Czech Economic Review 6 (2012) 139–154
Acta Universitatis Carolinae Oeconomica

Analysis of Regional Aspects of Voting Behaviour: The Case of Polish Presidential Election

Mariusz Mazurkiewicz*

Received 3 January 2011; Accepted 23 April 2012

Abstract Transition of votes is observed when a voter who voted for a given candidate in the first round of election, votes for another candidate in the runoff round of election. There are two types of transitions. Transition is obvious when preferred candidate is a loser in the first round. In other case transition also may happen. Exact information about the scale of transitions is usually unavailable. There are many opinion pools, even exit pools, but exact transition's data is not collected during election. Ecological regression techniques give an opportunity to obtain quantitative description of electoral behaviour from aggregated data. Aggregated data set is published after election. Ecological regression approach gives reliable results under homogeneity assumption. Homogeneity in this case is considered in term of electoral behaviour. Homogeneity assumption if applied for the whole country is usually not held. Regional decomposition of estimation process, used for the description of voters' behaviour, extends application the ecological regression to large regions or even for the whole country. In this analysis, maximum likelihood approach and regional decomposition of voting results is used. The last presidential election 2010 in Poland is used as an empirical example. Presidential election in Poland consists of two rounds if none wins more than 50% of the votes, there is runoff round. The interval between rounds of election in Poland is two weeks.

Keywords Ecological regression, electorate flows, transition of votes, homogeneity of electorate, decomposition
JEL classification C39

1. Introduction

Voting behaviour expressed by voting results may be generally described by voting data at the individual level. Anonymous voting eliminates availability of this type of statistical information. On the other hand, aggregated data available from published statistics, contains the information about individuals' voting behaviour. This information is given as a results of group decision making process. There exists a common approach in social sciences aimed at using aggregated data to analyse individual behaviour of voters. Data aggregation is based usually on geographic units such as countries and constituencies.

The main goal of the article is to estimate the proportions of voters who voted for the same candidate or who changed their preferences in a candidate's choice between

* Wroclaw University of Technology, Institute of Organization and Management, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland. Phone: +48713204276, E-mail: mariusz.mazurkiewicz@pwr.wroc.pl.

the two stages of elections. Ecological regression model (Goodman 1959) is one of the statistical tools which gives solution of this problem. The method is popular and well known especially in the case of a two-party system. There are also many improvements, modifications and extensions of ecological regression method. Extensions were introduced mainly with respect to municipal election (Füle 1994), and new estimation techniques (King 1997) and application in multi-party, instable system (Mazurkiewicz et al. 2006). In the actual approach the presidential election is considered. The main task in this case is to estimate individual-level transition coefficients from electoral data aggregated over various voting districts.

2. The 2 × 2 case for consecutive elections

The main idea of ecological regression is illustrated on two parties and two consecutive elections approach (see Table 1). The replacement of candidates instead of parties makes this approach proper to use for personal election, e.g. presidential. The simplest case in ecological regression approach is the 2 × 2 case. This case occurs when there are only two competitive parties in two consecutive elections: party 1 and party 2. Let N_{ij} $(i, j = 1,2)$ describe the number of those who vote for party i in the first election and party j in the second election. The situation is presented in a more convenient way where $N_{\cdot j}$ and $N_{i\cdot}$ are marginal values (sum in rows or columns e.g. $N_{1\cdot} = N_{11} + N_{12}$) for the first and the second election respectively: $N_{\cdot j}$ – part of electorate, which votes for the party j $(j = 1,2)$ in the second elections (it doesn't matter which party was supported by this part of the electorate in the first election), and $N_{i\cdot}$ – part of electorate voting for the party i in the first election (it doesn't matter which party will be supported by this part of the electorate in the second elections).

Table 1. Distribution of the electorate between two parties in two consecutive elections

N_{11}	N_{12}	$N_{1\cdot}$
N_{21}	N_{22}	$N_{2\cdot}$
$N_{\cdot 1}$	$N_{\cdot 2}$	

For two parties and two consecutive elections all marginal values are known as a result of elections and all cellular values (N_{11}, N_{12}, N_{21}, N_{22}) could be obtained as a solution of the system of equations. However, the 2×2 case is a simplification; it is not too difficult to show that even such a model is useful in practical analysis. If party 1 is a fraction of electorate taking part in the elections and party 2 stands for the part of the electorate not participating in these elections, the above model is proper to use for evaluation of electorate's flows from absence to participation and vice versa. The participation or non-participation is frequently under investigation of policy makers. The quantitative description of electorate's flows gives complementary information about the dynamic of the electorate in the sense of a politically active part of the society.

