The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

NICEP

Nottingham Interdisciplinary Centre for Economic and Political Research

NICEP Working Paper: 2016-10

The Volatility of Volatility: Assessing Uses of the Pedersen Index to Measure Changes in Party Vote Shares

Fernando Casal Bértoa

Kevin Deegan-Krause

Tim Haughton

Nottingham Interdisciplinary Centre for Economic and Political Research
School of Politics, The University of Nottingham, Law & Social Sciences Building,
University Park, Nottingham, NG7 2RD
ISSN 2397-9771

The Volatility of Volatility: Assessing Uses of the Pedersen Index to Measure Changes in Party Vote Shares

Fernando Casal Bértoa, Kevin Deegan-Krause and Tim Haughton

NICEP Working Paper Series 2016-10

October 2016

ISSN 2397-9771

Fernando Casal Bértoa

The University of Nottingham

fernando.casal.bertoa@nottingham.ac.uk

Kevin Deegan-Krause

Wayne State University

kdk@wayne.edu

Tim Haughton

University of Birmingham

t.j.haughton@bham.ac.uk

THE VOLATILITY OF VOLATILITY:

Assessing Uses of the Pedersen Index to Measure Changes in Party Vote Shares

Fernando Casal Bértoa*
Nottingham Research Fellow
School of Politics and International Relations, University of Nottingham
University Park Campus
NG72RD, Nottingham, United Kingdom
Tel: +44 115 95 14856
Fernando.Casal.Bertoa@nottingham.ac.uk
*corresponding author

Kevin Deegan-Krause
Associate Professor
Department of Political Science, Wayne State University
2040 Faculty Administration Building
Detroit, MI 48220
Tel: 313-577-2630 / Fax: 313-993-3435
kdk@wayne.edu

Tim Haughton
Associate Professor/Reader
Centre for Russian, European and Eurasian Studies (CREES)
Department of Political Science and International Studies
University of Birmingham
B15 2TT, Birmingham, United Kingdom
Tel: +44 121 414 6360
t.j.haughton@bham.ac.uk

Acknowledgements:

Numerous colleagues have commented on various iterations of papers produced as part of this project over the past few years at a number of conferences including APSA, CES, EPOP and the ECPR Joint Sessions. We are particularly grateful to participants at the Party System Stability and Electoral Volatility workshop held at the University of Nottingham in February 2015. Finally, we are deeply indebted to Davide Vampa for his major statistical contribution, as well as to Marko Klasnja, David Pupovac and Jonathan Rose for their helpful advice on statistical techniques.

THE VOLATILITY OF VOLATILITY:

Understanding and Recalibrating Measures of Change in Party Vote Shares

Abstract

Volatility is a widely used term in political science, but even the most widely used measure of volatility, Pedersen's index, can mask as much as it reveals. His simple and elegant calculation has become part of the political science toolbox, but scholars employing this tool have tended to produce distinctly different results thanks to a series of decisions about measurement and classification. Using examples from Central Europe the critical role of decisions related to party continuity and threshold of inclusion are identified. The article not only unpacks the underlying questions addressed by different uses of Pedersen's index, but offers standards for choosing particular methods over others and outlines the principles that must be followed in creating a useable successor to Pedersen.

Introduction

Explaining continuities and change lie at the heart of political science. The political world—and particularly the world of political parties—appears to be in the middle of a period of intense, perhaps unprecedented change, but we cannot be sure unless we can have confidence that our measurements capture the type and degree of change. Such confidence stems from identifying and challenging the assumptions of existing measures and developing new, more comprehensive indicators that fit an era in which party system change is possibly more complex and in which data on these changes are certainly more readily available.

Few indicators in political science are more widespread than Mogens N. Pedersen's Index of Electoral Volatility (1979). His simple and elegant calculation became part of the political science toolbox four decades ago and has remained constantly in use.[1] Nevertheless, scholars employing this tool have tended to produce distinctly different results thanks to a series of decisions about measurement and classification. Cooks following the same recipe should produce similar cakes, but only if that recipe specifies *all* of the important choices. Slightly different ingredients mixed in different ways can yield cakes which look and taste markedly different. Differences in volatility scores, however, are not just a matter of taste. Measurements of volatility matter because they are used as proxies in research on dozens of other political and economic questions. Just as there is constant discussion about the assumptions and tools used to measure a country's unemployment rate or its growth, so we need to think carefully about the assumptions that underpin volatility scores.

Not all scholars have been satisfied with Pedersen's tool. Work by Birch (2003), Mainwaring and colleagues (2009; 2016) and Powell and Tucker (2009; 2014) have produced newer and more sophisticated measures of volatility, but even these build on the foundations laid down by Pedersen in the 1970s. Pedersen's index remains the starting point for most attempts to capture electoral volatility and party system change, and any attempt to forge a better and more accurate measure must begin from analysis of the strengths and weaknesses of Pedersen's formula and its application.[2]

This article seeks to answer four main questions. Firstly, why does volatility matter for a polity and for political science? We argue that volatility measures serve as useful indicators of political change and that understanding the phenomenon requires a

[1] Pedersen's 1979 article had 728 citations in Google scholar as of 26 May 2016.
[2] It is worth noting that Pedersen was not the first to seek to measure volatility in this way. See Delruelle et al. (1970), Fraeys (1977) and Przeworski (1975).

broadly applicable tool that is both valid and reliable. Secondly, why do authors using the Pedersen's index, the simplest and most widely adopted indicator of volatility, produce wildly different scores for identical cases? We assemble evidence of the differences in results and classify the wide variety of subtle but significant differences in method that authors introduce in applying the index. Thirdly, how does each of these methods affect the calculation of volatility scores alone and in interaction with one another? We use examples from Central Europe, a region notable for both high and varied levels of volatility, to identify the critical role of decisions related to party continuity and threshold of inclusion. Finally, how can Pedersen's index be recalibrated to ensure better validity and reliability? We unpack the underlying questions addressed by each method, offer standards for choosing particular methods over others and outline the principles that must be followed in creating a useable successor to Pedersen.

Why volatility matters

There is a lot at stake in our measurement of volatility. Scholars agree that volatility in political party systems matters for democracy, and they use calculations of party vote shares as an indicator of a wide variety of phenomena and as an independent variable in studies about the health of democracy.

Scholars are nearly unanimous in seeing a strong link between changes in the party political landscape and the process of democratization, and they argue that a stable party system is a key element in overall democratic stabilization (e.g. Tavits 2005) because they "foster more effective programmatic representation and reduce uncertainty" (Mainwaring and Zoco 2007, 157). In their respective landmark studies of new democracies in Southern Europe and Latin America both Morlino (1998) and Mainwaring and Scully (1995) maintain that party system stability is a necessary (but not sufficient) condition for the consolidation of democracy, and many subsequent studies link high and persistent levels of volatility with democratic weakness. Volatility measures matter not only for the survival of new democracies but also for the health and development of long-standing democracies. As established democracies witness the growth of new, anti-establishment parties and experience "earthquake" elections, they also come to need a more accurate understanding of the magnitude of changes in voting patterns (Bos and Segert 2007; Casal Bértoa 2013; Haughton and Deegan-Krause 2015).

High standards for measuring volatility are particularly important because the Pedersen index is used not only as an indicator of changes in party vote share—the phenomenon that it directly measures—but also as a proxy for many other developments including voter movement (Epperly 2011; Dassonville and Hooghe 2015), government alternation (Mair 2007), party system institutionalization (Weghorst and Bernhard 2014; Chiaramonte and Emanuele 2015), elite change (Ishiyama 2013), declining partisanship (Lupu and Stokes 2010) and regime stability (Bielasiak 2001). Volatility may not necessarily suffice as a valid stand-in for all of these phenomena mentioned above, but as long as scholars use it for a wide range of applications in a wide range of countries, it is essential to 'pin down numbers' and make the measure as reliable as possible (Bartolini and Mair, 1990).

Why volatility is so volatile

The elegance and simplicity of Pedersen's index of volatility has allowed it to emerge as the dominant measure of change in party systems but those same qualities have not prevented scholars from disagreeing sharply about volatility scores for specific countries during specific periods. The index goes to the heart of the question of political change by looking at changes in party vote share over time. In its simplest form, it calculates the total amount of change experienced by all individual entities in a closed system. For each entity it calculates the net change of a particular characteristic between two time periods, then takes the absolute value of this change (to prevent positives and negatives from cancelling out), and divides the result by the total amount of the characteristic in the system at the first and second time periods. Phrased in terms of specific variables, it calculates the absolute value of the net change of a particular characteristic (P), for every entity (i) between two time periods (t and t+1) divided by the sum of the same characteristic (P) at both time periods (t and t+1)

$$V = \frac{\sum_{i=1}^{n} \left| \mathrm{P}_{i,t+1} - \mathrm{P}_{i,t} \right|}{\left(\sum_{i=1}^{n} \mathrm{P}_{i,t+1} + \sum_{i=1}^{n} \mathrm{P}_{i,t} \right)}$$

As used by Pedersen, it is the sum of the net change of party vote shares (before *to* after) divided by the sum of all party votes (before *and* after) (Pedersen 1979, 4). Since most calculations of this nature employ the vote share of the full party system, the sum of all values of P is equivalent to 1 at both t and t+1. The formula can therefore be simplified to

$$V = \frac{\sum_{i=1}^{n} \left| \mathrm{P}_{i,t+1} - \mathrm{P}_{i,t} \right|}{2}$$

which can be rephrased simply as:

$$V = \frac{\sum_{i=1}^{n} |\Delta P_i|}{2}$$

The parsimony of the formula does not, however, translate immediately into harmony among those scholars who apply it to complex party systems. Studies of volatility involving Central Europe[3] offer a useful starting point because concern about the region's democratic development has led many scholars to calculate its volatility scores. The rapid changes in the vote shares of parties in Central Europe are unusual by traditional Western Europe standards (Lane and Ersson 2007; Lewis 2006; Millard 2004; Tavits 2005, 2008), but they are not overwhelmingly different from those of Latin America, Africa and Asia and or indeed from recent electoral periods in Western European countries such as the Netherlands, Belgium, Greece, Italy, and Spain.

