The usefulness of ecological inference for the estimation of voter transition rates
Ioannis Andreadis
Department of Political Sciences
Aristotle University Thessaloniki

Abstract

Although in US based political science journals there are many examples of published papers which use ecological inference methods, it seems that for the European political science journals the use of ecological inference is faced with skepticism. This paper indicates that using ecological inference for the estimation of voter transition rates in the European countries can be a promising path. Simulation results presented in this paper indicate that the total fraction of the electorate misclassified by the used ecological inference algorithm ranges from 1.5% up to 13% with a mean value of about 6%. These figures show that ecological inference could and should be used for the estimation of voter transition rates.

Introduction

During the Elections, Public Opinion and Parties 2008 Annual Conference, I had a discussion with another participant who told me that a paper he had submitted to a European political science journal was rejected, because his findings were based on ecological inference methods. On the contrary, in US based political science journals there are many examples of published papers which use ecological inference methods. (King, Rosen, Tanner and Wagner, 2008; Coan and Holman, 2008; Imai and King, 2004).

This paper is an attempt to change some of the perceptions about ecological inference. Most of these perceptions are based on Robinson's famous paper (Robinson, 1950) which serves as the first example of "ecological fallacy". The output of ecological inference methods, as with every other inference method, strongly depends on the way these methods are applied. The frequently quoted: "garbage in – garbage out" applies here too.

In order to show that ecological inference can be used as a method for the estimation of voter transition rates, this paper includes findings from tests on the validity of these estimates. Since the actual voter transition rates are not known, the assessment of the estimates is based on their closeness to "real" values provided by the simulation of electoral procedures. The paper includes the assessment of both general (for the entire country) and regional estimates. The assessment indicates that, when applied carefully, ecological inference can provide useful estimates of voter transition rates.

The remaining sections of this paper are as follows: Firstly there is a section dedicated to the methodology used including a description of the ecological inference algorithm used and the simulation settings, followed by a section presenting the simulation results. After the simulation results there is a section about the use of ecological inference for the estimation of voter transition rates in the 27 Member states of the European Union including available data and suggestions for the selection of the electoral units in each case. The paper concludes with suggestions for further research.
**Methodology**

Ecological inference algorithm - VTR

Andreadis and Chadjipadelis (2008) have presented a comparison of methods for the estimation of voter transition rates in Elections, Public Opinion and Parties 2008 Annual Conference. In their paper they have shown that the average voter transitions estimation error using the Markov Chain Monte Carlo Hierarchical Method, a Multinomial-Dirichlet model, proposed by Rosen et al. ((Rosen, Jiang, King and Tanner, 2001)) is about seven times greater than the corresponding error using VTR algorithm (Andreadis and Chadjipadelis, 2009). Thus in this paper VTR is the only method that is being tested for the validity of the provided estimates.

This algorithm VTR consists of five tasks: i) initial data preparation, ii) the selection of the best pair of parties, iii) ecological inference for 2x2 tables, iv) extraction of the estimated quantities of interest and data preparation for the next cycle and v) final calculations. A short outline of the method is described as follows: Matrices $X$ (results of the first election per electoral unit), $T$ (results of the second election per electoral unit) and vector $N$ (electorate per electoral unit) are used as input to the algorithm. Each cycle begins with the selection of the "best pair" of parties ($i$, $j$) where party $i$ has run for the first election and party $j$ has run for the second election. "Best pair" choice is essentially based on the correlation coefficients. For this pair of parties a 2x2 ecological inference method is applied to estimate the voter transitions from party $i$ to party $j$ in each electoral unit $k$. Next, the remaining electorate is calculated (by subtracting the estimated voter transitions) and the algorithm proceeds to the next cycle. For a detailed presentation see Andreadis and Chadjipadelis (2009).

**National and Local estimates**

Usually the validity of ecological inference voter transition estimates for the entire country can be easily verified by comparing them against exit-poll estimates which are usually available on national level. For instance, Chadjipadelis and Andreadis (2009; 2007; 2004) have provided evidence that their ecological inference estimates of voter transition rates in Cyprus, Greece and France are very close to exit-poll estimates.

