WINNING PREFERENCES.
THE EFFECTIVENESS OF PREFERENCE VOTES
IN THE OPEN-LIST PROPORTIONAL SYSTEM[1]

Adam Gendźwiłł*
University of Warsaw

Jarosław Flis**
Jagiellonian University

Dariusz Stolicki***
Jagiellonian University

Abstract: The paper discusses the use of preference votes in the open-list proportional representation system (OLPR). The rules of the system determine the degree to which votes are effective, i.e. whether voters' preferences are reflected in the distribution of seats. In the OLPR system, 'wasted votes' may refer to both party representation and personal representation that frequently expresses distinct territorial identities. Using the data from elections to the Polish Sejm and local government bodies, the authors analyse the use of preference votes. In the last two decades the share of effective preference votes (cast for elected representatives) was only slightly higher than the share of votes cast for losing candidates in Sejm elections; in the case of local government elections, the latter category of votes visibly prevailed. The authors analyse in detail the differences between the elections, parties and Sejm electoral districts representing various characteristics. Non-parametric regression analysis indicates several factors

that increase the effectiveness of preference votes, in line with theoretical expectations. The smallest share of effective preference votes can be observed in smaller districts with high levels of party and territorial fragmentation.

Keywords: *electoral system, proportional representation system, proportionality, wasted votes, electoral behaviour.*

ZWYCIĘSKIE PREFERENCJE. SKUTECZNOŚĆ GŁOSÓW PREFERENCYJNYCH W SYSTEMIE PROPORCJONALNYM Z LISTAMI OTWARTYMI

Streszczenie: *Artykuł przedstawia problematykę wykorzystania głosów preferencyjnych w proporcjonalnym systemie wyborczym z otwartymi listami (OLPR). Reguły systemu warunkują, w jakim stopniu głosy wyborców są skuteczne, tzn. czy preferencje wyborców przekładają się na mandaty przedstawicielskie. W systemie OLPR „głosy zmarnowane" mogą dotyczyć zarówno reprezentacji partyjnej jak i reprezentacji personalnej wyrażającej często tożsamości terytorialne. Na przykładzie wyborów do Sejmu i organów samorządu terytorialnego w Polsce przedstawiona została analiza wykorzystania głosów preferencyjnych. W ciągu ostatnich 20 lat w wyborach sejmowych głosów oddanych na ostatecznych zwycięzców (zdobywców mandatów) było tylko nieznacznie więcej niż na przegranych; w przypadku wyborów samorządowych głosy oddane na kandydatów bez mandatu wyraźnie przeważały. Autorzy analizują dokładniej różnice pomiędzy wyborami, partiami i okręgami sejmowymi o różnych charakterystykach. Nieparametryczna analiza regresji wskazuje na kilka czynników sprzyjających większemu wykorzystaniu głosów preferencyjnych, zgodnych z teoretycznymi przypuszczeniami. Najmniejsze odsetki skutecznych głosów preferencyjnych obserwuje się w mniejszych okręgach, o wysokim poziomie fragmentacji partyjnej oraz terytorialnej.*

Słowa kluczowe: *system wyborczy, system proporcjonalny, proporcjonalność, głosy zmarnowane, zachowania wyborcze.*

INTRODUCTION

In the Polish variant of the proportional representation system, every constituent needs to indicate their personal preference explicitly, i.e. select a specific candidate from a list by casting a so-called preference vote. After the seats are divided between electoral lists, the seats are apportioned according to the order of candidates on the list based on the number of votes obtained by them. This is a brief summary

of the open-list proportional representation (OLPR) system, although scholars tend to disagree on the names of its variants[2]. What most of them agree on is that a proportional system with electoral lists proves effective – more effective than majority systems – in reflecting party preferences of the electorate and leads to a markedly lower percentage of 'wasted votes', i.e. votes cast for political groupings that ultimately won no seats (Reynolds et al., 2008; Anckar, 2003). However, this reflection in proportional systems may be problematic, as the final distribution of seats tends to prompt parties to seek a coalition necessary to form the government and the coalition's platform is only a resultant of the platforms advanced by its constituent parties (Kamiński, 2016). Proportional representation, defined as the similarity between the distribution of votes and distribution of seats, is naturally limited by statutory election thresholds, the method of converting votes into seats, as well as the district magnitude (the number of seats available in a district determines the level of the so-called natural, or effective, election threshold).

However, proportionality at the party level is only part of the equation. In the electoral system used for elections to the Sejm (the Polish Parliament), the European Parliament and a significant part of local government bodies, voters have the right of 'double expression', i.e. voting in favour of an electoral list (party), as well as a specific candidate who possesses certain personal vote-earning attributes (Shugart et al., 2005). Those presumed criteria of personal choice seem to include attributes that are independent of party affiliation, e.g. personal popularity or credibility (a result of holding a public office), but also gender, age and even famous name (like the name of another well-known politician). Most crucially, such attributes cannot be institutionalised in a legitimised way. An example of such institutionalisation is the now-defunct class-based electoral system or reserved political positions for ethnic minorities that are in use in some systems nowadays.

Territorial representation is the prime example of a legitimised distinction of one of a candidate's attributes (Shugart et al., 2005; Tavits, 2010; Crisp et al., 2013). In most electoral systems, a territorial unit is divided into smaller parts (districts) between which the seats are apportioned in order to ensure fair representation of communities so distinguished. In many cases, however, the resulting districts are inconsistent and composed of several units with separate territorial identities. Previous studies have shown that the localness of a candidate is of particular importance in the case of open lists. On the one hand, it is a phenomenon associated with social communication, known as friends-and-neighbours voting and observed in other systems as well (Gimpel et al., 2008; Górecki, Marsh, 2012; Górecki, 2015). On the other hand, the expression of collective identities is its distinct component – a phenomenon also known from other types of voting (Ginsburgh, Noury 2008).

---

[2]   The Polish system, where preferential voting is compulsory, is sometimes also called a flexible list system.

To conclude, preference voting in a proportional system makes personalised and localised (territorial) voting possible. However, it is of particular importance to examine how the rules of the system determine the nature of voting and transform personal preferences into seats in representative bodies.

Naturally, one may easily imagine a situation where two preferences, the partisan and personal one, are at variance. In such cases, some voters prefer a personal choice, as shown by the clear connection between the strength of local candidates and deviations in the parties' results in various elections (Flis 2014, 2016). Similarly, Poles consistently maintain in various opinion polls that personal choice is of the utmost importance to them (Gendźwiłł, Raciborski, 2014). Even if those declarations do not faithfully reflect the heuristic the voter applies on the election day, the answers to the poll question provide a general insight as to what the public regards as a socially expected heuristic. Experimental studies into the personalisation of voting preferences in the Netherlands, i.e. a country with fairly stable party identification, indicate that a part of the voters (approximately 9%) are inclined to vote for any other party, provided that their preferred candidate is put on its list (van Holsteyn and Andeweg, 2010).

It is sometimes argued that whenever a voter cannot avoid a personal choice (e.g. the voter does not know any of the candidates or cannot distinguish between them but has to cast a preference vote), a vote for the first candidate on the list is a de facto vote for the party and does not reflect personal preference. While the concentration of votes at the top of the ballot is evident (Gendźwiłł, Raciborski, 2014), election results do not provide conclusive evidence about what share of votes for list leaders is a consequence of a conscious personal choice and what is only a vote for the party. However, the role of the 'list leader' is undoubtedly a special one and used by both voters and the list designers: being the first on the list does not guarantee yet greatly increases the likelihood of winning a seat.

A notable feature of the Polish variant of the proportional system is that a preference vote has real, rather than purely symbolic, significance. Adam Gendźwiłł and Jacek Raciborski demonstrated that an average of 20-30% of seats in the Polish parliament (Sejm) were secured on account of preference votes (the candidates would not have won them if a closed-list system had been in place; Gendźwiłł, Raciborski, 2014), more than in similar electoral systems adopted in Belgium, Slovakia and the Czech Republic (Andre et al., 2017: 592; Renwick, Pilet, 2016). However, Jarosław Flis argues that such situations, known as a 'reshuffle', usually result from small differences in support between candidates listed farther down the ballot and are commonly a consequence of intraparty politics intended to weaken rivals within a list. The patterns of those political games often overlap with territorial divisions of

electoral districts. The fact that electoral districts consist of different local communities served by different politicians contributes to greater dispersion of preference votes.

Consequently, one may conclude that a proportional system with compulsory preferential voting gives the voters a kind of a promise of taking into account personal preferences regarding candidates. With the way the system is devised, it is basically impossible to avoid expressing them; Poles value voting 'for a person' more highly than voting 'for a party' and personal choice is also an opportunity to express strong and significant territorial identities within electoral districts.

All the more important is the question of what happens to preference votes and, in particular, to what degree they are 'wasted'. The concept of 'wasted votes' has been employed in electoral studies mainly to denote votes cast for political parties that did not secure representation in any representative bodies (Anckar, 1997). This is a direct reference to Duverger's thesis on the psychological effect of electoral laws, where the author points out that it is the fear of wasting their vote that may prompt voters to vote for parties and candidates who stand a chance of winning a seat (Duverger, 1959). Further in this paper, we argue that in the case of the Polish electoral system, the concept of 'wasted votes' can be also applied to preference votes.

Following a strategy of maximising their election result in the open-list proportional system (OLPR), political parties should put the maximum permissible number of candidates on their electoral lists. Unlike in the single non-transferable vote system (SNTV), parties are generally not harmed by competition between the candidates (Bergman et al., 2013). Incidentally, parties are known to attempt to control and structure it in various ways in order to mobilise a greater number of voters and influence personal allocation of seats (Flis, 2015).

As observed by Bergman et al.,

> 'Votes are wasted on the intraparty dimension if a party has more candidates on a list than can possibly win. Under OLPR we expect more votes to be wasted in this way than under SNTV... If voters are indifferent to candidates, then these wasted votes may not matter. Yet the very presumption of having preference voting is that voters are not indifferent, and benefit from intraparty choice.' (Bergman et al., 2013: 323)

The authors cited above compare the so-called logic models of the SNTV and OLPR systems, demonstrating that their very structure dictates that more preference votes are wasted, i.e. given to candidates who fail to win a seat, and the disproportions in the support for candidates within a single list is greater in the OLPR system. Due to the non-transferability of votes in the SNTV system, parties

make sure to nominate candidates with as balanced support as possible in every district (Bergman et al., 2013: 329).

This article analyses the use of preference votes based on the Polish Sejm (parliamentary) elections held in 2001-2015 as well as elections to local government councils, namely voivodeship or provincial assemblies (*sejmiki województw*), rural county councils (*rady powiatów ziemskich*) and some city councils or communal councils (*rady miast* or *rady gmin*) in 2002-2014[3].