Detailing disagreements among volatility calculations

The many calculations of Pedersen volatility index scores in Central Europe offer tend to agree in finding high average levels of volatility in the region, but a comprehensive examination of the results shows that they disagree sharply on the scores of particular countries in particular electoral periods, disagreements that have gone largely undiscussed. Table 1 displays the results of thirteen major studies generating Pedersen volatility index scores for Central European countries for electoral periods between 1990 and 2015.[4] The sheer number of independent studies calculating the same index for the same countries offers the first indication of the lack of coordination in the field. An even stronger indication comes from the significant differences among authors for the same countries across identical electoral periods. In

[3] Defined for the purpose of this paper as the 10 states from the region which joined the EU in 2004 and 2007: Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia

[4] Tavits´ (2005, 285) seminal study on the sources of party system stability in Eastern Europe draws its figures for volatility directly from Birch (2003).

one representative example, Poland's first electoral period, thirteen scholars produced thirteen different results with a gap of 29.5 points between the lowest score (19.4) and the highest (47.4). This reflects both the table's average gap of 29.0 points per electoral period, and the almost perfect absence of agreement among the authors. Indeed, out of 1,094 potential pairwise comparisons among authors for particular countries in particular periods, the authors reached identical scores in fewer than one percent of the cases (7) and came within one point more or less in only ten percent of cases (110). The largest gap between volatility assessments for any single country was 58.1 points for Poland's third electoral period, but all ten countries had at least one gap of more than 25 points and eight of the ten countries had at least one gap of more than 35 points. It is striking that even scholars with strong knowledge of party politics in their native country differ by significant margins. The gap between Gwiazda and Markowski in Poland, for example, averages 5 points and is never closer than 4.6 points in any of the five election periods in which they both calculate volatility.

[Table 1 around here]

Were this simply a question of each author having identical patterns but a different baseline or a different scale, it might be easy to harmonize them, but for many countries the relative positions on these scales differ sharply across authors and time. Some countries do exhibit consistent evaluations: Bulgaria is near the bottom of every author's scale of volatility for the first election pair and the Czech Republic is at the bottom of nearly every scale for the second and third periods; Latvia and Lithuania, by contrast, are almost universally at the top of authors' scales for the first three elections pairs. For other electoral periods there are no such harmonious patterns: the Czech Republic's first election pair, for example, appears near the bottom of

Mainwaring's and Bielasiak's ranks but near the top of Sikk's; Estonia's second election pair is near the bottom of Sikk's but near the top of Mainwaring, España and Gervasoni's and Powell and Tucker's; Slovenia's positions in the second period from near the bottom of Lane and Ersson's to near the top of Tóka's; and for Poland and Hungary in the first three election pairs there is virtually no pattern of consistency in the country's relative rankings among various authors.

Since all of these authors are working with more or less the same electoral data, the reason for the differences must lie elsewhere. A closer examination shows that there are significant differences in the methods that authors use to incorporate the election data into the Pedersen index calculations. The differences emerge because those who would calculate volatility index scores must wrestle with the interrelated problems of small parties and party continuity. For many authors, especially those engaged in large-scale comparative enterprises, identifying the twists and turns of small parties becomes an overwhelming burden. Identifying continuity is difficult enough with large parties, especially where these face party splits and engage in mergers and name changes. Pedersen's index actually provides several different options for dealing with these problems, but not all authors use the options in the same way, and the various combinations allow Pedersen's simple formula produce a remarkably complex array of results. The three main options for adapting Pedersen's formula to deal with size and continuity are rooted directly in the rules for linking party data cross time, for aggregating party data into a single number, and for recalculating individual data points.

Linkage between entities: What counts as a connection between $P_{i,t}$ and $P_{i,t+i}$

Perhaps the most complicated aspect of volatility calculations is linking party values over time in the face of party name changes, mergers, splits and various combinations thereof. Careful scholarship on ‘genuinely new’ parties (Sikk 2005) reveals that the category is smaller than it might at first look, and a large number of the entrants in any given electoral period involve partial continuity. To date, the type and extent of that continuity has attracted relatively little academic research. Many scholars of volatility respond with case-by-case judgments or with rule-based formulae for deciding which party $P_{i,t}$ to link with which party $P_{i,t+i}$. Methods that use a stricter interpretation of continuity between t and t+1 err on the side of discontinuity and tend to code name changes, splits and mergers and new parties that represent sources of volatility. Methods emphasizing continuity may opt instead for a more relaxed understanding of linkage and pair a party at time t with whatever party it is most closely related—either by name, organizational structure or vote total—at time t+1. Even in these relatively regularized, rule-based systems, however, there may still be considerable human judgment in the assessment of which parties are to be considered within the set of plausible successors. Such decisions become even more sensitive when applying Mainwaring, España and Gervasoni (2009; 2016) and Powell and Tucker’s (2009; 2014) intra- v. extra-system distinction that depends heavily on whether parties are considered as successors or as new.

Aggregation of entities: How many parties are included in a single P_i

For those scholars who are reluctant to declare a single predecessor in a merger or a single successor in a split, the Pedersen formula includes a second implicit option: the possibility of aggregating the values of two or more parties at either t or t+1 or both before linking them. For example, if Party A merges with Party B to form Party C, the aggregation method sums Party A and B into a single data point at time t and compares it with party C at time t+1. If, at the same time, party D splits off from party A, the method would compare the sum of Party A and B with the sum of Party

C and D. This approach eliminates the need for choice among potential links and emphasizes the continuity that may exist between the original party and multiple successors (or predecessors). Of course in the process it diminishes the number of entities (i) and therefore makes the indicator itself less sensitive.

Another common application of this method involves the aggregation of small parties at a particular time period into a single data point commonly referred to as "Other." Which parties are thus aggregated usually depends on rules for including parties in the main calculation.

Inclusion of entities: Which parties count as P?

Limits of time and information mean that it may not be practical (or even possible) to gather data for small parties. Even though new data storehouses and national websites provide extensive electoral results for most all countries, questions of linkage over time are still extremely difficult and time-consuming for very small parties, especially since such parties often face splits by disgruntled members and seek to improve their fortunes through name changes and mergers. Although some authors make ad hoc exclusions based availability of data, most seek to avoid selection bias problems by imposing a numerical criterion for inclusion usually based on the magnitude of a party's election result. Even with a clear threshold, additional questions remain, particularly whether to apply the threshold in a blanket fashion and include values for a particular party if it exceeds the threshold at any point in time, or to include only individual data *points* that exceed the threshold. The former option does little to solve the data collection problem since researchers may still need to find obscure data for parties that only later crossed the threshold; the latter option, by contrast, severs the continuity of party results by relegating some data points—those below the threshold—to the "other" category or dismissing them completely. An intermediate strategy involves the use of transitional pairs of data points to capture each period in which a party's support crosses the threshold, but does not calculate data points that are not adjacent in time to periods above the threshold. Table 2 offers a hypothetical example of how each of these strategies could be applied and how the application would affect volatility scores:

[Table 2 around here]

Using Transitional pairs thus involves some recalculation of data points (and the possible transfer of those values to the "Other" category), but not as much as using only Individual points. The Blanket inclusion method, by contrast, either includes all values or omits a party entirely.

Questions of inclusion are also at the heart of recent modifications to Pedersen's index which seek to distinguish changes in established parties from changes introduced by the entrance and exit of parties from the party system. Birch's (2003) formula, for example, is a restricted subset of Pedersen's full equation:

$$V = \frac{\sum_{i=1}^{n} \left| \mathrm{P}_{i_c,t+1} - \mathrm{P}_{i_c,t} \right|}{\left(\sum_{i=1}^{n} \mathrm{P}_{i_c,t+1} + \sum_{i=1}^{n} \mathrm{P}_{i_c,t}\right)}$$

(where, as above, V is volatility, $P_{(ic,\ t)}$ is the vote share of continuous party i at the first election and $P_{(ic,\ t+1)}$ is the vote share of continuous party i at the second election) in which calculations are restricted to include only parties with nonzero values in both time periods. Birch's method thus provides a 'measure of the changes of electoral fortunes of *existing* players in the political game' (Birch 2001:4, italics added), but the author herself recognized that this refinement did not sufficiently capture the changes in the range of parties on offer, particularly the entries, exits and alterations in the composition of coalitions, and supplemented it with a separate measure of party replacement which she calculated as the share of votes for parties at t+1 that did not exist at t.

$$R = \sum_{i=1}^{n} (\mathrm{P}_{i_n,t+1})$$

where i_n are parties that did not receive votes in the previous election.

Later authors extended Birch's work by elaborating on Pedersen in a way that integrates the changes among existing parties that Birch associates with "volatility" as well as the changes related to party entry and exit that she associates with "party replacement." In an instance of parallel invention, two groups of scholars—Mainwaring, España and Gervasoni (2009, 2016) and Powell and Tucker (2009; 2014) —simultaneously devised a formula for disaggregating Pedersen that builds on Birch's work:

$$V = \frac{\sum_{i=1}^{n}|\mathrm{P}_{i_c,t+1}-\mathrm{P}_{i_c,t}|+\sum_{i=1}^{n}(\mathrm{P}_{i_n,t+1}+\mathrm{P}_{i_x,t})}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1}+\sum_{i=1}^{n}\mathrm{P}_{i,t})} = \left(\frac{\sum_{i=1}^{n}|\mathrm{P}_{i_c,t+1}-\mathrm{P}_{i_c,t}|}{\sum_{i=1}^{n}(\mathrm{P}_{i,t+1}+\sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}(\mathrm{P}_{i_n,t+1}+\mathrm{P}_{i_x,t})}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1}+\sum_{i=1}^{n}\mathrm{P}_{i,t})}\right)$$

Within-system/Type B   Extra-system/Type A

where i_c refers to continuous parties with nonzero values in both t and t+1, i_n refers to new parties with zero values at t but nonzero values at t+1, and i_x refers to parties that exited the political scene yielding non-zero values at t and zero values at t+1. Since i_n by definition has no predecessors and i_x has no successors, these can stand alone in the calculation, implicitly subtracting the previous support or successive support (equal to zero). By then looking only at the left or right hand of the equation, these authors can calculate the total contribution to volatility represented by continuous parties and by those entering or leaving the system and provide distinct figures for within-system and extra-system volatility. The new formula incorporates Birch's insight about the distinct dynamic of change within the continuous parties but does not require a distinct measure of party replacement and does not differ depending on the share of the continuous parties. The scores generated by later authors for within-system volatility can also be added to the extra-system scores to produce an overall volatility score that is equivalent to Pedersen's formula.