On the other hand, it is usually more difficult to check the validity of regional estimates of voter transition rates. Exit-poll samples are usually too small to provide valid regional estimates of voter transition rates. As an example, consider an exit poll sample of 4000 respondents. Let's say that party $i$ which has got 30% in the first election lost 10% of its voters towards another party $j$ in a uniform way across all regions. Suppose that we want to estimate the voter transition rate from party $i$ to party $j$ in region $k$ with population that is 5% of the total electorate. Thus, in our sample we have 1200 voters of party $i$ in the first election. From these voters, 120 have moved to party $j$ in the second election. From this group, 6 people live in region $k$. Small figures like number 6 in our example make it very difficult for exit-poll analysts to announce regional estimates.
Simulation settings

Andreadis and Chadjipadelis (2008) have used simulated voter transition rates between the first and second round of the French Presidential elections that took place in 2007. They have used a matrix with four rows and three columns. The rows correspond to the top four candidates who have received most of the votes in the first round (i.e. SARKOZY, ROYAL, BAYROU and LE PEN). The three columns correspond to the three possible alternatives each voter had in the second round ((i.e. SARKOZY, ROYAL and "not valid + abstention"). They have used the 96 French departments as the electoral units for their analysis.

As an accuracy index of the estimation of the general transition matrix they have used the formula

$$\sum_{i=1}^{4} \sum_{j=1}^{3} |\hat{p}_{ij} - p_{ij}|$$

where $\hat{p}_{ij}$ is the estimated national voter transition rate from party $i$ to party $j$ and $p_{ij}$ is the "real" national transition rate calculated from the simulated values. The absolute deviation calculated by the aforementioned formula represents the total number of voters (as a fraction of the electorate) who were misclassified by the method. In the same paper they have demonstrated that the average absolute deviation of the VTR algorithm national level estimates is 10.89%.

This paper follows a similar approach on an improved version of VTR algorithm. The same 4x3 matrix is used as the template for the simulation and the same absolute deviation index is used to check the accuracy of the estimates. But in addition to presenting results from an improved version of VTR, this paper presents the accuracy index of both national and regional estimates.

The simulation is set in the following way: In each cycle suppose that the real 4x3 national transition matrix is given in Table 1. The values $p_{ij}$ in this table correspond to the number of voter transitions divided by the size of the total electorate.

**Table 1 General transition matrix**

<table>
<thead>
<tr>
<th></th>
<th>SARKOZY</th>
<th>ROYAL</th>
<th>Non Valid + Abstention</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARKOZY</td>
<td>0.3503</td>
<td>0.0036</td>
<td>0.0072</td>
</tr>
<tr>
<td>ROYAL</td>
<td>0.0015</td>
<td>0.2901</td>
<td>0.0044</td>
</tr>
<tr>
<td>BAYROU</td>
<td>0.0874</td>
<td>0.0852</td>
<td>0.0459</td>
</tr>
<tr>
<td>LE PEN</td>
<td>0.0797</td>
<td>0.0174</td>
<td>0.0274</td>
</tr>
</tbody>
</table>

Since the electoral unit used is the French department, for each of the 96 departments the simulation draws a sample from a multinomial distribution $\text{Multinomial}(n_k, p^*)$ with 12 possible outcomes, where $n_k$ is the size of the electorate in department $k=1, 2, \ldots, 96$ and $p^*_i \sim \text{Normal}(p_{ij}, p_{ij}/6)$. Thus, for each department we have $n_k$ independent trials from a multinomial distribution, where each trial results in exactly one of the 12 possible outcomes, with probabilities $p^*$. Also note that in each district the probabilities are different and the only assumption is that if the national transition rate from party $i$ to party $j$ is $p_{ij}$, then in department $k$ the corresponding transition rate $p_{ijk}$ will follow a Normal distribution with mean value $p_{ij}$ and standard deviation $p_{ij}/6$.