In real life, the size of electorate is changing slightly for two consecutive elections. To simplify, the assumption that $N_{\cdot 1} + N_{\cdot 2} \cong N_{1\cdot} + N_{2\cdot} \cong N$ is sufficient. Under this assumption, the process of searching cellular values from Table 1 is quite easy. Let x_i denote proportion of votes obtained by party i in the first election. Thus, $x_i = N_{i\cdot}/N$, and $y_j = N_{\cdot j}/N$, where y_j denotes proportion of votes obtained by party j in the second election. Let t_1 denote the proportion of voters who voted for the same party in the first election and in the second election. Let t_2 denote the proportion of voters who switched to another party between two consecutive elections. The results of the above transformation are presented in Table 2.

Table 2. Transition of the electorate between two parties in two consecutive elections

		Election 2		
		Party 1	Party 2	Total
Election 1	Party 1	t_1x_1	$(1-t_1)x_1$	x_1
	Party 2	t_2x_2	$(1-t_2)x_2$	x_2
	Total	y_1	y_2	1

According to the above notation, the proportion of voters who vote for party 1 in the second elections is a linear combination of x_1 (called ecological regression):

$$y_1 = t_1x_1 + t_2x_2 = t_1x_1 + t_2(1-x_1) = (t_1-t_2)x_1 + t_2$$

The solution of such a classical model is well known (Duncan and Beverley 1953; King 1997; Groofman and Merril 2002).

Generally, the relation between proportions of votes in the first elections and in the second elections is described by the system of the following regression functions:

$$\begin{array}{l} y_1 = t_{11}x_1 + t_{12}x_2 + \varepsilon_1 \\ y_2 = t_{21}x_1 + t_{22}x_2 + \varepsilon_2 \end{array},$$

where ε_1, ε_2 are unobserved disturbances. Thus, normalized 2×2 case of transition of votes between elections is presented in Table 3.

Table 3. Coefficient of transition of votes between parties in two consecutive elections

Party 1	t_{11}	t_{12}
Party 2	t_{21}	t_{22}
Total	1	1

3. The $n \times 2$ case for consecutive elections

Voting results in two consecutive elections can be expressed as a cross-tabulation of election data. Let N denote the size of the whole electorate (see Table 4).

Table 4. Distribution of the electorate in multi-round election (consecutive elections)

		Round II				Total
		Candidate 1	Candidate 2	$\cdots$	Candidate q	(round I)
Round I	Candidate 1	N_{11}	N_{12}	$\cdots$	N_{1q}	$N_{1\cdot}$
	Candidate 2	N_{21}	N_{22}	$\cdots$	N_{2q}	$N_{2\cdot}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	Candidate p	N_{p1}	N_{p2}	$\cdots$	N_{pq}	$N_{p\cdot}$
Total (round II)		$N_{\cdot 1}$	$N_{\cdot 2}$	$\cdots$	$N_{\cdot q}$	N

Contingency table notation is used in Table 4. Marginal frequencies in the last row correspond to aggregated election results of the second round. Marginal frequencies in the last column correspond to aggregated election results for the first round. Cellular frequencies N_{ij} denote the number of voters who voted for candidate i during the first round and for candidate j during the second round.

Results of both rounds are also described by the vector of votes secured by parties taking part in the first or the second election. The aim is to describe a transition-voting process in two consecutive elections with the probabilities of changes of political preferences. In this case, estimating procedure for transition probabilities is necessary when the behaviour of the whole electorate is analysed. A classical ecological approach is based on the assumption that the results of the second election are a linear function of the results of the first election:

$$TX + \varepsilon = Y, \tag{1}$$

where T is a transition matrix, X is a vector of results of the first round, Y is a vector of results of the second round, and ε is a vector of unobserved disturbances.