Classifying authors' choices of method

Table 3 shows the combination of choices made by the authors in the major areas above. Among the fourteen author methods cited here, there are eleven different choices and significant variation in each category: ten without specific thresholds and four with thresholds (ranging from 1% to about 5%, one using recalculation and three using pairwise exclusion); seven without other categories and seven with other categories (all including the other category in calculations); seven using the aggregation methods, four using the linkage method (two using maximum differences, two using minimum difference), and three using a mix or other variation. Only two pairs of authors share common choices in all of these areas: Sikk and Bielasiak with no threshold, an included other category, and the aggregation method, and Mainwaring and Tóka (who deliberately emulates Mainwaring in calculating his "raw" scores) with the minimum difference method of linkage and no threshold (so no other category).

[Table 3 around here]

How method choices affect volatility index scores

How significant a role do these different choices play in producing the different volatility results? Not all authors look at all countries in all electoral periods and so an overall comparison of differences is impossible, but it is possible to calculate pair-wise differences between each author-country pair and then average these to assess the relative magnitude of each author's volatility assessments. Comparing these results to choices about threshold shows no clear pattern. A slightly stronger pattern emerges with regard to the “other” category, with higher volatility scores concentrated among those who do not include “other” in their calculations. Finally, there appears to be a strong pattern related to the choice between aggregation and linkage. As Figure 1 indicates, with the exception of Birch (who in addition to these choices does not include new parties), the authors choosing the aggregation method have the lowest volatility scores, while those choosing the linkage methods are higher. It is also noteworthy that the authors using the linkage methods also have a far greater range of scores than those using aggregation, which is to be expected since the results of the linkage methods are more affected by differences in binary choices of predecessor and successor that might differ from author to author even in the best of circumstances.

[Figure 1 around here]

With enough countries and enough combinations of methods, it should theoretically be possible to separate out the effect of particular choices on volatility scores, but at present there are far more variable combinations than there are authors using those combinations. Figuring out the actual impact of each of these variables therefore requires a far more tightly focused method. Again Central European cases can provide the raw material for a closer analysis that lets us control for specific differences in inclusion, aggregation and linkage by applying every possible combination of those methods to the same sets of full electoral data. The time- and knowledge-intensiveness of this process is quite substantial (and helps to explain the use of selective inclusion methods). For many countries the demands of the process are prohibitive, but for three countries—the Czech Republic, Hungary and Slovakia—we were able to assemble accounts of even the smallest parties (some with fewer than

1,000 votes) and to determine linkage patterns among these parties across each time period. These three cases are also particularly useful because they exhibit a range of volatility characteristics: by nearly all estimates Slovakia experienced relatively high volatility throughout its first 20 years of democracy whereas Hungary's volatility was relatively low in contrast to other Central European cases; the Czech Republic's volatility was also low until the late 2000s but jumped sharply with the 2010 and 2013 elections. These differences both in level and consistency—one low, one high, one low then high—provide a solid initial basis for testing the various permutations of method.

To operationalize every one of the theorized variants of the Pedersen index on these three cases, we performed a full set of calculations for each permutation of four methods.

<u>Inclusion threshold size</u>: We calculated Pedersen scores first with no threshold and then with thresholds at ten 0.5% additional increments up to 5.0% (the highest level used by any author in the sample above).

<u>Inclusion method:</u> We calculated scores using three threshold-based inclusion methods in three ways, two of which are used by authors in the sample and a third potential option:

1) Individual points: This method includes only data points that stand above the threshold and recalculates others as zero.
2) Transitional pairs: This method includes all at points that stand above the threshold and data points from the periods immediately before and after the point above the threshold.

3) Blanket Inclusion: Including all parties whose maximum vote share rose above the threshold at any time period again.

Excluded data method: For parties that were excluded from the calculations above, we performed volatility calculations with and without an aggregated "other" category.

1) Aggregated and linked other. We produced an aggregated "other" for each election consisting of the sum of data points not included elsewhere, and treated this as a distinct unit, whose changes over time contribute to volatility calculations like those of any other party.
2) Uncalculated. Data points not included elsewhere were not included in volatility calculations.

Continuity method: we calculated volatility in each country according to three methods that affected linkage.

1) Linkage method. We made binary choices about the successors and predecessors in the case of splits, mergers and party name changes. There is a wide range of options in such cases. We calculated two distinct subsets of linkage according to different guidelines:
    a. Strict linkage: With even more focus on discontinuity than Birch (2003) or Powell and Tucker (2009; 2014), we calculated any split or merger involving a name change or other significant alteration from t to t+1 as a new party with no formal linkage to predecessors or successors.
    b. Relaxed linkage: Following Mainwaring, España and Gervasoni (2009; 2016) we linked parties that merely changed names and in the case of split or merger we linked parties to the largest successor or predecessor

while treating any other offspring or parents as unlinked across time periods (see Mainwaring, España and Gervasoni (2009; 2016)).

2) Inclusive aggregation method. In the case of a party split or merger we aggregated all predecessor and successor parties as a single party in each time period.

The combination of three continuity methods, three threshold-inclusion methods, two "other"-aggregation methods and 10 threshold levels produces 180 possible variations per election period in each country. Figures 2a through 2c show the average results of each of these combinations applied to electoral data for the Czech Republic, Hungary and Slovakia. For context, it also shows the baseline data with thresholds set at zero and the full Birch method which includes only continuous parties (this method is incompatible with the aggregation method which by definition incorporates non-continuous components in handling party splits and mergers.

Bivariate Results

In analyzing the results, we begin in reverse order because the effects of some variations vastly outweigh the effects of others (and the interaction among the variations becomes important).

<u>Continuity method:</u> The method of addressing continuity has an extremely strong impact on the measurement of volatility. With threshold set at zero, over seven electoral periods in Slovakia the Strict linkage method produced an average volatility score of 47 whereas the Relaxed linkage method averaged just over 31 points and the Inclusive aggregation method averaged just under 20 points. The results were similar in the Czech Republic and Hungary, though at lower overall levels and with smaller overall differences. A look at calculations that comprise these results explains this

difference: Strict linkage allows for the fewest connections across elections and therefore means that fewer parties have counterparts from t to t+1; Relaxed linkage, by contrast, seeks out relationships between t and t+1 but still excludes linkage from all but one designated predecessor or successor. The inclusive aggregation method nearly always produces lower volatility results because it adds together entities that would otherwise be separate, minimizing the differences between units across time periods. Indeed, the inclusion method sometimes goes so far in this direction that it loses the capacity to make meaningful distinctions when parties split from one source and then merge with another. Figure 3 shows an example of split-merger combinations in which the inclusive aggregation cannot detect real change.

[Figure 3 around here]

When this “Z-effect” repeats across the party system, it significantly reduces the basis for calculating volatility. In the example above, if Party D also lost a splinter, Party E, that merged with Party B and this process continued with Party F and beyond, the whole of the party system could be aggregated on each side of the calculation producing zero volatility even in the case of a wildly changing system. This is more than just a theoretical possibility: in the case of Slovakia between 1992 and 1994, mergers and splits produced a single aggregate that stretched across the political spectrum and accounted for nearly 50% of all votes in both years (Kopeček 2007; Haughton, 2014).

<u>Threshold Size:</u> A change in the threshold can also produce a significant change in volatility, but its effect is only partially predictable and depends on other factors. As Figures 2a-c show, volatility tends to decrease with an increasing threshold for both the Relaxed linkage and Strict linkage methods of continuity. In the example of Slovakia, the addition of a 5% threshold to the Strict linkage method results in a 10-

point drop in measured volatility. This finding fits expectations since in any party system with a significant number of small parties, a higher threshold forces the calculations to ignore changes among smaller parties and thus reveals apparently lower levels of volatility. Thresholds do not produce this effect in the inclusive aggregation method. The underlying party-level calculations show why: where small parties join with larger ones the elimination of small parent or offspring parties from one side of the equation because of thresholds actually decreases the measure's ability to recognize continuity and may produce higher volatility scores rather than lower ones.

Threshold inclusion method: The method of applying the threshold matters as well. As Figures 2a-c show, the dotted line representing "blanket inclusion" usually produces a smaller drop in volatility than either the pairwise or recalculation methods since this method expands the number of parties included in the calculations and therefore mitigates the threshold's overall effect.