First, we propose a classification of preference votes according to their significance in the election process. We analyse their diversity in various types of elections and different years, comparing votes cast for candidates of major parliamentary parties in the period analysed. Further in this paper, we focus on the votes that translated into seats in representative assemblies, regarding them as a significant indicator of the quality of representation that the proportional system with preference voting ensures. Our quantitative analyses examine what systemic parameters determine the percentage of preference votes used. We analyse how this share is connected with the magnitude and fragmentation of the district and the direct consequence of those parameters, i.e. the number of seats obtained by a specific list. The aim is to verify to what extent those factors – on which the voters have no direct impact – may influence the effectiveness of preference votes assessed from the perspective of a voter.

PREFERENCE VOTE CATEGORIES

In the electoral system under analysis, a part of valid votes translates into seat allocation. In the case of votes cast for candidates who won a seat, the substance of both party and personal preference of voters is reflected in the representative body. Given the distinctive nature of being listed first on the ballot and a clear domination of list leaders in districts, our further analyses make a distinction between **the candidates who obtained seats while running as listed first (seat – 'leaders')** and **those who obtained seats while running listed farther down the list (seat – other)**.

Another part of valid votes translates into seat allocation only to a certain degree, counting as votes cast for a list and affecting the division of seats between the lists which occurs first, but not providing enough support to secure a seat for the candidates selected. Jarosław Flis vividly describes them as '**beaters**', pointing out that votes they receive constitute a major portion of votes cast in Sejm elections (Flis,

---

[3] Proportional representation was the applicable system in all *gminy* (*communes*) with a population of over 20,000 in 2002-2010 and only in *miasta na prawach powiatu* (cities with a powiat status, city counties) in 2014.

2014); in English, the term 'list pushers' is sometimes used in this context (Bergman et al., 2013).

Finally, there are valid votes that do not directly translate into seat allocation and, in this sense, are compared to votes wasted in majority systems (FPTP or block voting). Such votes belong to one of the two categories: they are either **given to candidates from the lists that did not exceed the statutory threshold** that entitles the lists to be included in seat allocation or given to candidates from the lists that did exceed the statutory threshold but received insufficient support in a given electoral district to win a seat (in other words, they **did not exceed the effective election threshold**). These two categories of votes are 'wasted votes' in the sense in which they most often appear in empirical studies on electoral system proportionality (Anckar, 1997).

Consequently, one may assume that those preference votes that translate into seats in a representative assembly are used to the fullest extent, as this is when the electoral system takes into account both party and personal preferences expressed by the voters. We define such votes as effective preference votes, which complement 'wasted preference votes', i.e. votes that did not translate into seats for candidates.

The percentage of efsective preference votes, calculable at the list or district level, indicates to what degree an electoral system reflects personal preferences. The bigger the percentage, the better the reflection of 'double expression' of preferences encouraged by the Polish electoral system.

What happens to preference votes

First, let us examine what happens to preference votes in elections at various levels held under the proportional representation system with preference voting. This is illustrated by the diagram in Figure 1.

The comparison of elections at various levels shows both clear similarities and significant differences in existing patterns. The share of votes cast for lists that failed to obtain a seat in all types of elections in a given electoral district falls within the range of 10 and 20%. However, there are different proportions between ones that did not obtain seats due to the statutory threshold and ones for which it was the effective threshold that proved unreachable. The share of votes cast for lists that fell below the statutory threshold is higher in elections to the Sejm and provincial assemblies whereas the second type is increasingly prevalent in elections to lower tiers of local government. Presumably, support for marginal groupings would be lower on the local political scenes. Votes cast for lists that did not receive seats tend to occur due to a smaller magnitude of electoral districts.
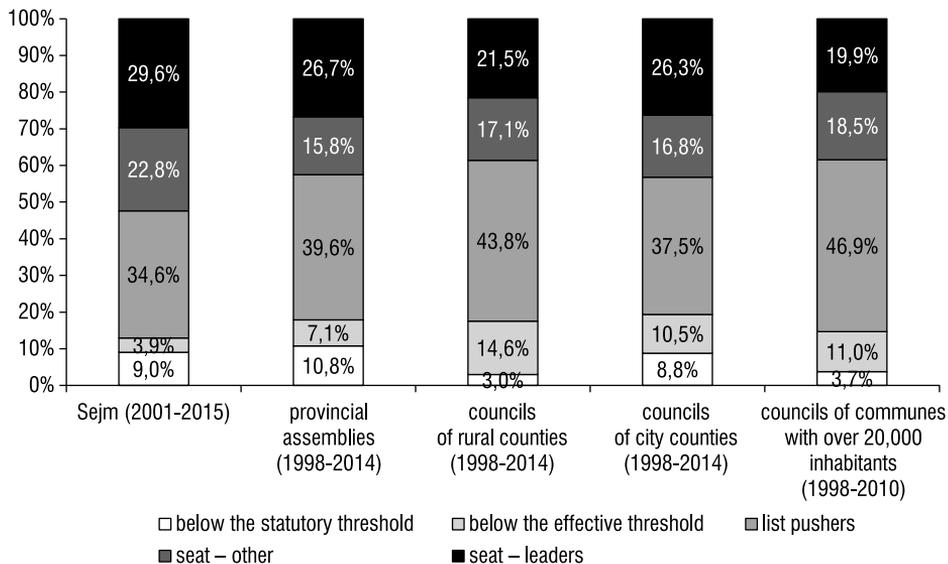
Figure 1. Preference votes in elections at various levels held in Poland under the proportional representation system with preference voting

Votes belonging to individual categories were added up for the entire period under analysis.
Source: PKW (State Electoral Commission), authors' calculations.

The share of votes cast for victorious candidates listed first falls within the range of 20 and 30% for all types of elections. The concentration of preference votes in the list leaders is the greatest in Sejm elections, smaller in elections to provincial assemblies and councils of city counties, smaller still in elections to councils of rural counties and the smallest in elections to communal councils. That difference between elections at various levels becomes even more noticeable if the share of district lists leaders among elected representatives is compared, which is related to the district magnitude; list leaders account for approximately 35% of Sejm deputies and an average of more than 50% of local councillors. This concentration of preference votes in list leaders may be associated with a smaller share of votes cast in local government elections for parties represented in the Sejm with a clearly collective identity. It is also possible that smaller administrative units have fewer candidates with high media exposure who in the case of Sejm elections are responsible for the concentration of votes in list leaders (Flis 2014; Peszyński 2011; Rakowski, 2012).

Between 15 and 25% of all votes cast are given to winning candidates listed farther down the list. While total votes cast for the victorious candidates slightly exceed 50% in the case of Sejm elections, they are definitely below that mark in local elections.

Accordingly, while one may say that voting for eventual winners is slightly more frequent than voting for losers in Sejm elections, votes cast for candidates without a seat (a total of wasted party votes and wasted preference votes) in local elections are in the majority. The majority is very significant in the case of rural counties and communes in rural counties. This is determined by numerous votes cast for defeated candidates from victorious lists. These dominate in each of the categories discussed, constituting from one-third of all votes in Sejm elections to more than 40% in county elections and nearly 50% in communal elections.

Figure 2 presents the distribution of votes cast in Sejm elections for selected nationwide parties, i.e. the four most stable political groupings in the period under analysis (we consider LiD and ZL to be the 'successors' of SLD).

Markedly higher percentages of effective preference votes (cast for seat winners) are typical of PiS and PO (exceeding 60% since 2005), with a significant share of votes cast for list leaders, in particular in the case of PO; those accounted for as many as over 40% of total votes cast for PO in 2007. These two parties are essentially unaffected by the problem of 'wasted' votes in the party context, the only exception being the case of PiS in 2001, when the party failed to win a minimum of one seat in every electoral district. Since then, the support for these two parties in districts has never declined below effective election thresholds.

The status of the two main parties in the Polish party system that PO and PiS achieved after the 2005 election has led to a peculiar balance between three categories of votes, namely votes cast for list leaders, other deputies and defeated candidates. Those proportions have evolved over time, albeit within a limited scope. The share of each of the three categories has always exceeded 20% but has never been higher than half of the total votes received. This diversity stems from evolving support, and hence the number of seats obtained by the lists in districts, but presumably also from changes in list design strategies. It is worth mentioning that very similar proportions between the three categories of votes are also observed in SLD lists in the 2001 election, when the party enjoyed support comparable to that of PO and PiS at the moments of their victories.

The fate of preference votes cast for PSL and the Left (in its various incarnations), smaller parties, is noticeably different. A fall in support for the two parties in the period under analysis is easily perceptible. In the extreme case of Zjednoczona Lewica in 2015, the entire list fell below the statutory electoral threshold, which was higher for coalitions. In the same year, slightly over 40% votes for PSL were cast in districts in which the party eventually failed to win any representation. PSL is also characterised by, relatively, the highest share of votes cast for 'list pushers' and the lowest share of votes cast for list leaders. Lately, PSL has won single seats in some

districts; however, the candidates who won those seats do not enjoy above-average support compared to other candidates, as is the case with PiS and PO lists.

PiS

PO

PSL

SLD / LiD / ZL
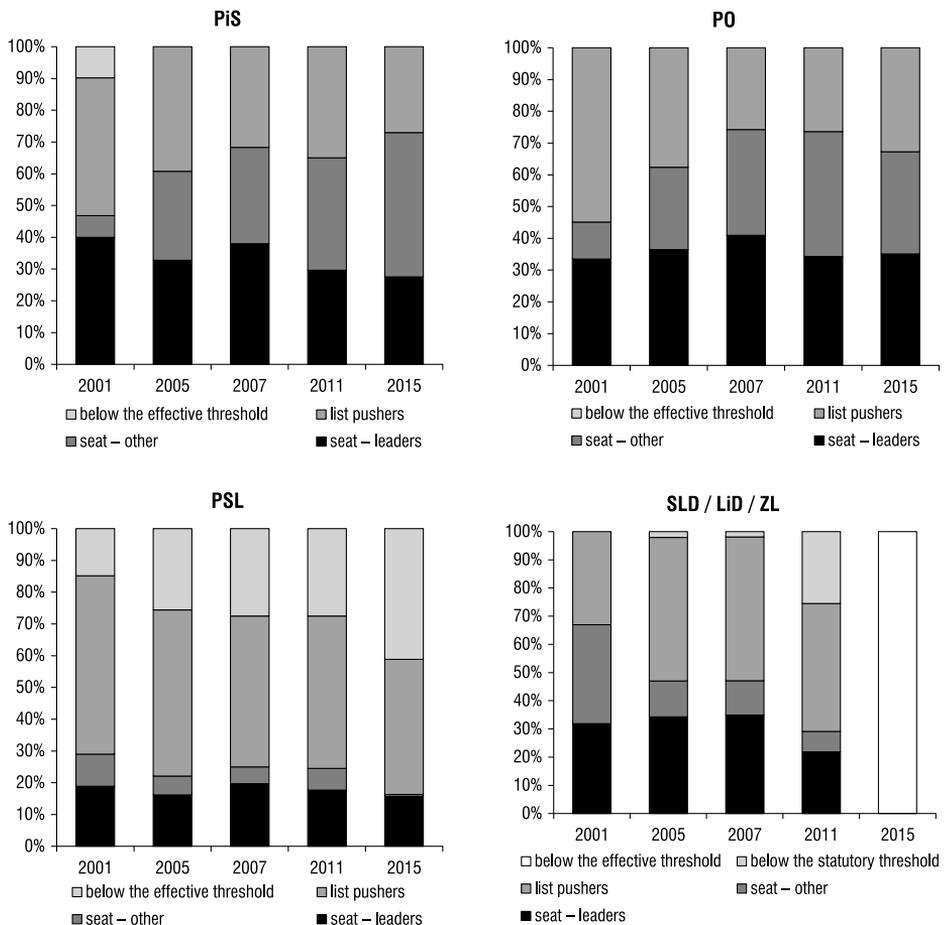
Figure 2. **Preference votes for candidates of selected political parties in Sejm elections in 2001-2015**
Source: PKW, authors' calculations.