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1k+1} \\ t_{21} & t_{22} & \cdots & t_{2k+1} \\ \vdots & \vdots & \ddots & \vdots \\ t_{k+11} & t_{k+12} & \cdots & t_{k+1k+1} \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k+1} \end{bmatrix} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{k+1} \end{bmatrix},$$

where:

x_j	–	share of votes for candidate j obtained during the first round;
y_j	–	share of votes for candidate j obtained during the second round;
k	–	total number of candidates, $k = \max(p,q)$;
x_{k+1}	–	abstention during the first round;
y_{k+1}	–	abstention during the second round;
t_{ij}	–	transfer coefficient of votes transferred from candidate j during the first round to votes for candidate i in the second round, $i, j = 1,2,\ldots,k$;
t_{k+1j}	–	transfer of votes for candidate j in the first round to abstention in the second round;
t_{ik+j}	–	transfer of votes from abstention in the first round to votes for candidate i in the second round.

Extended form of general equation (1) is given by the system of equations:

$$\begin{array}{rcl} t_{11}x_1 + t_{12}x_2 + \cdots + t_{1k+1}x_{k+1} + \varepsilon_1 & = & y_1 \\ t_{21}x_1 + t_{22}x_2 + \cdots + t_{2k+1}x_{k+1} + \varepsilon_2 & = & y_2 \\ & \vdots & \\ t_{k+11}x_1 + t_{k+12}x_2 + \cdots + t_{k+1k+1}x_{k+1} + \varepsilon_{k+1} & = & y_{k+1} \end{array}, \tag{2}$$

where coefficients t_{ij} fulfill conditions

$$\sum_{i=1}^{k+1} t_{ij} = 1 \quad \text{for all } j = 1,2,\ldots,k+1.$$

In the system of equations (2), transfer coefficients are unknown. Transfer coefficients have to be estimated from voting results for considered aggregation level. In this approach, complex knowledge about voting results is assumed.

First of all, any estimation approach needs a statistical sample. In the case of ecological regression there exists a simple way to obtain a statistical sample by using aggregated voting results. Usually, the voting results data is available, even on basic electoral districts level. Assumptions that data from every electoral district is homogenous are obviously not held. There is a necessity to divide electoral districts into homogenous groups—in the sense of electoral behaviour. This approach is called decomposition. The main idea of decomposition is to construct it with respect to reasonable assumption that small regions are more homogeneous than large regions. This approach is very convenient, because usually electoral data set is divided into geographically selected voting districts. Homogeneity simply means that transfer coefficients are the same or roughly the same for selected voting districts. This definition of homogeneity is difficult to evaluate in the sense of appropriateness, because transfer coefficients are unknown. Thus, the problem of formal definition of homogeneity of transfer electorate coefficients is avoided.

The procedure of estimation transfer coefficients is based on decomposition approach and the assumption of electoral behaviour homogeneity in small voting district. For instance the maximum likelihood methodology, which guarantees high statistical quality of results, can be used (Mazurkiewicz et al. 2010).

4. Decomposition approach

The homogeneity assumptions play crucial role in the estimation procedure. A simplified approach to use every electoral district's results as a sample is obviously not proper. There is necessity of increasing the level of homogeneity (Mazurkiewicz et al. 2004). The main part of this idea is to divide a given region into m more homogeneous sub-regions. Let vector D contains shares of total number of votes for each sub-region.

$$D = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}, \text{ where } d_r \geq 0 \text{ and } \sum_{r=1}^{m} d_r = 1 \ \forall r = 1, 2, \ldots, m.$$

In this approach, shares d_r are constant for two rounds of election. In each sub-region there exists a unique transition matrix. Matrices: $T^{(1)}, T^{(2)}, \ldots, T^{(m)}$ are defined for each sub-region separately. Let $X^{(1)}, X^{(2)}, \ldots, X^{(m)}$ denote vectors of results from the first round divided into sub-regions, and $Y^{(1)}, Y^{(2)}, \ldots, Y^{(m)}$ vectors for the second round. The main classical regression assumption that transition coefficients t_{ij} are constant in the sense of conditional expectation $E(T^{(k)}/X^{(k)})$, where k is the number of sub-region, plays a crucial role here.

