Influence of non-voters on electoral outcomes using the Kohler-Rose-Index*

Ulrich Kohler Steffen Zittlau Lukas Hoettges

version 1.2, December 20, 2017

* We wish to thank Trond Kvamme from the EED for his collaboration. We also like to thank Christoph Thewes for careful assistance.

Contents

1 Introduction **3**

2 Short reference **3**
2.1 Formulas 3
2.2 Data sources 4

3 More on the Kohler-Rose-Index **4**
3.1 Derivation 4
3.2 Further assumptions used 5
3.3 Likelihood of change 5

4 Data **6**
4.1 Requirements 6
4.2 Data for Europe 7
4.3 Data for Germany 8

References **8**

1 Introduction

Non-voters are often the largest group in the electorate and this is considered a problem in the public and also by many social scientists (Lijphart 1997). An analytic question raised by the debate on the importance of non-voting is stated by Bernhagen and Marsh (2007: 548): 'If turnout was 100 percent, would it affect the election result?' One answer to this question can be given by a statistic proposed by Kohler and Rose (2010). The index calculates the proportion of votes the runner-up party would have to win among the non-voters in order to change hands with the party who gains the plurality of votes (hereafter: the "winner"). If this Kohler-Rose-Index is higher than 100%, it would be impossible for the runner up to catch up the winner by gaining votes from the non-voters.

The following text documents details of the numbers provided on the web-site `http://www.wzb.eu/de/content/influence-of-non-voters` and gives further references.

2 Short reference

2.1 Formulas

The Kohler-Rose-Index P_{KR} is

$$P_{\mathrm{KR}} = \frac{W - R + L(1 - \frac{o}{w+o})}{L(2 - \frac{o}{w+o})} \tag{1}$$

with W and R being the absolute numbers of valid votes for the winning and runner-up party respectively, L being the absolute numbers of non-voters that could have participated in the election ("Leverage"), o being the relative share of valid votes for other parties who are neither the winner nor the runner-up, and w being the winning party's relative share of valid votes.

The Leverage L is

$$L = T_{\text{Max}} - V \tag{2}$$

with T_{Max} being the number of voters who could possibly participate at the election on the election day, and V being the number of valid votes cast. T_{Max} has been set to 96% of the registered electorate.

The absolute difference D_{abs} is

$$D_{\text{abs}} = P_{\text{KR}} - \frac{R}{V} \quad . \tag{3}$$

Likewise, the relative difference is D_{rel} is

$$D_{\text{rel}} = \frac{P_{\text{KR}} - \frac{R}{V}}{\frac{R}{V}} \quad . \tag{4}$$

2.2 Data sources

The data for calculating the Kohler-Rose-Index for the German federal elections and the German state elections has been taken from Statistisches Bundesamt (2005) and the website of the German Federal returning officer (http://www.bundeswahlleiter.de/en/index.html).

The data for calculating the Kohler-Rose-Index for the 32 European countries has been taken from the European Election Database (EED) on http://www.nsd.uib.no/european_election_database.

3 More on the Kohler-Rose-Index

3.1 Derivation

Let W and R be the absolute numbers of valid votes of the winning and runner-up party respectively, and L be the absolute numbers of non-voters that could have participated in the election (i.e. the Leverage). Assume that the (unknown) proportion of the non-voters who vote for the runner up party would be p. Then, the new number of votes for this party would be

$$R^* = R + p \cdot L \quad . \tag{5}$$

After subtracting pL from the number of nunvoters, the remaining absolute number of non-voters is then $(1-p)L$. In order to distribute these remaining non-voters between the winning parties and all other parties (except the runner-up party), denote the proportion of

votes for the winning party in the actual election to be w, and the proportion of votes for all other parties to bo o. Now, assume that both, the winning party and the other parties, keep their relative proportions, i.e. the winner party gets a poportion of $\frac{w}{w+o}$ percent of $(1-p)L$, and the other parties get $\frac{o}{w+o}$ of $(1-p)L$. In that case the new number of votes for the winning party would be

$$\begin{aligned} W^* &= W + L - pL - \frac{o}{w+o}(1-p)L \\ &= W + \left(1 - p - \frac{o}{w+o}(1-p)\right)L \quad . \end{aligned} \tag{6}$$

Setting the difference between (6) and (5) to zero, and solving for p leads to

$$p_{\mathrm{KR}} = \frac{W - R + L(1 - \frac{o}{w+o})}{L(2 - \frac{o}{w+o})} \quad . \tag{7}$$

The statistic p_{KR} has been first proposed by Kohler and Rose (2010), and we will use the term "Kohler-Rose-Index" throughout this paper to refer to it.