The individual point method creates a different sort of problem by simply zeroing out any figure below the threshold. The effect has sometimes unpredictable effects, however, because the recalculation produces two countervailing effects: on one hand, it reduces the measured level of volatility by eliminating all pairs of party results that are both below the threshold (a party's rise from 1.0% to 2.0% represents a 1.0 contribution to volatility, but this is not included in volatility calculations when the threshold is set at 2.5% and both points are set to zero); on the other hand it exacerbates volatility in cases where one result in an electoral period is above the threshold and the other is below (a party's fall from 4.0% to 1.0% actually reflects a 3.0 contribution to volatility score but the recalculation method drops the second of the two figures to 0.0%, producing an apparent contribution of 4.0, a "volatility subsidy" equal to the size of the smaller figure in the electoral period). Pedersen's original index can be expanded to isolate particular sets of entities, some with values above the threshold in both time periods i_{sa}, some with values below the threshold in both time periods, i_{sb}, and some with values on either side of the threshold, i_{sac} *and* i_{sbc}. This effect can be represented mathematically as

$$V = \left(\frac{\sum_{i=1}^{n}\left|\mathrm{P}_{i_{sa},t+1} - \mathrm{P}_{sa,t}\right|}{\sum_{i=1}^{n}(\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left|\mathrm{P}_{i_{sb},t+1} - \mathrm{P}_{i_{sb},t}\right|}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left(\mathrm{P}_{i_{sac},t+1} - \mathrm{P}_{sbc,t}\right)}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left|\mathrm{P}_{i_{sbc},t+1} - \mathrm{P}_{sac,t}\right|}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})}\right)$$

Supra-threshold Sub-Threshold Cross-Threshold(down) Cross-Threshold (up)

The exclusion method sets all i_{sb} to zero, producing

$$V = \left(\frac{\sum_{i=1}^{n}\left|\mathrm{P}_{i_{sa},t+1} - \mathrm{P}_{sa,t}\right|}{\sum_{i=1}^{n}(\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}|0 - 0|}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left(\mathrm{P}_{i_{sac},t+1} - 0\right)}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left|0 - \mathrm{P}_{i_{sac},t}\right|}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})}\right)$$

Supra-threshold Sub-Threshold Cross-Threshold (down) Cross-Threshold (up)

which can be simplified as

$$V = \left(\frac{\sum_{i=1}^{n}\left|\mathrm{P}_{i_{sa},t+1} - \mathrm{P}_{sa,t}\right|}{\sum_{i=1}^{n}(\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left(\mathrm{P}_{i_{sac},t+1}\right)}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})} + \frac{\sum_{i=1}^{n}\left|\mathrm{P}_{i_{sac},t}\right|}{(\sum_{i=1}^{n}\mathrm{P}_{i,t+1} + \sum_{i=1}^{n}\mathrm{P}_{i,t})}\right)$$

Supra-threshold Cross-Threshold (down) Cross-Threshold (up)

Raising the threshold has two effects. On one hand, a higher threshold may decrease the number of parties whose volatility counts in the overall volatility calculation (because it is happening below the notice of the threshold). At the same time, a higher threshold may increase the number of cases in which the result at either t or t+1 is zero, and thus may increase the net amount of cross-threshold volatility. These two countervailing influences are only tangentially related (constrained by the overall party system) and depend heavily on the relative sizes and shifts of party support within a party system during a given time period. Considerable change among sub-threshold parties will increase the totals for i_{sb} and lower volatility while activity that crosses the threshold will have the opposite effect.

How these effects interact will depend on what actually happens in the political sphere. The results of Figure 2a-c suggest that the cross-threshold effect tends to predominate in the countries studied here, causing increases in the threshold to increase the measured volatility. The recalculation method may thus actually produce results that more closely resemble a threshold of zero, but this occurs only because of the artificial cross-threshold counter-effect that is largely accidental and cannot be relied upon to consistently push the volatility figure toward the level it would reach without a threshold. It produces something close to the right answer but not necessarily for the right reason and therefore does not represent either a valid or reliable shortcut.

The transitional pair method solves some of the problems above. It excludes data in only those instances where both characteristics in a pair of time periods are below the threshold but permits those where only one is below. The advantage of this is to eliminate cross-threshold effects. In the example above, Party A's fall from 4.0% to 1.0% is counted only as a 3.0 contribution to volatility while the subsequent rise from 1.0% to 2.0% does not enter into the calculations at all. In practice this means a lower level of volatility than if the cross-threshold effects were included, but since those effects are artificial and more erratic, it actually produces a smoother line - more directly proportional to the level of the threshold - and a more theoretically justifiable basis for calculation. It is true that it requires slightly more effort in finding data for the time period before and after a party crossed the threshold, but this is more sustainable than gathering linkage information for a small party's entire existence, and it is precisely the kind of information more likely to appear in the popular press ("Party A rose to 4.0% from its previous performance of 1.0%").

<u>Excluded data method:</u> Finally, there is the question of how to handle data that is not included by the threshold exclusion methods above (Ocaña 2007). It is possible simply to omit the data from calculation (in effect recalculating it as zero), but some authors opt for an aggregation method and sum all excluded data into a single "other" category at the bottom of the data table. This aggregation method raises questions about linkage between the aggregated "other" values from one time period to the next. Since the aggregated "other" category may have different composition at each time point, and since, because each item aggregated into the "other" category may be (by definition) below the notice threshold of the scholar gathering the data, circumstances mean that there is no formal way to establish the degree of linkage between two successive "others." As with other questions above, there are several options with relatively clear effects on volatility: to create no linkage between the "other" results and calculate them as new, one-time parties (in effect adding the size of the "other" category to volatility), to refrain from including "other" at all, and to link the "other" results from one time-period to the next. While theoretically possible, the first of these approaches is not a useful option because it assumes that smaller parties have no linkage with any other entity, past or future, thus magnifying volatility by the entire size of the other category. The second of these, by contrast, presumes that there is *no* volatility among the unmeasured entities and must produce a level that is equal to or lower than the actual volatility level. The final method, which seems to offer a middle way, creates difficulties of another variety. The total size of the "other" from election point to election point is only distantly related to the "other" volatility.

The difference in the size of the “other” category theoretically depends very little on differences related to the overall level of volatility or on the volatility that is lost to calculation through exclusion of data or cases. The grey lines of Figure 1 shows that in practice in exclusion methods, linking data points in the “other” category restores about half the amount of volatility to the calculation that is removed through the imposition of thresholds, but the effect is not strictly linear and appears to increase disproportionately as the threshold approaches 5.0%. The volatility contributed to the overall result by the aggregated other categories, as with the subtractions in the recalculation method of the threshold, is an accidental effect. If all parties under the threshold move in the same direction (they all get smaller or larger) then the “other” category is mathematically identical to the aggregation of the individual elements. If, however, parties move in opposite directions, the volatility is masked and the “other” category obscures the actual pattern. The “other” category may add volatility at the same time that is removed by the threshold but it does not necessarily put back the *same* volatility. In the cases used here it tends to follow the overall curve, but there is no mathematical basis for thinking that it will do so in a reliable fashion and no theoretical justification for aggregating the parties below the threshold into a single, continuous “other” (unlike the more solid basis for aggregating parties involved in mergers and splits in the aggregation alternative to linkage).

Multivariate results

Looking at threshold, linkage and exclusion separately highlights how these individual ingredients can change the size and shape of the volatility, but what also matters is their relative effects and how they interact. To that end we subjected the data from the three countries to a multivariate regression (table 4). Given the continuous character of our dependent variable we use ordinary least squares with robust standards errors. In order to control for possible country- and election-effects, we use dummy variables for each election per country (not included in the results table for reasons of space).[5]

[Table 4 around here]

In the basic model, the continuity method has large and statistically significant impact (at .001 level). Shifting from the Inclusive aggregation method to the Relaxed linkage method increases volatility by nearly five points (4.8), and from there to the Strict linkage increases volatility by another 10 points for a total of fifteen points above the inclusive aggregation method (15.2), controlling for the other independent variables. Threshold size also has a significant impact: each one percentage point increase in the threshold reduces the amount of volatility by nearly one point (0.81). Including an "other" category is also significant, though the effect of including such a category is relatively small, accounting for approximately a one point increase in volatility (1.3). The Threshold inclusion method has no partial effect on the dependent variable, with no significance difference between methods that use individual points, the transitional-pairs, and blanket inclusion.

Since Figures 2a-2c suggest that thresholds do not function the same way in all circumstances Model 2 uses interaction terms for Threshold size and the other methods. The choice of Threshold inclusion method again has no effect, but all of the other methods do interact with the Threshold size in statistically significant ways. Model 3 removes the non-significant methods and interaction terms for cleaner results. We will therefore rely on Model 3, rather than Model 2, for the interpretation of the results. The interaction effect between the continuity method and threshold size are substantially and statistically significant. With the Inclusive aggregation

[5] The reference category for the dummies is the first Hungarian election.

method, raising the threshold by one percentage point reduces volatility by only 0.5 points, but with the Relaxed linkage method each one percentage point increase in the threshold produces a 1.3 unit drop in volatility and with the Strict linkage method the result is a 1.4 unit drop, provided that the interaction with "other" is kept at 0. A similar, but less statistically significant (at .01 level) interaction effect occurs between Threshold size and the Aggregated linked other method of dealing with excluded data. Indeed, provided that the interaction with method is kept at 0, if other is at 0 the model predicts that, holding the other variables constant, each additional threshold point will produce a 0.46 point decrease in volatility. On the other hand, the Aggregated linked other method would produce only a 0.02 decrease.

How to evaluate and improve volatility calculation methods

The purpose of the calculations above is not merely to explain the diversity of volatility results found above and the relative impact of specific variations of Pedersen's Index, but also to encourage more accurate and reliable measurement. Moving from least to most important we offer three specific recommendations and a point for further consideration.

Firstly, since significantly different variations of the Threshold inclusion method have no significant impact on volatility scores, the recommendation must rest on considerations other than results. The blanket inclusion approach is both the most demanding in terms of data demands and the most problematic in terms of the constant need to recalculate volatility based on party performance in subsequent years. The individual points method is a better candidate since it the easiest to apply, but its underlying mathematics—zeroing out anything below the threshold—may not in every case produce the balance between eliminating one realm of volatility while introducing another that we see in these three cases. Slightly preferable, though also slightly harder to apply, middle ground of the Transitional pair approach which captures important cross-threshold changes without the need for massive additional data collection.

Secondly, and with only slightly greater impact is the question of the right method for excluded data; while the decision has statistically significant consequences, it alters volatility scores by only a single point. Since the use of an Aggregated linked other category introduces as much noise as it removes (like the Individual point method of Threshold inclusion discussed above) and thereby introduces the possibility (albeit small) of more significant errors, it may be preferable simply to omit the category entirely and acknowledge that doing so will result in a small underestimate of volatility.

Thirdly, and more significant is the question of the threshold itself. The ideal Threshold size is obviously zero (especially since this eliminates the need for threshold inclusion and excluded data methods), but there is a clear trade-off between accuracy and the costs of gathering information. The experience of data collection for this study suggests that efforts to lower the threshold are accompanied by an almost exponential increase in information costs. For most larger parties the election data and continuity over time was easy to assemble, but as we pushed the threshold below 1.0% and especially below 0.5% parties became almost impossible to trace without archival searches and interviews. This difficulty may ease over time with the development of historical databases (especially if scholars gather information about small parties as each election happens), but with the current state of data for most countries, it is probably adequate to use a 1% threshold along with the acknowledgement that this may underestimate volatility by one to two percentage points in the case of the Relaxed linkage and Strict linkage methods and by a somewhat smaller amount with the Inclusive aggregation method.