Therefore, for all sub-regions conditions $T^{(r)}X^{(r)} + \varepsilon^{(r)} = Y^{(r)}$ are fulfilled for $r = 1, 2, \ldots, m$, where

$$T^{(r)} = \begin{bmatrix} t_{11}^{(r)} & t_{12}^{(r)} & \cdots & t_{1k+1}^{(r)} \\ t_{21}^{(r)} & t_{22}^{(r)} & \cdots & t_{2k+1}^{(r)} \\ \vdots & \vdots & \ddots & \vdots \\ t_{k+11}^{(r)} & t_{k+12}^{(r)} & \cdots & t_{k+1k+1}^{(r)} \end{bmatrix}; X^{(r)} = \begin{bmatrix} x_1^{(r)} \\ x_2^{(r)} \\ \vdots \\ x_{k+1}^{(r)} \end{bmatrix}; Y^{(r)} = \begin{bmatrix} y_1^{(r)} \\ y_2^{(r)} \\ \vdots \\ y_{k+1}^{(r)} \end{bmatrix}.$$

Decompositions of rounds of the election results fulfill conditions

$$X = \sum_{r=1}^{m} d_r X^{(r)}, \quad Y = \sum_{r=1}^{m} d_r Y^{(r)}$$

for the first and the second round.

In this approach, the common transition matrix T, describing general voting transitions is a function of sub-regional matrices

$$T = f(T^{(1)}, T^{(2)}, \ldots, T^{(m)}),$$

where elements of matrix T are weighted averages

$$t_{ij} = \frac{\sum_{r=1}^{m} d_r t_{ij}^{(r)} x_j^{(r)}}{x_j}, \quad i, j = 1, 2, \ldots, k+1, \ x_j = \sum_{r=1}^{m} d_r x_j^{(r)}.$$

Decomposition of the transition matrix extends area of application of the ecological regression estimators. Such decomposition is sufficient to build a common transition matrix for the whole country or a big region with respect to homogeneity assumption. From statistical point of view, quality of statistical sample in the event of decomposition is in general increasing.

5. Plurality with runoff scheme (the two-round system)

Presidential elections which took place in Poland since 1991 consists of one or two rounds. If in the first round one of the candidates achieved more than 50% of votes, he or she is elected and there is no second round. If none wins more than 50% of votes in the first round, then two most successful candidates are passing to the runoff round and the candidate, who obtains simple majority in the second round, is the winner. A more interesting situation can be observed when two rounds are necessary to elect the president and this type of situation is being analysed. Time period between rounds is two weeks. It means that voters who voted for lost candidates in first round should transfer vote in second round to another candidate or decide to skip runoff round.

To build a mathematical model as a version of classical ecological regression equation let assume that the total number of candidates in presidential election is equal to n. The first round is the pre-selection type procedure and "winners" of the first round are candidates: candidate 1, who received the biggest number of votes in the first round, candidate 2, who received the second biggest number of votes.

Mathematical model of votes' transfers is described by a system of equations. This system of equations contains three equations:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n + a_{1n+1}x_{n+1} \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n + a_{2n+1}x_{n+1} \\ y_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n + a_{3n+1}x_{n+1} \end{aligned} \quad ,$$

where:

x_1 – share of votes obtained by candidate 1 in the first round;
x_2 – share of votes obtained by candidate 2 in the first round;
x_3 – share of votes obtained by candidate 3 in the first round;
$\vdots$
x_n – share of votes obtained by candidate n in the first round;
x_{n+1} – abstention in first round;
y_1 – share of votes obtained by candidate 1 in the second round;
y_2 – share of votes obtained by candidate 2 in the second round;
y_3 – abstention in the second round;
a_{ij} – transfer coefficients.

Results of the first round fulfill the conditions

$$x_1 + x_2 + x_3 + \cdots + x_n + x_{n+1} = 1, \qquad \frac{x_i}{1 - x_{n+1}} \le \frac{1}{2}\ \forall i = 1,2,3,\ldots,n.$$

Candidates 1 and 2 are qualified to the second round. In terms of shares of votes conditions $x_1 > x_i$ and $x_2 > x_i$ are held for every $i = 3, 4, \ldots, n$.

For the second round, condition

$$y_1 + y_2 + y_3 = 1$$

is obviously held.

Transfer coefficients fulfill the following conditions:

$$\begin{aligned} a_{11} + a_{21} + a_{31} &= 1 \\ a_{12} + a_{22} + a_{32} &= 1 \\ a_{13} + a_{23} + a_{33} &= 1 \\ &\vdots \\ a_{1n} + a_{2n} + a_{3n} &= 1 \\ a_{1n+1} + a_{2n+1} + a_{3n+1} &= 1 \end{aligned}$$

$$T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & a_{1n+1} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & a_{2n+1} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & a_{3n+1} \end{bmatrix}$$

and $a_{ij} \geq 0$ for all $i = 1, 2, 3$ and $j = 1, 2, 3, \ldots, n+1$.

