Jahrbücher f. Nationalökonomie u. Statistik (Lucius & Lucius, Stuttgart 2011) Bd. (Vol.) 231/5+6

When Does the Second-Digit Benford's Law-Test Signal an Election Fraud?

Facts or Misleading Test Results

By Susumu Shikano and Verena Mack, Konstanz*

JEL C63; D72

Election fraud; Benford's Law; German federal election; simulation.

Summary

Detecting election fraud with a simple statistical method and minimal information makes the application of Benford's Law quite promising for a wide range of researchers. Whilst its specific form, the Second-Digit Benford's Law (2BL)-test, is increasingly applied to fraud suspected elections, concerns about the validity of its test results have been raised. One important caveat of this kind of research is that the 2BL-test has been applied mostly to fraud suspected elections. Therefore, this article will apply the test to the 2009 German Federal Parliamentary Election against which no serious allegation of fraud has been raised. Surprisingly, the test results indicate that there should be electoral fraud in a number of constituencies. These counterintuitive results might be due to the naive application of the 2BL-test which is based on the conventional χ^2 distribution. If we use an alternative distribution based on simulated election data, the 2BL-test indicates no significant deviation. Using the simulated election data, we also identified under which circumstances the naive application of the 2BL-test is inappropriate. Accordingly, constituencies with homogeneous precincts and a specific range of vote counts tend to have a higher value for the 2BL statistic.

1 Introduction

Recently, a growing number of studies have applied tests based on Benford's Law to detect electoral fraud. According to Benford's Law, leading digits of naturally occurring tables of numerical data are not uniformly distributed (Hill 1995), instead they follow a specific distribution originally found by Newcomb (1881). Hence, manipulated election results should show a significant deviation from the distribution. The Benford's Law-test has an important advantage that one only needs corresponding electoral results while other methods in general require further information like past election results, demographic information etc. Walter Mebane is the first in applying this method to detect electoral fraud. More specifically, he applied the "second-digit Benford's Law-test" (the 2BL-test) to US-presidential elections (Mebane 2006b, 2007a, 2008a) and sub-

* A part of this paper's result was presented in the first author's inaugural lecture at the University of Konstanz and in his fellow lecture at the Hanse Institute for Advanced Study. We thank the participants of both lectures for their helpful comments and feedbacks. We are also indebted to Walter Mebane for providing his programs and to Nadine Meidert for her comments on an earlier version of the manuscript. Further, the first author is grateful to the Hanse Institute for Advanced Study for the provided support.

sequently also analyzed further countries' elections including Mexico, Indonesia, Russia, and Iran among others (Mebane 2007b, Mebane/Kalinin 2009, Mebane 2010). Besides Mebane's work, we can find further applications of Benford's Law-test to different political systems (e.g. Pericchi/Torres 2004: Roukema 2009).

While there is an increasing number of application of this method, some scholars are skeptical about the validity of the Benford's Law-test to detect fraud. According to Deckert, Myagkov and Ordeshook (2010), applying Benford's Law for detection of fraud lacks any theory or model. They further criticized that Benford's Law-tests can indicate electoral frauds in some innocent cases. In this regard, Diekmann and Jann (2010) who applied Benford's Law to linear regression estimates, emphasize the Benford distribution is inappropriate to discriminate manipulated from non-manipulated estimates. Instead of Benford's distribution itself, Mebane (2007c) is rather skeptical about using the χ^2-test to decide whether election data has been manipulated. Therefore, he utilized some simulations to obtain an alternative distribution to χ^2. His latest study concludes "the evidence is strong that departures from 2BL, which also occur frequently, are related both to normal political phenomena and to serious election anomalies" (Mebane 2010: 26).

This academic disputes concerning Benford's Law-test, have an important deficit. Most studies applied Benford's Law to authoritarian systems or non-established democracies, which are suspected of electoral fraud prior to the test.[1] The United States are indeed categorized as an established democracy, but the validity of its elections has been questioned at least since the 2000 presidential election. To fill this research gap, this paper applies the 2BL-test to the national-level election result from a developed West-European democracy. More concretely, we analyze the 2009 German Federal Parliamentary Election (Bundestagswahl) against which no serious doubt of electoral frauds has yet been raised. The naive application of the 2BL-test shows, to our surprise, that digit frequencies do not follow the Benford's distribution in an over-proportionately high number of electoral districts. If we, however, use an alternative distribution to χ^2 as Mebane (2007c) did, the number of deviating electoral districts was reduced to the level of coincidence. This indicates that the naive applications of the 2BL-test is inappropriate for many German electoral districts. Which circumstances make the 2BL-test appropriate or inappropriate? This is the question which we are going to tackle.

In the remainder, we proceed as follows: After introducing Benford's Law in the next section, we present the results of the naive application of the 2BL-test using German election data. Subsequently, we introduce an improved 2BL-test and apply it to German election data. The fifth section investigates the factors which can boost the 2BL statistic. The concluding section summarizes the findings and gives implication for the use of the 2BL-test in different contexts. We also suggest further research agenda.

