Online appendices for "ecolRxC: Ecological inference estimation of R×C tables using latent structure approaches"*, Political Science Research and Methods.*

Jose M. Pavía and Søren Risbjerg Thomsen

Appendix I: The `ecolRxC` function

In addition to print, summary and plot methods, the `ecolRxC` package features a main function with the same name: `ecolRxC`. The description of its most important parameters are listed below:

- `votes.election1`: A data frame (or matrix) of order R×U with the votes gained by the parties competing on election 1.
- `votes.election2`: A data frame (or matrix) of order C×U with the votes gained by the parties competing on election 2.
- `scale`: A character string indicating the type of transformation to be applied to the observed vote fractions. Only `'logit'` and `'probit'` are allowed. Default, `'probit'`.
- `method`: A character string indicating the algorithm to be used for making congruent with the observed margins the initial crude probabilities. Only `'Thomsen'` and `'IPF'` (iterative proportional fitting) are allowed. Default, `'Thomsen'`.
- `census.changes`: A character string informing about the level of information available in `votes.election1` and `votes.election2` regarding new entries and exits in the election censuses between the two elections or indicating how their sum discrepancies should be handled. This argument admits nine values: `'adjust'`, `'raw'`, `'regular'`, `'ordinary'`, `'simultaneous'`, `'enriched'`, `'semifull'`, `'full'` and `'gold'`. Default, `'adjust'`.
- `reference`: A vector of two components indicating the parties to be used with `method = 'Thomsen'` as reference in election 1 and 2, respectively. When `reference = NULL`, the solution is constructed as a weighted average of all the congruent solutions achieved after considering as reference all the possible combinations. Default `NULL`.
- `confidence`: A number between 0 and 1 to be used as level of confidence for the confidence intervals of the transition rates. If `confidence = NULL`, confidence intervals are not computed. Default `NULL`.
- `Yule.aprox`: A `TRUE/FALSE` argument indicating which equation, either (4) or (3), is to be used to estimate the cross-proportions. Default `FALSE`, equation (3).

The value of the function is a `ecolRxC` class object, which is a list of length 16 that contains the main outputs of the ecological inference process as well as some additional relevant by-product features. The main components of the `ecolRxC` output are described below:

- `VTM`: A matrix of order R×C with the estimated proportions of the row-standardized vote transitions from election 1 to election 2.
- `VTM.votes`: A matrix of order R×C with the estimated vote transfers (counts) from election 1 to election 2.
- `VTM.lower`: A matrix of order R×C with the estimated lower limits of the confidence intervals for the proportions of the row-standardized vote transitions from election 1 to election 2. When `confidence = NULL`, this is a `NULL` object.

- `VTM.upper`: A matrix of order R×C with the estimated upper limits of the confidence intervals for the proportions of the row-standardized vote transitions from election 1 to election 2. When `confidence = NULL`, this is a `NULL` object.
- `VTM.units`: An array of order R×C×U with the estimated proportions of the row-standardized vote transitions from election 1 to election 2 attained for each unit.
- `VTM.votes.units`: An array of order R×C×U with the estimated transfer of votes from election 1 to election 2 attained for each unit.
- `VTM.lower.units`: An array of order R×C×U with the estimated lower limits of the confidence intervals for the proportions of the row-standardized vote transitions from election 1 to election 2 corresponding to each unit. When `confidence = NULL`, this is a `NULL` object.
- `VTM.upper.units`: An array of order R×C×U with the estimated upper limits of the confidence intervals for the proportions of the row-standardized vote transitions from election 1 to election 2 corresponding to each unit. When `confidence = NULL`, this is a `NULL` object.
- `Correlations`: A matrix of order R×C with the crude estimated correlations.
- `reference.outputs`: A list with three components: `vjk.averages`, `vjk.by.reference` and `vjk.units.by.reference`. The first component `vjk.averages` is a R×C×8 array with the eight different global solutions of transfer matrix of votes attained after combining with different weights, as detailed in the previous section, each of the solutions reached considering as reference all the possible pairs of a row and a column. The second component `vjk.by.reference` is a R×C×(RC) array with the RC different global solutions of transfer matrix of votes attained after choosing as reference all the possible combination of a row and a column. The third component `vjk.units.by.reference` is a R×C×U×(RC) array with the local solutions linked to `vjk.by.reference`.