Finally and most importantly there is the question of continuity method. Unlike the other choices discussed above, this one offers no quick solution. Each method actually represents a distinct and coherent way of thinking about the notion of change in party systems.

The default assumption of the Inclusive aggregation method is continuity: splits and mergers do not change voter preferences, and voters can navigate the institutional change to follow leaders or party factions they prefer, even if their favourites move out of an existing party and into a new splinter, or merge into a larger entity. In this light, voters who chose to switch their votes from a small party into its merged successor (or from a big party to one of its splinters) may be engaging in an act of *continuity* rather than change. This method is therefore at its most useful revealing when continuity is *absent*. When scores produced by the Inclusive aggregation method are high (by the standards of past elections or neighbouring countries), it signals voter shifts so significant that they fall outside of existing institutional boundaries.

By contrast, the default assumption of the Relaxed linkage and Strict linkage methods is change: splits and mergers can make it impossible for voters to stick with their past choices, and even choosing to follow a party absorbed by a merger or to follow favored leaders who split off to form a party of their own represents a rupture with past choices. This method is therefore at its most useful revealing when continuity is *present*. When scores produced by the linkage methods are low, they signal that few voters are shifting despite any pressure imposed by splits and mergers.

Neither method, then, is *intrinsically* better than another. One option is to follow the path taken by Tóka and Henjak in measuring *both* raw (Strict linkage) and adjusted (Inclusive aggregation) scores to identify the upper and lower bounds of the possible values. Taking this double calculation one step further produces a heretofore unnoticed advantage: a *comparison* of the results produced by both methods offers significant insight into the nature of volatility in particular countries and particular elections. The comparison in Figure 4 of Inclusive aggregation and Strict linkage volatility scores in the three countries used in this study accurately reflects the overall history of party change in the two countries. All three alternate between higher and lower amounts of overall volatility among party units and these stand at the upper right and lower left corners of the diagonal line, but certain elections in certain countries also involve an amount of institutional volatility—splits and mergers—over and above the voter shift. Shifts due to split or merger were relatively rare in Hungary and the Czech Republic and their data points tend to hug the diagonal (except for a slightly above-average shift related to the splintering and reformation of Hungary's left in 2014 and a major shift following the breakup of the anti-Communist Civic Forum in the Czech Republic in 1992), Slovakia, by contrast, experienced three big merger- and split-related institutional shifts in 1992, 1998 and 2002 which stand out sharply in a comparison of Slovakia's its Inclusive aggregation and Strict linkage volatility scores.

[Figure 4 around here]

This multiple-method approach is admittedly less useful for those who need a single volatility score as a dependent or an independent variable in a multivariate regression model. In such cases, scholars may need to make an explicit choice of method depending on whether the question at hand refers primarily to voter continuity or institutional change and acknowledge the potential impact of that choice on results. It should also be possible at present to create distinct regression models for volatility based on results from the linkage and aggregation methods and then examine the degree to which the results agree or disagree. In the longer run, however, the better option is to expand and improve volatility measurement so that the partial shifts related to mergers and splits are fully included in the formula and become an integrated part of overall volatility scores.

Conclusion

Pedersen's formula for calculating volatility has been widely used and for good reason. It offers a simple and straightforward way of capturing the extent of electoral change. But it also has its limits and its application can yield a wide variety of results depending on a series of secondary methodological assumptions about inclusion, aggregation and linkage that scholars use to address problems of limited party data and lack of clear party continuity. The relatively minor differences related to party size can be resolved without much controversy, but the much larger differences related to continuity are not so simple. From a practical standpoint it is unfortunate that there are no easy answers to the biggest question, but it is in only the process of thinking consciously about party continuity that we can strengthen Pedersen's Index against the bigger challenges that are likely to arise.

The critical innovation introduced by Pedersen and others during the 1970s was to look at the net amount of change in party vote share and to understand this as an indicator of overall change in the party system and as a reasonable proxy for change in voter preferences (Achen 2014). In the 2000s a new group of scholars saw that this measurement could provide additional clarity if it were subdivided into changes among existing parties and changes caused by party entry and party exit. Both of these insights, however, depend on a party system composed of whole, stable units. As soon as parties merge and split and morph, the task of measurement devolves into disagreements about continuity and the meaning of volatility scores becomes less and less clear. Having accommodated the question of party birth and death, the Pedersen's Index's biggest challenge is the accelerating rate of party *change*. It is not only in Central Europe where party splits and mergers have begun to dominate elections—systems around the globe are moving in the same direction—and unlike the emergence of new parties, the problem of party change undermines the use of binary judgements such as "party successor or not?" or "loyal voter or not?"

In such cases the best answer may be to rely on other data and other methods. Questions about shifts in voter party preference and its relationship to underlying policy preferences or socio-demographic position—demand-side shifts—cannot be easily answered by Pedersen's index where there are frequent party splits and mergers. In such cases it may be preferable to turn directly to voters. Here the methods are well-established: public opinion surveys with panel data where possible (and retrospective preference questions where not) and ecological analysis of disaggregated election data can offer strong evidence for changes in electoral behaviour and may actually perform better answering Pedersen's initial questions about voter shift tan does Pedersen's own index.

Questions about the dealing with splits and mergers at the institutional level—the supply side—also require data and methods that go beyond Pedersen's index. Unlike the demand side, however, the elements of supply-side analysis are not yet available. The elements of a mathematical formula are actually quite straightforward. If we understand political parties not as unitary entities but as aggregates of smaller party-like components (factions, branches, members of parliament or any number of other units), then it is possible to see a party split simply as a separation of pre-existing components which can be recombined in new ways and whose change can be calculated using a formula based on Pedersen's model.

The hard part is defining those components and quantifying them but this too may be possible. Efforts to gather data on party institutional change on a cross-national basis has only recently become subject to quantification (Ibenskas 2012), and most current efforts focus on counting and analysing the splits and mergers themselves rather than analysing the relative strength of the component parts that split apart or merge. That task can seem overwhelming both because it requires in-depth local knowledge and because it is difficult to assess the relative weights of components within political parties that may exist on the ground, in public office and in central office (Katz and Mair's 1993), but scholars and practitioners regularly assess the relative share of assets (many of them intangible) in corporate mergers and spinoffs, and political scientists regularly complete expert surveys on questions of party ideological positions that are similarly complex. A detailed, open database that pools and coordinates the experiences of country experts regarding the flow of assets in party splits and mergers can significantly improve our understanding of volatility at the level of party institutions and help to resolve the ambiguities that threaten to undermine the usefulness of Pedersen's elegant formula.

References

Achen, Chistopher. 2014. Electoral Volatility: Why the Pedersen Index Succeeds. Annual Meeting of the Midwest Political Science Association, April 3-6.

Bågenholm, Andreas & Andreas Johansson Heinö. 2009. Still not stable after all these years: The logic of path dependence in party system stability in Central and Eastern Europe, Paper prepared for the ISA Annual Convention, New York, February 15-18

Bakke, Elisabeth, and Nick Sitter. 2005. "Patterns of Stability. Party Competition and Strategy in Central Europe since 1989." *Party Politics* 11 (March) 243-263

Bartolini, Stefano, and Peter Mair. 1990. *Identity, Competition, and Electoral Availability: the Stabilization of European Electorates 1885–1985*. Cambridge: Cambridge University Press.

Bielasiak, Jack. 2001. "On the Institutionalization of Party Regimes in Emerging Democracies", *Working Paper*, n. 351, Glasgow: University of Strathclyde.

Bielasiak, Jack. 2002. "The Institutionalization of Electoral and Party Systems in Postcommunist States." *Comparative Politics* 34 (January): 189-210.

Bielasiak, Jack. 2005. "Party Competition in Emerging Democracies: Representation and Effectiveness in Post-communism and Beyond." *Democratization* 12 (June): 331-56

Birch, Sarah. 2003. *Electoral System and Political Transformation in Post-Communist Europe*. New York: Palgrave Macmillan.

Bos, Ellen and Dieter Segert. 2007. "Osteuropa als Trendsetter? Parteiensysteme in repräsentativen Demokratien unter mehrfachem Druck". In *Osteuropäische Demokratien als Trendsetter?*, eds. Ellen Bus and Dieter Segert. Opladen: Verlag Barbara Budrich: 323-338.

Casal Bértoa, Fernando. 2013. "Post-communist Politics: on the Divergence (and/or Convergence) of East and West." *Government and Opposition* 48 (July): 398-

433.

Chiaramonte, Alessandro, and Vincenzo Emanuele. 2015. “Party System Volatility, Regeneration and De-institutionalization in Western Europe (1945-2015).” *Party Politics* First Online (August)

Dassonneville, Ruth, and Marc Hooghe. 2011. “Mapping electoral volatility in Europe. An analysis of trends in electoral volatility in European democracies since 1945.” *Paper presented at the 1st European Conference on Comparative Electoral Research (December)*

Dassonneville, Ruth and Marc Hooghe. 2015. “Economic indicators and electoral volatility: economic effects on electoral volatility in Western Europe, 1950–2013.” *Comparative European Politics* First Online (February)

Delruelle, Nicole, René Evalenko and William Fraeys. 1970. *Le Comportement Politique des Électeurs belges. Enquête de Sociologie Électorale. La Rupture de 1965. Ses Prolongements en 1968*. Bruxelles: Institut de Sociologie de Bruxelles.

Epperly, Brad. 2011. “Institutions and Legacies: Electoral Volatility in the Post-communist World.” *Comparative Political Studies* 44 (July): 829-853.

Fraeys, William. 1977. “Les Élections Législatives du 17 Avril 1977: Analyse des Résultats.” *Res Publica*, 19:497-513.