---

1 For the code used see the Appendix.
The algorithm stores the sample values to a 4x3x96 array with elements $n_{ijk}$ where $i=1, 2, 3, 4$, $j=1, 2, 3$ and $k=1, 2, ..., 96$. The values of $n_{ijk}$ represent the number of voters who have moved from party $i$ to party $j$ in department $k$. Summing for all $k$ we get the 4x3 matrix with elements $n_{ij}$, where $n_{ij}$ is the number of voters who have moved from party $i$ to party $j$ in all departments. Summing for all $j$ we get the 96x4 matrix $X$ with elements $n_{i,k}$, where $n_{i,k}$ is the number of voters of the first round of the elections who have voted for party $i$ in department $k$. Summing for all $i$ we get the 96x3 matrix $T$ with elements $n_{j,k}$, where $n_{j,k}$ is the number of voters of the second round of the election who have voted for party $j$ in department $k$. Then matrices $X$ and $T$ along with vector $N=(n_1, n_2, ..., n_k)$ are used as input to VTR algorithm. The aforementioned procedure was replicated for 100 times and the findings are presented in the following section.

**Simulation Results**

Table 2 includes summary statistics of absolute deviations for VTR algorithm after 100 replicates of the simulation procedure. National level estimates (first row of the table) are very close to simulated values. The average total number of voters (as a fraction of the electorate) who were misclassified by VTR is 6.18%. The maximum absolute deviation of the voter transition rates estimates from the simulated (true) values is 13.01%. For half of the cycles the absolute deviation is in the range (4.63%, 7.37%)

<table>
<thead>
<tr>
<th>Estimates</th>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>National</td>
<td>1.56%</td>
<td>4.63%</td>
<td>6.16%</td>
<td>6.18%</td>
<td>7.37%</td>
<td>13.01%</td>
</tr>
<tr>
<td>Regional</td>
<td>0.67%</td>
<td>4.73%</td>
<td>6.37%</td>
<td>6.52%</td>
<td>7.94%</td>
<td>16.37%</td>
</tr>
<tr>
<td>Department</td>
<td>0.88%</td>
<td>5.17%</td>
<td>6.94%</td>
<td>7.30%</td>
<td>9.08%</td>
<td>22.71%</td>
</tr>
</tbody>
</table>

Estimating department level voter transition rates using as electoral units the same departments is a more difficult task. The average percentage of the electorate misclassified by VTR is 7.30% on the department level. The maximum absolute deviation of the voter transition rates estimates from the simulated (true) values on the department level is 22.71%.

The selection of electoral units is a very important factor for the successful estimations of quantities of interest. Andreadis and Chadjipadelis (2009) argue that as electoral units we should use the most detailed level of available electoral outcomes for which the electorate has remained unchanged between the first and the second election. Even if we are not interested in presenting our results on such a detailed level we should use this level and then combine the corresponding estimates and use the weighted means of the low level estimates to produce the estimates for the level we want to present. This approach offers the advantage of having more available data that will provide more deterministic information for the feasible range of values of quantities of interest.

Thus, according to Andreadis and Chadjipadelis, if we want to produce estimates of voter transition rates for regions defined by a division of the electorate, it is better to use electoral units defined by a subdivision of this division. For instance, in France if we want to estimate voter transition rates for the 22 Regions, it is better to use the 96
Departments as electoral units in VTR. The benefit is obvious from the second row of Table 2. The average percentage of the electorate misclassified by VTR is 6.37% on the regional level. The maximum absolute deviation of the voter transition rates estimates from the simulated (true) values on the regional level is 16.37%.

The maximum absolute deviation is different for each region and it is related with the number of departments included in the regions. In France there are regions with only two departments and other regions with up to eight departments. The maximum absolute deviation presented in Table 2 is among all regions. Restricting the searching of maximum only to the regions of France which have eight departments the corresponding value is 14.24%.