2 Benford's Law

Benford's Law stems from the empirical observation that leading digits of numerical data are often not uniformly distributed. The non-uniform distribution of digits was first dis-

[1] Exceptionally, Beber and Scacco (2008) compare Swedish election results with Nigerian results. However, they do not use the 2BL-test but another test which focuses on the last digit. Breunig and Goerres (Forthcoming) apply the 2BL-test to the German federal elections from 1990 to 2005. In their research question, however, they observe East Germany as one of transitional democracies with possible electoral mismanagement and/or fraud.

covered by Newcomb (1881) and formalized as "*[t]he law of probability of the occurrence of numbers is such that all mantissa [base10] of their logarithms are equally probable*" (Newcomb 1881: 40). About 60 years later Benford (1938: 553) gave a wide range of data, as the area of rivers, the population of the USA or death rates, from which digits follow approximately the law. Both Newcomb and Benford give a mathematical formalization of their empirical observation, but no theoretical explanation why this phenomenon occurs so often in our surrounding. According to Benford's Law, the 1 occurs in 30.1 % and the 9 only in 4.6 % of the time as the leading digit. The probability with which digits 1 to 9 as the first digit and digits 0 to 9 as the second digit accrue is displayed in Figure 1(a). This might be surprising as we would intuitively expect the same probability for a 1 or a 9 as the leading digit. The theoretical explanation was controversially discussed by scholars over many years. According to Weaver (1963: 277), the "law for first integers is a built in characteristic of our number system".[2] This idea seems plausible if we consider integers between 1 and 20. While the probability to have 1 as the leading digit is $p = 11/20$, we only have a probability of $p = 2/20$ for 2 and $p = 1/20$ for the digits 3 to 9. The incompleteness of this argumentation is shown by Raimi (1976). The characteristics of our number system are not enough to produce digits in data that follow Benford's Law. Raimi (1976) attributes it rather to a mixture of distributions. Finally, Hill (1995) gives the formal proof that "if probability distributions are selected at random and random samples are then taken from each of these distributions in any way so that the overall process is scale (or base) neutral, then the significant-digit frequencies of the combined sample will converge to the logarithmic distribution" (Hill 1995: 360).[3]

Over the last years many scholars have shown that certain digits in election data from different countries approximately follow Benford's Law (Pericchi/Torres 2004; Mebane 2006a, 2007b,a 2008a; Beber/Scacco 2008; Mebane/Kalinin 2009; Roukema 2009; Mebane 2010). The theoretical approach was provided by Mebane (2006b). Like Raimi (1976) and Hill (1995), he argues that a mixture of some distributions will make digits follow Benford's Law (Mebane 2006b: 4). Such a mechanism could be produced by voters' preferences and mistakes in submitting and counting the vote. Both voters preferences and mistakes vary across precincts because of differences in partisanship, economic class, mobilization campaigns, administrative rules and other details (Mebane 2008a). In his latest work Mebane (2010) finds that deviations from the 2BL distribution can be both, a result of serious election anomalies as well as normal political phenomena. This indicates that we have to depart from the question whether election data follows Benford's Law and what mechanisms are responsible for such a distribution. Instead, we should ask the question when the distribution of our election data satisfies the assumptions of Benford's Law. As Fewster (2009) states, "data from *any* distribution will tend to be 'Benford', as long as the distribution spans several integers on the $\log_{10}$ scale – several orders of magnitude on the original scale – and as long as the distribution is reasonably smooth". We will analyze our data in regard to this aspect in section 4.2.

3 Naive application of the 2BL-test to the 2009 German Federal Election

If digit frequencies of the total number of votes for each party within each precinct follow Benford's Law, we should be able to find such a pattern in the 2009 German Federal

[2] A similar explanation is given by Goudsmit and Furry (1944).

[3] Similar approaches can be found by Janvresse and De La Rue (2004), Miller and Kontorovich (2005) and Berger (2005).

Parliamentary Election for the five most important parties (CDU/CSU, SPD, FDP, die Linke and Bündnis 90/Die Grünen).[4] Germany has a mixed member proportional representation election system. Each citizen has a district vote (Erststimme) and a list vote (Zweitstimme). Since the list vote is crucial for seat allocation it should be more vulnerable to election fraud. Therefore, we focus on the list vote. As unit of analysis we chose constituencies (Wahlkreise) with their precincts (Wahlbezirke). At the 2009 German Federal Parliamentary Election, Germany was divided into 299 constituencies which in average consisted out of 297 precincts whereas the smallest contained 156 and the largest 586 precincts. If any, we suspect election fraud to be most likely at the lowest level of vote aggregation. Thus, we apply the 2BL-test to each constituency separately using vote counts at the precincts level.[5]

Figure 1(b) displays the frequency of the second digit in vote counts, aggregated over the five parties and all precincts. The figure supports the naive assumption that the second digit in vote counts follow Benford's Law. The curve displays the expected distribution of second digits, the histogram gives the empirical distribution of the second digits in vote counts.

We focus on the second digit because Brady (2005) and Mebane (2006b) reasonably argue that the more or less constant precinct size can cause significant deviation from Benford's Law. Imagine a situation with about 1000 voters per precinct and voter preferences between 40 % and 50 % in all precincts of the analyzed constituency. This would result in an over proportional number of 4s as first digits without any intervention of fraud. Further support for using the second digit is given by Diekmann (2007), who shows with an experimental approach that the first digit does not detect fraudulent regression coefficients as effective as all other digits. If the first digit is not adequate to detect fraudulent data, using the second digit seems to be the best option in order to capture fraud within this given limitation. While Figure 1(b) can only give the information that German vote counts approximately follow Benford's Law, we need to test whether empirical deviations from the theoretical prediction are statistically significant. The corresponding test statistic was introduced by Mebane (2006b) and is based on the Pearsons χ^2 statistic.

$$\chi^2_{2BL} = \sum_{i=0}^{9} \frac{(d_i - dq_i)^2}{dq_i}, \tag{1}$$

where q_i denotes the expected relative frequency of i at the second digit (Figure 1(a)). d_i is the empirical frequency of second digit i in a constituency.[6] Comparing the statistic with a χ^2 distribution with 9 degrees of freedom leaves us with a critical value of 16.9, using a significance level of 5 %.

