_i s_i^2}\sqrt{\sum_i v_i^2}}$
Mixture D'Hondt `"mixture"`	$\pi^*_{DH} = 1 - \frac{1}{\max_i s_i / v_i}$

References

Chessa, Michela, and Vito Fragnelli. 2012. "A note on 'Measurement of disproportionality in proportional representation systems'". *Mathematical and Computer Modelling* 55 (3): 1655–1660.

Eurostat. 2017. "Nomenclature of territorial units for statistics".

Gallagher, Michael. 1991. "Proportionality, disproportionality and electoral systems". *Electoral Studies* 10 (1): 33–51.

Grilli di Cortona, Pietro, et al. 1999. *Evaluation and Optimization of Electoral Systems*. SIAM.

Jeworutzki, Sebastian. 2016. *cartogram: Create Cartograms with R*. R package version 0.0.2. `https://CRAN.R-project.org/package=cartogram`.

Karpov, Alexander. 2008. "Measurement of disproportionality in proportional representation systems". *Mathematical and Computer Modelling* 48 (9): 1421–1438.

Kestelman, Philip. 2005. "Apportionment and proportionality: A measured view". *Voting Matters* 20:12–22.

Loosemore, John, and Victor J Hanby. 1971. "The theoretical limits of maximum distortion: some analytic expressions for electoral systems". *British Journal of Political Science* 1 (4): 467–477.

Marcelino, Daniel. 2016. *SciencesPo: A tool set for analyzing political behavior data*. R package version 1.4.1. `http://CRAN.R-project.org/package=SciencesPo`.

Monroe, Burt L. 1994. "Disproportionality and malapportionment: Measuring electoral inequity". *Electoral Studies* 13 (2): 132–149.

Rae, Douglas W. 1967. *The Political Consequences of Electoral Laws*. New Haven: Yale University Press.

Sainte-Laguë, André. 1910. "La représentation proportionnelle et la méthode des moindres carrés". In *Annales scientifiques de l'École Normale Supérieure*, 27:529–542.

Stewart, Jay, et al. 2006. "Assessing alternative dissimilarity indexes for comparing activity profiles". *Electronic International Journal of Time Use Research* 3 (1): 49–59.