Appendix II: Code of the example described in section 4

```
> # Data preparation
> library(ei.Datasets)
> data <- merge_small_options(ei_NZ_2017,
+                             min.party = 3,
+                             min.candidate = 3)
> district <- which(data$District == "Northland")
> party.margins <- data$Votes_to_parties[[district]]
> candidate.margins <- data$Votes_to_candidates[[district]]
> party.margins <- party.margins[, -c(1, 2)]
> candidate.margins <- candidate.margins[, -c(1, 2)]
>
> # Transfer matrix estimation
> library(ecolRxC)
> set.seed(123)
> transitions <- ecolRxC(party.margins, candidate.margins,
+                        confidence = 0.95, B = 100)
>
> # Transfer matrix plot
> plot(transitions, size.labels = 3.5,
+      size.numbers = 8, size.margins = 5)
```

Appendix III: Example of estimated confidence intervals

In this appendix, interested readers can find the estimated confidence intervals obtained for the row-transfer estimates corresponding to the example developed in Section 4 of the paper. The data have been obtained after applying `ecolRxC` (version 0.1.1-10) with arguments *`confidence = 0.95`* and *`B = 100`*. The values in the table are presented in percentages.

Confidence intervals ($1 - \alpha = 0.95$) for the vote transfer estimates in Figure 1 (section 4).

	Candidate				
Party	HUGHES	PRIME	PETERS	KING	Others
Green Party	19.1 – 21.6	46.6 – 49.7	14.2 – 17.7	13.3 – 15.8	0.8 – 1.3
Labour Party	7.1 – 7.5	51.8 – 52.7	32.9 – 33.6	5.6 – 6.1	1.8 – 2.1
National Party	0.8 – 1.0	2.3 – 2.8	23.7 – 24.5	71.3 – 72.0	0.8 – 1.0
NZ First Party	0.5 – 0.9	3.4 – 5.0	79.7 – 81.4	12.9 – 14.6	1.1 – 1.6
Other parties	11.7 – 13.8	28.4 – 31.4	42.5 – 46.4	5.5 – 7.3	5.8 – 6.9

Source: compiled by the authors after applying `ecolRxC` (version 0.1.1-10) with arguments *`confidence = 0.95`* and *`B = 100`* to the voting data recorded in the electorate of Northland during the 2017 New Zealand general elections.

Appendix IV: Averages of EPW and EQ errors by group of elections

Here, interested readers may find the equivalent to Table 2 in the paper for the case of EPW and EQ errors.

Table 2B. Averages of EPW errors by group of elections.

Country Year	NZ 2002	NZ 2005	SCO 2007	NZ 2008	NZ 2011	NZ 2014	NZ 2017	NZ 2020	NZ+SCO
logit transformations									
VTR-Yule	10.03	8.15	20.88	7.78	8.80	10.10	8.84	9.20	10.53
VTR-local-Yule	9.94	8.04	20.74	7.67	8.69	9.99	8.74	9.10	10.42
ecol	9.28	7.26	17.65	6.84	7.72	8.80	7.50	8.03	9.19
ecol-biN	7.96	6.14	14.69	5.56	6.20	6.85	5.28	5.84	7.34
VTR	7.65	6.14	17.53	5.25	6.02	6.76	5.24	5.86	7.60
VTR-local	7.62	6.06	17.51	5.19	5.98	6.75	5.25	5.87	7.57
ecolRxC-Yule	7.94	6.58	15.82	6.10	6.76	7.90	6.68	7.24	8.17
ecolRxC	6.06	5.28	12.45	4.55	4.88	5.36	3.92	4.60	5.91
probit transformations									
VTR-Yule	9.96	8.03	20.72	7.66	8.65	9.97	8.70	9.02	10.39
VTR-local-Yule	9.87	7.93	20.59	7.56	8.55	9.87	8.60	8.93	10.29
ecol	9.19	7.16	17.50	6.75	7.59	8.76	7.40	7.87	9.07
ecol-biN	7.90	6.09	14.55	5.51	6.13	6.75	5.22	5.74	7.27
VTR	7.60	6.08	17.38	5.19	5.92	6.66	5.14	5.72	7.50
VTR-local	7.58	6.00	17.36	5.14	5.88	6.65	5.16	5.74	7.48
ecolRxC-Yule	7.89	6.50	15.68	6.02	6.65	7.80	6.59	7.10	8.07
ecolRxC	6.03	5.24	12.31	4.53	4.85	5.31	3.88	4.53	5.86

Table 2C. Averages of EQ errors by group of elections.

