**Predicting the Outcome of the 2018 Hungarian General Elections**

**Tactical Voting Summary and Election Predictions using UNS and GLM**

**1 Election system and tactical voting overview**

In 2012, a new election system has been introduced in Hungary. In the new system 106 seats are distributed through single seat constituencies (OEVK) and 93 seats are proportionally shared based on party list votes. In addition, a new “winner bonus” system has been added to this system where the winning candidates in OEVKs receive carry some votes to the party list where the seats are distributed between parties. This new bonus means that those parties that win the majority of seats will have a stronger majority than in the former, more proportional election system.

Because of this change in the system, opposition parties had to coordinate their efforts in seats to win as many seats as possible. This coordination was quite limited with only 43 candidates withdrawing from the race in 29 seats[[1]](#footnote-1). Withdrawals mainly happened in key districts in Budapest, but even in the capital, coordination was quite limited between parties. To increase the number of opposition seats, voters had to opt for tactical voting in many districts where at least two (but often more) opposition candidates stood against the governing party, Fidesz.

Since voters had to choose from many potential opposition candidates, taktikaiszavazas.hu (meaning “tactical voting”) was launched with the list of constituencies and candidates in all 106 districts with suggestions on who the strongest candidate is in each district. Similar sites have also been created such as kireszavazzunk.hu (“who should we vote for?”) where small sample (500 respondents) telephone polling was carried out in all districts. This site also provided recommendation on the strongest candidates.

The election predictions on taktikaiszavazas.hu used a mixture of constituency-level polling in 20 OEVKs carried out by Medián and Závecz Research, two polling agencies, and the 2014 election data on which a uniform national swing (UNS) model was used. The polling was ordered by Közös Ország Mozgalom (Common Country Movement or KOM), an NGO group that facilitated political discussion between parties, and they also tried to increase voter turnout among opposition voters. Using the constituency polls and UNS, it was possible to identify 40-45 districts (about 40% of all seats) where opposition parties had some chance to win the seat by either standing head to head (with the strongest candidate) against Fidesz, or by encouraging voters to vote tactically for a certain candidate.

In the following sections, I will explain the following aspects of this election prediction project:

* Different data sources and solutions used to run the website and create the predicitions
* Prediction models for pre-election recommendations
* Election results and the analysis of deviations from predictions
* Conclusions and further research potential

**2 Technical solutions and data sources**

**2.1 Running the taktikaiszavas.hu website**

The main website, taktikaiszavazas.hu, was set up as a static website running as GitHub page. The reason to use this approach was to reduce the impact of DDos (denial of service) attacks. Each visitor was also redirected to Cloudflare to further mitigate the overload attacks potential.