Another interesting category is the three 'new parties', i.e. the newcomers to the parliamentary scene: Ruch Palikota in 2011, and Kukiz'15 and Nowoczesna Ryszarda Petru in 2015. All three emerged on the political scene shortly before the election and were established based on the media popularity of their nationwide leader. The share of votes received in districts where those parties failed to exceed the effective threshold is a resultant of the degree to which their support differed in individual

districts as well as subtle differences in the functioning of the proportional system which occur if the nationwide support falls within the range of 8 and 10%. While parties supported by approximately 10% of voters win seats in nearly all districts, parties supported by anywhere near or even below 8% come short of the effective threshold in a significant part of the districts (Flis 2014). The new parties have the highest share of votes cast for list leaders and a marginal share of votes given to other seat winners. This is presumably a result of lacking locally recognised candidates and the way party preference is shaped based on a mediagenic leader, who is nonetheless a candidate in only one electoral district.
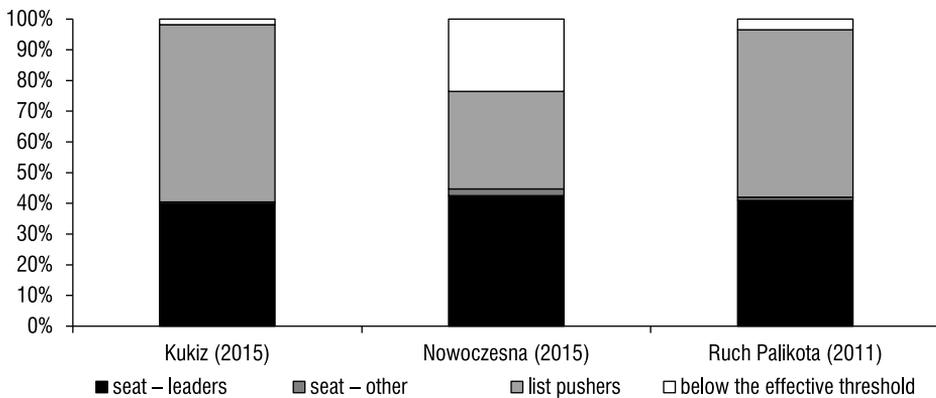
Figure 3. **Preference votes for candidates of selected new Sejm parties in the first parliamentary election in which they obtained seats**

Source: PKW, authors' calculations.

This analysis shows that the fate of preference votes differs systematically in elections at various levels and from one party to another. Presumably, the patterns discussed largely depend on the universal impact of the electoral system in districts of various magnitude and electoral lists of various nature. In order to reconstruct this impact further in this paper we are going to discuss the case of Sejm elections.

WHAT DETERMINES THE EFFECTIVE USE OF PREFERENCE VOTES

Is it possible to predict the percentages of wasted and effective preference votes or determine on what factors and mechanisms they depend by merely looking at the structure of the electoral system? Vote effectiveness is definitely influenced by the

distribution of support for individual parties in the districts, the method of converting votes to seats, as well as the mechanics of the statutory and effective electoral threshold (Taagepera, 1998). In the case of candidate lists that won no seats in a district, it is essentially irrelevant which candidates received the votes. This is known to occur in the case of lists that enjoy relatively weak support and small electoral districts where the party system is highly fragmented.

More can be said about effective preference votes in the case of lists that won a seat. The maximum possible share of effective preference votes is 100%, which may happen if the votes are concentrated only in those candidates who win a seat. The minimum possible share of effective preference votes, i.e. votes that translate into seats, is denoted by the following formula:

$$SGP_{min} = \frac{100\,m}{k}$$

where k is the number of candidates put forward by the party, and m the number of seats won by the party. Most preference votes are wasted when support for candidates within a list is distributed evenly (tie-breaking procedures are required to allocate seats). More seats won by the party demonstrably increase the effectiveness of preference votes but, at the same time, additional candidates put on the list, especially those who exceed the number of available seats in a district, decrease the effectiveness of preference votes. It should be noted that the extreme case defined by the formula may only occur if k is the divisor of the total number of votes cast for the list; in general, due to the indivisibility effect, the minimum may be slightly higher.

The m/k ratio included in the formula is usually defined by the rules of the electoral system. In Poland, until recently, it could vary from $\frac{1}{2\,M}$, where M is the maximum number of available seats in the district, to 1 (when a party wins all seats available in the district and the number of candidates it put forward equals the number of seats available in the district). Following an amendment to the Election Code, the minimum m/k ratio is $\frac{1}{M+2}$, as the maximum length of election lists was shortened. In practice, this means that the share of effective preference votes for a list put up in the largest Sejm electoral district (M=20) could vary from 2.5% to 100% in the previous election and from 4.55% to 100% after the lists were shortened to M+2.

The extreme values occurring empirically in the Sejm elections in 2001-2015 varied from 11.91% to 96.69%, thereby using up a significant part of the range of possible values. These extreme values are a particularly striking illustration of the anomalies involved in the open-list proportional system. In 2001, Andrzej Lepper received 44,814 votes in electoral district no. 40 (Koszalin), which represented 95.63% of total votes cast for Samoobrona RP's list. The next Samoobrona candidate

who became a deputy, Jan Łączny, obtained merely 498 votes; however, his victory was a consequence of adding up votes cast for all candidates from the list. At the other end of the spectrum is the case of district No. 7 (Zamość) in the 2015 Sejm election. The only seat from Polskie Stronnictwo Ludowe's list was won by Genowefa Tokarska, who received fewer than 12% of votes cast for PSL's list of 24 candidates, the maximum length possible.

Consequently, theoretical speculations lead to the conclusion that the percentage of effective preference votes within a list depends on the number of seats the list wins in a district (party magnitude) as well as the number of candidates on the list. However, a simple empirical verification of this hypothesis based on the case of 890 district lists in the 2001-2015 Sejm elections (all lists that won at least 1 seat were taken into account) shows that the indicator in question is in fact determined in any significant way only by party magnitude (Fig. 4).
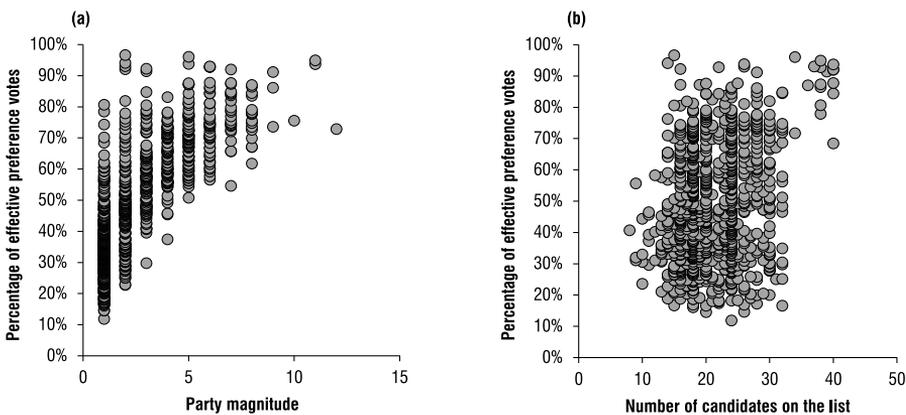
Figure 4. **The connection between the percentage of effective preference votes and the number of seats won by a list (A) and the number of candidates on a list (B) – Sejm elections, 2001-2015**

One dot represents a district list. The diagram takes into account only the lists that obtained seats in districts. Source: PKW, authors' calculations.

This pattern of relationship can be explained by the typical skewness of preference vote distributions within a list. Studies into the effects of candidate name order clearly demonstrate that preference votes are usually heavily concentrated in candidates with a better position on the list (Gendźwiłł, Raciborski, 2014). Expanding lists by additional candidates has a limited impact on the distribution of preference votes between the highest-scoring candidates. Incidentally, modelling distributions of preference votes within a list is a truly interesting research subject; in this context,

the logical models estimating vote shares of the first and last winner within a list in the OLPR system proposed by Bergman et al. (2013) may be a good starting point.

Previous electoral studies in Poland have shown that the distributions of preference votes within lists can be explained by several factors. The existence of effects of candidates' order on the list is well documented; however, those effects are fairly universal. What may differentiate one vote distribution from another is the number of popular candidates, in particular candidates seeking re-election (incumbents). Presumably, whereas one experienced and recognisable candidate accumulates preference votes, a greater number of them leads to greater dispersion.

Previous studies have also shown that the distributions of preference votes within lists differ between districts with different levels of territorial fragmentation. The territorial fragmentation of districts or, in other words, the division of a district into smaller territorial communities with local identities and particular interests, prompts constituents to the so-called friends-and-neighbours voting. Expressed through preference votes, personal choice is therefore also a display of loyalty to a specific territorial community within a district. Evidence for the existence of such voting patterns in the Polish system was presented by Flis (2015) and Górecki (2015), among other scholars.

The impact of those additional factors determining the use of preference votes is illustrated by the charts in Fig. 5 and 6.

The cases that were selected are two district lists of two major parties (PO and PiS) in the 2015 Sejm elections in two electoral districts with different characteristics. District no. 9 (łódzki) is an example of a metropolitan district with little territorial fragmentation. Dominated by the city of Łódź, it also consists of the suburban łódzki wschodni rural county and the small brzeziński rural county. District no. 37 (koniński), on the other hand, is a typical case of a non-metropolitan district, with high territorial fragmentation. It consists of Konin city county, and 8 rural counties. These include the koniński rural county, which is strongly tied to the city of Konin, and the gnieźnieński rural county, which includes the second largest city in the electoral district.

