Computing Voter Transitions:

The Elections for the Catalan Parliament, from 2010 to 2012

Albert Corominas, Amaia Lusa, M. Dolors Calvet

Universitat Politècnica de Catalunya (Spain)

albert.corominas@upc.edu, amaia.lusa@upc.edu, m.dolors.calvet@upc.edu

Abstract:

Purpose: To estimate the transition rates corresponding to the 2010 and 2012 elections to the Catalan Parliament for the four constituencies in which Catalonia is divided for this purpose. The main features of the results, which are obtained by means of mathematical programming, are commented.

Design/methodology/approach: Mathematical programming optimization models are formulated in order to find the transition rates that yield a better adjust between the actual results in 2012 and those computed applying the transition rates to the 2010 results. The transition rate matrices are estimated for each one of the four constituencies, since the set of options is not the same for all them. No other assumptions that those of numerical consistency are adopted.

Findings and Originality/value: The transition rate models provide satisfactory goodness of fit. Mathematical programming turns out to be an easy-to-use tool for estimating the transition rates and, at the same time, very flexible, since, if necessary, it allows incorporating the constraints corresponding to additional assumptions.

Originality/value: The transition rates from 2010 to 2012 in Catalonia are particularly interesting, since 2012 results implied a significant change in the composition of the Catalan Parliament. To the best of our knowledge, no other scientific journal paper has dealt with this

question. Our results are available to the researchers in order to interpret the change and try to foresee future flows of voters.

1. Introduction

When considering the results of two consecutive elections in the same electoral area, a usual way of trying to interpret the results of the later polls is seeing them as the consequence of voter transitions from the options they preferred in the former. Politicians, the media, political scientists and most citizens are interested in the changes of the preferences of people having the right to vote.

According to Hawkes (1969) the historian Trevor Lloyd was the first of stating the question. Before 1970, it was dealt with in some scarce works (Benewick, Birch, Blumler & Ewbank, 1969; Berrington, 1965; Butler, 1952, 1953; Butler & King, 1966;).

Formally, given the results of two consecutive (or even simultaneous) elections for each one of the *divisions* (constituencies, municipalities, polling stations or any other partition) of an electoral area, the problem is to find the matrix of transition rates from the options available in the first election (rows) to the options in the second one (columns). Of course, if one considers only aggregate results of the whole territory or of any set of constituencies, in general there are infinitely many solutions to the matrix. On the other hand, applying a unique transition rate matrix to diverse constituencies or groups of constituencies, the computed results will not always coincide with the real ones. Clearly, the elements of the matrix must be nonnegative and those belonging to any given file must sum up to 1 (these two conditions imply that the elements must be less than or equal to 1).

A survey can be used to estimate the elements of the transition rate matrix. However, the results are highly unreliable, because of many reasons that are discussed, for instance, in Brown and Payne (1986) and in Van der Ploeg, Van de Pol and Kampen (2006). Moreover, unless the number of elements in the sample is very high, many elements of the matrix (those corresponding to small values of the transition rate) will be equal to zero. Therefore, at the expense of greater modelling and computing efforts, the use of the results of both elections is a more reliable way to obtain the matrix.

Hawkes (1969) tried "to estimate the number of people voting for a particular party at one election who subsequently vote for another specified party at the next election". In order to do this, proposed three methods. However, these did not guarantee that the results fulfil the stated above conditions. Therefore, the author concluded that although "the attempt has not been as successful as one would wish, some useful results are obtainable". Miller (1972) and

Upton (1977) adopt an approach that may be inscribed in the same stream that Hawkes (1969). Other related works are Hayes (1976) and Moores (1987).

Instead, Irwin and Meeter (1969) and McCarthy and Ryan (1977) use quadratic programming to estimate the transition rates, thus guaranteeing from the outside the fulfilling of the above mentioned conditions, which are imposed by means of constraints in the mathematical programming. Tziafetas (1986) uses an approach similar to that of McCarthy and Ryan (1977) and a variant of it, which consists in using the absolute value of the deviations instead of the squares of them.

Upton (1978), concerning McCarthy and Ryan's approach, observes that it gives a high proportion of zeroes in the matrix and concludes that overestimates the proportion of *stayers* (voters that do not change their preferences between the to elections –actually, electors that vote for options whose name is the same in both elections–, contrasting with *movers*). This criticism, which we do not deem fully justified, has been assumed by other authors, as, for instance Johnston and Hay (1983).