Let us assume that the share of votes for candidate 1 in the first round is bigger than share for candidate 2, $x_1 > x_2$.

Transfer coefficients have specific interpretation in two rounds election. For instance coefficient a_{11} is a share of voters who supported the same candidate 1 in both rounds, similar situation is observed for coefficient a_{22} and candidate 2. Both coefficients are recognized as a measures of electorate stability. Coefficients a_{12} and a_{21} contain information about the shares of voters who completely change their preferences in two weeks period between rounds. It is probably the influence of very intensive electoral campaign in last two weeks before the final round. Let us say it is a measure of efficiency of last two weeks campaign. Coefficients $a_{13}, a_{14}, \ldots, a_{1n}$ for candidate 1 and $a_{23}, a_{24}, \ldots, a_{2n}$ for candidate 2 describe "pure" transfers from lost candidates to the winners of the first round.

6. Statistical data

The basic administrative division of the Republic of Poland is based on voivodships division (see Figure 1). Each voivodship is moreover divided into elementary voting constituencies (see Table 5). The total number of elementary voting constituencies is equal to 25,773. In the calculation, the data coming from 263 elementary voting constituencies is excluded. These 263 elementary voting constituencies were established as pooling stations for voters being abroad in days of election. Number of votes coming from 263 excluded elementary voting constituencies is arbitrary recognized as non significant in the term of results for whole election. In the first round of election it was less than 1% of votes, in second round less than 1.2% of votes.

Table 5. Statistical data: presidential election 2010

No.	Voivodship (sub-region)	Elem. voting constit.	People entitled to vote	Votes 1st round	Votes 2nd round
1	Dolnośląskie	1,826	2,279,205	1,244,829	1,213,014
2	Kujawsko-pomorskie	1,402	1,622,392	859,574	832,984
3	Lubelskie	1,806	1,730,376	908,864	947,013
4	Lubuskie	665	796,241	406,598	393,385
5	Łódzkie	1,717	2,022,693	1,140,991	1,130,665
6	Małopolskie	2,270	2,552,790	1,491,017	1,530,285
7	Mazowieckie[1]	3,293	4,301,592	2,475,296	2,452,615
8	Opolskie	812	817,546	383,456	381,530
9	Podkarpackie	1,684	1,668,328	900,709	932,837
10	Podlaskie	878	941,976	492,285	520,710
11	Pomorskie	1,309	1,718,400	1,015,184	1,056,302
12	Śląskie	2,716	3,649,473	2,032,279	1,980,800
13	Świętokrzyskie	948	1,038,362	514,413	546,264
14	Warmińsko-mazurskie	1,011	1,125,908	557,948	570,403
15	Wielkopolskie	2,120	2,666,486	1,497,318	1,437,626
16	Zachodniopomorskie	1,053	1,339,612	717,694	725,002
	Total	25,510	30,271,380	16,848,831	16,638,455

Source: www.pkw.gov.pl.

In terms of reliability is significantly better to assume homogeneity of electorates' behaviour for small region than for whole country. In this analysis homogeneity assumptions is valid for sub-regions of Poland called voivodships. Decomposition approach requires to estimate transition coefficients for electorates in each voivodships separately. According to this condition the number of wards in voivodships is a size of sample used in estimation process.

7. Polish presidential election 2010

In the 2010 presidential election in Poland 10 candidates took part in first round (exact results in Appendix). Winners (position 1st and 2nd) of the first round achieved 78% of votes.

The decomposition approach was used to calculate common transition matrix for the whole country. In the first step transition matrices for each voivodships were calculated separately, in the second step common transition matrix for whole country was calculated from voivodships results. The simple administrative division was used as a base of regional decomposition.

[1] Usually results from voting wards from abroad are joined to results from Mazowieckie voivodships. In this case votes from abroad elementary voting constituencies are excluded.