The charts show the number of votes won by each candidate in each county. The size of the circles is proportional to the number of votes obtained. The top row shows the total result of the list in each county and the rightmost column shows the total result of the candidate in all counties. The winners are marked in bold and incumbents circled in black. For Łódź, two parts of the city were intentionally distinguished based on the division of Łódź into two electoral districts in Senate elections, No. 23 and 24.

Both charts illustrate the more even inflow of votes from individual counties for list leaders and, albeit to a lesser extent, for incumbents running from farther down the ballot (Flis 2014). In the case of candidates placed farther down the list in the koniński district, there is a noticeable concentration of votes in their home counties. Similar territorial differences in support for strong candidates, especially incumbents, were not observed in the łódzki district. In the case of weak candidates listed lower on the ballot, the difference is bigger yet far from what can be observed in the case of certain candidates in the koniński district. In particular, the PO list in Łódź shows a perceptible concentration of preference votes in the incumbent deputies. This can also be observed in the PiS list in the koniński district, where it is nonetheless highly limited by competition from local candidates from different counties that form the electoral district.

The percentage of wasted preference votes is also an interesting characteristic of an electoral district. It may be a good indicator of the quality of political representation ensured by an electoral system with a strong element of voting personalisation. Again, the extreme indicators of preference vote use give an interesting insight into the differentiation of electoral districts in this regard. In districts no. 12 (chrzanowski; 7 seats at the time) and no. 22 (krośnieński; 11 seats) in the 2001 Sejm elections, as many as 73% of preference votes were cast for candidates who failed to win a seat. The most effective use of preference voting occurs in the atypical – in terms of magnitude, candidate profiles and voting patterns – district of Warsaw. In 2007, almost 87% of preference votes were cast for candidates who won seats, compared to slightly more than 85% in 2011. In the period of 2001-2015 under analysis, more votes were cast for losing candidates than seat winners in almost exactly half of the districts. Similar analyses for local elections showed that an even smaller share of votes is cast for seat winners in elections to local representative bodies, in particular rural county councils (cf. Flis, Gendźwiłł, 2017).

Presumably, a significant factor behind the effectiveness of preference votes at the electoral district level is its magnitude (measured as the number of seats available) and party fragmentation. Both factors affect the effective election threshold (and thus 'wasting' votes cast for lists that fail to win a seat) as well as the number of seats won by individual lists. Although candidate lists are longer in larger districts and may consequently disperse preference votes, larger electoral districts should be conducive to a greater use of preference votes as the effective election thresholds are lower there and larger parties usually receive more than one seat, which promotes a greater use of preference vote. At the same time, if the party system is heavily fragmented, even the lists that participate in the allocation of seats obtain few of them, if more than
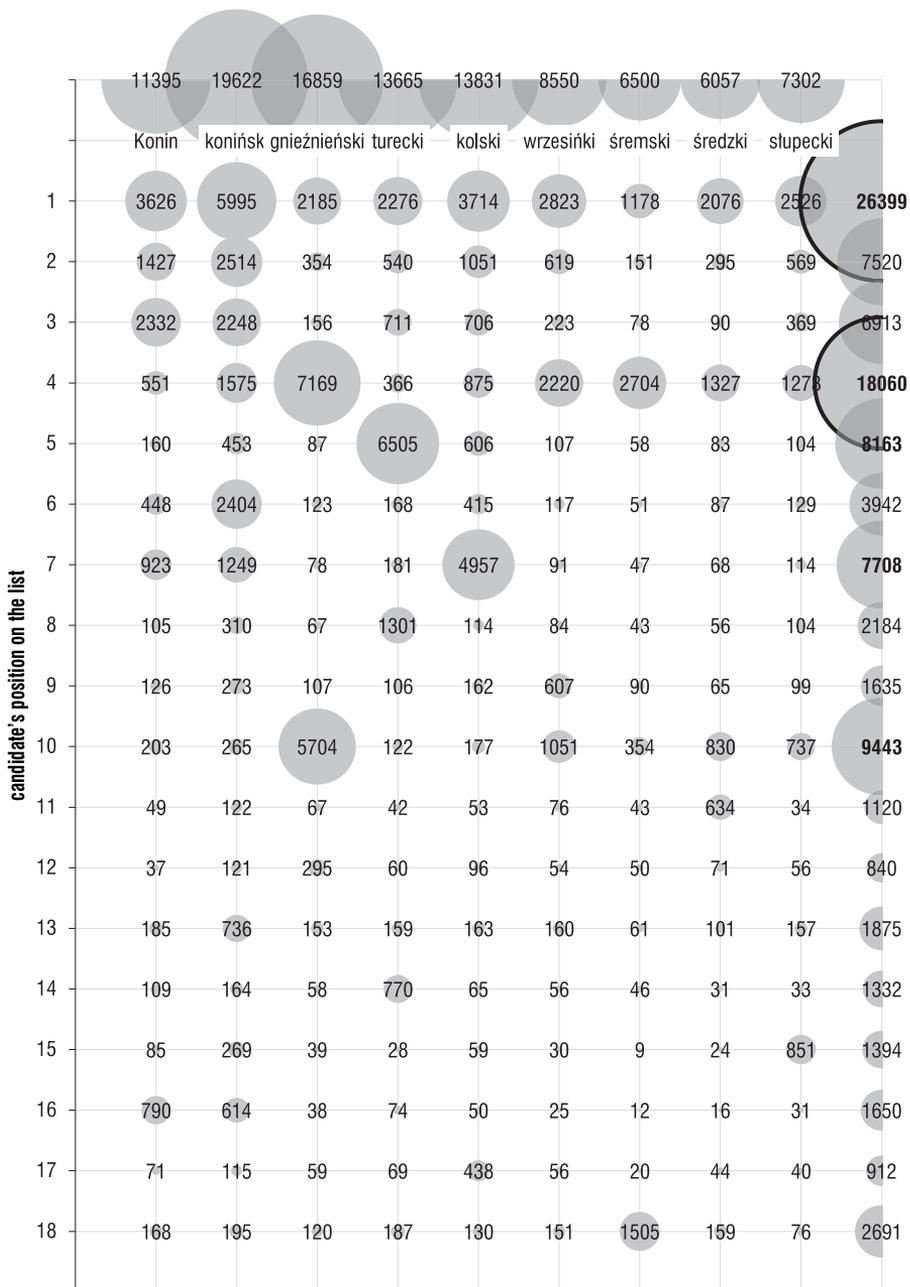
	Konin	konińsk	gnieźnieński	turecki	kolski	wrzesińki	śremski	średzki	słupecki	
	11395	19622	16859	13665	13831	8550	6500	6057	7302	
1	3626	5995	2185	2276	3714	2823	1178	2076	2526	**26399**
2	1427	2514	354	540	1051	619	151	295	569	7520
3	2332	2248	156	711	706	223	78	90	369	6913
4	551	1575	7169	366	875	2220	2704	1327	1273	**18060**
5	160	453	87	6505	606	107	58	83	104	**8163**
6	448	2404	123	168	415	117	51	87	129	3942
7	923	1249	78	181	4957	91	47	68	114	**7708**
8	105	310	67	1301	114	84	43	56	104	2184
9	126	273	107	106	162	607	90	65	99	1635
10	203	265	5704	122	177	1051	354	830	737	**9443**
11	49	122	67	42	53	76	43	634	34	1120
12	37	121	295	60	96	54	50	71	56	840
13	185	736	153	159	163	160	61	101	157	1875
14	109	164	58	770	65	56	46	31	33	1332
15	85	269	39	28	59	30	9	24	851	1394
16	790	614	38	74	50	25	12	16	31	1650
17	71	115	59	69	438	56	20	44	40	912
18	168	195	120	187	130	151	1505	159	76	2691

candidate's position on the list

Figure 5. **The distribution of votes cast for PiS candidates in electoral district no. 37, 2015 Sejm elections**

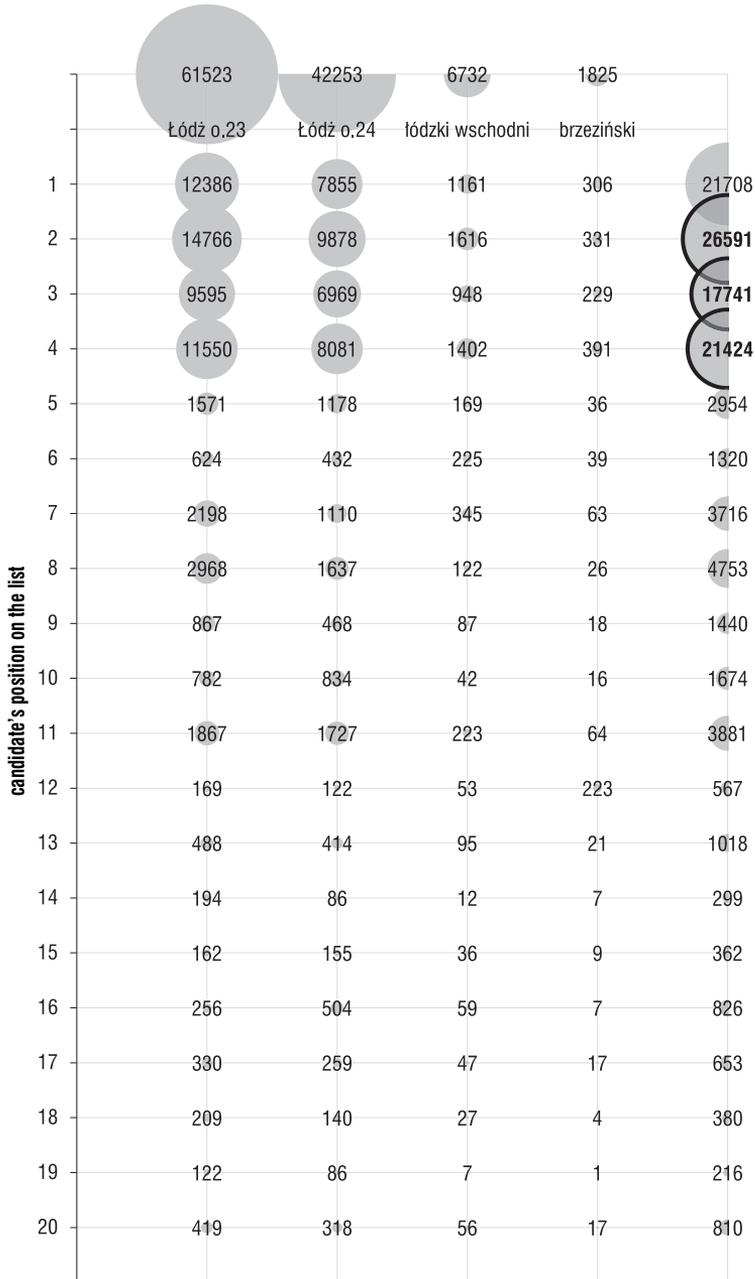The winners' results are in bold. The emboldened circles represent incumbents.
Source: PKW.

