Indications of vote manipulation in the 2011 Russian elections

Alex Melnik,† John Melnik† and Nikolay Zhelev
†Samarcan Analytics, SamarcandAnalytics.com

Introduction
Several Russian commentators\(^1\) have noted unusual voting patterns in the publicly-available preliminary election results that are posted to the Russian election commission web site.\(^2\) Here, we look at this data and discuss a number of unusual features such as a correlation between voter turnout and percent voting for the United Russia party and a very large voter turnout with almost total support for United Russia in some unexpected regions.

If you have any questions about this work or notice any errors, please email us at at web@samarcandanalytics.com.

Voter turnout and support for United Russia
We define voter turnout as the ratio of all ballots cast (both valid and invalid)\(^3\) to eligible voters.\(^4\)

This plot shows a large number of polling stations which have near perfect turnout. Of ninety-five thousand polling stations, almost seven thousand reported a turnout above 97%. The support for United Russia has a very strong dependence on turnout, while a simple model would predict that these two should be largely independent. Similarly, the support for the Communist Party declines proportionally with increasing turnout.

---

1. [http://eugenyboger.livejournal.com/4514.html](http://eugenyboger.livejournal.com/4514.html) originally looked at and posted this data
2. [http://cikrf.ru](http://cikrf.ru), we have also posted a compiled collection of the data to [http://SamarcandAnalytics.com/election_data](http://SamarcandAnalytics.com/election_data)
4. Число избирателей, внесенных в список избирателей - Number of voters included in voters list
Many of the polling stations with nearly full turnout and essentially all votes being cast for United Russia are located in Russia's ethnic republics. Five republics with unusual voting patterns are shown below, with Moscow City added for comparison.

The results for Chechnya, in which United Russia won 99.5% of the vote with 99.5% of eligible voters voting, seem particularly implausible, although it is comparable to previous results published for the region.\(^5\)

**Votes captured with increasing turnout**

Rather than looking at the percentage of the votes cast, it is often instructive to look at each party's share of the electorate; the ratio of votes for each party versus the number of eligible voters at each polling station (V/E).\(^6\) In a homogenous region where party preference is uncorrelated with turnout, each party should capture an increasing portion of the eligible votes as the turnout increases. We note that under no circumstances should we expect any party to capture more than a single additional vote on average when the turnout increases by one person. This can be quantified by fitting a line to V/E with respect to the turnout at each polling station. If this regression coefficient (slope of the fit) is larger than one then it is a strong indication of fraud.

If a region's voting patterns are not homogenous, aggregation errors can lead to false positives or negatives. This could happen in a region where there are urban and rural polling stations with different voting patterns, or in a region with spatially segregated ethnic groups. To address this, we look carefully at each region and the statistical significance of the fit.

We performed fits to V/E verses T for each of the 81 electoral districts with more than 100 polling stations for both United Russia and the Communist Party. Of these, one district had a regression

---


\(^6\) This measure is discussed in detail in Myagkov, Ordenshook & Shakin, *The Forensics of Election Fraud*. Cambridge University Press (2009).
coefficient larger than 1 for the Communist Party and 26 regions had a regression coefficient larger than 1 for United Russia.

We first look at the republic of Kabardino-Balkaria, the only region to show a regression coefficient larger than 1 for the Communist Party.

![Figure 3: Share of electorate verses turnout for Kabardino-Balkaria](image)

The R$^2$ value is very low for both parties and neither appears to follow a linear trend, which may indicate aggregation errors. It is not possible to draw any conclusions for this region based on this data, although the very large turnout is suspicious.

The region with the largest regression coefficient for United Russia, the Republic of Mordovia, displays a strongly linear relationship between share of electorate and turnout.

![Figure 4: Share of electorate verses turnout for Mordovia](image)

On average, for each additional voter who arrived at a polling station, the Communist Party lost 0.38 votes and United Russia gained 1.72 votes. In addition, a large number of polling stations reported almost full turnout with essentially all votes being cast for United Russia, as noted above. These two indicators, combined, strongly indicate widespread fraud in Mordovia.
However, this type of analysis is not always helpful. For example, in Khabarovsk neither of these parties show a regression coefficient outside of the normal range, and while a large number of polling stations show nearly full turnout, these stations don't appear to predominately favor either of these parties.

This does not provide any evidence for vote manipulation in this region.

While a full discussion of each region is beyond the scope of this work, we note that several regions show very clear evidence of vote manipulation by this metric, with many more demonstrating suspicious behavior.

**Distribution of voter turnout**

A simple model for voting trends suggests that voter turnout should have a binomial distribution around a central value. This should be true both in small homogenous areas, and also in large heterogeneous regions unless there are well-defined spatial differences in voting trends. For example, small differences can still be seen in voter turnouts between the former states of East and West Germany.