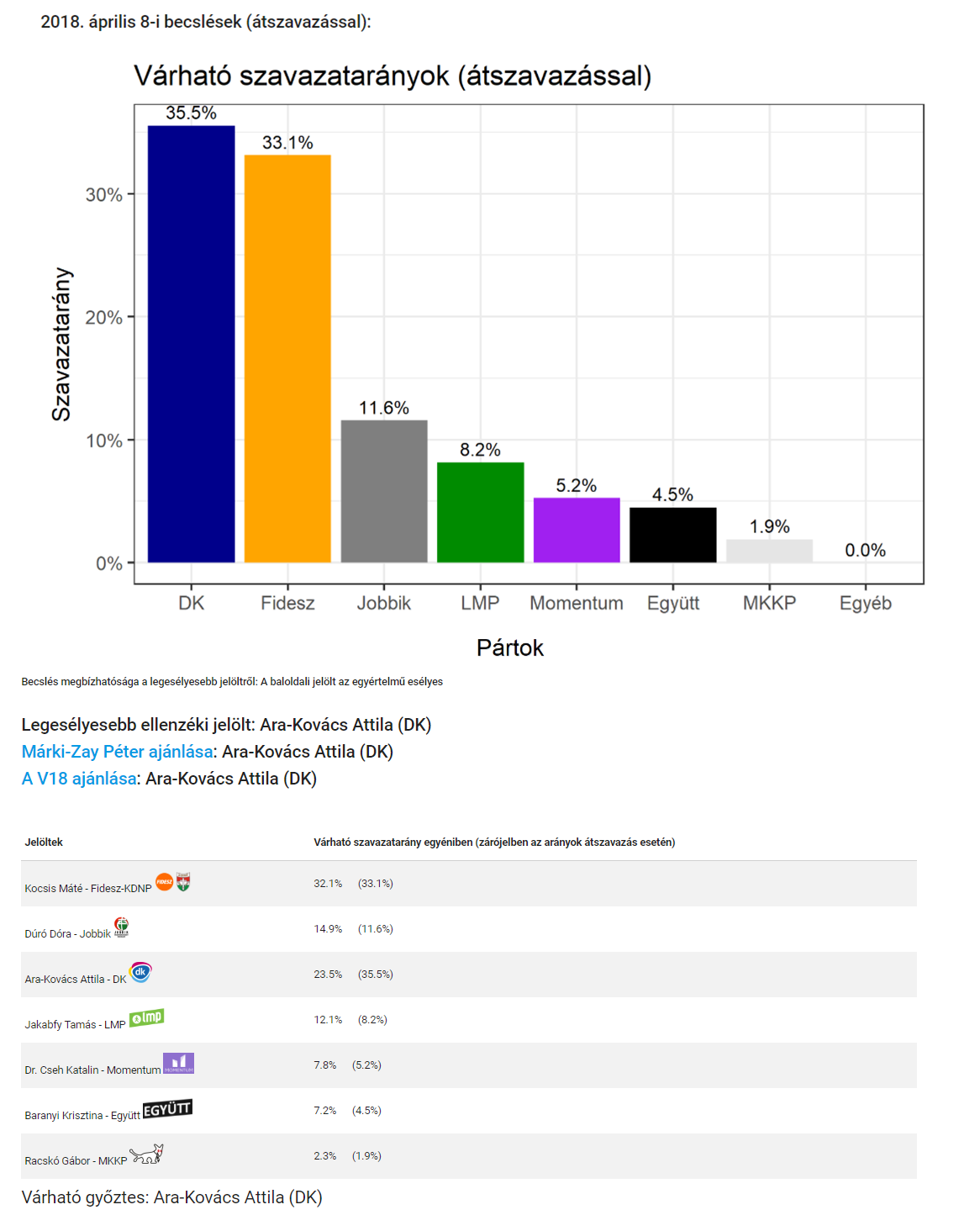
The website is divided into different sections most of them focusing on single seat races.

The different tables and graphics have been predefined in the page template that was fed from two different sources.

The table are pulled from a csv file that with a Perl script. The script automatically amends values and generates new single static pages (one for each single seat race). It was relatively easy to update the website every week when new, adjusted predictions were published.

The graphics showing the graphical representation of the expected vote shares in the table were generated in ggplot. They also corresponded to the auto-generated webpage names so all of them could be inserted to the website without any manual editing.

The other parts of the website were also partially auto-generated through the Perl script such as the list of all expected results by constituency in a summary page. Overall, managing the website required some effort when images such as the constituency maps had to be updated, but the calculated values and related visuals were mostly automated.



**Typical view of a webpage showing expected vote shares in one constituency**

**2.2 Data sources used for predictions**

To create predictions for the 2018 election, data from the previous 2014 election was used. The data was shared by Gábor Tóka, who was an advisor and analyst for KOM during the constituency-level polling. Also, I combined this data with different demographics data acquired from the Hungarian Central Statistical Office (KSH). The 2018 election data was also downloaded and shared online in the coming days after the election. Fortunately, all data sets were in .csv format, so the data cleansing required less effort. The main issue was with the encoding of constituency and town names as some of them were not read correctly in utf-8 format. Both in R and Excel, I created a set of golden data sources, where all the fields are formatted correctly and can be used for analysis. These files are the following:

