

US elections are non-Normal, just like Russian elections

Mikhail Simkin

United Russia's reported election results are so improbable as to violate Gauss's groundbreaking work on statistics.

The Washington Post

(http://www.washingtonpost.com/world/europe/russians-scoff-at-medvedev-election-inquiry/2011/12/11/gIQAmBR8nO_story.html)

Last week I [discussed](#) the allegation that the results of Russian parliamentary elections violate mathematics. The figure below shows the distribution of the percent of the vote for parties among election precincts. It traveled over hundreds of blogs during past couple of weeks. It was accompanied by comments (The originally linked article at <http://www.themarketfinancial.com/russian-electionsstatistical-bias-between-10000000-and-20000000-votes/128218> went down the memory hole – M.S., 3/26/2018) that the distribution for United Russia party violates a fundamental law of nature, because it is non-Gaussian. In the previous [article](#), I argued that there is no reason for the distribution to be Gaussian. However, a commentator challenged me to show non-Gaussian distributions in USA elections. I took up the challenge.

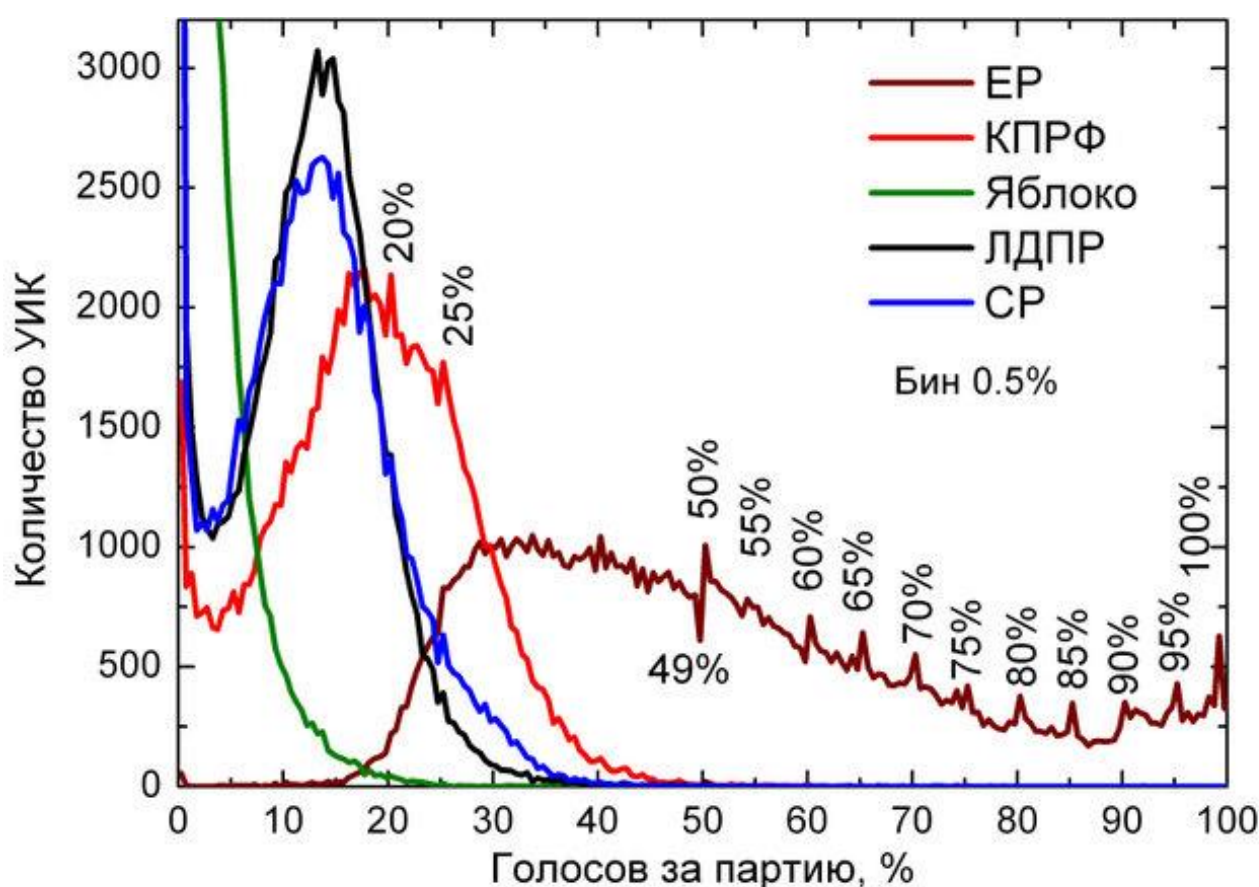


Figure 1. The distribution of the percent of the vote for parties among election precincts. The x-axis shows percentage of votes for the party, y-axis – number of precincts. The bin is 0.5%. Brown line is for United Russia, red – Communist party, green - Russian United Democratic Party "Yabloko," black – Liberal Democratic Party, blue - A Just Russia.

I decided to look at [2008 Republican primaries](http://uselectionatlas.org/). Mostly because this was the last election, I voted in. The primaries differ from national elections, because different states hold vote on different dates. Some candidate drop out during the race and it complicates the analysis. However, 21 states held elections the same day, called Super Tuesday. Since almost half of the nation votes the same day, it looks like a National primary. The most complete elections results database I could find is Dave Leip's Atlas of U.S. Presidential Elections (<http://uselectionatlas.org/>). It does not have precinct level results for the election in question, but has results by county for 19 out of 21 Super Tuesday states (except for Alaska and North Dakota). I computed the distribution of the percent of the vote for four major candidates among 1,162 counties. See Table 1 and Figure 2.