---

4 As we collected data no federal-wide election results at precincts levels were available. Therefore, the "Statistisches Landesamt" of each federal state provided us with their election result of the 2009 German Federal Parliamentary Election. Recently, the whole election data got available at the "Statistische Bundesamt". In this paper we will refer to the five listed parties only as "CDU", "SPD", "FDP", "Linke" and "Grüne".

5 A specific description how the election result in Germany is allocated can be found in "Bundeswahlordnung" § 70 and § 71.

6 More precisely, d can be smaller than the total number of precincts in a constituency if the party reached less than 10 votes in at least one precinct. This formula can be found in Mebane (2008b: 179).

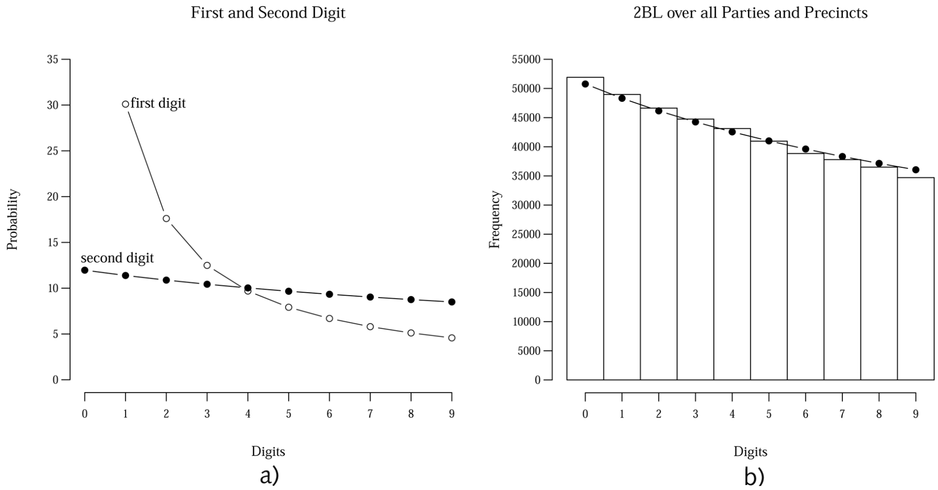


Figure 1 Expected and empirical distribution

The results for applying the 2BL-test separately for each party in each of the 299 constituencies of the 2009 German Federal Parliamentary Election are displayed in Figure 2. The grey line marks the critical value of 16.9 for $\alpha = 5\%$. Accordingly, the percentage of constituencies with a 2BL statistic larger than 16.9 is much higher than 5 %. Such constituencies are, however, expected to be 5 % or less, if the election is not fraudulent and tests are independent from each other. This is a surprising result if we consider that there

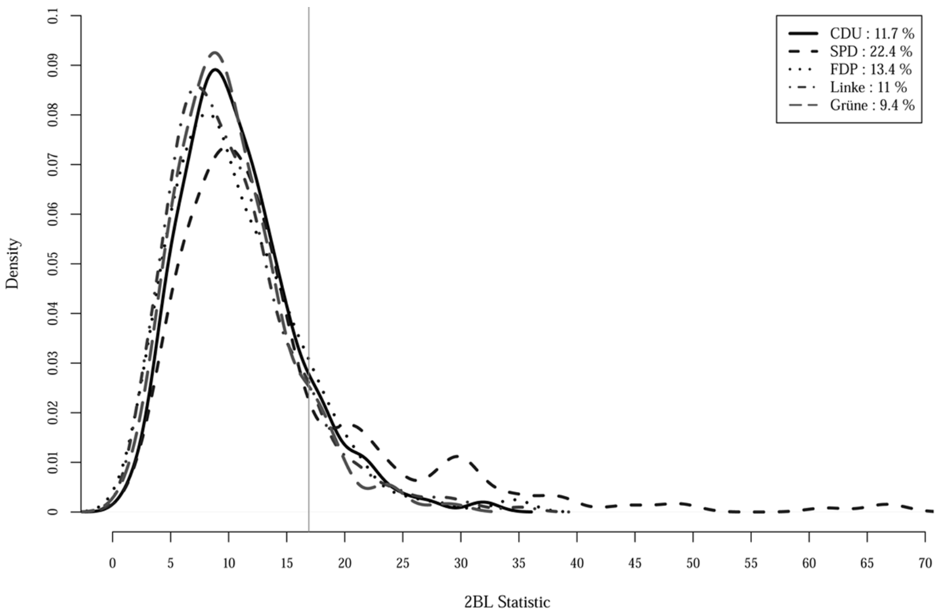


Figure 2 Density distribution of the 2BL statistic

has been no allegation of fraud, and this year's election supervisors did not find any irregularities (OSCE/ODIHR 2009). Our findings imply two possible answers: The 2009 German Federal Parliamentary Election is fraudulent which has not been detected yet or the 2BL-test statistic is misleading.

To investigate the second possibility, we address the following question: What kind of the 2BL statistic do we obtain when simulating election data, and do we need to adjust our critical value according to it?

4 Modifying the 2BL test

So far we used the χ^2 distribution to test if our empirical frequency distribution of the second digits in vote counts follows Benford's Law. This implicitly assumes that digits were independently generated according to a multinomial distribution, given a large number of precincts (Mebane 2007c). It is an unsolved issue if second digits of vote counts fulfill this assumption. As suggested by Mebane (2007c), we will use the frequency distribution of second digits in simulated vote counts to construct an alternative distribution to test the 2BL statistics. This makes the test results independent from the formal χ^2 distribution which might be inappropriate. Mebane (2007c) uses the information about precincts and recorded election vote counts to simulate the mechanisms which produce digits in vote counts that follow Benford's Law. We follow this modeling strategy of Mebane, however, with some mofidications.

