
Detecting Fraud in Russian Federal Duma Elections

Anastasiia Alekseeva
Matrikelnummer 5994775
anastasiia.alekseeva@student.uni-tuebingen.de

Abstract

The main hypothesis of the article is that the results of the latest Federal Duma election held in 2021 in Russia were partially falsified by ballot box stuffing and "drawing" the final results in the official reports at polling stations. To detect the aforementioned two kinds of fraud, I have examined the distribution of vote-shares, their relationship with turnout, and compared the results with the fair 2021 parliamentary election in Germany. Moreover, I have estimated the number of anomalous integer numbers in each region. As a result of this study, the main hypothesis was corroborated. Github page of the project.

1 Introduction

Elections in the electoral authoritarian regimes are used for many purposes[1]: two of them are to enforce legitimacy of an autocrat as well as by doing so the regional elites show loyalty to the federal one. Usually, the election results in such regimes look very impressive. One of the methods to achieve them is by mass falsifications in the polling stations. There are two major kinds of fraud widely used: 1) ballot box stuffing and 2) adjusting election results in the official reports.

In this regard, Russian parliamentary elections are of great interest. Russia is considered an electoral authoritarian regime where from 2000 to nowadays many illegitimate manipulations of the election results were applied and identified not only thanks to decent observers at some polling stations but via statistical methods as well[2][3]. The latest Russian parliamentary elections, held in September 2021, was not an exception as well: 5789 messages about possible violations of election procedure¹ were received by Golosinfo organization. Apart from that, the elections were organized during 3 days, many opposition candidates were not allowed to participate, many independent observers were forced to leave the polling stations when election Committees were counting the votes.

Taking all that into account, I hypothesise that the 2021 State Duma election was partially falsified in favor of the ruling party United Russia. Moreover, the level of falsifications depends on the region.

2 Data

I use two datasets in the analysis:

1. Polling station data for Russian Federal elections: Dmitry Kobak's repository. 96,325 polling stations, 109,204,662 registered voters. Only traditional voting was included in the analysis. The observations (polling stations) with zero registered voters and zero given ballots were removed.
2. Polling station data for German Federal elections: the Federal Returning Officer's website. 94668 polling stations, including voting by post, 61,181,072 eligible voters. The

¹Map of violations: <https://www.kartanarusheniy.org/2021-09-19>

observations (polling stations) with zero registered voters and zero given ballots were filtered out. The results from postal voting were excluded from the analysis as the dataset did not allow matching the polling station and the voters used post, and consequently, calculate turnout for them correctly.

These datasets include only the results of voting for political parties, rather than for candidates as they provide the samples large enough for the analysis. For Russia, only the first six parties received more than 3% in total were included in the analysis: United Russia, CPRF, LDPR, Just Russia and New People. As for Germany, I included only the six largest parties such as the union of CDU and CSU, SPD, AfD, Die Linke, Grüne, FDP.

3 Analysis

3.1 Ballot Box Stuffing

The fingerprints of ballot box stuffing can usually be traced in the positive and significant correlation between turnout and vote-share in favor of the pro-government party. As can be seen in Figure 1, the vote-shares received by the United Russia at each polling station is positively correlated with turnout (74.15% Pearson correlation). While the vote-shares of the opposition party CPRF, which has taken a second place in the election, are negatively correlated with turnout (-61.18% Pearson correlation). As a counter-example, the vote-shares for German political party SPD do not show any linear relationship with turnout.

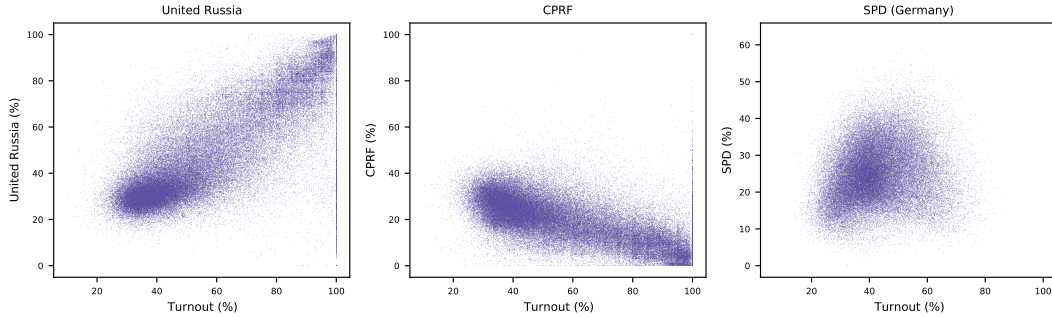


Figure 1: Scatter plots showing the relationship between turnout and the result of a party at each polling station. Parties from the left to the right of the figure: United Russia (Russia), CPRF (Russia) and SPD (Germany).