It should be noted that the absolute deviation index presented so far is the sum of absolute deviations for all 12 cells of the transition matrix. Another useful index is the average absolute deviation which is calculated by the formula: 
\[
\frac{1}{12} \sum_{i,j} \frac{1}{4} \sum_{\hat{p}_{ij} - p_{ij}}
\]
This index shows the absolute average deviation of the estimated value from the simulated value in each of the 12 cells of the general transition matrix. This index is included because it is helpful to remember that a 6% total absolute deviation in our case with a 4x3 transition matrix means that the value of average absolute deviation is 0.5%, i.e. for the average cell the difference between the estimates and the real values is 0.5%.

Available data in EU

Estimation of voter transition rates with ecological inference methods is done using electoral results from two consecutive elections. Thus, it is useful to know what the available data are for each of the 27 member states of the European Union. The numbers of first and second level administrative divisions for which there are available electoral results are shown in Table 3 for each European Union country. According to the elements of this table the average unit of the first level electoral division of a European country consists of about 21 second level electoral division units. Spain, with 17 autonomous communities and 52 provinces, is the country which according to this table presents the smaller number of second level units per first level unit (3.1). On the other hand, UK presents the largest number with an average of 161.5 constituencies per country.

Table 3. Availability of electoral results

<table>
<thead>
<tr>
<th>Member</th>
<th>1st Level name</th>
<th>2nd Level name</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>9 States</td>
<td>121 Districts²</td>
<td>13.4</td>
</tr>
<tr>
<td>Belgium</td>
<td>11 Constituencies</td>
<td>208 Cantons</td>
<td>18.9</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>28 Provinces</td>
<td>260 Municipalities</td>
<td>9.3</td>
</tr>
<tr>
<td>Cyprus</td>
<td>5 Districts</td>
<td>334 Municipalities</td>
<td>66.8</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>14 Regions</td>
<td>91 Territorial units</td>
<td>6.5</td>
</tr>
<tr>
<td>Denmark</td>
<td>5 Regions</td>
<td>98 Municipalities</td>
<td>19.6</td>
</tr>
<tr>
<td>Estonia</td>
<td>15 Counties</td>
<td>261 Municipalities</td>
<td>17.4</td>
</tr>
<tr>
<td>Finland</td>
<td>15 Electoral districts</td>
<td>416 Municipalities</td>
<td>27.7</td>
</tr>
<tr>
<td>France³</td>
<td>22 Regions</td>
<td>96 Departments</td>
<td>4.4</td>
</tr>
</tbody>
</table>

² In Austria there are 98 (Politische Bezirke and Statutarstädte) and Vienna which is divided into 23 districts
<table>
<thead>
<tr>
<th>Country</th>
<th>Levels</th>
<th>Subdivisions</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>16 States</td>
<td>299 Districts</td>
<td>18.7</td>
</tr>
<tr>
<td>Greece</td>
<td>56 Electoral districts</td>
<td>1034 Municipalities</td>
<td>18.5</td>
</tr>
<tr>
<td>Hungary</td>
<td>20 Regions</td>
<td>173 Sub regions</td>
<td>8.7</td>
</tr>
<tr>
<td>Ireland</td>
<td>43 Constituencies</td>
<td>192 Local Electoral Areas</td>
<td>4.5</td>
</tr>
<tr>
<td>Italy</td>
<td>27 Districts</td>
<td>110 Counties</td>
<td>4.1</td>
</tr>
<tr>
<td>Latvia</td>
<td>5 Districts</td>
<td>34 Regions</td>
<td>6.8</td>
</tr>
<tr>
<td>Lithuania</td>
<td>71 Constituencies</td>
<td>2034 Districts</td>
<td>28.6</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>12 Cantons</td>
<td>118 Municipalities</td>
<td>9.8</td>
</tr>
<tr>
<td>Malta</td>
<td>13 Districts</td>
<td>55 Localities</td>
<td>4.2</td>
</tr>
<tr>
<td>Netherlands</td>
<td>12 Provinces</td>
<td>465 Municipalities</td>
<td>38.8</td>
</tr>
<tr>
<td>Poland</td>
<td>16 Provinces</td>
<td>379 Counties</td>
<td>23.7</td>
</tr>
<tr>
<td>Portugal</td>
<td>20 Districts</td>
<td>308 Municipalities</td>
<td>15.4</td>
</tr>
<tr>
<td>Romania</td>
<td>42 Counties</td>
<td>312 Colleges</td>
<td>7.4</td>
</tr>
<tr>
<td>Slovakia</td>
<td>8 Regions</td>
<td>79 Districts</td>
<td>9.9</td>
</tr>
<tr>
<td>Slovenia</td>
<td>8 Constituencies</td>
<td>88 Districts</td>
<td>11.0</td>
</tr>
<tr>
<td>Spain</td>
<td>17 Autonomous communities</td>
<td>52 Provinces</td>
<td>3.1</td>
</tr>
<tr>
<td>Sweden</td>
<td>21 Counties</td>
<td>290 Municipal assemblies</td>
<td>13.8</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>4 Countries</td>
<td>646 Constituencies</td>
<td>161.5</td>
</tr>
</tbody>
</table>