4.1 Simulation model

The basic idea to derive an alternative distribution for the 2BL-test statistics is the same as that of Mebane (2007c). Based on the empirical vote count distribution we estimate a multivariate distribution for each constituency. From this estimated multivariate distribution, we randomly draw electoral results and calculate the 2BL-statistic for each of them. Based on this 2BL-statistic, we construct the alternative distribution. Differently from Mebane's model which contains only two parties and abstention, our model can treat the results of multi-party elections. In particular, we follow Katz and King (1999) who model multi-party elections using the multivariate t-distribution. They denote the proportion of vote for party j in precinct k by V_{jk}:

$$V_{jk} \in [0,1] \text{ for all } j \text{ in } k, \tag{2}$$

The vote proportions for all parties in a precinct add up to the sum of one.[7] Votes for the last party J can be calculated depending on the votes of other parties:

$$V_{Jk} = 1 - \sum_{j=1}^{J-1} V_{jk}. \tag{3}$$

Let Y_k be the vector of $J-1$ log-ratios $Y_{jk} = \ln(\frac{V_{jk}}{V_{Jk}})$, for party j ($j = 1, ..., J-1$) relative to party J. For the German Federal Parliamentary Election, the vector Y_k contains the log-ratio of CDU, SPD, FDP, Linke, Grüne and not-voters relative to "other parties"

[7] Our notation deviates from that of Katz and King (1999) without losing equivalence. This is because not-voters do not exist in the absentee vote precincts.

(sonstige Parteien), which means that Y_{Jk} is fixed to 0. Assume now that Y_j is multivariate Student t distributed.

$$Y_{jk} \sim f_{mt}(\mu, \Sigma, \upsilon), \tag{4}$$

the expected value of Y_{jk} are μ_{jk} and its variance is $\Sigma\upsilon/(\upsilon-2)$, and $\upsilon(\upsilon > 0)$ is the degrees of freedom parameter. To estimate these parameters for each constituency, we only need the vote proportion of party j in precinct k in the corresponding constituency (V_{jk}). For estimation, we used the Bayesian method without informative priors.[8]

Having estimated the parameters of multivariate Student t distribution for each constituency, we draw 1000 Y_{jk} in each of 299 constituencies of the 2009 election. With the following formula, we calculate simulated vote percentage based on the draws of Y_{jk} in a constituency:

$$V_{jk} = \frac{exp(Y_{jk})}{1 + \sum_{j=1}^{J-1} exp(Y_{jk})}. \tag{5}$$

By multiplying V_{jk} by the number of precincts, we can obtain the simulated vote counts for which we can calculate the 2BL statistic. Thus, 1000 draws give us 1000 2BL statistics which constitute the alternative distribution to the χ^2 distribution. Its use makes our test robust against the influence of specific characteristics of the vote count distribution. We conduct the 2BL-test for the simulated vote counts and use its 95 % quantile as the new critical value which we denote as χ^2_{sim95}.

4.2 Results

To make the results more comparable with those of the former naive 2BL-test, we transformed the empirical results of Figure 2 by using the new scale based on the simulated 2BL statistic.[9] The transformation is done as following:

$$\chi^2_{trans} = \frac{\chi^2_{2BL} * 16.9}{\chi^2_{sim95}}, \tag{6}$$

whereas χ^2_{2BL} is the empirical 2BL statistic, 16.9 the critical value according to our formal χ^2 distribution and χ^2_{sim95} gives the 95 % quantile of our simulated 2BL statistics. The results are displayed in Figure 3. In contrast to Figure 2 the proportion of the 2BL statistic which exceeds the critical value is below 5 % for all parties. Hence there is no indication of election fraud in the German Federal Election 2009.

5 Why do we get inflated 2BL statistics?

The results above highlight the importance of adjusting the critical value based on the simulated 2BL statistic, but it doesn't solve the question why we get misleading test results based on the inflated 2BL statistic of the naive application. Hence, this section focuses on this aspect.

---

8 The parameter estimation as well as the simulation of vote counts is separately conducted for absentee votes and election day votes.

9 The deviation between simulated and empirical vote counts does not exceed 10 % of the voters. In most constituencies this ratio is even below 5 %, which is a satisfying result for our simulated data.