3.2 Further assumptions used

In the formula for the Kohler-Rose-Index there is the unknown quantity of available non-voters, L. Conceptually, the number of available non-voters is equal to the maximum turnout minus the observed valid turnout, with maximum turnout being the highest possible turnout. Kohler and Rose (2010) argue that the highest possible turnout cannot be 100% for a number of reasons. In this article, the estimation of maximum turnout was based on the average number of absent electors and invalid votes in compulsory voting countries with strong enforcement (Australia, Belgium, Luxembourg), which led to 86.1% of the electorate as maximum turnout. While setting up the data base for this web-page it has turned out, however, that several elections had a higher turnout than 86.1. Maximum turnout has been therefore set to 96%, which is slightly above the highest valid turnout observed in the data (Malta 1990: 95.5%). From the standpoint of the original publication, the Kohler-Rose-Index given on the web-site is therefore biased downwards, or, in other words, the potential influence of non-voters might be too overstated.

3.3 Likelihood of change

Assuming that other parties stay as they are, the Kohler-Rose-Index calculates the proportion of the available non-voters the runner-up party must win to close the gap to the winner in the counter-factual situation of maximum turnout. A value of, say, 40% would mean for example that ...

> ... the runner up party would have needed 40% of the non-voters to close the gap to the winner party if all non-voters had participated in the election.

A Kohler-Rose-Index of more than 100% would mean that the runner-up party would need more votes than actually available to bridge the gap to the winner. In this case, the positions of winner and runner-up changing hands would be considered impossible.

In the more general case, the question whether it is realistic that the runner up party gain 40% of the non-voter's votes cannot be answered by the Kohler-Rose-Index alone. It should be clear, however, that winning these 40% would be more likely, if the runner-up party had won 35% in the real world election than if the runner-up had won just around 15%. The difference between the real world election result of the runner-up party and the Kohler-Rose-Index is therefore a way to operationalize the likelihood of change due to non-voters.

Based on these considerations, we propose two derived statistics for quantifying the likelihood of change due to non-voters: First the absolute difference D_{abs} of the Kohler-Rose-Index and the actual vote proportion of the runner-up party, i.e.

$$D_{\text{abs}} = P_{\text{KR}} - \frac{R}{V} \quad . \tag{8}$$

A value of $D_{\text{abs}} = 40\%$ would mean, that the runner up party would need a vote proportion among the non-voters that is 40 percentage points higher, than the actual vote-proportion among the voters. The higher this number, the less likely the change due to non-voters.

The second derived statistics expresses D_{abs} as relative change, i.e.

$$D_{\text{rel}} = \frac{D_{\text{abs}}}{\frac{R}{V}} = \frac{P_{\text{KR}} - \frac{R}{V}}{\frac{R}{V}} \quad . \tag{9}$$

A value of $D_{\text{rel}} = 100\%$ would mean that the runner-up party would have to double their real world vote-proportion among the non-voters . Again, the higher this number, the less likely the change due to non-voters.

It should be noted that, regardless of the values of D_{rel} and D_{abs}, other conditions of the real world election might also effect the likelihood of change. To help you to come up with an informed guess about the likelihood of change, the figures on the web-site also show the names of the respective runner-up parties.

A more thorough estimation of the likelihood of government change due to non-voters can be found in Kohler (2011).

4 Data

4.1 Requirements

The potential influence of one single non-voter decreases with the size of the electorate. It is therefore not possible to arrive at a sensible statistic for the influence of non-voters from election results expressed in percentage points. The derivation of the Kohler-Rose-Index therefore started from absolute numbers, and the values of the Kohler-Rose-Index cannot be calculated using percentage points as input. The datasets to calculate the Kohler-Rose-Index therefore require information about

- the absolute numbers of votes received by individual parties participating in elections,
- the total number of valid votes, and
- the size of the electorate.