**Initial sources (csv)**

vote\_counts\_precincts\_2b (2014 data)

2018\_egyeni\_listas\_29May (2018 data)

6 additional files from KSH data (tstar data)

**Generated and cleaned output files (csv)**

2018\_UNS\_baseline (used as the prediction base)

2018\_by\_oevk\_clean (2018 results for each OEVK)

oevk2014\_clean\_new (2014 results for each OEVK)

oevk2014\_telepules (2014 results by town)

2014\_2018\_tstar\_all\_by\_telep (2014-2018 results combined with KSH data by town)

2018\_by\_telep (2018 results by town)

2018\_by\_szavazokor (2018 results by precincts)

2018\_egyeni\_becsles\_elteres (Prediction error between pre-election predictions vs. 2018 results)

Some of the manual changes in these files included adding some additional data about OEVKs such as the typical profile of the district (based on the types and numbers of town making up that voting area). These extra metrics were used during the pre-election prediction (for example, by calculating with higher than expected number of votes cast on smaller parties in major cities).

In addition to the 2014 data and other population data from KSH, the summary of 20 constituency-level polls were also used for predictions together with national polls conducted by different polling agencies between January-March 2018. The constituency level polls included the summary of potential vote shares of each party, the vote shares of each candidate and the expected tactical voting behavior of people. The last of these aspects was very important to create a simple but plausible approach to tactical voting, which was hard to predict with any prior pattern.

**3 Prediction model using UNS with additional factors**

**3.1 Polling data by Common Country Movement**

The pre-election predictions have been updated on taktikaiszavazas.hu on a weekly basis. The constituency-level polls, which were commissioned by KOM, had been released gradually making the available sample bigger on UNS predictions. Because of that, the weekly predictions showed big swings first, but then by week 5 and 6, there was little or no change in the predictions.

The OEVK-based polls focused on key districts in Budapest (8 out of 20 polls) and 12 other battleground seats (8 of which were OEVKs in major cities). The polling showed the expected vote share of party lists as well as candidates, so the results were useful to calculate both tactical voting (taking into account voters’ behavior when choosing between multiple candidates) as well as see the relative strengths of each party in these districts. Extrapolating the results from 20 polls to 106 OEVKs provided the numbers that I could use for the UNS predictions.

**3.2 Three steps of the UNS prediction**

The UNS prediction was constructed the following way:

1. Compute the expected number of votes in 2018 for each party using the 2014 national vote shares multiplied by the ratio given by the UNS vote share. For example, if Fidesz got 44.1% in 2014 and their expected vote share in 2018 is 43%, then their vote share ratio will be 97.5% (44.1 \* (43 / 44.1)), so they are likely to lose 2.5% of their voters from 2014.
2. Increase the expected vote share by candidate strength that is calculated differently for each party candidate. The calculation gave an additional 1 to 3% of all votes for each candidate. Strong candidates were defined as either being the incumbent MP (Fidesz and MSZP/DK candidates), someone holding an office in local government (such as mayors in mid-size towns), a candidate who was featured in national media, or a candidate who was in the top 5-10 places in a given party’s party list.
3. Calculate the vote shares of smaller parties (Momentum, MKKP, and Együtt) that had no records from 2014 as they did not exist or did not contest the election alone. The expected vote shares have been calculated from LMP’s 2014 results as these parties were likely to aim for LMP voters too.

These three steps in each OEVK returned the strongest candidate which was usually a candidate of MSZP/DK or Jobbik. In the final set of predictions, 62 districts had a Jobbik candidate as the strongest candidate while in 44 districts, MSZP/DK, LMP, Együtt, or independent candidates were deemed to stronger.

The expected difference between candidates was used for the last step of the calculation, which was tactical voting done by voters.