The estimation of voters' transition rates is often seen as a particular case of the ecological inference problem, i. e., to deduct individual behaviours from aggregate data, a problem which was deemed impossible, with the methods available at that time, in Robinson (1950). This notwithstanding, many methods have been proposed for dealing with it. Some of them assume, perhaps implicitly, that the behaviour pattern is the same or very similar in all the areas (ecological regressions; see: Goodman, 1953; Goodman, 1959; concerning transition rates: Fülle, 1994; van der Ploeg et al., 2006). Others, consider that the behaviour pattern may depend on the areas and usually adopt a probabilistic approach (ecological inference; see: Glynn & Wakefield, 2010; Greiner & Quinn, 2009; Grofman & Merrill, 2004; King, 1997; concerning transition rates –probabilities–: Andreadis & Chadjipadelis, 2009; Antweiler, 2007; Brown & Payne, 1986; Johnston & Hay, 1983).

The purpose of the present paper is to determine, for the 2010 and 2012 elections to the Parliament of Catalonia, a transition matrix for each constituency that (i) fulfils the nonnegative and sum-to-one constraints; (ii) applied to the aggregate results of the first election give exactly the results of the second one for each one of the available options in this later and (iii) minimises a function of the discrepancies between the results obtained with the matrix and those given by the count of votes in every division. Note that we neither formulate any assumption about the differences or coincidences between the behaviour patterns of the electors corresponding to diverse polling stations nor adopt a probabilistic point of view.

Therefore, the problem can be stated as a mathematical programming model, which for some kinds of discrepancy functions is easy to solve. Even though our purpose is to find transition matrices for any set of divisions (and for different kinds of divisions), without introducing any a

priori assumption about the values of the transition rates, we will comment some results given by the models in order to facilitate their interpretation.

The layout of the rest of the article is as follows. Section 2 presents the problem and its mathematical programming formulations. The data and the obtained results are presented and commented in Section 3. Section 4 ends the paper with some short conclusions.

2. Statement of the Problem and its Mathematical Programming Formulation

It is assumed that we have the results corresponding to two elections in the same electoral area, such that the first one happened at t and the second one at $t'(\geq t)$.

When t' is very close to t (or even equal to t) it may be that the censuses corresponding to both elections are identical. If actually they are not, however, it is usual to circumvent this difficulty assuming that the behaviour of the electors not belonging to the intersection of both censuses is not different from those that belong to it. In fact, this is equivalent to assume that both censuses are identical and this assumption is reasonable when t' is not far from t (Brown & Payne, 1986; Hawkes, 1969; McCarthy & Ryan, 1977), as happen with the two elections considered in this paper, separated by only two years (2010, 2012).

The electoral area is partitioned into constituencies and these, at the end, into polling stations. Therefore the results are available for all the polling stations of the electoral area. It may happen that the polling stations belonging to a given constituency do not coincide from one election to another (because some are created, suppressed or divided) and in this case one can only compare the results corresponding to the polling stations common to both elections.

On another hand, the available options (including blank vote, null vote and abstention) may be different from one constituency to another (in the case of the elections for the Catalan Parliament, they are, since the candidates to win the seats are different, even for the options with the same name). Therefore, the constituencies have to be considered separately.

Hence, the data must refer always to a given constituency or a subset of polling stations belonging to a given constituency. The considered polling units can be grouped to form divisions (therefore, a division is a set of one ore more polling stations; every considered polling station must belong to one and only one division). The divisions may be, for instance, municipalities, districts or any sets of polling units that be convenient for the analysis.

The notation that we use for the data are as follows:

m Number of divisions (common to both elections).

n, n' Number of options at t and t', respectively.

p_{ik} Proportion of votes obtained at t by the option k in the division i ($i = 1, …, m$; $k = 1, …, n$).

p_{ij} Proportion of votes obtained at t' by the option j in the division i ($i = 1, …, m$; $j = 1, …, n'$).

c_i Census (i.e., number of people having the right to vote) of division i at t', assumed to be equal to that of t ($i = 1, …, m$).

And for the decision variables:

r_{kj} Transition rate from option k at t to option j at t' ($k = 1, …, n$; $j = 1, …, n'$).