Figure 2 presents the number of ballots cast for each party depending on the turnout (in 1% bins) at the polling stations. The curve for the United Russia deviates strongly from other parties. Furthermore, anomalously many votes were received by the United Russia at the polling stations with the turnout rate larger than 60%, which might be evidence of the falsifications[3]. Besides, as can be seen in Table 1, the United Russia received 20.55% of its votes at the polling stations, with turnout greater than 90%. In contrast, in Germany, the proportion of votes for each party does not change that drastically, with the change in turnout.

	(0,10]	(10,20]	(20,30]	(30,40]	(40,50]	(50,60]	(60,70]	(70,80]	(80,90]	(90,100]
UR (%)	0.00	0.04	2.11	13.40	14.51	10.55	10.10	11.79	17.01	20.55
Ballots (%)	0.00	0.07	3.51	20.99	19.87	11.60	8.71	9.05	11.53	14.67

Table 1: Vote-shares of the United Russia and the percentage of ballots cast in 10% bins of turnout.

It is worth noting that the positive correlation between turnout and vote-shares for the United Russia may be caused not only by ballot box stuffing but by mobilization of employees at workplace to vote for the regime[4] and other technologies like carousel voting could contribute as well.

Some considerable deviations of the distribution of vote-shares from hypothetical distribution might be a sign of ballot box stuffing as well. In the homogeneous population, under democratic regime

and without pronounced strategic voting, the distribution of vote-shares should be unimodal and look similar to the normal distribution. The distribution of votes in the latest election in Germany may serve a good example of how the curve should look, especially for SPD and the union of CDU and CSU, which is provided in Figure 3. In a sharp contrast, the kernel density for the United Russia is strongly skewed to the right.

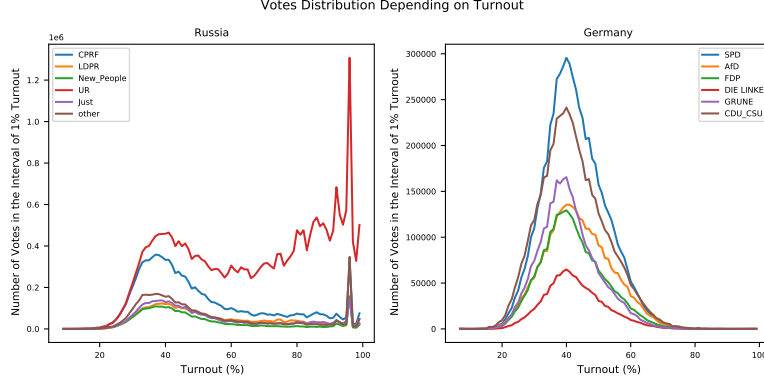


Figure 2: The number of ballots cast for each party depending on the turnout (in 1% bins) in Russia and Germany at the federal level.

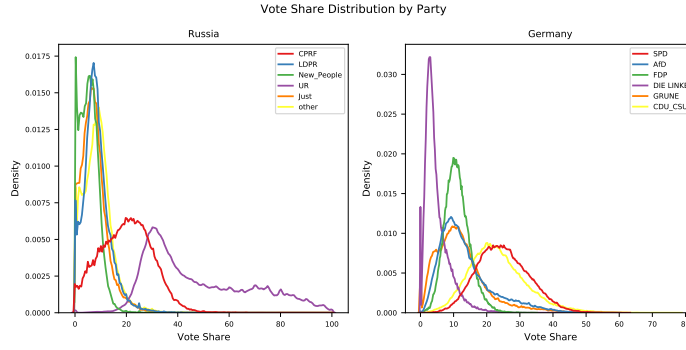


Figure 3: Kernel densities of vote-share distribution in Russia and Germany at the federal level.

3.2 Anomalous Integers

Another commonly used fraud is adjusting the results in the official reports at the polling stations, so that they comply with the instruction from the upper-level authorities. This fraud might be recognized by the high frequency of integers in the vote-share distribution. I apply the method devised by D. Kobak et al.[2] at a regional level to find the regions that clearly "draw" the results of election. The integer in this analysis is considered the percentage that is $\pm 0.05\%$ percentage points from an integer number. I_r will denote the number of integer-result polling stations in a region r .

Note: all polling stations with less than 100 registered voters were disregarded from the analysis.

In order to model the voting behavior at each polling station, I assume that each citizen votes independently and the votes for a party follows binomial distribution $Binom(N_i, \pi_i = n_i/N_i)$ where N_i is the number of cast ballots at a polling station i , n_i is the number of votes given for the party of interest at i , π_i is n_i/N_i or the probability of each voter to vote for the party of interest at a polling station.