**4 Tactical voting predictions**

**4.1 Tactical voting levels by different parties**

The hardest part of the prediction was to estimate the number of voters who would vote tactically to support the strongest candidate in their OEVK. For this estimation, a simple calculation was used based on a national poll conducted by Závecz Research[[2]](#footnote-2) in March 2018:

1. There was a standard highest tactical voting share between different parties if the difference between the strongest and second strongest candidate is more than 7%:
   * 20% of Jobbik voters would support the stronger left-wing candidate
   * 30% of MSZP/DK voters would support Jobbik candidates
   * 30% of LMP and Momentum voters would support either candidate
   * 40% of Együtt voters would support either candidate
2. If the difference was less than 7%, then only 1-7% of Jobbik or MSZP/DK voters would support the slightly stronger candidate while the numbers remained the same for other parties.

The idea behind this approach was that people were more likely to vote tactically if they knew that the race was clearly between Fidesz and a stronger opposition candidate, then they were more likely to vote tactically.

**4.2 Formal withdrawals and the effect on tactical voting numbers**

In 29 of 106 seats, one or more opposition candidates withdrew from the race before election day. In these districts, it was important to recalculate the expected vote shares of candidates since some voters could not choose their own party’s candidate anymore. In these cases, 80% of voters were reallocated to the strongest candidate showing that a clear withdrawal from the race will push most voters to the strongest candidate.

**4.3 The final pre-election prediction model**

Overall, the following prediction model was used to publish the list of strongest candidates and potential target seats for the opposition parties:

* UNS vote shares calculated from the 2014 votes, and slightly adjusted based on the individual strength of each candidate
* Tactical votes reallocated between candidates in all seats and further increased where a candidate had formally withdrawn from the race

The final prediction used the following vote shares for each party:

|  |  |  |
| --- | --- | --- |
| **Party** | **Vote share %[[3]](#footnote-3)** | **Expected number of seats** |
| Fidesz | 42% | 82 |
| Jobbik | 22% | 9 |
| MSZP | 12.5% | 8 |
| LMP | 8.5% | 2 |
| DK | 7.5% | 4 |
| Momentum | 3% | 0 |
| Együtt | 2% | 1 |
| MKKP | 1% | 0 |
| Independents | - | 0 |

Turnout: 65% (5.2 million votes)

This calculation was expecting that roughly 360,000 voters would vote tactically, and this was labelled as the realistic scenario. A more optimistic scenario was expecting 550,000 tactical voters with a 70% turnout. In this case, an additional 18 seats were taken by the opposition parties. However, 24 wins out of 106 was considered a more reliable estimation.