Secondly, the original data should stem from trusted sources such as official electoral commissions and preferably be collected by a singular entity to assure that the data is, formally and with regards to content, comparable.

Thirdly, as we want to answer whether non-voters matter overall, we appreciate the data to include a large number of elections. This translates into maximizing the number of political systems included and maximising the covered period of time.

4.2 Data for Europe

We find all criteria to be best fulfilled by the data provided by the European Election Database (EED)[1]. The EED project started out under the European Sixth Framework Research Programme and was sponsored by the European Commission with the purpose of collecting election results across Europe. The project is now overseen and constantly updated by Norwegian Social Science Data Services (NSD). The EED tracks elections in 35 countries since 1990. The countries include all EU member countries, as well as Iceland, Norway, Switzerland, Liechtenstein, Croatia, Montenegro, Macedonia and Turkey. The recorded data includes the absolute number of votes each party received in an election, but also aggregate information such as the size of the electorate and the total number of valid and invalid votes. For the majority of elections, the data is available in statistical software-ready format and can be easily collected.

While the EED data nicely fits our purposes, additional notes are necessary for a number of cases:

- The EED does not provide structured data for Italy.
- The electoral system of Liechtenstein and Luxembourg provide voting individuals with multiple votes, each of which is recorded as a separate vote in the official records, leading to the situation that the number of votes cast exceeds the number of voters. For Liechtenstein and Luxembourg we have therefore used the number of possible votes for defining T_{Max} instead of the number of voters.
- The EED data does not include data on the size of the electorate for the UK, which is rectified by inserting data from official British sources, i.e. the Electoral Commission.
- For a number of country-elections (IE, RO, LV, LT) we use the International Institute for Democracy and Electoral Assistance's (IDEA) dataset on Voter Turnout to

1 http://www.nsd.uib.no/european_election_database

recover missing data, mainly for the size of the electorate and the number of valid votes. As we came across some significant discrepancies between the IDEA and EED data, we only used data from the IDEA for countries where we did not find substantial mismatches for elections where both sources provide data. We fail to recover the necessary data only for two elections.

- To further improve the integrity of the collected data, we checked whether electoral alliances consisting of more than one party were jointly counted as one party in the EED data. The ParlGov project Döring et al. (2010)[2] coded the relevant information for a broad range of countries, which we were able to employ to identify electoral alliances in 28 of the 32 countries included in our data. For the remaining four countries (ME, MK, TR, HR) we conducted a thorough web research to make sure that the member parties of electoral alliances were not counted as separate parties.

Following the above cited guidelines, we yield a data set that covers 180 parliamentary elections from 34 European countries between 1990 and 2011.

4.3 Data for Germany

References

Bernhagen, P., Marsh, M. (2007): The partisan effects of low turnout: Analyzing vote abstention as a missing data problem. *Electoral Studies*, 26, 548–560, download: kostenpflichtig.

Döring, H., Manow, P. . P., government composition database (ParlGov): An infrastructure for empirical information on political institutions Version 10/02. (2010): Parliament and government composition database (ParlGov): An infrastructure for empirical information on political institutions - Version 10/02, http://dev.parlgov.org/.

Kohler, U. (2011): Estimating the Potential Impact of Nonvoters on Outcomes of Parliamentary Elections in Proportional Systems with an Application to German National Elections from 1949 to 2009. *Electoral Studies*, 30, 497–509.

Kohler, U., Rose, R. (2010): Under what circumstances could maximizing turnout alter an election result? *Representation*, 46(2), 119–138.

Lijphart, A. (1997): Unequal participation: Democracy's unresolved dilemma - Presidential Address 1996. *American Political Science Review*, 91(1), 1–14, http://www.jstor.org/stable/pdfplus/2952255.pdf?acceptTC=true.

Statistisches Bundesamt (2005): Endgültiges Ergebnis der Wahl zum 16. Deutschen Bundestag am 18. September 2005. *Wirtschaft und Statistik*, 11/2005, 1153–1167.

2 http://dev.parlgov.org/ The used data set is still in the development stage.