Electoral units for national estimates

The simulation results using as template the first and second level administrative (and electoral) French divisions indicate that the number of units is a significant factor for the accuracy of estimates of voter transition rates. Considering this, it seems that to produce national level estimates of voter transition rates in Romania, Ireland, Greece and Lithuania, (which all are divided to more than 40 areas using their first level division) we can use as electoral units the areas defined by the first level division. As the number of areas gets smaller we can be less confident for the estimates. For instance, it would not be wise to use the 5 districts of Cyprus or Latvia, the 5 regions of Denmark, or the 4 countries of United Kingdom to produce general estimates.

For the rest of the member states, researchers have to decide considering the level of accuracy they require and the special characteristics of each country. For instance, in Spain there are 17 autonomous communities and a lot of local political parties which run for the elections only in selected areas. An analysis on Spanish elections would be better if the researcher would use the 52 provinces or even a more detailed division. In fact, some preliminary results on the application of VTR in Spain indicate that it is better to apply the algorithm in each autonomous community separately. This could be also true for other member states of the EU (i.e. federal states comprising a number of partially self-governing regions) and United Kingdom consisting of four countries and having Scottish, Welsh and Northern Ireland parties although VTR can handle similar cases to some degree.

---

3 France is divided into 22 regions and four departments which lie overseas, the latter departments are also counted as regions totalling the number of France regions to 26.
Regional estimates
Apart from Spain and UK which were discussed above, according to the elements of Table 3, in the rest of the EU Members the average unit of the first level electoral division includes a number of second level electoral division units which number ranges from about 4 (Italy, Malta, France) up to more than 66 (Cyprus). As the simulation results indicate when the ratio between first and second level divisions is about 1 to 4 and the second level division units are used as electoral units, the average percentage of the electorate misclassified by VTR on the fist division level is 6.37% and the corresponding maximum value is 16.37%.

Thus, it seems that the interested readers could use ecological inference for the estimation of voter transition rates on both national and regional levels in all the 27 European Union countries.

Conclusions
In the aftermath of an election, political parties need to analyse their performance at the maximum possible detailed level. If they have raised their power, they need to know previous electoral behaviour of their new voters. If they have suffered from defections, they need to know the political parties which have earned from these defections. Ecological inference can serve as a useful tool for these cases because it provides a general table of voter transitions for the entire country and a detailed array of voter transitions for each of the departments included in this country.

The simulation findings presented in this paper indicate that the VTR algorithm provides estimates which are close to reality and it suggests that the method could and should be used in all the member states of the European Union. This paper includes preliminary results on the validity of ecological inference for the estimation of voter transition rates on both national and regional levels. Further research should include more extensive simulations, i.e. with more than 100 replicates/cycles and with larger than 4x3 tables. A full proof of the usefulness of ecological inference for the estimation of voter transition rates for all the 27 European Union Members could be given by two ways. The first one would be by using a more general simulation experiment with random selection of rows and columns and random selection of the general voter transition matrix. The second way would be to run a separate simulation experiment for each one of the 27 European Union Members using a national specific general transition matrix for each one of them.