The matrix R made up of the transition rates r_{kj} must belong to the set F defined by the following constraints:

$$\sum_{j=1}^{n'} r_{kj} = 1 \qquad\qquad k = 1,…,n \qquad\qquad (1)$$

$$\sum_{i=1}^{m} c_i \cdot \sum_{k=1}^{n} p_{ik} \cdot r_{kj} = \sum_{i=1}^{m} c_i \cdot p_{ij}' \qquad j = 1,…,n' \qquad\qquad (2)$$

$$r_{kj} \geq 0 \qquad\qquad k = 1,…,n;\ j = 1,…,n' \qquad\qquad (3)$$

Equation 1 impose that the transition rates from an option at t to every other at t' must sum up to one; Equation 2, that the total number of votes obtained in the constituency by an option at t' equals the number that results when applying the transition rates to the number of votes corresponding to t (under the assumption that the census at t and at t' are the same); Equation 3 enforce the obvious non-negativity condition.

The mathematical programming models proposed in McCarthy and Ryan (1977) and Tziafetas (1986) include constraints (1) and (3). Constraints (2) are similar to those proposed in Johnston and Hay (1983).

These constraints define a set of matrices having generally infinitely many elements. One way to select one of these elements is to minimise the discrepancies between the actual results of elections at t' and those resulting from the application of the transition rates. Of course, the selected matrix depends on the used measure of the discrepancies.

This way we define the following four models:

$$\text{M1 } \textit{minimise } \sum_{i=1}^{m}\sum_{j=1}^{n'}\left(c_i \cdot p'_{ij} - \sum_{k=1}^{n}\left(c_i \cdot p_{ik} \cdot r_{kj}\right)\right)^2 = \sum_{i=1}^{m}c_i^2\sum_{j=1}^{n'}\left(p'_{ij} - \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right)\right)^2 \quad \text{s.t. } R \in F$$

$$\text{M2 } \textit{minimise } \sum_{i=1}^{m}\sum_{j=1}^{n'}\left(p'_{ij} - \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right)\right)^2 \quad \text{s.t. } R \in F$$

$$\text{M3 } \textit{minimise } \max_{\substack{i=1,\dots,m\\ j=1,\dots,n'}} c_i \cdot \left| p'_{ij} - \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right)\right| \quad \text{s.t. } R \in F$$

$$\text{M4 } \textit{minimise } \max_{\substack{i=1,\dots,m\\ j=1,\dots,n'}} \left| p'_{ij} - \sum_{k=1}^{n}\left(p_{ik} \; r_{kj}\right)\right| \quad \text{s.t. } R \in F$$

The objective functions of the models are (M1) the sum of the squared discrepancies between actual and modelled number of votes, (M2) the sum of the squared discrepancies between actual and modelled proportions, (M3) the value of the maximum discrepancy between actual and modelled number of votes and (M4) the value of the maximum discrepancy between actual and modelled proportions.

Since the constraints that define the set F are linear, M1 and M2 are quadratic programs. For their part, M3 and M4 can be reformulated as linear programs as follows:

M3 $\textit{minimise } \Delta$

$$\text{s.t. } R \in F, \; \Delta \geq c_i \cdot \left(p'_{ij} - \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right)\right), \Delta \geq c_i \cdot \left(-p'_{ij} + \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right)\right) \quad i=1,\dots,m; \; j=1,\dots,n'$$

M4 $\textit{minimise } \delta$

$$\text{s.t. } R \in F, \; \delta \geq p'_{ij} - \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right), \delta \geq -p'_{ij} + \sum_{k=1}^{n}\left(p_{ik} \cdot r_{kj}\right) \quad i=1,\dots,m; \; j=1,\dots,n'$$

As it is known, however, *minmax* problems have usually multiple optimums, since the objective function does not take into account the values of the discrepancies that are strictly less than the optimum value. Therefore, after solving M3 and M4, more satisfactory solutions may be obtained using a second criterion (the sum of the absolute values of all the discrepancies); this leads to the following two quadratic programming models:

M3'
$$\text{minimise } \sum_{i=1}^{m}\sum_{j=1}^{n'}\left(c_i\cdot p'_{ij}-\sum_{k=1}^{n}\left(c_i\cdot p_{ik}\cdot r_{kj}\right)\right)^2=\sum_{i=1}^{m}c_i^2\sum_{j=1}^{n'}\left(p'_{ij}-\sum_{k=1}^{n}\left(p_{ik}\cdot r_{kj}\right)\right)^2$$

s.t. $\qquad\qquad R\in F$

$$\Delta^*\ge c_i\cdot\left(p'_{ij}-\sum_{k=1}^{n}\left(p_{ik}\cdot r_{kj}\right)\right),\Delta^*\ge c_i\cdot\left(-p'_{ij}+\sum_{k=1}^{n}\left(p_{ik}\cdot r_{kj}\right)\right)\qquad i=1,\dots,m;\ j=1,\dots,n'$$

M4'
$$\text{minimise } \sum_{i=1}^{m}\sum_{j=1}^{n'}\left(p'_{ij}-\sum_{k=1}^{n}\left(p_{ik}\ r_{kj}\right)\right)^2$$

s.t. $\qquad\qquad R\in F$

$$\delta^*\ge p'_{ij}-\sum_{k=1}^{n}\left(p_{ik}\cdot r_{kj}\right),\delta^*\ge -p'_{ij}+\sum_{k=1}^{n}\left(p_{ik}\cdot r_{kj}\right)\qquad i=1,\dots,m;\ j=1,\dots,n'$$

Where Δ^* and δ^* are, respectively, the optimum values of the objective function corresponding to M3 and M4.

3. Data and results

The data sets used in the computational experiment correspond to the elections to the Parliament of Catalonia held on 28 November 2010 and 25 November 2012.

The constituencies coincide with the four provinces of Catalonia: Barcelona, Girona, Lleida and Tarragona.

Each province is divided into *comarques,* and, for its part, each *comarca* is divided into municipalities. Catalonia has 41 *comarques* and 947 municipalities.

The municipalities, for electoral purposes, are divided into sections, so that in each section the number of registered electors belongs to the interval 500-2,000 (with the possible exception of small relatively isolated villages, in which the number of registered electors may be less than 500). According to this, because of the changes in population, from one election to the next a section can be divided into two or more or can be incorporated to another. Of course, the data used in the computational experiment are only those corresponding to sections that remain unchanged.

The values of *m*, *n* and *n'* corresponding to the four constituencies in the used data set are included in Table 1.

Constituency	Number of divisions (m)	Number of options in 2010 (n)	Number of options in 2010 (n')
Barcelona	3573	32	19
Girona	523	29	19
Lleida	390	32	19
Tarragona	523	33	19

Table 1. Sizes of the data sets used in the computational experiment. In all cases, the number of options includes null and blank votes and abstention

To solve the six mathematical programming models described in the preceding section for each one of the constituencies the software CPLEX 12.2 was used. The computing times were not significant in any case. The transition matrices obtained with models M3 and M4 can be disregarded, since those corresponding to models M3' and M4', respectively, are always preferable; only the optimum values of their respective objective functions are used (as inputs for models M3' and M4', respectively).

In order to evaluate the goodness of fit of the solutions provided for the different models we will use the following criteria, respectively related to the objective functions of models M1, M2, M3' and M4':

- G_v^2, coefficient of determination corresponding to the numbers of votes, defined as follows:

$$G_v^2 = 1 - \frac{\sum_{j=1}^{n'}\sum_{i=1}^{m}\Delta_{ij}^2}{\sum_{j=1}^{n'}\sum_{i=1}^{m}\left(c_i \cdot p_{ij}' - c_i \cdot \bar{p}_j'\right)^2}$$

where $\bar{p}_j' = \dfrac{\sum_{i=1}^{m} c_i \cdot p_{ij}'}{\sum_{i=1}^{m} c_i}$

(i.e., the proportion of votes obtained globally by the option j in the set of divisions).

- G_p^2, coefficient of determination corresponding to the proportions:

$$G_p^2 = 1 - \frac{\sum_{j=1}^{n'}\sum_{i=1}^{m}\delta_{ij}^2}{\sum_{j=1}^{n'}\sum_{i=1}^{m}\left(p_{ij}' - \bar{p}_j'\right)^2}$$

- Δ_{max}, the maximum discrepancy between actual and modelled numbers of votes.

- δ_{max}, the maximum discrepancy between actual and modelled proportions of votes.