**5 Election results and the effect of tactical voting**

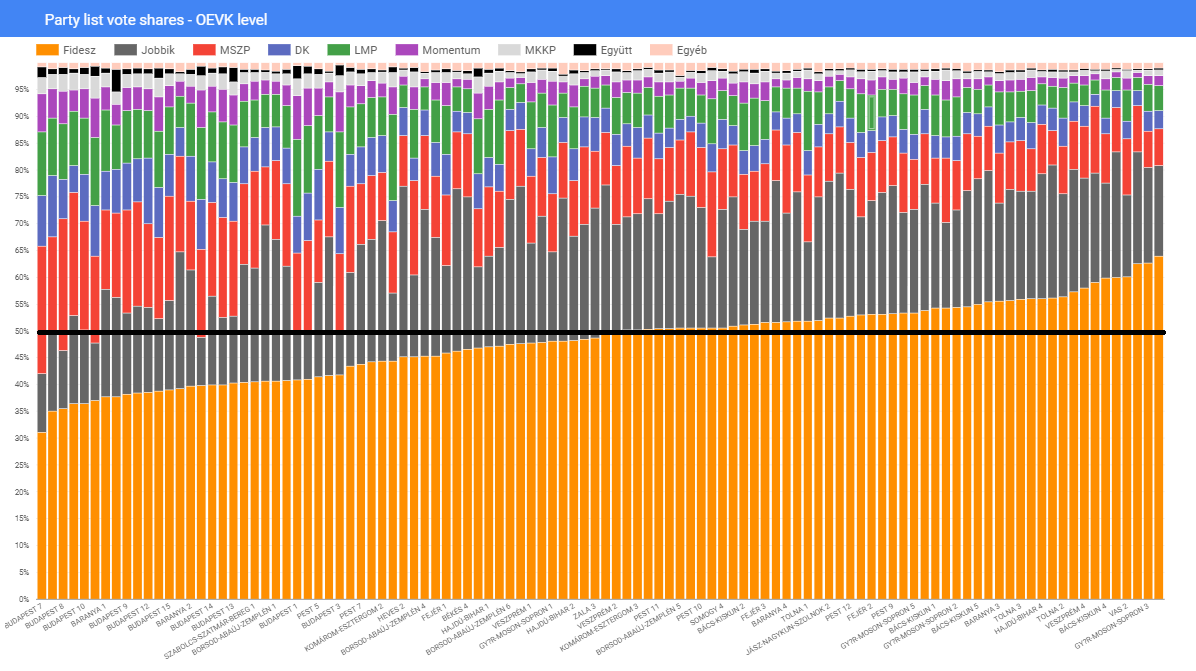
**5.1 Election results**

On 8 April 2018, the election was held. To the surprise of many people, Fidesz had performed much better than the OEVK-level polls had shown, and they received almost 48% of all votes domestically. The following number of votes were cast on party lists (domestic votes only):

|  |  |  |
| --- | --- | --- |
| **Party** | **Vote share %** | **Deviation from prediction** |