References


Appendix Simulation code in R language

source("VTR.R") #load VTR algorithm (download from http://polres.polsci.auth.gr)

# Initialize values which should be changed for each country
MaxRepl<-100 # number of replications
DimX<-4 # number of parties in e1
DimT<-3 # number of parties in e2
NRegions<-22 # number of regions
NDeps<-96 # number of Departments

nfrance<-read.csv2("nfrance.csv") # file with NDeps rows and two columns: nk values and region codes
regions<-factor(nfrance$reg)
p<-c(0.3503,0.0015,0.0874,0.0797,0.0036,0.2901,0.0852,0.0174,0.0072,0.0044,0.0459,0.0274)

# Define arrays to store the values
x.ijk<-array(0, dim=c(DimX,DimT,NDeps)) # the number of people who voted party i in e1 and party j in e2 in department k
tr.ijk<-array(0, dim=c(DimX,DimT,NDeps)) # transition rate of voters of party i who voted for party j in department k
vtr<-array(0, dim=c(DimX,DimT,MaxRepl)) # the estimated general transition matrix for each replicate (as a fraction of power of party i)
vtr96<-array(0, dim=c(NDeps,DimX,DimT,MaxRepl)) # estimated voter transitions as a fraction of power of party i in department k
vtrtotal96<-array(0, dim=c(NDeps,DimX,DimT,MaxRepl)) # estimated voter transitions as a fraction of the total electorate in department k
vtrtotal22<-array(0, dim=c(NRegions,DimX,DimT,MaxRepl)) # estimated voter transitions as a fraction of the total electorate in each region
devanal<-array(0, dim=c(DimX,DimT,MaxRepl)) #absolute deviations from the simulated "true" values
devanal96<-array(0, dim=c(NDeps,DimX,DimT,MaxRepl)) #absolute deviations from the simulated "true" values in each department
devanal22<-array(0, dim=c(NRegions,DimX,DimT,MaxRepl)) #absolute deviations from the simulated "true" values in each region

vtrdev<-array(0,MaxRepl) #sum of absolute deviations
vtrdev96<-array(0, dim=c(NDeps,MaxRepl)) #sum of absolute deviations in each department
vtrdev22<-array(0, dim=c(NRegions,MaxRepl)) #sum of absolute deviations in each region
disturb<-6  # define the standard deviation as 1/6 of the mean value

# Replication starts
for (repl in c(1:MaxRepl)) {
  cat(repl)
  print(date())
  set.seed(repl)  # to keep track of the simulated values and check outliers

  # Draw values in each department
  for (k in c(1:NDeps)) {
    # in each department the voter transition rate follows a Normal distribution with mean value p and sd p/disturb
    # (it needs to be reformed to work with any number of dimensions)
    p2<-c(rnorm(1, mean = p[1], sd = p[1]/disturb),
          rnorm(1, mean = p[2], sd = p[2]/disturb),
          rnorm(1, mean = p[3], sd = p[3]/disturb),
          rnorm(1, mean = p[4], sd = p[4]/disturb),
          rnorm(1, mean = p[5], sd = p[5]/disturb),
          rnorm(1, mean = p[6], sd = p[6]/disturb),
          rnorm(1, mean = p[7], sd = p[7]/disturb),
          rnorm(1, mean = p[8], sd = p[8]/disturb),
          rnorm(1, mean = p[9], sd = p[9]/disturb),
          rnorm(1, mean = p[10], sd = p[10]/disturb),
          rnorm(1, mean = p[11], sd = p[11]/disturb),
          rnorm(1, mean = p[12], sd = p[12]/disturb))
    p2[p2<0]<-0.00001  # if a simulated value is negative put a small value for the corresponding probability
    nk<-nfrance$nk[k]  # size of electorate in each department
    nij<-rmultinom(n=1,size=nk,prob=p2)  # draw from a multinomial distribution
    # (it gives the number of voters who moved from party i to party j in department k)
    dim(nij)<-c(DimX,DimT)  # change the dimensions to fit the dimensions of the voter transition matrix
    x.ijk[,k]<-nij  # the simulated "true" number of people who voted party i in e1 and party j in e2 in department k
    tr.ijk[,k]<-x.ijk[,k]/rowSums(x.ijk[,k])  # transition rate of voters of party i who voted for party j in department k
  }
}