Coefficients of determination, which are useful as complementary information to assess the goodness of fit of the models, may be defined for each option as well:

$$g_v^2(j) = 1 - \frac{\sum_{i=1}^{m} \Delta_{ij}^2}{\sum_{i=1}^{m} \left(c_i \cdot p_{ij}^{'} - c_i \cdot \bar{p}_j^{'}\right)^2} \; ; \; g_p^2(j) = 1 - \frac{\sum_{i=1}^{m} \delta_{ij}^2}{\sum_{i=1}^{m} \left(p_{ij}^{'} - \bar{p}_j^{'}\right)^2}$$

$$j = 1, \dots, n'$$

Of course, models M1 and M2 will yield always the best values of G_v^2 and G_p^2, respectively (since the denominators in the definition of these criteria are constants and the respective numerators are the objective functions of M1 and M2). In a similar way, models M3-M3' and M4-M4' will be the best for Δ_{max} and δ_{max}, respectively. The behaviour of the mentioned models with regard to the other criteria is difficult to forecast, excepting that M1 and M2 are more robust than M3' and M4' in relation to outliers.

The transition rates matrices and the values of all the criteria described above for Barcelona, Girona, Lleida and Tarragona, can be found in:

https://dl.dropboxusercontent.com/u/6741065/RESULTS%20BARCELONA.xlsx,

 https://dl.dropboxusercontent.com/u/6741065/RESULTS%20GIRONA.xlsx,

https://dl.dropboxusercontent.com/u/6741065/RESULTS%20LLEIDA.xlsx and

https://dl.dropboxusercontent.com/u/6741065/RESULTS%20TARRAGONA.xlsx, respectively.

Table 2 shows the values of the four criteria corresponding to the four models M1, M2, M3' and M4' and to the four constituencies.

One can see that the values of G_v^2 and G_p^2 are good (high). The latter, with M2, reaches 0.97 for Tarragona, while the best value of G_v^2 is 0.86 (Barcelona and Girona). Concerning these two criteria, the results are very similar for both M1 and M2, although, of course, M1 gets the best values for G_v^2 and M2 for G_p^2. The worst value, from those given by M1, of G_v^2 is 0.77 (Tarragona) and the worst of G_p^2, from those given by M2, is 0.75 (Lleida).

For their side, the values of Δ_{max} and δ_{max} are high, even for M3' and M4' and, although they are not very much better than those obtained with M1 and M2, when the objective of minimising these criteria is imposed (M3' and M4') the values of G_v^2 and G_p^2 deteriorate.

		G_v^2		G_p^2		Δ_{max}		δ_{max}
M1	BAR	**0.86**	BAR	0.85	BAR	289.01	BAR	0.30
	GIR	**0.86**	GIR	0.83	GIR	123.60	GIR	0.19
	LLE	**0.81**	LLE	0.74	LLE	139.52	LLE	0.18
	TAR	**0.77**	TAR	0.97	TAR	175.80	TAR	0.27
M2	BAR	0.86	BAR	**0.85**	BAR	289.50	BAR	0.30
	GIR	0.85	GIR	**0.83**	GIR	129.16	GIR	0.18
	LLE	0.80	LLE	**0.75**	LLE	153.24	LLE	0.19
	TAR	0.76	TAR	**0.97**	TAR	177.38	TAR	0.26
M3'	BAR	0.83	BAR	0.83	BAR	**220.79**	BAR	0.30
	GIR	0.85	GIR	0.82	GIR	**112.06**	GIR	0.19
	LLE	0.80	LLE	0.73	LLE	**119.60**	LLE	0.18
	TAR	0.74	TAR	0.97	TAR	**151.23**	TAR	0.27
M4'	BAR	0.82	BAR	0.82	BAR	363.75	BAR	**0.28**
	GIR	0.84	GIR	0.82	GIR	140.13	GIR	**0.15**
	LLE	0.70	LLE	0.65	LLE	181.51	LLE	**0.14**
	TAR	0.63	TAR	0.96	TAR	223.96	TAR	**0.23**

Table 2. Values of the four criteria corresponding to the four models and the four constituencies: Barcelona (BAR), Girona (GIR), Lleida (LLE) and Tarragona (TAR). The best values of the criteria, for each constituency, are highlighted in bold. Of course, they are, for G_v^2, G_p^2, Δ_{max} and δ_{max} those obtained, respectively, with models M1, M2, M3' and M4'.

Since the 16 transition rate matrices obtained for the four constituencies with the four models, even though they show some common characteristics, are fairly different, it is not possible to analyse them here in detail and we will limit ourselves to some general comments and to deepen a little more into the transition rate matrix given by M1 for the constituency of Barcelona.