| Fidesz | 47.71% | +5.71% |
| Jobbik | 19.95% | -2.05% |
| MSZP | 12.5% | -0.04% |
| LMP | 7.36% | -1.14% |
| DK | 5.62% | -1.88% |
| Momentum | 3.19% | +0.19% |
| Együtt | 0.68% | -1.32% |
| MKKP | 1.80% | +0.80% |
| Other minor parties | 1.23% | - |

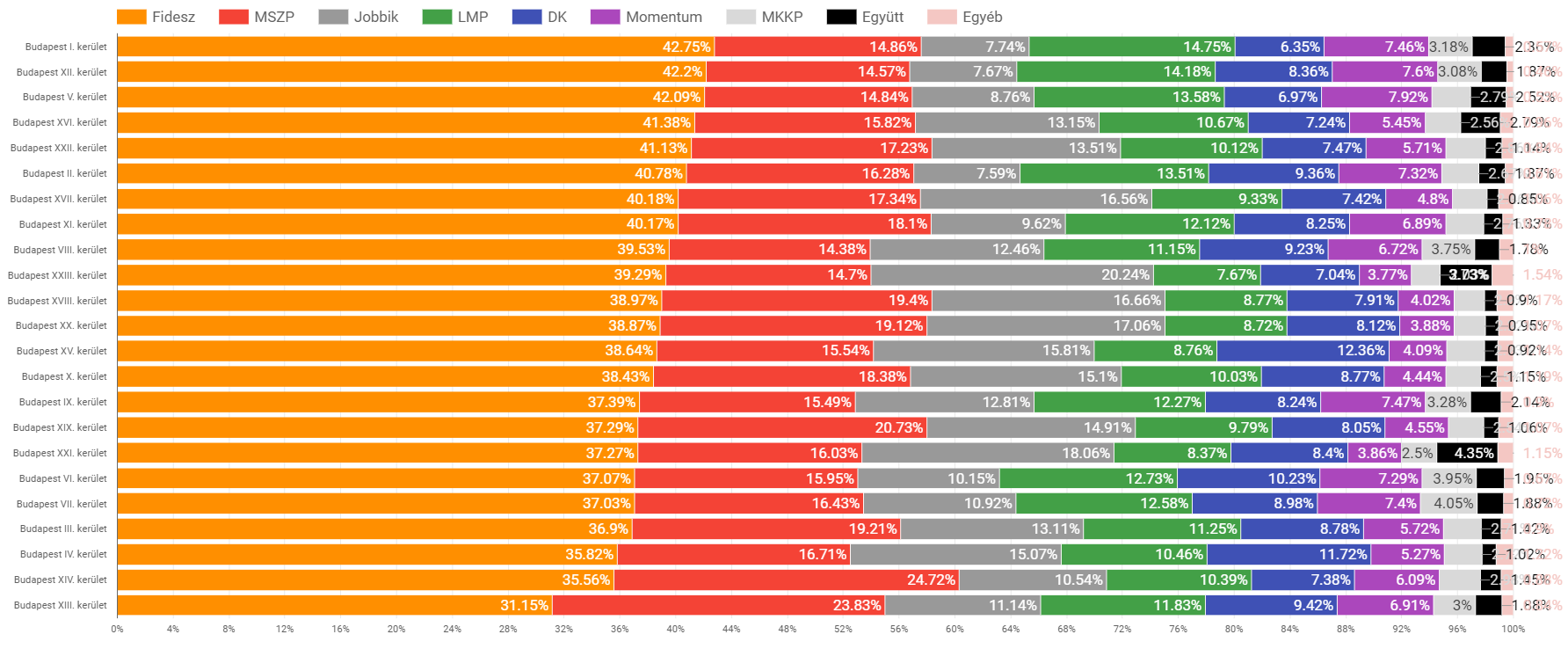
As the table shows, Fidesz was massively underpredicted while Jobbik, LMP, DK and Együtt all performed worse than expected, while MKKP was performing much better among minor parties not making the 5% parliamentary threshold.