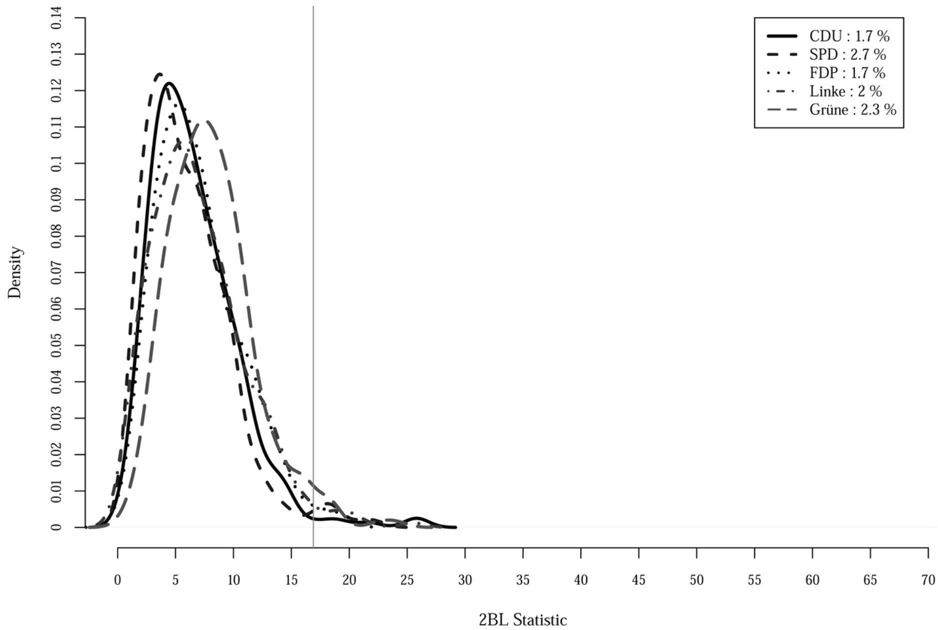


Figure 3 Transformed density distribution of 2BL statistic

Our starting point of the investigation is Fewster (2009: 28). Accordingly, if Benford's Law is applied to data, the distribution of the data needs to span several orders of magnitude on the original scale and be reasonably smooth. That is, the underlying distribution should be heterogeneous to a certain degree and cover a certain range. We approximate both by maximum and mode of the vote count distribution. The maximum value of the density distribution of vote counts gives information about the homogeneity of the distribution (Figure 4). With an increasing value for the maximum, the shape of the distribution gets thinner, which means its homogeneity increases. The 2BL statistic can also

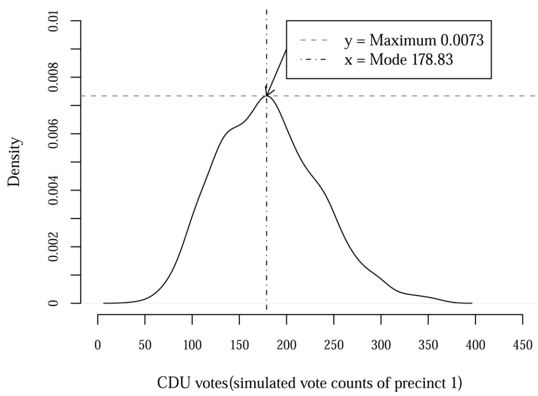


Figure 4 Maximum and modus of the density distribution

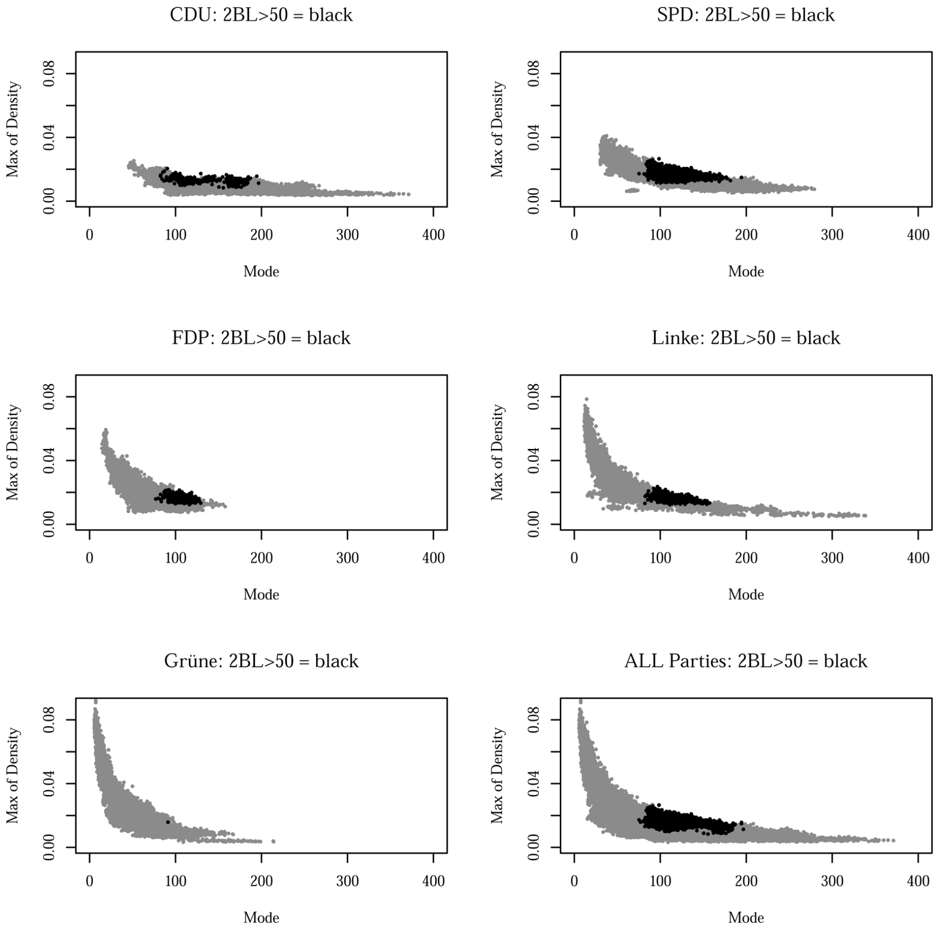


Figure 5 Distribution of the maximum and mode of the simulated vote counts

be influenced by the mode (Figure 4). Since we are interested in the second digits in vote counts, we have to keep in mind that an additional vote between 9 and 99 changes the second digit while between 100 and 999 an additional vote would have a one in ten chance to change the second digit. Hence we expect more deviation from Benford's Law when having a homogeneous distribution and a lot of observation above 100. One might ask why we analyzed simulated data and not our empirical data in this section. However, empirical data often do not give enough insight in the mechanism behind the interesting phenomena due to limited number of cases. To this point, we will return below.

Figure 5 displays the distribution of the maximum value and the mode of the simulated vote counts.[10] The black dots mark all points whose 2BL statistic is higher than 50. Fig-

[10] The maximums, mode and the 2BL statistics were calculated separately for each of the 1000 simulations over all precincts in a constituency. This graphic displays only every 25^{th} point out of efficiency reasons.