	Łódź o.23	Łódź o.24	łódzki wschodni	brzeziński	
	61523	42253	6732	1825	
1	12386	7855	1161	306	21708
2	14766	9878	1616	331	**26591**
3	9595	6969	948	229	**17741**
4	11550	8081	1402	391	**21424**
5	1571	1178	169	36	2954
6	624	432	225	39	1320
7	2198	1110	345	63	3716
8	2968	1637	122	26	4753
9	867	468	87	18	1440
10	782	834	42	16	1674
11	1867	1727	223	64	3881
12	169	122	53	223	567
13	488	414	95	21	1018
14	194	86	12	7	299
15	162	155	36	9	362
16	256	504	59	7	826
17	330	259	47	17	653
18	209	140	27	4	380
19	122	86	7	1	216
20	419	318	56	17	810

candidate's position on the list

Figure 6. **Distribution of votes cast for PO candidates in electoral district no. 9, 2015 Sejm elections**
The winners' results are in bold. The emboldened circles represent incumbents.
Source: PKW.

one. As shown before, the fewer seats a list wins, the fewer effective preference votes are cast for candidates from that list. Consequently, greater party fragmentation in a district should presumably lead to a smaller share of effective preference votes.

The chart in Figure 7 illustrates empirical relationships between the percentage of effective preference votes and district magnitude and party fragmentation in Sejm elections held in the period under analysis.
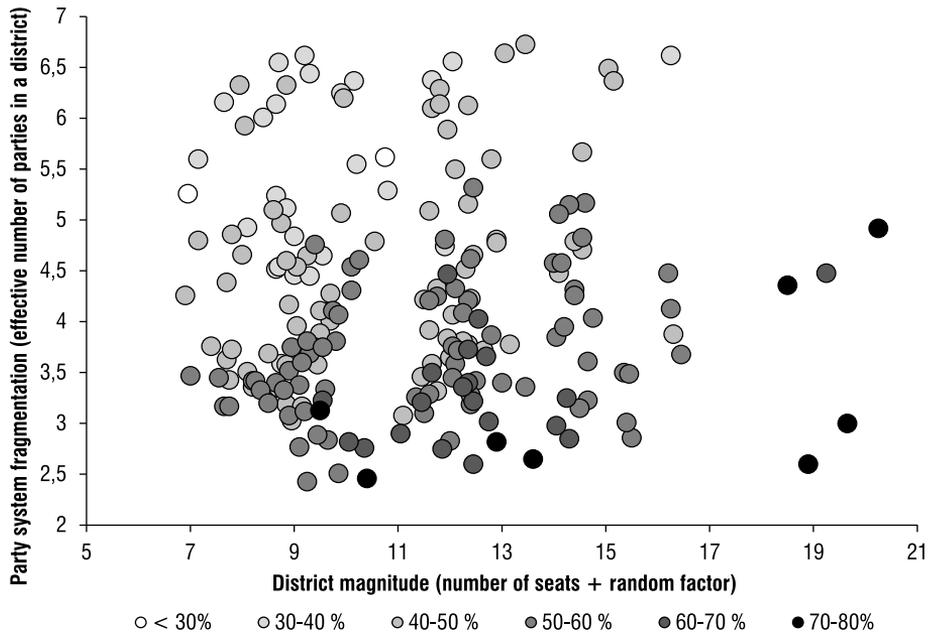

Figure 7. **Percentage of effective preference votes vs party system fragmentation and district magnitude – Sejm elections, 2002-2015**

One dot represents an electoral district. The colour of each dot depends on the percentage of preference votes used and its location within the coordinate system is determined by a combination of district magnitude and the effective number of parties in the district. To preserve clarity of the chart, a random factor – a randomly selected number within the [-0.5, 0.5] range – was added to X-axis values (district magnitude).

Source: PKW, authors' calculations.

The picture that emerges from Fig. 7 essentially confirms the theoretical assumption that the lowest percentages of votes that translated into representative seats generally occur in small districts with heavy party fragmentation (white and light grey dots) while the highest occur in large districts with the lowest level of fragmentation (black dots).

These relationships are further affected by variables that have a direct impact on the distributions of preference votes within electoral lists and were discussed above, namely territorial fragmentation of the district and participation of incumbents seeking re-election in the district.

MODELLING THE USE OF PREFERENCE VOTES

Data and methodology

In order to test factors that explain the effectiveness of preference vote use at the list and district level, we developed explanatory models based on empirical data, i.e. 2001-2015 Sejm election results announced by the State Electoral Commission (Państwowa Komisja Wyborcza, PKW).

Unfortunately, using classical methods of parametric statistics or Monte Carlo methods to model the effectiveness of preference vote is beset with complex problems of statistical nature. As observed before, the effectiveness of preference vote use directly depends on three factors only: the number of seats available; distribution of votes between the lists, and the distribution of votes within each winning list. The remaining explanatory variables, the impact of which we are going to analyse in this study, affect it only indirectly, by changing the parameters of the aforementioned distributions. Therefore, we are dealing with a number of hidden variables that additionally affect the observable variable in a non-deterministic manner. Consequently, we decided to employ non-parametric methods, i.e. Nadaraya-Watson kernel regression analysis (Nadaraya, 1964; Watson 1964). See the appendix for a detailed explanation of why a non-parametric regression was chosen.

Kernel regression analysis, much like parametric regression, answers the question regarding the expected value of a dependent variable with a given value of an independent variable. However, unlike parametric methods, it does not provide formulae describing the relationship but only numerical results for individual points of the explanatory variable that can be illustrated on a chart. From the technical point of view, kernel regression analysis is a generalisation of the well-established moving average; the value expected for every point in the space of independent variables is simply the average of values of other points in that space calculated with weights given using a predetermined kernel function (Racine, 2008)[4].

---

[4]   In this study, we employed the generalised product kernels method (Li, Racine, 2003), which allows combining continuous and categorical variables within one model. We adopted the second-order Gaussian kernel with a smoothing factor h varying depending on adaptive bandwidth selection (Breiman et al., 1977; Schucany, 1995) for the former and the Aitchison-Aitken kernel for the latter (Aitchison, Aitken, 1976). For calculations, we employed the R np package (Racine, Hayfield, 2017).

For the purpose of our analyses, we adjusted two models.

In the first model, we studied the effectiveness of preference vote use for every electoral list that won at least one seat in a district. The response variable was the ratio of effective votes to all votes cast for the list in a given district. We regarded it as a continuous variable within the range (0,1). In this model, we took into account seven explanatory variables:

(1) number of seats obtained by a given list (party magnitude)

(2) the number of candidates on the list

(3) the number of incumbents seeking re-election from a given list (this value was normalised by dividing it by the number of seats in the district in order to allow for the differences in the possible number of incumbents between districts of different magnitude)

(4) the effective number of districts as an indicator of territorial fragmentation of the district; this indicator is calculated analogously to the indicator of party system fragmentation, the effective number of parties (Laakso, Taagepera, 1979); its interpretation is highly intuitive, being the reverse of the probability that two randomly selected voters in an electoral district live in the same district.

(5) the existence of a dominant centre in the district (a county or city county whose population constitutes more than half of the total population entitled to vote in the electoral district)

(6) election year (to control time differences)

(7) the party that registered the list (to control fixed effects).

The first three variables were discrete, the effective number of counties was continuous and the latter three were non-ordinal categorical variables.

Introducing the effective number of counties into the model presents a certain difficulty as it is strongly correlated with district magnitude, which in turn translates into the number of candidates on the list, and the existence of a dominant centre. In order to take those relationships into account in the model, we employed a procedure known from path analysis by calculating (using non-parametric regression again) the expected effective number of counties E_{ELP} as the function of district magnitude and existence of a dominant centre in the district; then we included only the quotient of the actual and expected number of counties in the proper model as the indicator of territorial fragmentation.

In the second model, we analysed the effectiveness of preference vote use in the entire electoral district, i.e. the ratio of effective votes (cast for candidates who

won seats) to all votes cast in a given district. In this case, we took into account the following explanatory variables:

(1) district magnitude (the number of seats available)

(2) party fragmentation of the district (the indicator of the effective number of electoral lists calculated based on the number of votes won by the lists), according to the formula proposed by Laakso and Taagepera (1979)

(3) the number of incumbents seeking re-election in the district (normalised as in the previous model)

(4) territorial fragmentation of the district (defined using the effective number of districts, as in the previous model)

(5) the existence of a dominant centre in the district (a county or city county whose population constitutes more than half of the total population entitled to vote in the electoral district)

(6) election year[5].

The first three variables were discrete, territorial fragmentation was continuous and the latter two were non-ordinal categorical variables.

Preference vote effectiveness at the electoral list level

Presented below is the first model referring to the degree to which preference votes are used at the election list level. Table 1 presents the general parameters of the model (smoothing parameters, results of the test of the significance of the correlation coefficient (Racine, 1997; Racine et al., 2006) and goodness-of-fit measures); Table 2 presents conditional expected values for all levels of categorical variables, whereas the charts in Figures 8-9 show marginal relationships for variables with bootstrap confidence intervals[6].

The model indicates that the basic determinant of the effectiveness of preference votes cast for candidates from a given list is the number of seats the list wins; while an increase in the percentage of effective preference votes is visible with each additional seat, there is a certain saturation point at approximately 7 seats obtained in a district (which is relatively rare in Poland, with only the largest parties managing to achieve that in the largest districts) as the role of every additional seat past that mark becomes marginal. Perhaps this 'saturation' may be attributed to the fact that with a certain

---

[5]   What seems particularly important is to control the election year, given that although every election held since 2005 used the D'Hondt method to allocate seats, a modified version of the Saint-Laguë method was employed in the 2001 election. That difference affected the distribution of seats between parties and therefore had an impact on the level of effective election thresholds and effectiveness of preference vote use.

[6]   Due to the structure of the charts, the level of the remaining interval variables was determined at the median level and the categorical variables at their reference levels: district without a dominant centre, 2001, PO list.

number of seats, every major territorial community secures its representative among deputies elected in a given electoral district.