Gwiazda, Anna. 2009. “Poland’s Quasi-institutionalized Party System: the Importance of Elites and Institutions.” *Perspectives on European Politics and Society* 10 (September): 350-375

Haughton, Tim, 2014. “Exit, choice and legacy: explaining the patterns of party politics in post-communist Slovakia.” *East European Politics* 30 (June): 210-

229

Haughton, Tim, and Kevin Deegan-Krause, 2015. “Hurricane Season: Systems of Instability in Central and East European Party Politics”, *East European Politics and Societies* 29 (February): 61-80

Ibenskas, Raimondas. 2012. Explaining Party Splits and Mergers in Europe. PhD thesis. Trinity College, Dublin, Ireland.

Ishiyama, John. 2013. “Elite Volatility and Change in North Korean Politics: 1997-2010.” *Journal of Asian and African Studies* 49 (October): 570-583

Katz, Richard, and Peter Mair. 1993. “The Evolution of Party Organizations in Europe: The Three Faces of Party Organization.” In William Crotty (ed.) *Political Parties in a Changing Age*, special issue of *American Review of Politics*, 14: 329-45.

Kopeček, Lubomír. 2007. *Politické strany na Slovensku: 1989 až 2006*. Brno: Centrum pro studium demokracie a kultury.

Kreuzer, Markus, and Vello Pettai. 2003. “Patterns of Political Instability: Affiliation Patterns of Politicians and Voters in Post-Communist Estonia, Latvia, and Lithuania.” *Studies in Comparative International Development* 38 (Summer): 76-98.

Lane, Jan-Erik. 2008. *Comparative Politics: The principal-agent perspective*. London, New York: Routledge.

Lane, Jan-Erik, and Svante Ersson. 2007. “Party System Instability in Europe: Persistent Differences in Volatility between West and East?.” *Democratization* 14 (February): 92-110.

Lewis, Paul G. 2006 "Party Systems in Post-communist Central Europe: Patterns of Stability and Consolidation." *Democratization* 13 (August): 562-83.

Lupu, Noam and Susan Stokes. 2010. "Democracy, Interrupted: Regime Change and Partisanship in Twentieth-century Argentina." *Electoral Studies* 29 (March): 91-104

Mainwaring, Scott. 1999. *Rethinking Party Systems in the Third Wave of Democratization: The Case of Brazil*. Stanford, CA: Stanford University Press.

Mainwaring, Scott, Annabella España, and Carlos Gervasoni. 2009. "Extra System Volatility and the Vote Share of Young Parties." *Paper presented at the Annual Meeting of the Canadian Political Science Association*, May

Mainwaring, Scott, Annabella España, and Carlos Gervasoni. 2016. "Extra- and within-system electoral volatility." *Party Politics*. Posted 11 January.

Mainwaring, Scott, and Timothy Scully. 1995. *Building Democratic Institutions: Party Systems in Latin America*. Stanford, CA: Stanford University Press.

Mainwaring, Scott, and Edurne Zoco. 2007. "Political Sequences and the Stabilization of Inter-party Competition." *Party Politics* 13 (March): 155-178.

Mair, Peter. 2007. "Party Systems and Alternation in Government, 1950-2000: Innovation and Institutionalization". In Siri Gloppen and Lise Rakner (eds.) *Globalisation and Democratisation: Challenges for Political Parties*, Bergen: Fagbokforlaget.

Markowski, Radosław. 2008. "The 2007 Polish Parliamentary Election: Some Structuring, Still a Lot of Chaos." *West European Politics* 31 (September): 1055-1068.

Meleshevich, Andriy. 2007. *Party Systems in Post-Soviet Countries: a Comparative Study of Political Institutionalization in the Baltic States, Russia, and Ukraine*. New York: Palgrave MacMillan.

Millard, Frances. 2004. *Elections, Parties and Representation in Post-communist Europe*. New York: Palgrave Macmillan.

Morlino, Leonardo. 1998. *Democracy between Consolidation and Crisis: Parties, Groups, and Citizens in Southern Europe*. Oxford: Oxford University Press.

Ocaña, Francisco A. 2007. "An Approximation Problem in Computing Electoral Volatility." *Applied Mathematics and Computation* 192(September): 299–310.

Pedersen, Markus N. 1979. "The Dynamics of European Party Systems: Changing Patterns of Electoral Volatility." *European Journal of Political Research* 7 (March): 1-26.

Powell, Eleanor N., and Joshua A. Tucker. 2009. "New Approaches to Electoral Volatility: Evidence from Postcommunist Countries." Conference of the American Political Science Association, 13 August 2009, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1449112.

Powell, Eleanor N., and Joshua A. Tucker. 2014. "Revisiting Electoral Volatility in Post-communist Countries: New Data, New Results and New Approached." *British Journal of Political Science* 44 (January): 123-147.

Przeworksi, Adam. 1975. "Institutionalization of Voting Patterns, or Is Mobilization the Source of Decay?" *American Political Science Review* 6 (March): 49-67.

Sikk, Allan. 2005. "How Unstable? Volatility and the Genuinely New Parties in Eastern Europe." *European Journal of Political Research* 44 (May): 391-412.

Tavits, Margit. 2005. "The Development of Stable Party Support: Electoral Dynamics in Post-Communist Europe." *American Journal of Political Science* 49 (April): 283-298.

Tavits, Margit. 2008. "On the Linkage between Electoral Volatility and Party System Instability in Central and Eastern Europe." *European Journal of Political Research* 47 (August): 537-55.

Tóka, Gabor and Henjak, Andrija. 2007. "Party Systems and Voting Behaviour in the Visegrad Countries 15 Years after Transition." in Pavel Šaradín and Eva Bradová (eds.) *Visegrad Votes: Parliamentary Elections 2005-2006*. Olomouc: Palacky University Press, available online: http://www.personal.ceu.hu/staff/Gabor_Toka/Papers/TokaHenjak07.pdf.

Weghorst, Keith R., and Michael Bernhard. 2014. "From Formlessnesss to Structure? The Institutionalization of Competitive Party Systems in Africa." *Comparative Political Studies* 47 (October): 1707-1737.

Table 1. Electoral volatility in 10 EU post-communist democracies (1990-2010)

Country	**Author(s)**	**Average electoral volatility during election period**				
		1st electoral period	**2nd electoral period**	**3rd electoral period**	**4th electoral period**	**5th electoral period**
Bulgaria	Bielasiak	17.5	20.2	-	-	-
	Birch	10.0	22.0	-	-	-
	Dassonneville & Hooghe	17.8	21.8	31.3	32.3	35.8
	Lane	21.1	22.1	30.7	-	-
	Powell & Tucker	21.7	39.3	57.2	41.2	54.5
	Sikk	19.1	21.9	-	-	-
Czech Republic	Bågenholm	24.2	19.2	15.6	16.2	19.7
	Bakke & Sitter	24.2	20.3	17.8	-	-
	Bielasiak	18.5	13.0	-	-	-
	Birch	16.1	7.6	-	-	-
	Dassonneville & Hooghe	24.2	20.7	17.6	17.5	21.3
	Lane	27.0	18.2	17.5	-	-
	Mainwaring et al	31.4	25.0	21.9	-	-
	Powell & Tucker	38.5	16.7	19.4	17.3	-
	Sikk	27.0	21.4	-	-	-
	Tóka & Henjak (Raw)	72.4	31.4	18.6	30.9	28.5
	Tóka & Henjak (Adjusted)	21.3	29.2	16.4	16.6	20.9
Estonia	Bågenholm	39.6	30.9	31.9	29.5	-
	Bielasiak	28.4	25.9	31.4	-	-
	Birch	13.2	12.1	-	-	-
	Dassonneville & Hooghe	41.7	34.8	32.0	29.8	26.4
	Lane	27.9	30.7	32.2	-	-
	Lane	26.1	28.8	25.2	-	-
	Mainwaring et al	55.6	48.8	44.7	-	-
	Meleshevich	68.0	58.5	50.9	-	-
	Powell & Tucker	62.9	44.4	40.8	39.7	-

	Sikk	21.3	22.7	-	-	-
Hungary[a]	Bågenholm	22.3	26.4	24.1	19.1	21.8
	Bakke & Sitter	26.8	29.2	25.8	-	-
	Bielasiak	26.4	27.4	25.8	-	-
	Birch	23.9	29.4	19.1	-	-
	Dassonneville & Hooghe	44.5	38.1	32.6	26.5	27.6
	Mainwaring et al	28.9	31.3	29.2	-	-
	Powell & Tucker	40.9	26.2	28.3	10.4	-
	Sikk	23.7	28.2	-	-	-
	Tóka & Henjak (Raw)	28.3	33.6	54.3	49.5	-
	Tóka & Henjak (Adjusted)	25.8	31.7	18.2	8.4	-
Latvia	Bågenholm	39.8	41.8	42.8	38.9	-
	Bielasiak	33.5	29.0	36.4	-	-
	Birch	25.7	23.0	-	-	-
	Dassonneville & Hooghe	42.1	44.1	45.4	40.6	38.4
	Lane	37.9	41.5	43.0	-	-
	Mainwaring et al	47.2	51.9	52.0	-	-
	Meleshevich	55.1	57.2	62.3	-	-
	Powell & Tucker	54.8	58.3	62.3	26.6	-
	Sikk	36.6	40.9	-	-	-
Lithuania[a]	Bågenholm	30.4	35.2	40.1	38.0	-
	Bielasiak	28.9	39.3	-	-	-
	Birch	20.4	36.7	-	-	-
	Dassonneville & Hooghe	35.8	42.3	45.1	42.2	-
	Lane	37.4	42.9	45.2	-	-
	Meleshevich	64.0	69.1	-	-	-
	Powell & Tucker	73.1	71.8	85.1	47.0	-
	Sikk	35.9	42.2	-	-	-
	Bågenholm	24.4	20.0	25.4	27.6	26.9
	Bakke & Sitter	34.3	27.8	35.6	-	-