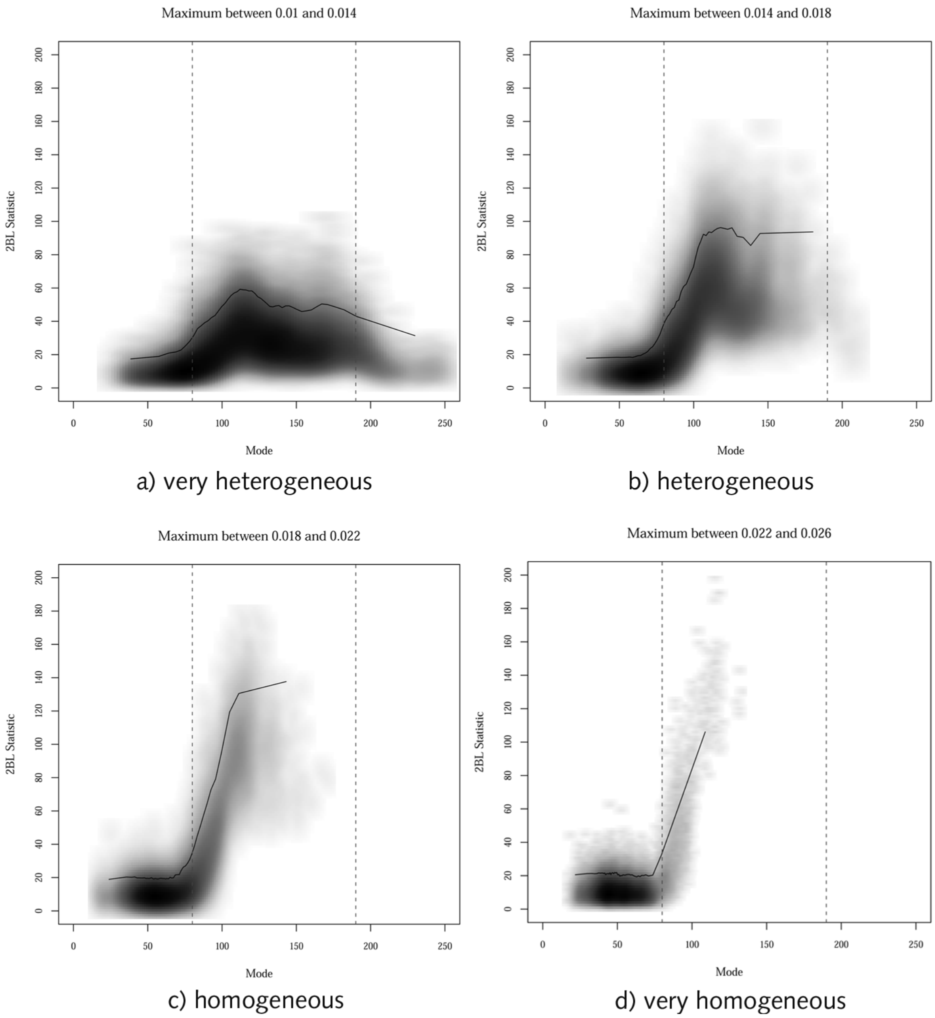


Figure 6 Impact of the central tendencies (mode) of vote distributions on the 2BL Statistics for different magnitudes of homogeneity

ure 5 shows that all parties except the Grüne have a cluster of black dots by a maximum of about 0.02 (plus and minus 0.01) AND a mode between about 90 and 200. Having a homogeneous density distribution is highly correlated with a small mode and a high mode is correlated with a very heterogeneous density distribution. This also gives the answer to the different performance of big and small parties. While small parties more often have a homogeneous density distribution, at the same time they have a small mode. Figure 5 stresses the relationship of both conditions. We can only expect an extremely high 2BL statistic if our maximum value of the density distribution is at least above 0.01 and the mode of this distribution is higher than 90.

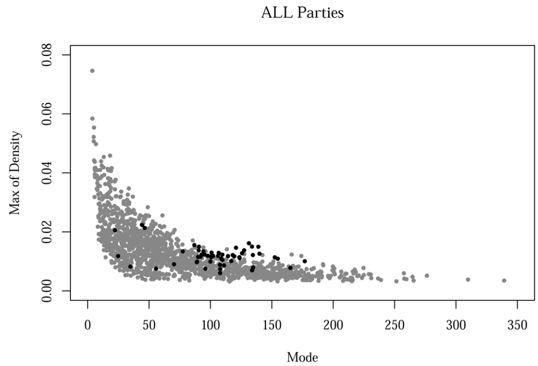


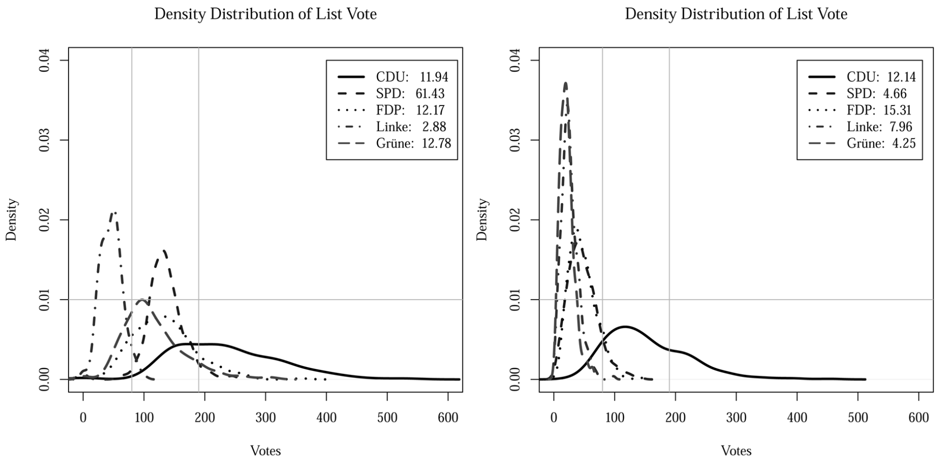
Figure 7 Empirical data: Distribution of the maximum and mode. We marked dots with a 2BL statistic $>=$ 25 black.