The voter turnout for all polling stations in Russia shows a distinctly irregular distribution with a second peak near 100% turnout. Much of this is due to the large turnout in several regions mentioned

---

7 [http://www.bundeswahlleiter.de/de/bundestagswahlen/BTW_BUND_09/ergebnisse/landesergebnisse/](http://www.bundeswahlleiter.de/de/bundestagswahlen/BTW_BUND_09/ergebnisse/landesergebnisse/)
previously.

Note that the Moscow City region also shows a second peak near 100% turnout, which could indicate either large scale ballot stuffing at several polling stations, or possibly the presence of a smaller geographic area with very different voting patterns inside of this region.

We can also compare the distribution of total votes cast and valid votes cast.

A large number of polling stations reported receiving almost as many ballots as there were eligible voters, although many of those were later ruled invalid.

**Distribution of votes by party**

As with voter turnout, the percentage voting for a given party is also expected to have a binomial distribution, centered around the average likelihood of voting for that party.
The results for the Communist Party roughly follow the expected distribution, although there is a small peak at zero percent. The results for United Russia, however, have a very long tail on the side of high support, and a peak near 100% that can't be explained by this model.

A careful observer will note that the histogram above demonstrates a number of unusual “spikes” in an otherwise smooth distribution. These spikes become particularly evident at a smaller bin size.

Russian commentators have referred to these spikes as “Churov's Comb” after Vladimir Churov, the head of the Central Election Commission. A closer look at the areas in which they occur reveal that they occur at very specific values.
Many of these peaks have a benign explanation due to the distribution of natural numbers. Consider all the polling stations that had close to 50% support for a given party. Of these, roughly half can possibly have exactly 50% support – namely the ones in which the total number of voters are divisible by two. Fewer polling stations will see exactly 49% support for such a party, since only the polling stations with a total number of voters evenly divisible by 100 can see exactly 49% support. Because of this, we expect to see spikes in the distribution at simple fractional values such as \( \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, \text{etc.} \) We expect that the peak at 50% should be roughly twice as large as the peaks at 25% and 75%, although this is modified by the party's overall distribution of support.\(^8\)

To provide a simple comparison, we have included a simulated party in the figure above, which received a random, and uniformly-distributed portion of the votes at each polling station. All of the peaks in this section of the histogram have a natural explanation.

---

This portion of the histogram has several peaks which are due to natural causes, but also a number of peaks which are clearly not. The peaks which occur at between 90% and 100%, and the peak at 65% are much larger than can be attributed to the effect above, and are likely due to fraud.

**Distribution of votes for polling stations sorted by turnout**

We noted above that polling stations with high turnout often vote differently than those with average turnout. To see the extent of this we can sort all polling stations by their turnout and then plot the distribution of votes for each quintile of turnout. The middle quintile represents a mean voter turnout of 61% while the top quintile has a mean voter turnout of 93%.
All quintiles show a relatively standard distribution except for the highest quintile of turnout which trends very strongly for voting for United Russia. Similarly, the portion voting for United Russia trends upwards for the 3rd and 4th quintiles, although it is almost the same for the first two.

The support for the Communist Party shows a relatively weak dependence on voter turnout, and has the expected distribution except for in the 5th quintile.

Because of the unusual voting seen in this quintile, it is instructive to look at what the voting totals would have been without its inclusion.

<table>
<thead>
<tr>
<th>Party</th>
<th>Percent of the vote</th>
<th>Percent without high-turnout polling stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Russia</td>
<td>49.3</td>
<td>42.7</td>
</tr>
<tr>
<td>Communist Party</td>
<td>19.2</td>
<td>21.2</td>
</tr>
<tr>
<td>A Just Russia</td>
<td>13.2</td>
<td>15.2</td>
</tr>
<tr>
<td>Liberal Democratic Party</td>
<td>11.7</td>
<td>13.3</td>
</tr>
<tr>
<td>Yabloko</td>
<td>3.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Patriots of Russia</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Right Cause</td>
<td>0.6</td>
<td>0.7</td>
</tr>
</tbody>
</table>

In this case United Russia is still the dominant party, but has lost 6.6%, while other parties have gained slightly. (Note that these values sum to less than 100% since this is the percent of total votes cast, which includes invalid ballots).

**Conclusions**

There is strong evidence for widespread vote manipulation in the 2011 Russian elections. While much of the press has focused on reports of voter fraud in large metropolitan areas, this analysis indicates...
that fraud may have occurred on a ever-wider scale in other areas. It is difficult to quantitatively estimate how much fraud occurred, but a simple estimate would be to see how many less ballots would have been cast for United Russia if the 20% of polling stations with the highest turnout voted along the same lines as the others. In this case United Russia would have received 3.54 million fewer votes.