Poland	Bielasiak	26.8	24.6	30.7	-	-
	Birch	17.9	12.3	30.0	-	-
	Dassonneville & Hooghe	34.3	32.7	37.0	39.7	36.7
	Gwiazda	28.5	22.4	29.8	30.3	28.5
	Lane	19.4	21.5	25.8	-	-
	Mainwaring et al	35.4	43.5	47.8	45.5	-
	Markowski	34.8	27.0	34.4	35.4	33.3
	Powell & Tucker	41.6	64.8	54.4	35.3	34.1
	Sikk	31.7	26.4	-	-	-
	Tóka & Henjak (Raw)	47.4	67.0	83.5	52.4	-
	Tóka & Henjak (Adjusted)	34.5	19.5	51.6	41.1	-
Romania	Bågenholm	14.4	22.4	21.5	21.1	-
	Bielasiak	12.4	16.7	-	-	-
	Birch	14.0	16.6	-	-	-
	Dassonneville & Hooghe	14.2	22.4	20.9	21.1	-
	Lane	16.8	25.8	21.2	-	-
	Mainwaring et al	29.1	38.4	37.6	-	-
	Powell & Tucker	49.2	52.6	30.9	49.6	-
	Sikk	14.3	21.7	-	-	-
Slovakia	Bågenholm	20.5	21.3	25.2	24.1	24.5
	Bakke & Sitter	20.6	20.5	23.7	-	-
	Bielasiak	13.3	16.7	-	-	-
	Birch	12.8	9.4	-	-	-
	Dassonneville & Hooghe	31.1	32.4	39.0	35.4	34.4
	Lane	13.7	17.6	21.3	-	-
	Powell & Tucker	45.0	61.0	68.9	39.6	-
	Sikk	13.6	16.9	-	-	-
	Tóka & Henjak (Raw)	55.7	43.9	53.9	63.1	33.0
	Tóka & Henjak (Adjusted)	19.6	23.8	21.4	29.9	32.9

Slovenia	Bågenholm	22.0	24.8	24.3	27.0	-
	Bielasiak	25.4	23.8	-	-	-
	Birch	-	17.1	-	-	-
	Dassonneville & Hooghe[b]	31.0	28.0	26.0	27.0	-
	Lane	24.4	21.6	21.8	-	-
	Powell & Tucker	39.9	34.4	52.2	41.3	-
	Sikk	23.8	21.3	-	-	-
	Tóka & Henjak	50.8	39.4	19.7	31.3	-

[a]Although it is not always clear from the published sources, our conversations with authors indicate that all of these measures refer to the proportional representation segments within these mixed systems.

[b]Interpolated from graphic presentation of data in Dassonneville and Hooghe (2011).

Sources: Bågenholm (2009); Bakke and Sitter (2005); Bielasiak (2002; 2005); Birch (2003); Dassonneville and Hooghe (2011; 2015); Gwiazda (2009); Lane and Ersson (2008); Mainwaring et al. (2009; 2016); Markowski (2008); Meleshevich (2007); Powell and Tucker (2009; 2014); Sikk (2005); Tóka and Henjak (2007) (original data files and answers to process questions provided in personal follow-up conversation)

Table 2. Volatility calculation Values for hypothetical Party A according to various threshold inclusion methods.

Threshold inclusion method	**Threshold example**	**Data category**	**Election 1**		**Election 2**		**Election 3**		**Election 4**	**Total contribution to volatility**
No threshold	0%	Actual Value	1%		5%		1%		3%	
		Change		*+4*		*-4*		*+2*		*10*
Individual points	2%	Adjusted value	0%*		5%		0%*		3%	
		Change		*+5*		*+5*		*+3*		*13*
	4%	Adjusted value	0%*		5%		0%*		0%*	
		Change		*+5*		*-5*		*0*		*10*
Transition pairs	2%	Adjusted value	1%		5%		1%		3%	
		Change		*+4*		*-4*		*+2*		*10*
	4%	Adjusted value	1%		5%		1%		0%*	
		Change		*+4*		*+4*		*-1*		*9*
Blanket inclusion	2%	Adjusted value	1%		5%		1%		3%	
		Change		*+4*		*-4*		*+2*		*10*
	4%	Adjusted value	1%		5%		1%		3%	
		Change		*+4*		*-4*		*+2*		*10*

*Value recalculated from actual election value due to threshold inclusion method.

Table 3. Method choices of authors in Pedersen Index calculations

Author	Dealing with size				Dealing with change		
	Inclusion standards and method				Continuity method		
	Threshold for inclusion?	Size of threshold	Threshold inclusion method	Excluded data method	How to determine successor in case of party split	How to determine predecessor in case of party merger	Categorization
Bågenholm	Yes	Election threshold	Individual points	Aggregated and linked other	Party of same name; if not, no successor	Sum of predecessors	Mixed method (Inclusive aggregation for predecessor, Relaxed linkage for successor)
Bakke	Yes	2%	Transitional pairs	Aggregated and linked other	Sum of successors	Sum of predecessors	Inclusive aggregation
Bielisiak	No	-	-	Aggregated and linked other	Sum of successors	Sum of predecessors	Inclusive aggregation
Birch	No	-	-	Aggregated and linked other	Party of same name; if not, no successor	Party of same name; if not, merger is new	Strict linkage
Dassonenville & Hooghe	Yes	1%	Individual points	Aggregated and linked other	Sum of successors	Sum of predecessors	Inclusive aggregation
Gwiazdka	No	-	-	None	Largest offspring party	Sum of predecessors	Mixed method (Inclusive aggregation for predecessor, Relaxed linkage for successor)

Lane & Ersson	No	-	-	Aggregated and linked other	Sum of successors	Sum of predecessors	Inclusive Aggregation
Mainwaring, España & Gervasoni	No	-	-	None	Party of same name; if not, largest offspring	Party of same name; if not, largest parent*	Relaxed linkage
Markowski	No	-	-	None	Sum of successors	Sum of predecessors	Inclusive aggregation
Meleshevich	No	-	-	None	Author chooses successor	Author chooses predecessor	Mixed linkage (Relaxed linkage and Strict linkage)
Powell & Tucker	Yes	2%	Individual points	Uncalculated	Party of same name; if not, no successor	All mergers new unless under 5%	Strict linkage
Sikk	No	-	-	Aggregated and linked other	Sum of successors	Sum of predecessors	Inclusive aggregation
Tóka (raw)	No	-	-	None	Party of same name; if not, largest offspring	Party of same name; if not, largest parent	Relaxed linkage
Tóka (adjusted)	No	-	-	None	Sum of successors	Sum of predecessors	Inclusive aggregation

*All authors use an inclusive aggregation approach for electoral coalitions except for Mainwaring, España and Gervasoni, who in this case adopt a relaxed linkage approach by using the largest predecessor party.

Table 4: The relative Importance of linkage, threshold and exclusion

Model		**Model 1**	**Model 2**	**Model 3**
Description		**Basic Model**	**Basic +Interactions**	**Only Significant**
Number of observations		3,960	3,960	3,960
R-squared		0.673	0.676	0.676
Root MSE		0.082	0.082	0.082
Constant		0.200***	0.188***	0.189***
		(0.006)	(0.007)	(0.006)
Continuity method (baseline = Inclusive aggregation)	Relaxed linkage	0.048***	0.070***	0.070***
		(0.003)	(0.005)	(0.005)
	Strict linkage	0.152***	0.174***	0.174***
		(0.004)	(0.007)	(0.007)
Threshold size		-0.809***	-0.358	-0.457**
		(0.082)	(0.197)	(0.159)
Threshold inclusion method (baseline = Individual points)	Transitional pairs	-0.003	0.002	-
		(0.003)	(0.006)	-
	Blanket inclusion	-0.002	0.001	-
		(0.003)	(0.006)	-
Excluded data method	Aggregated linked other	0.013***	0.002	0.002
		(0.003)	(0.005)	(0.005)
Interaction: Threshold size with Continuity method (baseline = aggregation)	Relaxed linkage	-	-0.821***	-0.829***
		-	(0.160)	(0.159)
	Strict linkage	-	-0.901***	-0.901***
		-	(0.228)	(0.227)
Interaction: Threshold size with Threshold inclusion method (baseline = Individual points)	Transitional pairs	-	-0.200	-
		-	(0.201)	-
	Blanket inclusion	-	-0.098	-
		-	(0.202)	-
Interaction: Threshold size with Excluded data method	Aggregated linked other	-	0.435**	0.435**
		-	(0.164)	(0.164)

Note: OLS with Robust Standard Errors (in brackets). ** $p < .01$; *** $p < .001$

Figure 1. Averages of pair-wise differences in election-period volatility results according to author.