Source: www.pkw.gov.pl

Figure 1. Administrative division of Poland, 2010

The electorate stability in terms of two main candidates (participants of the second round) is very high. Estimated parameters a_{11} and a_{22} (see Table 6) have almost the same values. Over 99% of electorate candidate 1 and 2 vote for them again in the second round. Over 90% of voters who didn't take part in the first round, didn't take

Table 6. Transfer coefficients for presidential election 2010 in Poland

i	a_{1i}	a_{2i}	a_{3i}
1	0.9942	0.0000	0.0058
2	0.0014	0.9965	0.0021
3	0.2541	0.5712	0.1747
4	0.6010	0.0441	0.3549
5	0.1699	0.7368	0.0933
6	0.4627	0.2293	0.3080
7	0.5033	0.1677	0.3290
8	0.6845	0.0337	0.2818
9	0.0925	0.8237	0.0838
10	0.3035	0.4433	0.2532
11	0.0200	0.0734	0.9066

part in the second round either. The electorate flow from abstention to participation is lower than 10% (coefficient a_{311}, Table 6). The remaining of parameters describe natural vote's flows from lost candidates to the second round participants, who are the winners of the first round.

Candidate 7 had third result in the first round. His electorate's behaviour is very significant in the second round. Transfer of voters, who supported him in the first round for whole country is described by coefficients a_{17}, a_{27} and a_{37}. Parameter a_{17} is share of voters of candidate 7 in the first round, who voted for candidate 1 in the second round, analogously a_{27} is a share of voters who voted for candidate 2 and a_{37} is a share of voters decided to skip the second round (see Figure 2).

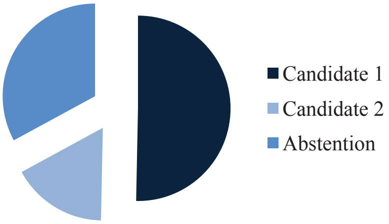


Figure 2. Flows of candidate 7 electorate

Table 7. Stability coefficients for sub-regions

Sub-region	a_{11}	a_{22}	a_{311}
1	1.0000	1.0000	0.9344
2	0.9775	1.0000	0.9424
3	1.0000	1.0000	0.8683
4	1.0000	0.8772	0.9502
5	0.9857	1.0000	0.8370
6	1.0000	1.0000	0.8872
7	0.9797	1.0000	0.8416
8	1.0000	1.0000	0.8474
9	1.0000	1.0000	0.8988
10	1.0000	1.0000	0.9028
11	1.0000	0.9997	0.9833
12	1.0000	1.0000	0.9591
13	1.0000	1.0000	0.8296
14	1.0000	1.0000	0.8874
15	1.0000	1.0000	0.9322
16	1.0000	0.9550	0.9985

The elements of common transition matrix are calculated as weighted averages of elements of regional transition matrices. The results of calculation (see Table 7) in regions are very similar for coefficients a_{11} and a_{22}, except the region number 4. In the case of coefficients a_{311} every value is bigger than 80% but more significant differences with respect to votes for candidates 1 and 2 appear.

Regional coefficients for candidate 7 show distinction in the sense of range. Voters who supported this candidate in the first round of election distributed their votes among candidates number 1 and number 2 in the second round of election (see coefficients a_{17} and a_{27}, Table 8). Some parts of the first round electorate decided not to attend in second round of election (coefficient a_{37}, Table 8).

Table 8. Flows of electorate of candidate 7

Sub-region	a_{17}	a_{27}	a_{37}
1	0.3265	0.3248	0.3487
2	0.5714	0.0897	0.3389
3	0.6574	0	0.3426
4	0.6424	0.2282	0.1294
5	0.3854	0	0.6146
6	0.4455	0.1735	0.3810
7	0.4540	0	0.5460
8	0.5210	0	0.4790
9	0.8419	0	0.1581
10	1.0000	0	0
11	0.6196	0.3804	0
12	0.5542	0.3346	0.1112
13	0.5224	0	0.4776
14	0.2055	0.2440	0.5505
15	0.2986	0.2470	0.4544
16	0.6169	0.3831	0

In the opinion of many analysts (Drozdowski 1997; Alberski 2002; Hołubiec et al. 2008; Riedel 2008), the problem of turnout is very significant in terms of final result. Common opinion in the circle of politicians' campaign strategy advisors based on the statement that the key to win election is in light poll for some candidate or heavy poll for others. From this point of view the nonvoting group of electorate in the first round and their behaviour in the second round is very significant. In 2010 election about 9.5% of the non-voting electorate in the first round decided to take part only in the second round. 2% of them voted for candidate 1, the winner, 7.34% voted for candidate 2.