Table 1
Factors explaining the percentage of effective preference votes at the list level – non-parametric regression model parameters

Variable	h smoothing parameter scale factor	p value	
seats won	1	< 0.001	***
candidates on the list	3	0.211	
incumbents	17	< 0.001	***
territorial fragmentation	49	< 0.001	***
existence of a dominant centre	0.0203	< 0.001	***
election year	0.7746	< 0.001	***
political party	0.3255	< 0.001	***
N	890		
R²	0.8563		
MSE	0.0052		

Table 2
Factors explaining the percentage of effective preference votes at the list level – conditional expected values for the response variable for different values of explanatory variables

Variable/level	Expected value	95% confidence interval	
district without a dominant centre	0.4242	0.3908	0.4577
district with a dominant centre	0.5205	0.4799	0.5611
2001	0.4242	0.3928	0.4556
2005	0.4304	0.4005	0.4604
2007	0.4313	0.3991	0.4635
2011	0.4286	0.3980	0.4593
2015	0.4295	0.3991	0.4599
list: Kukiz '15	0.3792	0.3604	0.3979
list: Liga Polskich Rodzin	0.3792	0.3336	0.4249
list: Mniejszość Niemiecka	0.3710	0.3322	0.4099
list: Nowoczesna	0.4052	0.3772	0.4333
list: Prawo i Sprawiedliwość	0.4327	0.3833	0.4821
list: Platforma Obywatelska	0.4242	0.3925	0.4559
list: Polskie Stronnictwo Ludowe	0.2981	0.2724	0.3239
list: Ruch Palikota	0.3846	0.3662	0.4031
list: Samoobrona	0.3791	0.3540	0.4042
list: Sojusz Lewicy Demokratycznej	0.4173	0.3841	0.4505

Similarly to the previously discussed bivariate analyses, a multivariate analysis shows that the impact of list length, i.e. the number of candidates put forward, on the percentage of effective preference votes is minimal. Every list has a considerable
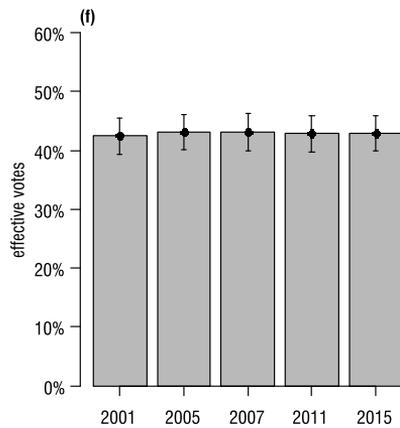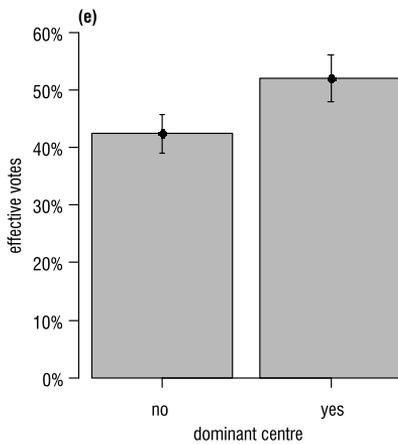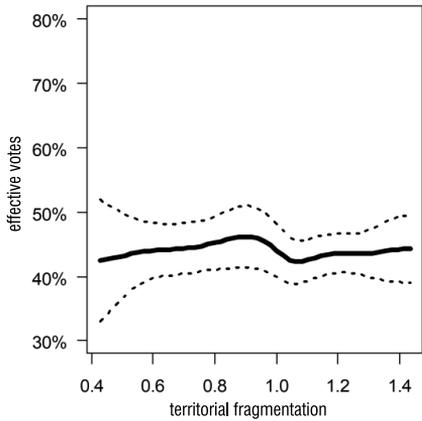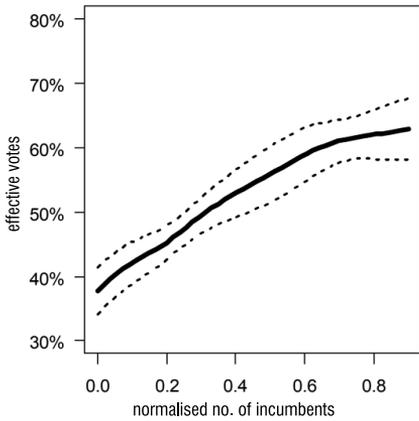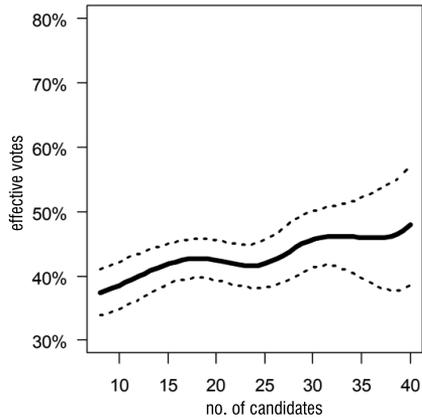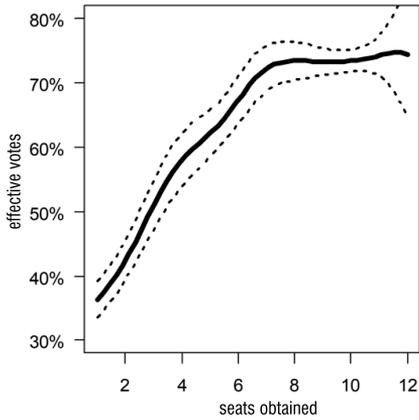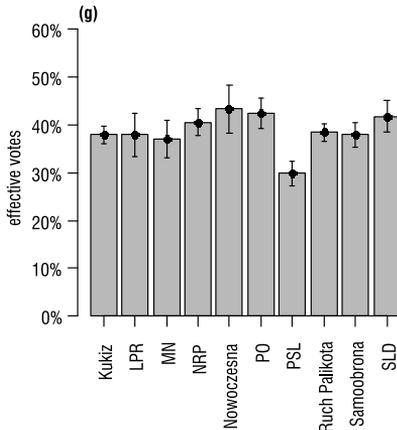
Figure 8. **Charts for marginal effects for explanatory variables – the share of effective preference votes depending on (a) the number of seats won by a given list; (b) the number of candidates on the list; (c) the normalised number of incumbents on the list; (d) the territorial fragmentation rate; (e) the dominant centre in the district; (f) the election year, and (g) the party**
Data: PKW, authors' calculations.

share of list pushers. It is essentially not worth it for parties to put forward the minimum permissible number of candidates.

Territorial fragmentation, as suggested by the initial hypotheses, has a significant impact on the use of preference votes within individual lists, although the impact is relatively little (however, one should note the complex relationship between territorial fragmentation and the number of seats available in a district and obtained by individual parties; internally divided districts usually have more seats). The percentage of effective preference votes is higher for lists competing in districts with smaller territorial fragmentation. The domination of one centre (districts with metropolitan areas in the case of Sejm elections) also favours smaller dispersion of preference votes and therefore a greater degree to which they are used. The question of explaining peculiar distributions of votes in those districts remains open; is it a result of greater homogeneity, less pronounced differences in territorial identities within a district, or a result of greater concentration of votes in the most recognisable candidates (list leaders in major cities are usually among leading politicians on the nationwide scene).

Also in line with expectations, the greater the share of incumbents on the list, the greater the degree of preference vote use.

The differences between individual elections prove minor and statistically insignificant. Instead, what is more distinctive is the differences between the parties analysed. Even with many other attributes controlled, they reveal different patterns of preference vote use (which are likely influenced by different patterns of designing lists and voting). Preference votes were more effective in the case of PiS and PO while being relatively the least effective in the case of PSL.

Preference vote effectiveness at the district level

The models explaining the use of preference votes at the district level are presented in the following tables: Table 3 presents the general parameters of the model; Table 4 presents conditional expected values for all levels of categorical variables, while the charts in Fig. 9 present the marginal effects for explanatory variables with bootstrap confidence intervals.

Table 3
Factors explaining the percentage of effective preference votes at the district level
– non-parametric regression model parameters

Variable	h smoothing parameter scale factor	p value	
district magnitude	1	< 0.001	***
political fragmentation	65	0.010	*
incumbents	22	0.115	
territorial fragmentation	93	0.003	**
existence of a dominant centre	0.0013	< 0.001	***
election year	0.1385	< 0.001	***
N	205		
R^2	0.9343		
MSE	0.0007		

Table 4
Factors explaining the percentage of effective preference votes at the district level – conditional
expected values for the response variable for various values of explanatory variables

Variable/level	Expected value	95% confidence interval	
district without a dominant centre	0.4451	0.4120	0.4782
district with a dominant centre	0.5978	0.5447	0.6510
rok 2001	0.4451	0.4133	0.4768
rok 2005	0.4653	0.4342	0.4965
rok 2007	0.5164	0.4856	0.5473
rok 2011	0.4984	0.4762	0.5206
rok 2015	0.4853	0.4604	0.5102