As you can see in Figure 1, the Huckabee distribution has two equal peaks at 15 and 35 percent. The drop between peaks is half peak's height. The McCain distribution has one peak at 35% and another at 80%. Between these peaks, the distribution drops almost to zero. Romney has one peak at 25% and another at 90%. Paul got an exponential distribution, which resembles the one of "Yabloko" in Russian elections. There is little Gaussian about these distributions. Apparently, American elections also "violate Gauss's groundbreaking work on statistics."

Another issue brought by the bloggers is that there are some spurious peaks at 50% and other multiples of 10 (see Figure 1). However, when you go to precinct-level results – you notice that in many precincts very few people voted, as little as 1 in some of them. When 2, 4, 6, 8, or 10 people voted you can get a 50% result, but never a 49% result. Can it explain it all? I plan to address this question in a future article.

Table 1. Results of 2008 Republican presidential primaries in 19 Super Tuesday states. Figure 2 shows the same data graphically.

Percent of vote	Number of counties			
	Huckabee	McCain	Romney	Paul
5%	56	11	47	723
10%	95	26	77	305
15%	138	34	122	51
20%	115	91	181	34
25%	72	99	203	21
30%	96	152	174	9
35%	140	175	142	6
40%	137	173	85	2
45%	113	147	22	4
50%	65	115	31	2
55%	39	60	18	0
60%	29	21	14	0
65%	20	1	7	2
70%	28	1	8	1
75%	14	13	1	0
80%	5	37	4	0
85%	0	6	4	0
90%	0	0	14	2
95%	0	0	8	0
100%	0	0	0	0

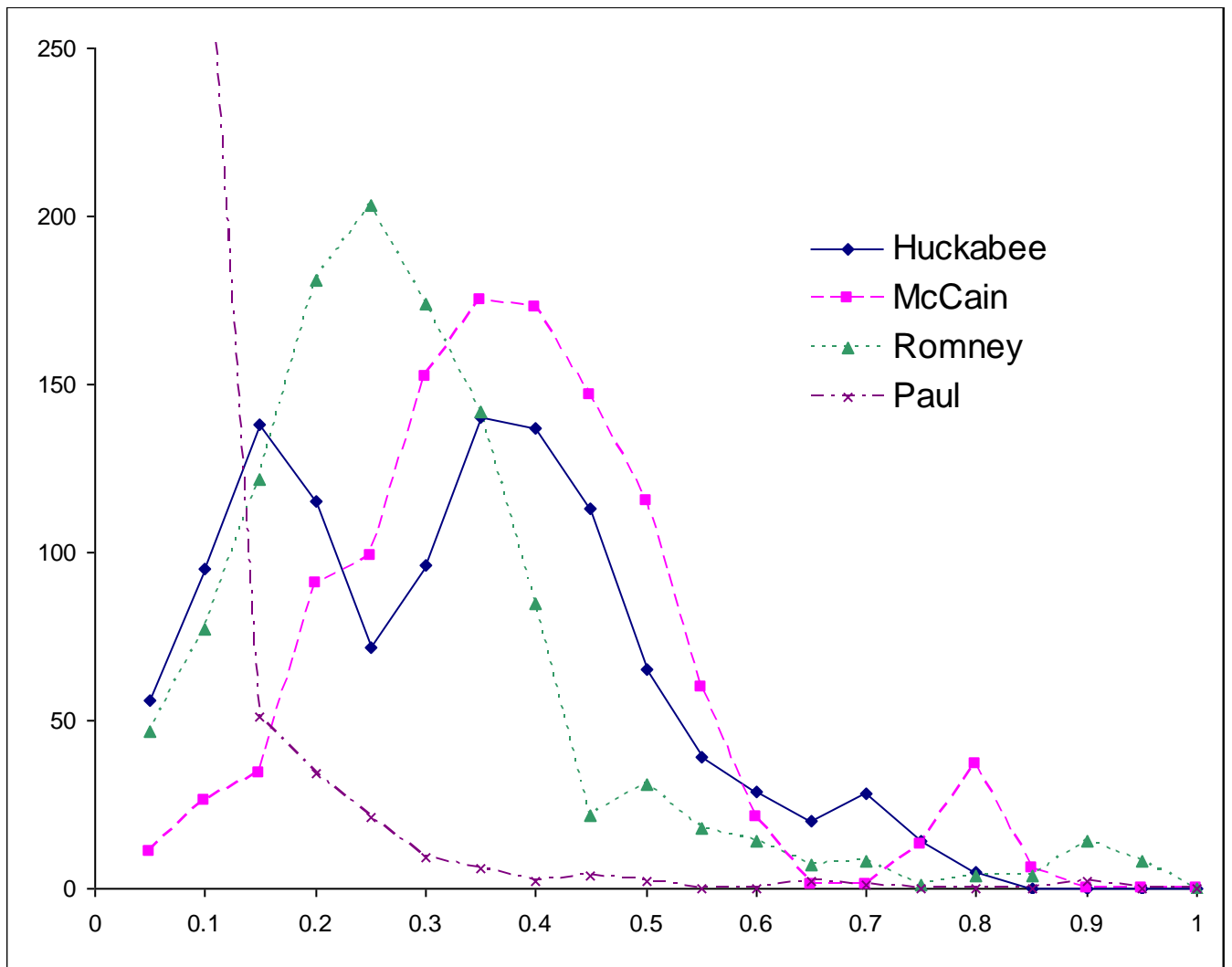


Figure 2. The distribution of the percent of the vote for four major candidates among 1,162 counties.