Detailed information about regional distribution of voters who did not attend in first round and decided to vote in second is given in Table 9. Information about stability of abstention plus information about flows of turnout from first round makes the picture of specific voters' behaviour in the context of the second round completed.

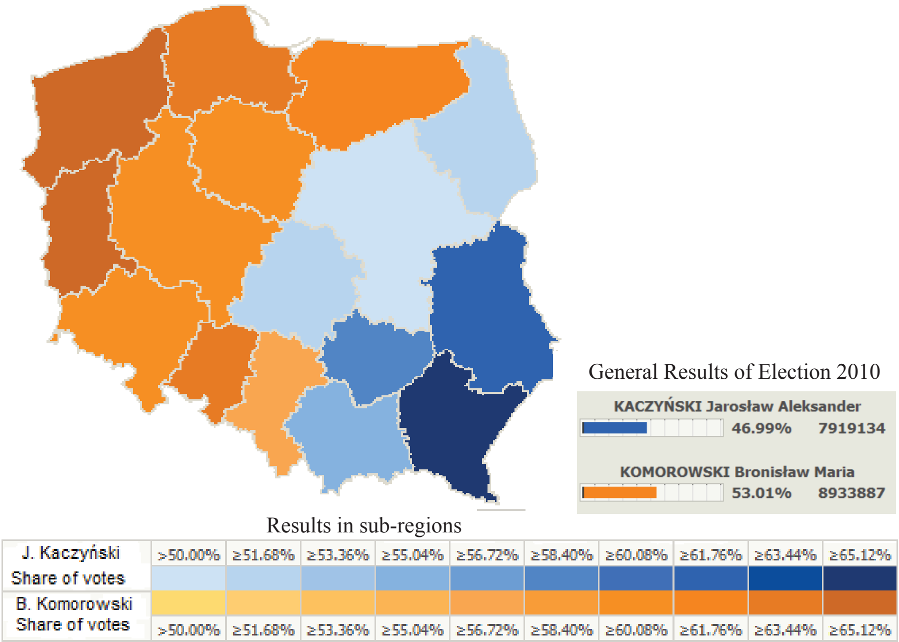
Table 9. Turnout and its flows – regional differences

Sub-region	a_{111}	a_{211}
1	0.0378	0.0278
2	0.0018	0.0558
3	0.0093	0.1224
4	0.0076	0.0422
5	0.0265	0.1365
6	0.0206	0.0923
7	0.0246	0.1338
8	0.0076	0.0422
9	0	0.1012
10	0.0093	0.0879
11	0.0167	0
12	0.0117	0.0292
13	0.0137	0.1567
14	0.0702	0.0425
15	0.0335	0.0343
16	0	0.0015

8. Conclusion

Extended information about presidential election could contain voting results from sub-regions, voting districts – for instance the voivodships. Instead of voivodships another division based on voting wards could be established. In each regional decomposition of results capacity of potential conclusion is significantly smaller than in case of analysis of transition matrix. Some examples of conclusion based on transfer coefficients were given above. Generally the voting behaviour of electorate does not vary in properties. Type of transfers measured by transfer coefficients is independent of regional voting results. In the ecological regression approach division of Poland with respect to presidential election results in not used. The estimation procedure does not depend on partial results in voivodships, approach is uniform in terms of assumption for whole country, however electorate is divided in the sense of political preferences (see Figure 3).

Information about electorate's flows is the key information for analysis of efficiency of electoral campaign on a local level. Generally the strategy of electoral campaign of a given candidate is the same for whole country, but there may be some differences, e.g. regional differences. Election results as the output of campaign are not the same in sub-regions (Figure 3). Investigation of transfer coefficient allows for deep analysis, clearly extends the base of this analysis of voting behaviour during election. At the same time this investigation may be used as a foundation for more sophisticated regional campaign strategy for future election and as a base of predictions' analysis for the next election.



Source: www.pkw.gov.pl

Figure 3. Presidential election 2010: final results (2nd round)

Acknowledgment The author thanks organizers of the research workshop "Voting, Power and Manipulation" at the Institute of Economic Studies, Charles University in Prague for invitation and all participants of this workshop for comments and remarks very helpful in improvement of this article.