In 2010, seven options won seats in the Catalan Parliament. In 2012, there were again seven the options present in the Parliament, but different from those of 2010, because SI (4 seats in 2010) did not obtain representation in 2012 and CUP, which did not present lists of candidates in 2010, won 3 seats in 2012. Table 3 shows, for the options that won seats, either in 2010 or in 2012, and for the abstention, the results they obtained, for the overall of the four constituencies, in both elections. For its side, Table 4 shows analogous data for the constituency of Barcelona.

OPTION	% 2010	Seats 2010	% 2012	Seats 2012	% 2012 - % 2010	Seats 2012 – seats 2010
CiU	22.91	62	21.16	50	-1.75	-12
PSC-PSOE	10.91	28	9.96	20	-0.95	-8
PP	7.35	18	8.95	19	1.60	1
ICV-EUiA	4.39	10	6.83	13	2.44	3
ERC	4.17	10	9.44	21	5.27	11
C's	2.02	3	5.22	9	3.20	6
SI	1.97	4	0.89	0	-1.08	-4
CUP	-	-	2.40	3	2.40	3
Other	4.05	-	3.10	-	-0.95	-
Blank	1.75	-	1.01	-	-0.74	-
Null	0.43	-	0.61	-	0.18	-
Abstention	40.05	-	30.43	-	-9.62	-

Note that the percentages are over the census (not, as it is more usual, over the valid votes for the lists of candidates)

Table 3. Main global results of the elections to the Catalan Parliament held in 2010 and 2012

OPTION	% 2010	Seats 2010	% 2012	Seats 2012	% 2012 - % 2010	Seats 2012 – seats 2010
CiU	22.00	35	19.45	26	-2.55	-9
PSC-PSOE	11.42	18	10.68	14	-0.74	-4
PP	7.67	12	9.21	12	1.54	0
ICV-EUiA	4.94	8	7.73	10	2.79	2
ERC	3.80	6	8.81	12	5.01	6
C's	2.30	3	5.85	8	3.55	5
SI	1.86	3	0.82	0	-1.04	-3
CUP	-	-	2.36	3	2.36	3
Other	3.94	-	3.44	-	-0.50	-
Blank	1.74	-	0.97	-	-0.77	-
Null	0.38	-	0.57	-	0.19	-
Abstention	39.95	-	30.11	-	-9.84	-

Note that the percentages are over the census (not, as it is more usual, over the valid votes for the lists of candidates)

Table 4. Main results, in the constituency of Barcelona, of the elections to the Catalan Parliament held in 2010 and 2012

In all constituencies, according to the solutions of the models, there is a positive flow from CiU to ERC, which, in Girona, Lleida and Tarragona is the balance that results from a bidirectional flow (from CiU to ERC and from ERC to CiU). Also in all constituencies PSC-PSOE losses votes in favour of a certain number of options (mainly, ICV-EUiA, C's and PP). A certain number of PP voters in 2010 changed to to C's in 2012. With the exception of Barcelona, ICV-EUIA shows a low proportion of stayers and the movers go to CUP and ERC. C's enjoys high proportions of stayers and gets votes from PSC-PSOE and PP. SI has no stayers in any constituency and its movers go mainly to CiU and to ERC as well; in Barcelona it receives a small proportion of the

votes got by CiU in 2010. The votes for CUP come mainly from ICV-EUiA and numerous minor options that won seat neither in 2010 nor in 2012; in Barcelona, moreover, CUP receives a significant amount of votes from CiU.