Country Year	NZ 2002	NZ 2005	SCO 2007	NZ 2008	NZ 2011	NZ 2014	NZ 2017	NZ 2020	NZ+SCO
logit transformations									
VTR-Yule	7.37	6.74	15.33	6.48	6.96	7.64	7.23	7.54	8.20
VTR-local-Yule	7.29	6.66	15.22	6.40	6.87	7.57	7.16	7.47	8.12
ecol	7.41	6.55	13.56	6.20	6.68	7.25	6.60	7.06	7.96
ecol-biN	7.03	6.11	12.11	5.51	5.98	6.24	5.31	5.72	6.77
VTR	6.12	5.63	13.60	4.90	5.26	5.57	4.79	5.22	6.42
VTR-local	6.07	5.56	13.55	4.85	5.21	5.55	4.78	5.21	6.38
ecolRxC-Yule	5.90	5.62	11.48	5.25	5.51	6.11	5.63	6.01	6.46
ecolRxC	5.04	4.96	9.52	4.41	4.48	4.63	3.83	4.36	5.17
probit transformations									
VTR-Yule	7.30	6.65	15.24	6.39	6.86	7.55	7.12	7.40	8.10
VTR-local-Yule	7.24	6.58	15.17	6.32	6.77	7.48	7.06	7.33	8.03
ecol	7.33	6.47	13.48	6.13	6.59	7.16	6.52	6.94	7.61
ecol-biN	6.98	6.06	12.03	5.47	5.93	6.17	5.27	4.32	6.72
VTR	6.09	5.59	13.52	4.86	5.21	5.52	4.74	5.12	6.36
VTR-local	6.04	5.52	13.47	4.81	5.16	5.50	4.74	5.11	6.33
ecolRxC-Yule	5.85	5.56	11.41	5.19	5.44	6.05	5.57	5.91	6.40
ecolRxC	5.04	4.93	9.45	4.39	4.47	4,61	3.82	5.64	5.15

Appendix V: Averages of EPW and EQ errors for the eight composite solutions

Here, interested readers may find the equivalent to Table 3 in the paper for the case of EPW and EQ errors.

Table 3B. Averages of EPW and EQ errors by group of elections for the eight composite solutions.

Country Year	NZ 2002	NZ 2005	SCO 2007	NZ 2008	NZ 2011	NZ 2014	NZ 2017	NZ 2020	NZ+SCO
EPW: discrepancy errors									
Mean	6.22	5.14	13.18	4.40	4.91	5.49	4.02	4.68	6.03
RCNV	6.43	5.69	11.63	4.99	5.40	5.69	4.15	4.56	6.09
SQRCNV	6.31	5.38	12.78	4.65	5.08	5.51	3.98	4.54	6.05
SQRM	6.25	5.44	12.61	4.64	5.08	5.53	4.02	4.61	6.05
AVCR	6.03	5.24	12.31	4.53	4.85	5.31	3.88	4.53	5.86
LRCNV	6.37	5.66	11.52	4.95	5.35	5.65	4.13	4.52	6.04
LSQRCNV	6.29	5.37	12.74	4.66	5.07	5.50	3.97	4.54	6.05
LSQRM	6.23	5.43	12.62	4.62	5.07	5.53	4.02	4.62	6.04
EQ: quadratic errors									
Mean	5.14	4.92	10.09	4.32	4.55	4.73	3.93	4.41	5.28
RCNV	5.41	5.30	9.09	4.69	4.91	4.90	4.01	4.37	5.35
SQRCNV	5.29	5.09	9.88	4.48	4.70	4.77	3.92	4.38	5.33
SQRM	5.20	5.09	9.58	4.45	4.65	4.75	3.91	4.38	5.27
AVCR	5.04	4.93	9.45	4.39	4.47	4.61	3.82	4.32	5.15
LRCNV	5.36	5.26	9.02	4.65	4.87	4.86	3.99	4.34	5.31
LSQRCNV	5.27	5.08	9.85	4.47	4.68	4.76	3.91	4.38	5.32
LSQRM	5.19	5.08	9.60	4.44	4.64	4.75	3.91	4.38	5.26

Source: compiled by the authors after applying the function `ecolRxC` with default options (`method = 'Thomsen'`, `scale = 'probit'`, `Yule.aprox = FALSE`) to the 565 datasets of the R package `ei.Datasets` (Pavía, 2022). The definition and acronyms of the different composite solutions are detailed in Section 3. The smaller the number, the better the accuracy.