To test whether a region has produced anomalously large amount of integers at polling stations I use parametric bootstrap with 1000 samples. The null hypothesis is $I_r(observed) = \frac{1}{1000} \sum_{s=1}^{1000} I_{rs}(simulated)$, the alternative hypothesis is $I_r(observed) > \frac{1}{1000} \sum_{i=1}^{1000} I_{rs}(simulated)$. The test is one-tailed as we are not interested in relatively small I_r . To avoid high Type I error, I use the confidence level $\alpha = 0.1\%$.

After having applied the parametric bootstrap procedure to the results of United Russia and its main competitor CPRF in all Russian regions, I have identified the regions where the null hypothesis has been rejected, or $I_r(\text{observed})$ is larger than 99.9% of obtained via simulation $I_r(\text{simulated})$. These are Volgogradskaya, Voronezhskaya oblast', Krasnodarsky kray, Bashkortostan, Tatarstan and Stavropol'sky kray. However, in the previous elections there were a lot more integer-result regions than in 2021 election. This might have been caused by the 3-day duration of the 2021 parliamentary election: the results could be "drawn" in the first two days, and in the third day they were blurred with real votes.

To examine this phenomenon a little bit closer, I have plotted vote-share distribution of United Russia and CPRF in two anomalous regions Krasnodarsky kray and Bashkortostan, and in Irkutskaya oblast' where the communist party (CPRF) has very strong support (Figure 4). In the first two plots, the kernel density of the percentage of votes for the United Russia has several pronounced integer peaks, while in the plot for Irkutskaya oblast', no significant peak is observed.

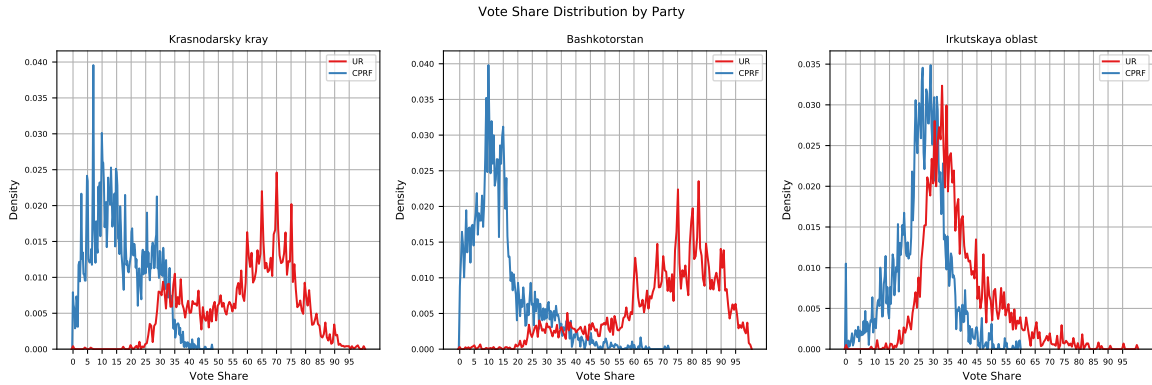


Figure 4: The result at the polling station (in 0.05% bins) in Russian national republic Bashkortostan and Krasnodarsky kray.

4 Conclusion

The performed analysis has provided some evidence confirming the main hypothesis of the study that the Russian Federal Duma election was not entirely fair. Firstly, the fingerprints of ballot box stuffing and adjusting the final results in favor of the pro-government party were identified via visualization and parametric bootstrap. Secondly, Russian regions are not identical in terms of the level of falsifications. Moreover, they may adapt to the changing reality and switch from old techniques of manipulation of election results to the new ones.

References

- [1] J. Gandhi and E. Lust-Okar. "Elections Under Authoritarianism". In: Annual Review Political Science. 2009. Vol.12, pp.403-422.
- [2] D. Kobak, S. Shpilkin and M. S. Pshenichnikov. "Integer Percentages as Electoral Falsification Fingerprints". In: The Annals of Applied Statistics. 2016, Vol.10, No. 1, 54-73.
- [3] D. Kobak, S. Shpilkin and M. S. Pshenichnikov. "Statistical anomalies in 2011–2012 Russian elections revealed by 2D correlation analysis". 2012. Preprint. Available at arXiv:1205.0741.
- [4] T. Frye, O.J. Reuter, and D. Szakonyi. "Political Machines at Work: Voter Mobilization and Electoral Subversion in the Work- place". In: World Politics. 2014, Vol 66, No 2, pp. 195–228. doi: <https://doi.org/10.1017/S004388711400001X>.