The Barcelona constituency stands out for the high proportion of stayers that have the abstention and the options that won seats both in 2010 and 2012 (from M1: CiU, 77%; PSC-PSOE, 82%; PP, 91%; ICV-EUiA, 94%; ERC, 100%; C's, 100%; abstention, 75%). Figure 1 shows the main flows, from 2010 to 2012, given by M1 in this constituency.
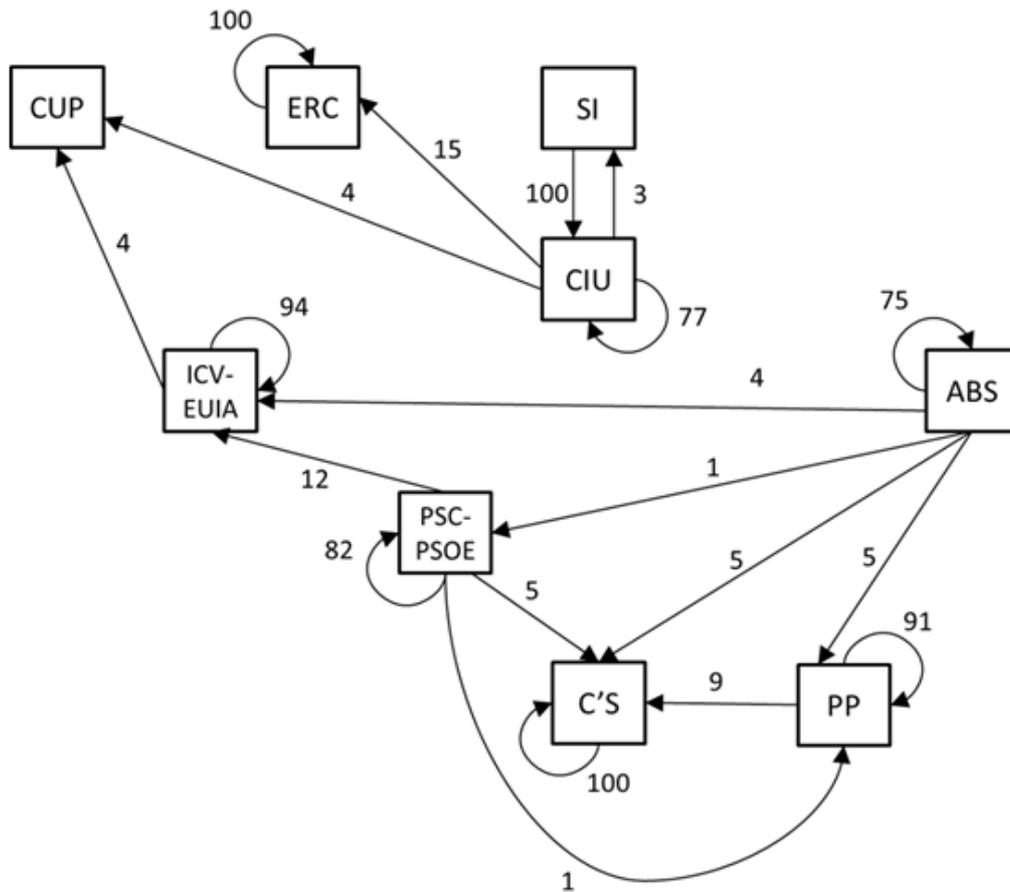
Figure 1. Main flows from 2010 to 2012, given by M1 for the constituency of Barcelona, for the elections to the Catalan Parliament. The numbers next to the arrows are, expressed as a percentage, the corresponding transition rates (those under 1% are omitted). All the options represented in the figure, including abstention and with the exception of CiU, also receive flows from a certain number of minor options

4. Conclusions

This paper describes the problem of computing the transition rates corresponding to two consecutive elections held in the same electoral area. Four mathematical programming models are proposed to deal with it and are applied, regarding the 2010 and 2012 elections to the Catalan Parliament, to the four constituencies in which Catalonia is divided for this purpose.

The models include constraints to guarantee that the total number of votes obtained in the constituency by an option at t' equals the number that results when applying the transition rates to the number of votes corresponding to t. The main features of the obtained results are commented. Mathematical programming reveals itself as a powerful flexible tool to calculate transition rates matrices.

Acknowledgement

References

Andreadis, I., & Chadjipadelis, T. (2009). A Method for the Estimation of Voter Transition Rates. *Journal of Elections, Public Opinion and Parties*, 19, 203-218. http://dx.doi.org/10.1080/17457280902799089

Antweiler, W. (2007). Estimating voter migration in Canada using generalized maximum entropy. *Electoral Studies*, 26, 756-771. http://dx.doi.org/10.1016/j.electstud.2007.07.005

Benewick, R.J., Birch, A.H., Blumler, J.G., & Ewbank, A. (1969). The floating voter and the liberal view of representation. *Political Studies*, 17, 177-195. http://dx.doi.org/10.1111/j.1467-9248.1969.tb00634.x

Berrington, H.B. (1965). The General Election of 1964. *Journal of the Royal Statistical Society*, A 128, 17-51. http://dx.doi.org/10.2307/2343436

Brown, Ph. J., & Payne, C.D. (1986). Aggregate Data, Ecological Regression and Voting Transitions. *Journal of the American Statistical Association*, 81, 394, 452-460. http://dx.doi.org/10.1080/01621459.1986.10478290

Butler, D.E. (1952). *The British General Election of 1951.* McMillan.