The models explaining preference vote effectiveness at the district level essentially confirm the hypotheses proposed before. The charts of marginal effects show the difference between smaller (up to 10-11 seats) and larger (approx. 12-15 seats) electoral districts. In larger districts, preference votes tend to be more effective, i.e. more votes ultimately translate to seats. It should be noted here that it is difficult to discern any patterns when interpreting the results from the 19-seat Warsaw district, which is significantly larger than the others.
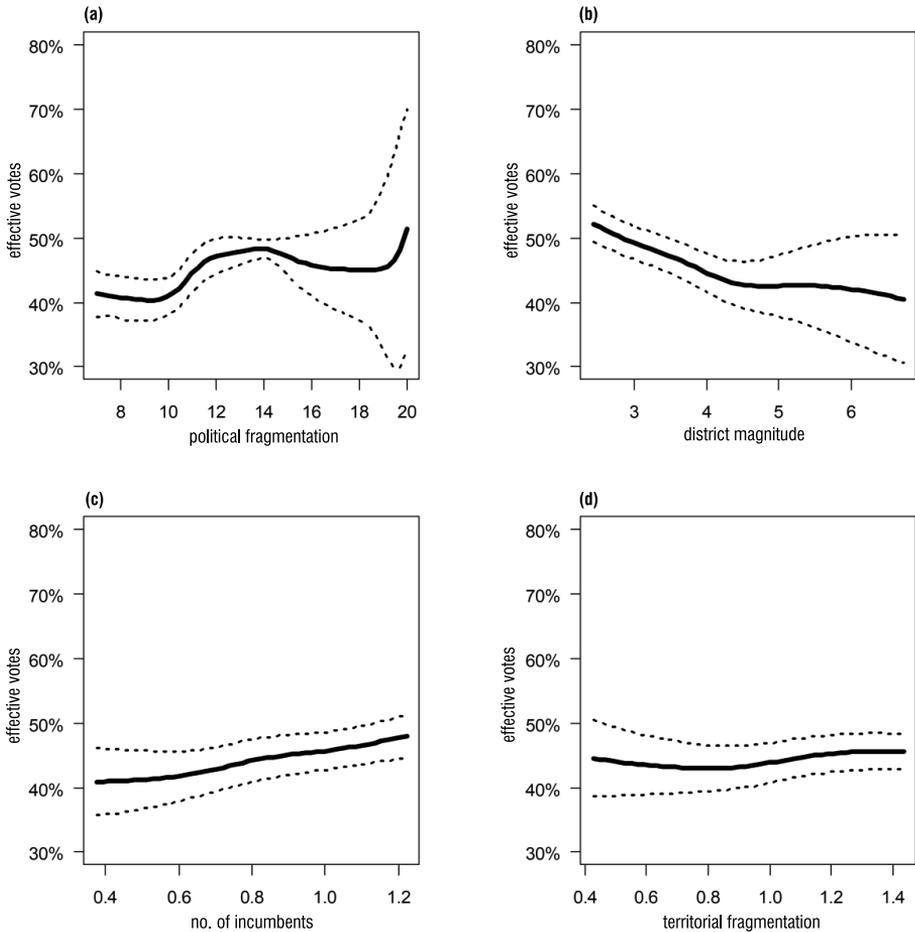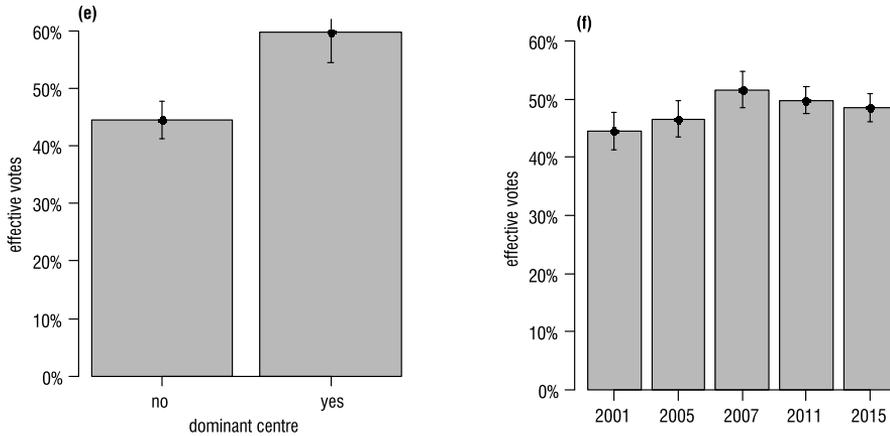
Figure 9. **Charts for marginal effects for explanatory variables – the share of effective preference votes depending on (a) district magnitude (the number of seats); (b) party fragmentation (the effective number of electoral lists); (c) the normalised number of incumbents in the district; (d) the territorial fragmentation rate; (e) the dominant centre in the district, and (f) the election year**
Data: PKW, authors' calculations.

In line with expectations, a greater level of party fragmentation in the district, provided that other factors considered in the model are controlled, decreases the effectiveness of preference votes. As party fragmentation increases, confidence intervals become very wide because such situations are poorly represented in the data analysed. Admittedly, the dependence of effective preference vote use on the presence of incumbents in the district takes an expected direction, but it proves relatively weak and statistically insignificant.

As regards the direct impact of a district's territorial fragmentation on preference vote effectiveness, the only noticeable systematic difference concerns 'metropolitan' districts with a visibly dominant territorial community; such districts produce more effective preference votes, with the difference in the share of effective preference votes estimated at approximately 15 pp. However, it may be presumed that, as far as non-metropolitan areas are considered, the limited importance of territorial fragmentation in this model stems from the fact that parties always try to adjust the placement of candidates on their lists to the territorial structure. For instance, every smaller county has one candidate of its own whereas if there are fewer counties in a district because they are larger, there are now two candidates apportioned for each county, where one of them represents the capital of the county and the second comes from the second largest town. Territorial competition may, therefore, lead to a similar dispersion of preference votes in both cases.

Analyses conducted at the electoral district level reveal more pronounced differences between subsequent parliamentary terms compared to analyses conducted at the list level; those differences are most likely associated with the patterns of seat distribution between electoral lists, although it is possible that the distribution of preference votes within the lists and the very design of the lists by dominating parties changed as well (this requires further analysis). The highest preference vote efficiency was recorded in 2007, which may have resulted from the fact it was a snap election (candidate lists were registered by fewer parties than usual) and a relatively high concentration of votes cast for parliamentary parties. The lowest percentage of effective preference votes was recorded in 2001, when the fragmentation of the Polish party system was significantly higher and the Sainte-Laguë method was the applicable seat allocation system.

CONCLUSION

The percentage of wasted votes, i.e. votes that do not translate to political representation according to the preference expressed by voters is a somewhat intuitive and well-established indicator of electoral system proportionality. Previously, it has practically been employed exclusively to compare the distribution of votes between parties (lists) participating in an election. There is a long-standing belief that fewer votes are wasted in proportional systems as opposed to majority systems, especially in the FPTP (first-past-the-post) system. In this paper, we propose to broaden the perspective on wasted votes (and the complementary effective votes) by analysing preference votes cast in the open-list proportional representation (OLPR) system.

We argue that in the electoral system adopted in Sejm elections preference votes often express a personal preference about a specific candidate or are a display of loyalty to a certain territorial community within an electoral district. An important aspect of assessing such an electoral system is the analysis of what happens to preference votes and, in particular, understanding what has an impact on their effectiveness. Under the Polish electoral system, voters are expected to define a double preference, towards a list and a specific candidate. However, whereas the first preference is of primary importance for the rules of the system, common beliefs and expectations of voters expressed in opinion polls indicate a substantial role of personal choice.

Our analyses show that the patterns of preference vote use differ between elections at various levels, which may be attributed to not only different patterns of the electoral system but also differences in electoral behaviour. In local elections, more votes are cast, on average, for lists that fall below election thresholds, both

the statutory and effective one, while markedly fewer votes are obtained by seat winners, including list leaders.

We propose to define effective preference votes as votes that translate to seats for candidates who obtained those votes. In the analyses discussed in this paper, we show the difference in the percentage of such votes at the list and electoral district level. The models explaining shares of effective preference votes indicate that the most important factor that explains the effectiveness of preference votes is the number of seats won by the list or the number of seats available in the electoral district. In an attempt to generalise the results of the analyses, one may say that larger shares of preference votes are used in larger electoral districts (above 10-11 seats) with a lower level of party fragmentation with one dominant, homogeneous metropolitan area (so-called 'metropolitan districts'). In the case of districts with higher territorial fragmentation resulting from the division of the electoral district into smaller political communities served by various representatives, the share of votes obtained by winning candidates decreases. It is those districts where territorial competition is especially important (Flis 2014). Analyses show that this effect is also strong when the factors unrelated or only indirectly related to the territorial structure, namely district magnitude, political fragmentation and the number of incumbents, are controlled.

The requirement to cast a preference vote can be seen not only as a possibility but also a promise made to voters through the rules of the electoral system. The question to what degree the promise is fulfilled seems particularly significant for the understanding of the quality of political representation ensured by elections.

There remain numerous other issues concerning the use of preference votes that require more detailed research. Most importantly, differences in the distributions of preference votes within lists require further analysis, in particular defining to what degree those distributions may be tactically shaped by list designers. Further attention should also be drawn to the role of territorial divisions within electoral districts. This issue should be revisited in the future.

References

Aitchison, J., & Aitken, C.G. (1976). Multivariate binary discrimination by the kernel method. *Biometrika*, *63*(3), 413-420.

Anckar, C. (1997). Determinants of disproportionality and wasted votes. *Electoral studies*, *16*(4), 501-515.

André, A., Depauw, S., Shugart, M.S., & Chytilek, R. (2017). Party nomination strategies in flexible-list systems: Do preference votes matter?. *Party Politics*, *23*(5), 589-600.

Barakat, H., Abdelkader, Y. (2004). Computing the Moments of Order Statistics from Nonidentical Random Variables. *Statistical Methods and Applications*, 13 (1), 15-26.

Bergman, M.E., Shugart, M.S., & Watt, K.A. (2013). Patterns of intraparty competition in open-list & SNTV systems. *Electoral Studies*, *32*(2), 321-333.

Breiman, L., Meisel, W., & Purcell, E. (1977). Variable kernel estimates of multivariate densities. *Technometrics*, *19*(2), 135-144.

Bukowski, M., Flis, J., Hess, A., & Szymańska, A. (2011). *Opcja czy osoba?: Upartyjnienie versus personalizacja w wyborach samorządowych*. Wyd. UJ.

Coleman, J. S. (1964). *Introduction to Mathematical Sociology*. London: Glencoe.

Crisp, B.F., Olivella, S., Malecki, M., & Sher, M. (2013). Vote-earning strategies in flexible list systems: Seats at the price of unity. *Electoral Studies*, *32*(4), 658-669.

Duverger, M. (1959*). Political parties: Their organization and activity in the modern state*. Methuen.

Eggenberger, F., Pólya, G. (1923). Über die Statistik verketteter Vorgänge. *Zeitschrift für Angewandte Mathematik und Mechanik*, 3 (4), 279-289.

Eggenberger, F., Pólya, G. (1928). Calcul des probabilités – sur l'interprétation de certaines courbes de fréquence. *Comptes rendus de l'Académie des Sciences*, 187, 870-872.

Flis, J. (2014). Głos i rozgłos. Medialne i systemowe determinanty indywidualnego wyniku wyborczego liderów list. *Zeszyty Prasoznawcze*, 2(218): 203-220.

Flis, J. (2015). Reprezentacja powiatów w sejmikach na przykładzie Małopolski. *Acta Politica. Zeszyty Naukowe Uniwersytetu Szczecińskiego*, 33, 32-57.

Flis, J. (2016). Reprezentanci polityczni: reprezentanci handlowi. W: M. Kułakowska, P. Borowiec, P. Ścigaj (red.), *Oblicza kampanii wyborczych 2015*. Kraków: Wyd. UJ.

Flis, J. & Gendźwiłł, A. (2017), *Przedstawiciele i rywale - samorządowe reguły wyboru*. Warszawa: Fundacja Batorego.

Gimpel, J.G., Karnes, K.A., McTague, J., & Pearson-Merkowitz, S. (2008). Distance-decay in the political geography of friends-and-neighbors voting. *Political Geography*, *27*(2), 231-252.

Ginsburgh, V., & Noury, A. G. (2008). The Eurovision song contest. Is voting political or cultural?. *European Journal of Political Economy*, *24*(1), 41-52.

Gendźwiłł, A., & Marcinkiewicz, K. (2018, forthcoming). Interventionism of voters: district size, level of government, and the use of preference votes. *Acta Politica*.

Gendźwiłł, A., & Raciborski, J. (2014). Jak głosują wyborcy w warunkach preferencyjnych list wyborczych: przypadek Polski. *Decyzje*, (22), 47-70.