References

Aleskerov, F., Gamze Avci, V. I. and Türem, Z. U. (2002). European Union Enlargement: Power Distribution Implications of the New Institutional Arrangements. *European Journal of Political Research*, 41(3), 379–394.

Alberski, R. (2002). Wybory prezydenckie w Polsce w latach 1989–2000 i ich funkcje polityczne (*in Polish*). In Antoszewski, A. (ed.), *Demokratyzacja w III Rzeczypospolitej*. Wrocław, University of Wrocław Publisher, 97–118.

Drozdowski, R. (1997). Regulatory decyzji wyborczych, Przyczynek do teorii zachowa wyborczych (*in Polish*). *Ruch prawniczy, ekonomiczny i socjologiczny*, 1, 159–188.

Duncan, D. and Beverley, D. (1953). An Alternative to Ecological Correlation. *American Sociological Review*, 18, 665–666.

Füle, E. (1994). *Ecological Inference on Voting Data*. Lund, University of Lund, Department of Statistics.

Goodman, L. (1959). Ecological Regression and the Behaviour of Individuals. *American Sociological Review*, 18(6), 663–664.

Groofman, B. and Merrill, S. (2002). Ecological Regression and Ecological Inference. Paper presented at the Ecological Inference Conference, 2002. Harvard University, Cambridge MA. Available at: http://course.wilkes.edu/merrill/.

Hołubiec, J., Szkatuła, G. and Wagner, D. (2008). Kampania prezydencka 2005 – modelowanie preferencji wyborców (*in Polish*). In Trzaskalik, T. (ed.), *Modelowanie preferencji a ryzyko' 08*. Katowice, University of Economics in Katowice Publisher, 45–59.

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

Mazurkiewicz, M., Mercik, J. and Turnovec, F. (2004). Regional Decomposition in Ecological Regression. Proceedings of the 15th International Conference on Systems Science, Wroclaw 2004, 487–493.

Mazurkiewicz, M., Mercik, J. and Turnovec, F. (2006). Transfer of Electorate's Preferences during 2001 and 2005 Parliamentary Elections in Poland *(in Polish)*. In Kacprzyk, J. and Budziński, R. (eds.), *Badania operacyjne i systemowe 2006 – Metody i techniki*. Warsaw, Exit, 175–183.

Mazurkiewicz, M., Mercik, J. and Turnovec, F. (2010). *Transition of Votes for Consecutive Elections – Mathematical Modelling.* Paper presented at the Annual meeting of the European Public Choice Society, Efez, April 2010.

Riedel, R. (2008). Charakterystyka przyczyn niskiej partycypacji politycznej obywateli podczas wyborów parlamentarnych i prezydenckich w Polsce (*in Polish*). In Wrobel, S. (ed.), *Polska demokracja: koncepcje, płaszczyzny, instytucje*. Toruń, Adam Marszałek Publisher, 161–183.

Appendix

Results of presidential election 2010 in Poland expressed as shares of population. Candidates in presidential election represent themselves. Some of them have strong party support. Information about main parties support (in some cases affiliation) is in the third column.

Table A1. Detailed results of presidential election 2010

Can.	Name	Party affiliation	Round I	Round II
1	Komorowski Bronisław Maria	PO, www.platforma.org	22.6571%	28.9742%
2	Kaczyński Jarosław Aleksander	PIS, www.pis.org.pl	19.8885%	25.6832%
3	Jurek Marek	party Prawica RP	0.5755%	–
4	Korwin-Mikke Janusz Ryszard	Wolność i Praworządność	1.3530%	–
5	Lepper Andrzej Zbigniew	Samoobrona RP	0.6966%	–
6	Morawiecki Kornel Andrzej	no party affiliation	0.0701%	–
7	Napieralski Grzegorz Bernard	SLD, www.sld.org.pl	7.4640%	–
8	Olechowski Andrzej Marian	no party affiliation	0.7868%	–
9	Pawlak Waldemar	PSL, www.psl.org.pl	0.9550%	–
10	Ziętek Bogusław Zbigniew	Trade Union “Sierpień 80”	0.0959%	–
11	Abstention		45.4575%	45.3426%
	Turnout		54.5425%	54.6574%

Source: www.pkw.gov.pl