Butler, D.E. (1953). *The Electoral System in Britain: 1918 – 1951.* Clarendon Press.

Butler, D.E., King, A. (1966). *The British General Election of 1966.* McMillan.

Füle, E. (1994). Estimating Voter Transitions by Ecological Regression. *Electoral Studies*, 13, 313-330. http://dx.doi.org/10.1016/0261-3794(94)90043-4

Glynn, A.N., & Wakefield, J. (2010). Ecological inference in the social sciences. *Statistical Methodology*, 7, 307-322. http://dx.doi.org/10.1016/j.stamet.2009.09.003

Goodman, L.A. (1953). Ecological Regressions and Behaviour of Individuals. *American Sociological Review*, 18(6), 663-664. http://dx.doi.org/10.2307/2088121

Goodman, L.A. (1959). Some Alternatives to Ecological Correlation. *American Sociological Review*, 64(6), 610-625. http://dx.doi.org/10.1086/222597

Greiner, D.J., & Quinn, K.M. (2009). *R* x *C* ecological inference: Bound, correlations, flexibility and transparency of assumptions. *Journal of the Royal Statistical Society*, 172, Part 1, 67-81. http://dx.doi.org/10.1111/j.1467-985X.2008.00551.x

Grofman, B., & Merrill, S. (2004). Ecological Regression and Ecological Inference. In King, G., Rosen, O, Tanner, M. A., eds., *Ecological Inference. New Methodological Strategies*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511510595.007

Hawkes, A.G. (1969). An Approach to the Analysis of Electoral Swing. *Journal of the Royal Statistical Society*, A 132, 68-79. http://dx.doi.org/10.2307/2343756

Hayes, M. (1976). Swing Analysis and Electoral Forecasting. *Operational Research Quarterly*, 27, 329-340. http://dx.doi.org/10.1057/jors.1976.60

Irwin, G., & Meeter, D. (1969). Building Voter Transition Models from Aggregate Data. *Midwest Journal of Political Science*, 13, 545-566. http://dx.doi.org/10.2307/2110071

Johnston, R.J., & Hay, A.M. (1983). Voter Transition Probability Estimates: An Entropy-Maximizing Approach. *European Journal of Political Research,* 11, 93-98. http://dx.doi.org/10.1111/j.1475-6765.1983.tb00045.x

King, G. (1997). *A Solution to the Ecological Inference Problem.* Princeton University Press.

McCarthy, C., & Ryan, T.M. (1977). Estimates of Voter Transition Probabilities from the British General Elections of 1974. *Journal of the Royal Statistical Society*, A 140, 78-85. http://dx.doi.org/10.2307/2344518

Miller, W.L. (1972). Measures of Electoral Change Using Aggregate Data. *Journal of the Royal Statistical Society*, A 135,122-142. http://dx.doi.org/10.2307/2345042

Moores, B. (1987). Some Consequences of a Three-Party Split in a UK General Election. *Journal of the Operational Research Society*, 38, 569-576. http://dx.doi.org/10.1057/jors.1987.99

Robinson, W.S. (1950). Ecological Correlation and the Behaviour of Individuals. *American Sociological Review*, 15, 351-357. http://dx.doi.org/10.2307/2087176

Tziafetas, G. (1986). Estimation of the Voter Transition Matrix. *Optimization*, 17, 275-279. http://dx.doi.org/10.1080/02331938608843128

Upton, C.J.G. (1977). A Memory Model for Voting Transitions in British Elections. *Journal of the Royal Statistical Society*, A 140, 86-95. http://dx.doi.org/10.2307/2344519

Upton, C.J.G. (1978). A Note on the Estimation of Voter Transition Probabilities. *Journal of the Royal Statistical Society*, A 141, 507-512. http://dx.doi.org/10.2307/2344485

Van der Ploeg, C., Van de Pol, F., & Kampen, J. (2006). *A Comparison of Different Estimation Methods with an Application in the Dutch National Elections of 2003 and 2006.* Discussion paper (09030). Statistics Netherlands, The Hague/Heerlen.