Górecki, M.A., & Marsh, M. (2012). Not just 'friends and neighbours': Canvassing, geographic proximity and voter choice. *European Journal of Political Research*, *51*(5), 563-582.

Górecki, M. (2015). „Głosowanie przyjacielsko-sąsiedzkie": Przyczynek do badań w kontekście polskim. *Studia Socjologiczno-Polityczne. Seria Nowa*, *1*, 63-74.

Kamiński M.M. (2016). *Ordynacje większościowe i JOW-y. Kompendium reformatora ordynacji wyborczej*. Warszawa: Wyd. Naukowe Scholar.

Katz, J.N., King, G. (1999). A Statistical Model for Multiparty Electoral Data. *American Political Science Review*, 93 (1), 15-32.

Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: a measure with application to West Europe. *Comparative political studies*, *12*(1), 3-27.

Li, Q., & Racine, J.S. (2003). Nonparametric estimation of distributions with categorical and continuous data. *Journal of Multivariate Analysis*, *86*(2), 266-292.

Nadarajah, S., Pal., M. (2008). Explicit Expressions for Moments of Gamma Order Statistics. *Bulletin of the Brazilian Mathematical Society, New Series*, 39 (1), 45-60.

Nadaraya, E.A. (1964). On estimating regression. *Theory of Probability & Its Applications*, *9*(1), 141-142.

Peszynski, W. (2011). Personalne aspekty rywalizacji wyborczej do Sejmu Rzeczypospolitej Polskiej: Analiza zjawiska „jeżdżenia na gapę" na wybranych przykładach. *Studia Wyborcze*, 11, 47-68.

Racine, J.S. (1997). Consistent significance testing for nonparametric regression. *Journal of Business & Economic Statistics*, *15*(3), 369-378.

Racine, J.S. (2008). Nonparametric econometrics: A primer. *Foundations and Trends in Econometrics*, *3*(1), 1-88.

Racine, J.S., Hart, J., & Li, Q. (2006). Testing the significance of categorical predictor variables in nonparametric regression models. *Econometric Reviews*, *25*(4), 523-544.

Racine, J.S. & Hayfield, T. (2017) *Nonparametric Kernel Smoothing Methods for Mixed Data Types*. https://cran.r-project.org/web/packages/np/np.pdf).

Rakowski, M. (2012). „Lokomotywy" czy „odkurzacze" – O wynikach liderów list wyborczych. Polemika z Wojciechem Peszyńskim. *Studia Wyborcze*, 13, 61-77.

Renwick, A., & Pilet, J.B. (2016). *Faces on the ballot: The personalization of electoral systems in Europe*. Oxford University Press.

Reynolds, A., Reilly, B., & Ellis, A. (2008). *Electoral system design: The new international IDEA handbook*. International Institute for Democracy and Electoral Assistance.

Schucany, W.R. (1995). Adaptive bandwidth choice for kernel regression. *Journal of the American Statistical Association*, *90*(430), 535-540.

Shugart, M.S., Valdini, M.E., & Suominen, K. (2005). Looking for locals: voter information demands and personal vote-earning attributes of legislators under proportional representation. *American Journal of Political Science*, *49*(2), 437-449.

Słomczyński, W., Stolicki, D. (2018, w druku). *Matematyczny model wyborów w systemie FPTP*. W: Flis, J., Słomczyński, W., Stolicki, D. (red.), *Skrzywione podziały. W poszukiwaniu gerrymanderingu w wyborach do rad gmin w 2014 r.* Kraków.

Taagepera, R. (1998). Effective magnitude and effective threshold. *Electoral studies*, *17*(4), 393-404.

Tavits, M. (2010). Effect of local ties on electoral success and parliamentary behaviour: The case of Estonia. *Party Politics*, *16*(2), 215-235.

Van Holsteyn, J.J., & Andeweg, R.B. (2010). Demoted leaders and exiled candidates: Disentangling party and person in the voter's mind. *Electoral Studies*, *29*(4), 628-635.

Watson, G.S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 359-372.

APPENDIX. WHY THE NON-PARAMETRIC MODEL WAS CHOSEN

Let n be the number of candidates on the list, s the number of seats won by the list, and $v_i v_i$ the normalised result of the ith candidate (i.e. the ratio of the number of votes cast for the candidate to the number of all votes cast for the list). Note that $\sum_{i=1}^{n} v_i = 1$. The vector $(v_1, ..., v_n)$ is denoted v. Let the candidates be arranged in descending order according to v_i and the result of the kth best candidate be denoted $v_{(k)}$. Note that preference vote use rate p is given by the following function:

$$p(n, v, s) = \sum_{k=1}^{s} v_{(k)}.$$

Now, suppose that the vector v is a realisation of a random variable V with the probability distribution D_n. The variables $V_1, ..., V_n$ are not independent because $\sum_{i=1}^{n} V_i = 1$, but one may find distributions $M_i, ..., M_n$ such that variables $X_1 \sim M_1, ..., Xn \sim M_n$ are independent but, at the same time, the equality $V_i = X_i / \sum_{j=1}^{n} X_j$ is true for each $i \in \{1, ..., n\}$. Using Barpat-Beg's theorem, we can obtain a joint probability distribution $X_{(1)}, ..., X_{(n)}$, which can in turn be used to obtain the distribution of sums $X_{(1)}, ..., X_{(s)}$ and $V_{(1)}, ..., V_{(s)}$ for the determined s. However, in order to simplify the following argument, suppose (cf. Coleman 1964, Katz and King 1999; see also Słomczyński and Stolicki 2018) that the distribution D_n is the Dirichlet distribution with the parameters $(\alpha_1, ..., \alpha_n)$ (see the Pólya-Eggenberger urn model (1923, 1928) for an example of a stochastic model with such a distribution). Consequently, the distributions $M_1, ..., M_n$ will be gamma distributions with the parameters $(\alpha_i, 1)$. Based on the results presented by Barakat and Abdelkader (2004) as well as Nadarajah and Pal (2008), we obtain the formula for the expected value $V_{(k)}$:

$$\mathrm{E}[V_{(k)}] = \sum_{j=k}^{n} (-1)^{j-k} \binom{j-1}{k-1} \sum_{1 \le i_1 \le \cdots \le i_n \le n} \int_0^{\infty} \prod_{t=1}^{j} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) dx.$$

Adding up elements following $k \in \{1, ..., s\}$, we obtain

$$\mathrm{E}[p(n, V, s)] = \sum_{k=1}^{s} \sum_{j=k}^{n} (-1)^{j-k} \binom{j-1}{k-1} \sum_{1 \le i_1 \le \cdots \le i_j \le n} \int_0^{\infty} \prod_{t=1}^{j} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) dx.$$

Note that dependent variables may affect the expected p value in three ways:

- by a discrete change of s
- by a discrete change of n
- by a continuous change of one or more parameters $(\alpha_1, ..., \alpha_n)$.

In the first case, the change is expressible as the difference quotient

$$\Delta_s \mathrm{E}[p] = \mathrm{E}[p(n, V, s+1)] - \mathrm{E}[p(n, V, s)] = \mathrm{E}\left[V_{(s)}\right] =$$

$$= \sum_{j=s}^{n} (-1)^{j-s} \binom{j-1}{s-1} \sum_{1 \le i_1 \le \cdots \le i_j \le n} \int_0^\infty \prod_{t=1}^{j} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) dx.$$

In the second case, it is expressible as the difference quotient

$$\Delta_n \mathrm{E}[p] = \mathrm{E}[p(n+1, V, s)] - \mathrm{E}[p(n, V, s)] =$$

$$= \sum_{k=1}^{s} \sum_{j=k}^{n} (-1)^{j-k} \binom{j-1}{k-1} \sum_{l=1}^{j} \sum_{1 \le i_1 \le \cdots \le i_{j-l} \le n} \int_0^\infty \left(1 - \frac{\gamma(\alpha_{n+1}, x)}{\Gamma(\alpha_{n+1})}\right)^l \prod_{t=1}^{j-l} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) dx +$$

$$\sum_{k=1}^{s} (-1)^{n+1-k} \binom{n}{k-1} \sum_{1 \le i_1 \le \cdots \le i_{n+1} \le n+1} \int_0^\infty \prod_{t=1}^{n+1} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) dx.$$

In the third case, it is expressible as a directional derivative that is a linear combination of partial derivatives that can be calculated applying Leibniz's rule for differentiation under the integral sign:

$$\frac{\partial}{\partial \alpha_i} \mathrm{E}[p] =$$

$$= \sum_{k=1}^{c} \sum_{j=k}^{n} (-1)^{j-k} \binom{j-1}{k-1} \sum_{1 \le i_1 \le \cdots \le i_j \le n} \int_0^\infty \prod_{t=1}^{j} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) (\phi(\alpha_i) - \ln x) \, dx -$$

$$\sum_{k=1}^{c} \sum_{j=k}^{n} (-1)^{j-k} \binom{j-1}{k-1} \sum_{1 \le i_1 \le \cdots \le i_j \le n} \int_0^\infty \prod_{\substack{t=1 \\ t \ne i}}^{j} \left(1 - \frac{\gamma(\alpha_{i_t}, x)}{\Gamma(\alpha_{i_t})}\right) \frac{G_{2\,3}^{3\,0}\left(x \left| \begin{matrix} 1,1 \\ 0,0,\alpha_i \end{matrix}\right.\right)}{\Gamma(\alpha_i)} \, dx,$$

where G is the Meijer G-function.

Note that the expected value of the response variable X in the generalised linear model is a composition of the link function f and the linear combination of the function $g_1, \ldots, g_v$ with the form $g_i(x_i) = \beta_i x_i$, where x_i is the ith response variable and β_i the corresponding coefficient. Given the chain rule, it follows that

$$\frac{\partial \mathrm{E}(X)}{\partial x_i} = \frac{\partial f\left(\sum_{j=1}^{v} g_j(x_j)\right)}{\partial g_i(x_i)} \frac{d g_i(x_i)}{d x_i} = \frac{\partial f\left(\sum_{j=1}^{v} \beta_j x_j\right)}{\partial a_i x_i} \beta_i,$$

where C is a constant, so, with an accuracy of β_i, the derivatives of the response variable following all explanatory variables have an identical form. As can be observed, the three difference operators do not meet this requirement; for nearly all values n, s and $\alpha_1, ..., \alpha_n$, there exists no coefficient c such that $\Delta_s \mathrm{E}[p] = c\Delta_v \mathrm{E}[p]$, $\Delta_s \mathrm{E}[p] = c\frac{\partial}{\partial \alpha_i} \mathrm{E}[p]$ or $\Delta_v \mathrm{E}[p] = c\frac{\partial}{\partial \alpha_i} \mathrm{E}[p]$. Consequently, this precludes the possibility of applying the generalised linear model to analyse the relationships between our variables.