To investigate more closely the conditions for a high 2BL statistic, we display in Figure 6 the relationship between the mode and the 2BL statistic for different ranges of the max density. The changing intensity of black/grey in Figure 6 indicates the amount of data points, where black indicates the highest and light grey the lowest density of data points. The black line marks the 95 %-quantile of the 2BL statistic for every 2 %-quantile of the mode. Lower values of the maximum goes along with a smaller and more dispersed 2BL statistic as can be seen in Figure 6(a). Given a maximum density above 0.01, the grey dashed lines mark the range of modes, for which the 2BL statistic can be high. In the next higher range for the maximum of the density distribution (Figure 6b), we can observe a higher 2BL statistic if the mode exceeds 80. This pattern is clearer if the distribution is more homogeneous (larger maximum). We find the peak of the 2BL statistic while having a mode of about 120.

So far we focused on explaining the boosted 2BL statistics under the condition of a homogeneous distribution (a high value for the maximum of the density distribution) and the position of the mode, which had to be between about 80 and 190. The maximum of the 2BL statistic is 50, if the distribution has a mode between 0 and 80. This might still seem too high in comparison with the critical value of 16.9. We, however, have to remind ourselves that 16.9 is a 95 %-quantile. Observing the 95 % quantile (black line) in Figure 6 we find that no more than 5 % of the simulated 2BL statistics exceed about 20. This is close to what was theoretically expected.

We argued above that simulated data are superior to empirical data since the latter have a limited number of cases. This makes it more difficult to find a specific pattern in such data. To show this we conduct the same analysis as in Figure 5 using the empirical data. The maximum and mode of the density distribution for each party are calculated using the real vote counts in all precincts of each constituency. Figure 7 displays the distributions' maximum and mode of each party in each constituency. If a party has a 2BL statistic greater than 25 in a constituency, the corresponding dot is marked black. Most of such dots can be found when the mode is above 80 and the maximum of the density distribution is higher than 0.01.[11] The unclear pattern in our empirical data is partly

[11] Separate observation of this distribution for each party reveals that these dots only appear for the SPD.



(a) Constituency 219 (Munich East) (b) Constituency 227 (Deggendorf)

Figure 8 Empirical example: How we can explain the high 2BL statistic of the SPD in constituency 219

the result of the chosen critical value of 25 for black dots. In Figure 5 we marked such dots grey while the black dots in Figure 5 have a 2BL statistic equal or greater than 50. Such high values barely exist in the empirical data. Hence, it is more difficult to find the combined impact of homogeneity and range (mode) of vote distribution on the 2BL statistic. When analyzing empirical data, we are often faced with a trade-off between an accurate operationalization and the number of observations.

To obtain a clear pattern, we need a more accurate operationalization which, however, can lead to loss of observations. To solve this problem, we generated data points based on a simulation model which is calibrated by empirical data.[12]

One might still be skeptical about our simulation-based results that a homogeneous vote distribution (with a maximum of the density distribution above 0.01) and with a mode between 80 and 190 could cause a high 2BL statistic without electoral fraud (Figure 5 and Figure 6). For validation, we apply those borders to the constituency 219 (Munich East) with a high 2BL statistic for the SPD (Figure 8a). The only two parties which satisfy the first condition (a higher maximum of the density distribution than 0.01) are SPD and Linke. If we now check the second condition whether the mode is between 80 and 190, only the SPD satisfies this condition. This finding also matches the empirical 2BL statistic. While the Linke has an extremely low 2BL statistic (2.88), the one of the SPD is extremely high (61.43). Figure 8b, which is the same figure for constituency 227 (Deggendorf), displays a similar distribution form for SPD as Figure 8a. However, the mode of the distribution is around 50 and the value of the 2BL statistic is much lower (4.66). Furthermore none of the parties in Figure 8(b) satisfy both condition and therefore none of the 2BL statistics exceed the critical value.

[12] We emphasize here that our analysis is an explorative one. If we regress the empirical 2BL statistic of individual constituencies on the mode/maximum of vote distribution and their interaction, we also find effects consistently with our results (available in online-appendix). Our point, however, is that we could hardly reach to this regression model without simulated data.

6 Conclusion

In this article we applied the second digit Benford's Law test to the 2009 German Federal Parliamentary Election, against which there has been no allegation of election fraud so far. We posed two questions which arose from this unique application of the 2BL-test. First, the test statistics were higher in more districts than random chance would predict. The answer to this first question is that χ^2 distributions are inappropriate for the 2BL-test statistics. Instead of χ^2 distributions, we derive distributions based on simulated electoral results in individual districts. Applying the new distributions reveals that the number of districts deviating from Benford's Law is less than random chance. This result, however, does not answer the question, in which circumstances the 2BL-test statistics are boosted. Our findings indicate that the second digit Benford's Law is for some vote count distributions inappropriate. An investigation of the 2BL-test statistics based on our simulated data shows that the test statistics can exceed 50 when the following both condition are complied: the mode of vote counts lies between 80 and 190 AND the distribution of vote counts has a certain degree of homogeneity. Hence, specific characteristics of election unites can cause significant deviations from the 2BL-test at the absence of election fraud. Based on these results the second digit Benford's Law can't be applied to all election results.