# Prepare data for input in VTR
n.ij <- apply(x.ijk,c(1,2),sum)  # general voter transition matrix (absolute values) (kept for testing)
n.ik <- apply(x.ijk,c(1,3),sum)  # number of voters of party i; matrix with dimensions(DimX, NDep)
n.jk <- apply(x.ijk,c(2,3),sum)  # number of voters of party j; matrix with dimensions(DimT, NDeps)
n.k <- apply(x.ijk,3,sum)  # number of voters in department k; vector with length(NDeps); it should be the same with nk: (kept for testing)
data<-as.data.frame(cbind(n.k,t(n.ik),t(n.jk)))  # prepare the data frame (N,X,T)
data[-1]<-data[-1]/n.k  # calculate the relative values (vote share for each party in each department)
MyX<-data[2:5]
MyT<-data[6:8]
MyN<-data[1]

# Run VTR
z<-multirate(MyN,MyX,MyT,0.001)
vtr[,,repl]<-z$Bb  # store the estimated general transition matrix for the current replicate (as a fraction of power of party i)
vtrtotal<-pij*z$Bb  # calculate voter transitions as a fraction of the total electorate
vtr96[,,repl]<-z$bb  # store the estimated transition matrix of each department for the current replicate (as a fraction of power of party i)
pij<rowSums(n.ik)/sum(n.ik)  # calculate the national relative values of parties in X (vote shares)
pij96<-t(n.ik)/n.k  # calculate the relative values of parties in X (vote share for each party in each department)
for (j in c(1:DimT)) { vtrtotal96[,,j,repl]<-pijd96*z$bb[,,j]  # calculate and store voter transitions as a fraction of the total electorate in department k }

#Calculate the estimated values for each region
n.22<-aggregate(n.k, list(regions), sum)
vtrtotal22temp<-aggregate(vtrtotal96[,,repl]*n.k, list(regions), sum)
vtrtotal22temp<-as.matrix(vtrtotal22temp[-1])
dim(vtrtotal22temp)<-c(NRegions,DimX,DimT)
vtrtotal22temp<-vtrtotal22temp[22]*n.22$x
vtrtotal22[,,repl]<-vtrtotal22temp

#Calculate the simulated "true" voter transition rates as a fraction of the total electorate
theta<-n.ij/sum(n.k)  # simulated "true" voter transition rates as a fraction of the total electorate
newx.ijk<-aperm(x.ijk, c(3,1,2))
theta96 <- newx.ijk/n.k  # simulated "true" voter transition rates as a fraction of the total electorate in each department
theta22 <- aggregate(newx.ijk, list(regions), sum)
theta22 <- as.matrix(theta22[-1])
dim(theta22) <- c(NRegions, DimX, DimT)
theta22 <- theta22/n.22$x  # simulated "true" voter transition rates as a fraction of the total electorate in each region

devanal[,,repl] <- abs(vtrtotal-theta)  # calculate absolute deviations from the simulated "true" values
vtrdev[repl] <- sum(devanal[,,repl])  # sum of absolute deviations

devanal96[,,,repl] <- abs(vtrtotal96[,,,repl]-theta96)  # calculate absolute deviations from the simulated "true" values in each department
vtrdev96[,repl] <- rowSums(devanal96[,,,repl])  # sum of absolute deviations in each department

devanal22[,,,repl] <- abs(vtrtotal22[,,,repl]-theta22)  # calculate absolute deviations from the simulated "true" values in each region
vtrdev22[,repl] <- rowSums(devanal22[,,,repl])  # sum of absolute deviations in each department

}  # End of replication