Figure 2a. Volatility in the Czech Republic, 1990-2013

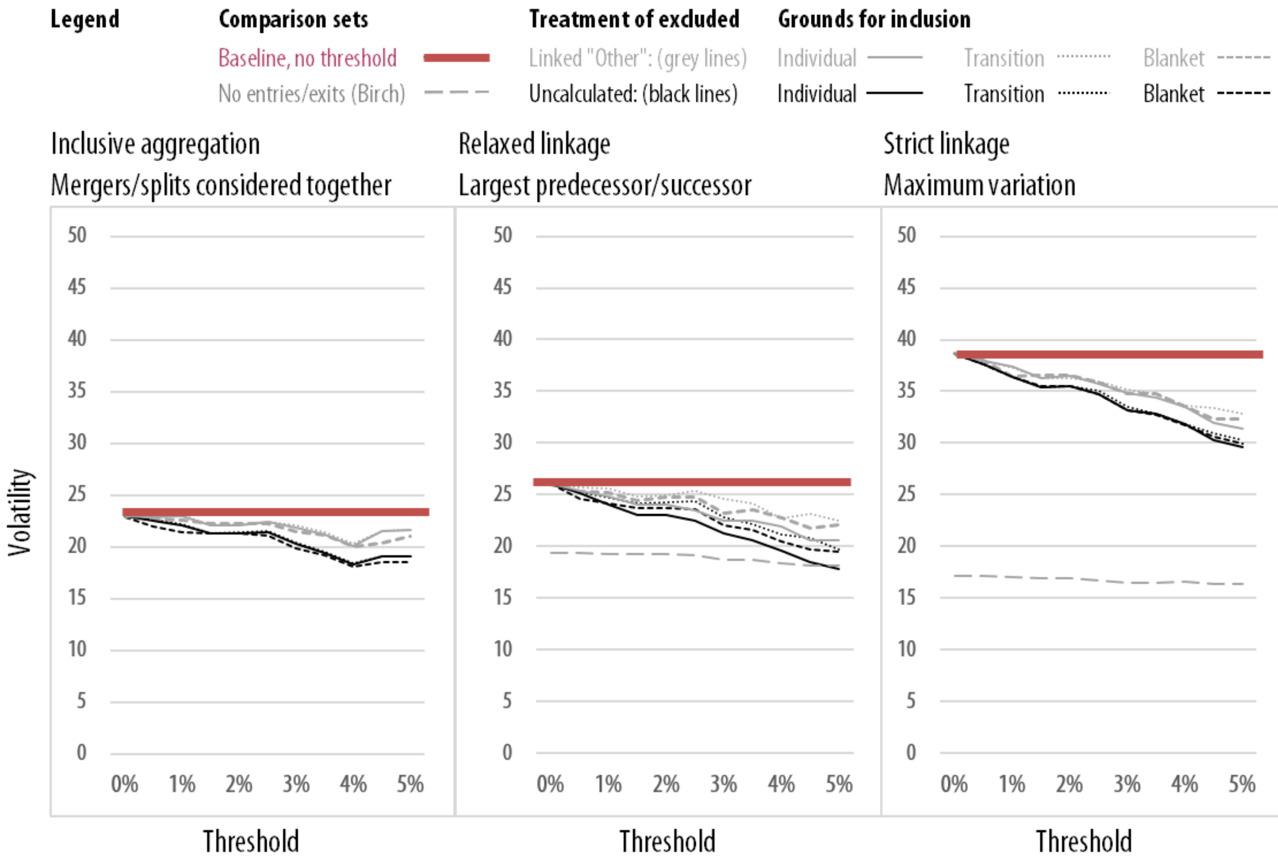

Figure 2b. Volatility in Hungary, 1990-2014

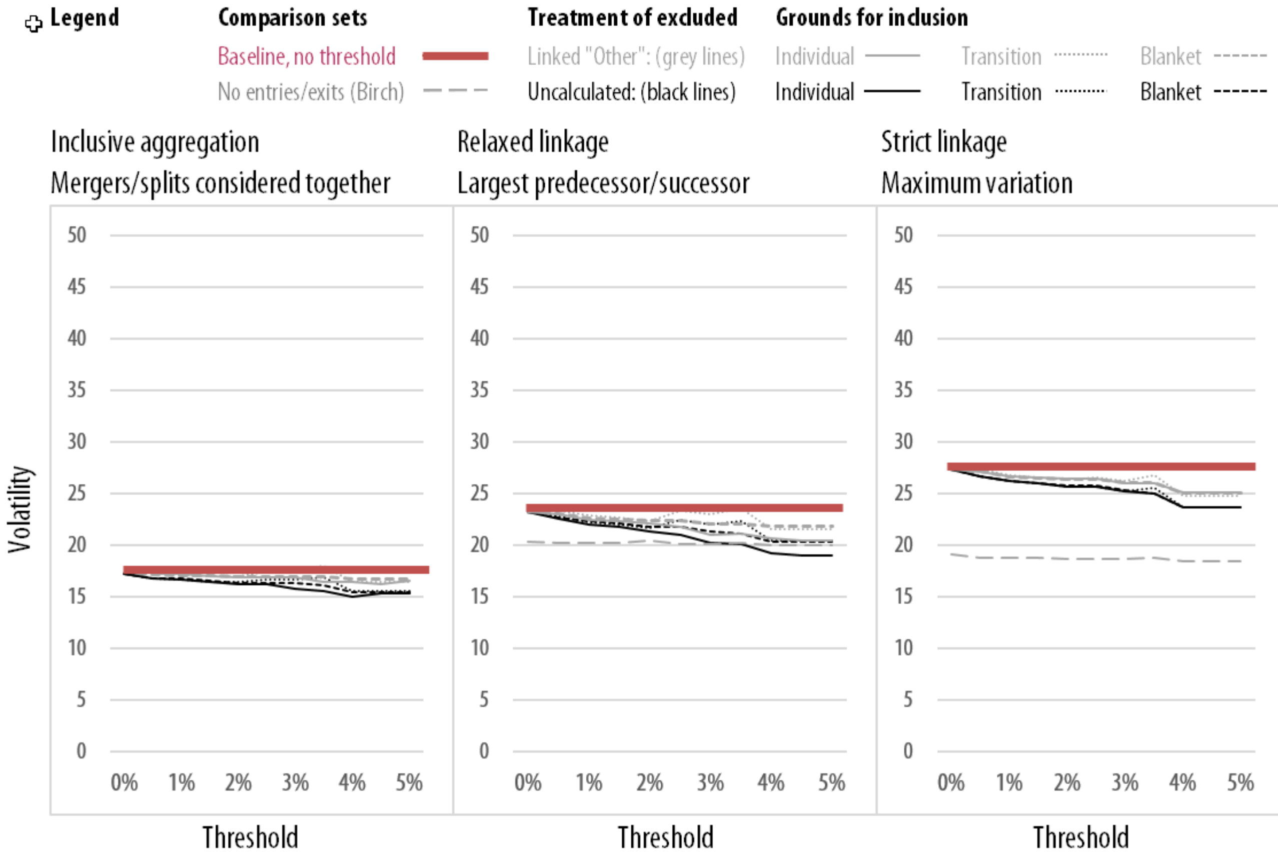

Figure 2c. Volatility in Slovakia, 1990-2012

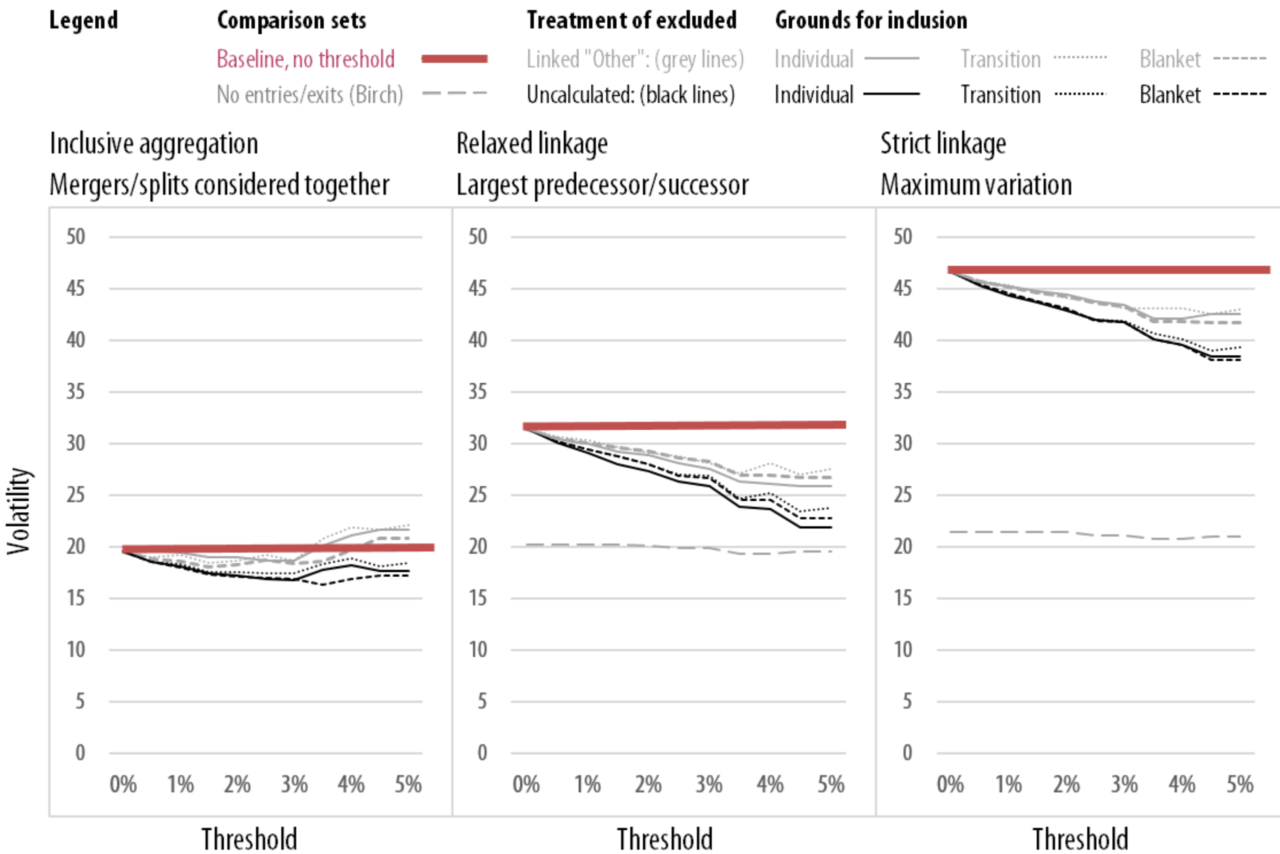

Figure 3: The Inclusive aggregation method's "Z-effect"

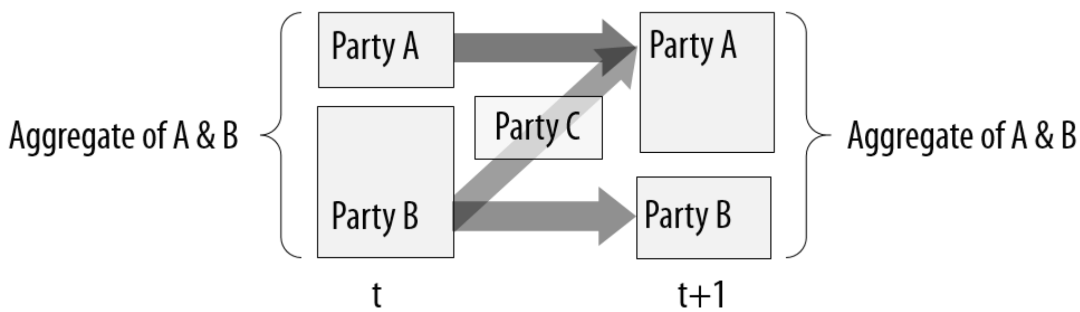

Figure 4. Volatility scores for Inclusive aggregation and Strict linkage methods in the Czech Republic, Hungary and Slovakia, 1990-2014.