One important assumption of our analysis above is that electoral frauds do not affect the simulation model, in particular its parameters. We estimated three parameters from empirical results: expected value, variance-covariance, and the degree of freedom for the propensity underlying the vote count. If these parameters are influenced by electoral frauds, our results in Figure 3 should underestimate the true electoral results. Testing this assumption is, however, not possible for our data since we have no serious doubt for the 2009 German Federal Election. Therefore, the next step should be to replicate our analysis for election results whose irregularity has already been known due to other sources. This would enable us to observe whether any systematic differences can be found in simulation models for districts with and without electoral frauds. Disentangling the mechanisms of triggering (simulation based) 2BL statistics will reveal more detailed assumptions which election data have to fulfill. This in turn will hopefully enable scholars in future to develop a more accurate and simply applicable Benford's Law based test.

References

Beber, B., A. Scacco (2008), What the Numbers Say: A Digit-Based Test for Election Fraud Using New Data from Nigeria. Annual Meeting of the American Political Science Association, Boston, MA.

Benford, F. (1938), The Law of Anomalous Numbers. Proceedings of the American Philosophical Society 78(4): 551-572.

Berger, A. (2005), Multi-Dimensional Dynamical Systems and Benford's Law. Discrete and Continuous Dynamical Systems (DCDS-A) 13(1): 219-237.

Brady, H. E. (2005), Comments on Benford's Law and the Venezuelan Election. Unpublished manuscript, Stanford University, January 19, 2005.

Breunig, C., A. Goerres (2011), Searching for electoral irregularities in an established democracy: Applying Benford's Law tests to Bundestag elections in Unified Germany. Electoral Studies, forthcoming.

Deckert, J., M. Myagkov, P. C. Ordeshook (2010), The Irrelevance of Benford's Law for Detecting Fraud in Elections. University of Oregon and California Institute of Technology. Working Paper.

Diekmann, A. (2007), Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data. Journal of Applied Statistics 34(3): 321-329.
Diekmann, A., B. Jann (2010) Benford's Law and Fraud Detection: Facts and Legends. German Economic Review 11(3): 397-401.
Fewster, R. M. (2009), A Simple Explanation of Benford's Law. The American Statistician 63(1): 26-32.
Goudsmit, S. A., W. H. Furry (1944), Significant Figures of Numbers in Statistical Tables. Nature 154(3921): 800-801.
Hill, T. P. (1995), A Statistical Derivation of the Significant-Digit Law. Statistical Science 10(4): 354-363.
Janvresse, É., T. De la Rue (2004), From Uniform Distributions to Benford's Law. Journal of Applied Probability 41(4): 1203-1210.
Katz, J. N., G. King (1999), A Statistical Model for Multiparty Electoral Data. American Political Science Review 93(1): 15-32.
Mebane, W. R. (2006a), Election Forensics: The Second-Digit Benford's Law Test and Recent American Presidential Elections. Election Fraud Conference, Salt Lake City, Utah.
Mebane, W. R. (2006b), Election Forensics: Vote Counts and Benford's Law. Pp. 20-22 in: Summer Meeting of the Political Methodology Society. UCDavis, July.
Mebane, W. R. (2007a), Election Forensics: Statistical Interventions in Election Controversies. Prepared for presentation at the Annual Meeting of the American Political Science Association.
Mebane, W. R. (2007b), Election Forensics: Statistics, Recounts and Fraud. Pp. 12-16 in: Annual Meeting of the Midwest Political Science Association. Chicago, IL, April.
Mebane, W. R. (2007c), Statistics for Digits. Pp. 18-21 in: Summer Meeting of the Political Methodology Society. Pennsylvania State University, July.
Mebane, W. R. (2008a), Election Forensics: Outlier and Digit Tests in America and Russia. American Electoral Process Conference, Center for the Study of Democratic Politics, Princeton University.
Mebane, W. R. (2008b), Election Forensics: The Second-Digit Benford's Law Test and Recent American Presidential Elections. Pp. 162-181 in: R. M. Alvarez, T. E. Hall, S. D. Hyde (eds.), Election Fraud. Detecting and Deterring Electoral Manipulation. Brookings Institution Press.
Mebane, W. R., K. Kalinin (2009), Comparative Election Fraud Detection. Prepared for presentation at the 2009 Annual Meeting of the American Political Science Association, Toronto, Canada, Sept 3-6.
Mebane, W. R. (2010), Election Fraud or Strategic Voting? Can Seconddigit Tests Tell the Difference? Summer Meeting of the Political Methodology Society, University of Iowa.
Miller, S. J., A. Kontorovich (2005), Benford's Law, Values of LFunctions and the 3x+ 1 Problem. Acta Arithmetica 120(3):269-297.
Newcomb, S. (1881), Note on the Frequency of Use of the Different Digits in Natural Numbers. American Journal of Mathematics 39-40.
OSCE/ODIHR (2009), Federal Republic of Germany. Election to the Federal Parliament (Bundestag). Election assessment mission report OSCE/ODIHR.
Pericchi, L. R., D. Torres (2004), La Ley de Newcomb-Benford y sus aplicaciones al Referendum Revocatorio en Venezuela. Reporte Técnico no-definitivo 2a. versión: Octubre 01, 2004.
Raimi, R. A. (1976), The First Digit Problem. American Mathematical Monthly: 521-538.
Roukema, B. F. (2009), Benford's Law Anomalies in the 2009 Iranian Presidential Election. http://en.scientificcommons.org/46955838.
Weaver, W. (1963), Lady Luck: The Theory of Probability. Doubleday.

Prof. Dr. Susumu Shikano, Chair of Political Methodology, Verena Mack, Chair of Political Methodology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany.
susumu.shikano@uni-konstanz.de; verena.mack@uni-konstanz.de