Fidesz performed very well in all areas getting as high as 64% in one district and over 50% in 51 out of 106 districts[[4]](#footnote-4). It is important to highlight that some OEVKs contain a mixture of cities and smaller towns so the vote share of Fidesz was dependent on the composition of a given district. There will be further analysis provided on this later.

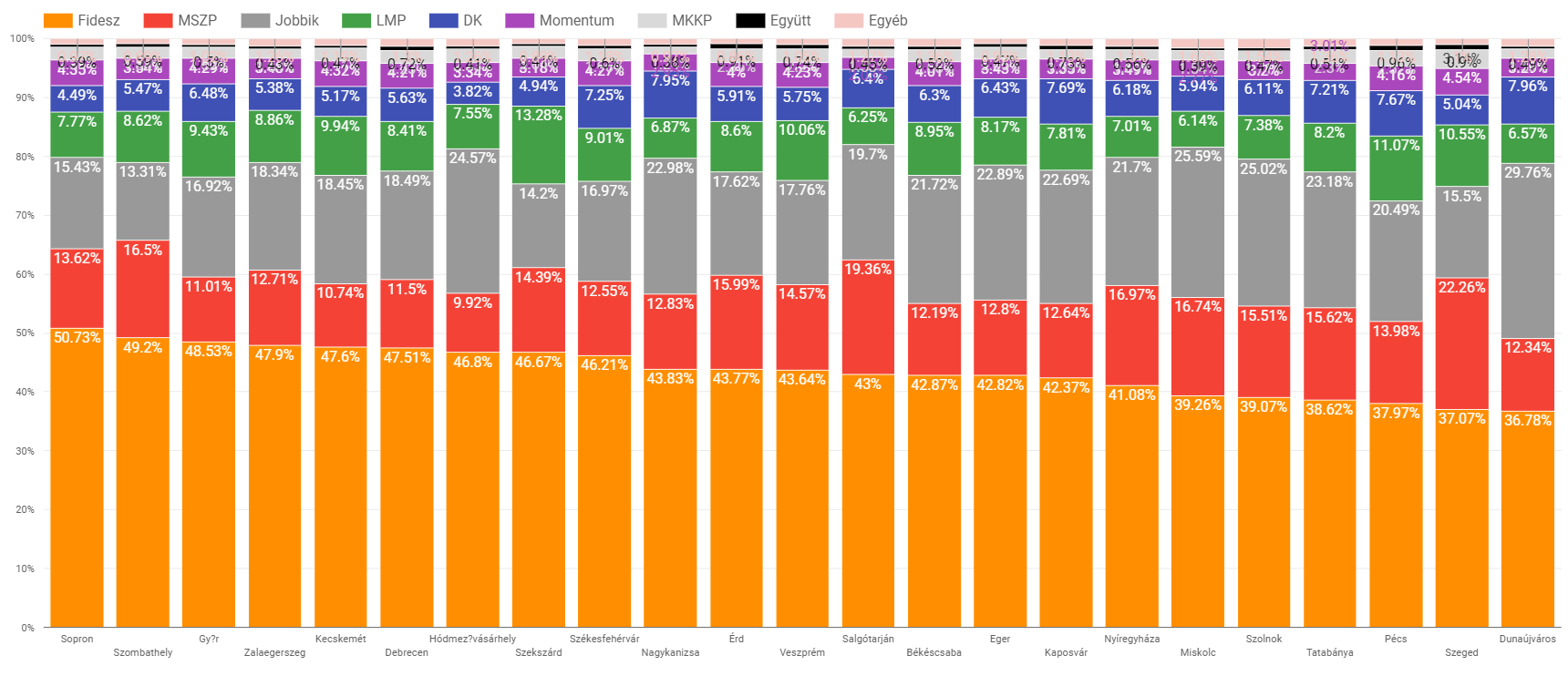


**List of districts by Fidesz party vote share (ascending)**

It was also interesting to see how Fidesz and other parties performed in major cities and in Budapest since these areas were the target seats for left-wing opposition parties. In Budapest, Fidesz only got 38% of all votes (compared to their national average of 47.71%) and they also performed worse in the 23 major cities of Hungary getting 43.35%. Opposition parties had strong areas such as Budapest district no. 13, 14 and Szeged for MSZP, Budapest district 1 and 2 for LMP, or Dunaújváros and Miskolc for Jobbik.



**Party vote shares in 23 districts of Budapest (Fidesz descending order)**



**Party vote shares in 23 major citites (Fidesz descending order)**

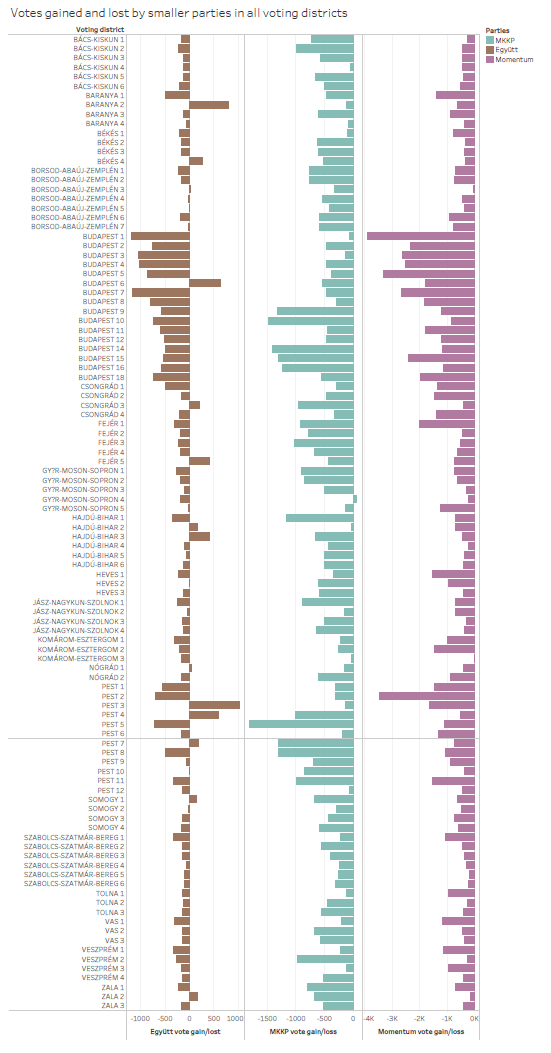
**5.2 Tactical voting numbers**

It was also interesting to see the difference between party list and party candidate votes. These numbers shown instantly the effect of tactical voting:

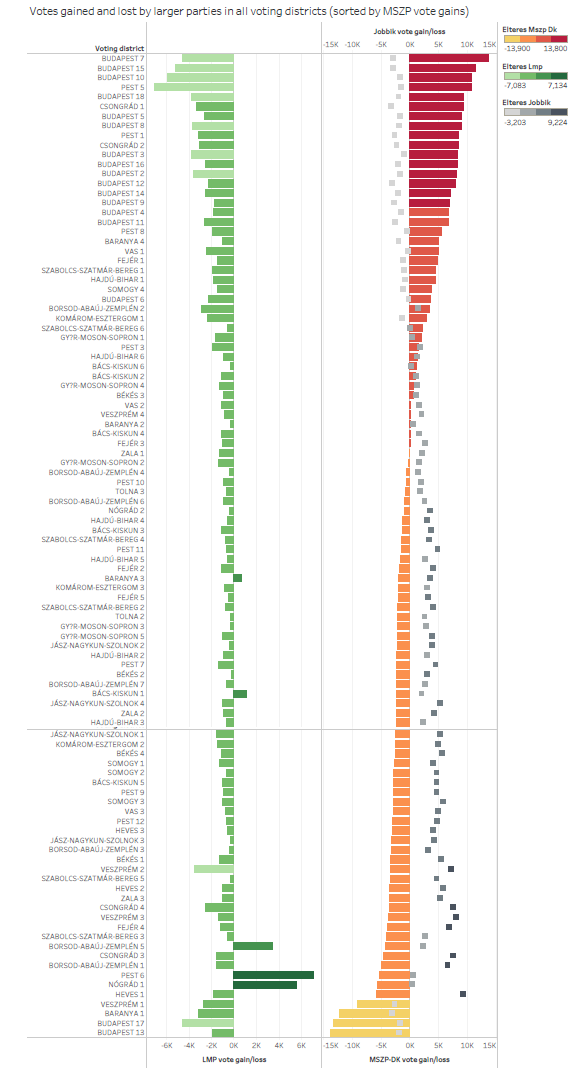
|  |  |  |  |
| --- | --- | --- | --- |
| **Party** | **Party list votes** | **OEVK candidate votes** | **Tactical voting gains/losses** |
| Fidesz | 2608086 | 2636203 | +28117 |
| Jobbik | 1090550 | 1276766 | +186216 |
| MSZP | 681358 | 622458 | -58900 |
| LMP | 402266 | 312731 | -89535 |
| DK | 307401 | 348178 | +40777 |
| Momentum | 174225 | 74906 | -99319 |
| Együtt | 37018 | 58565 | +21547 |
| MKKP | 98671 | 39704 | -58967 |

The table shows a mixed picture, but it is clearly visible that LMP, Momentum, and MKKP voters (the last of these was very surprising) voted tactically in favor of the stronger candidates. Együtt voters also followed this approach, but since they contested two seats, this is not apparent from this list. Also, MSZP and DK lost some votes, but they gained crucial votes from Jobbik and other smaller parties in key Budapest districts.

The following visualizations[[5]](#footnote-5) highlight how smaller and larger opposition parties voted in each OEVK (bars going to the left indicate number of tactical voters of a given party while bars going to the right are vote gains by OEVK candidates of given parties). Vote gains and losses are calculated based on the difference between party list vote and candidate votes.



**Difference in party list vote and candidate vote of smaller parties**



**Difference in LMP and MSZP-DK and Jobbik party list and candidate vote shares (MSZP/DK gains in descending order, MSZP/DK vs. Jobbik on joint axis)**

TB next: highlight interesting patterns (Együtt vote, LMP gains, best districts)

**5.3 Impact of tactical voting in OEVK**

As the charts show, voters of minor parties (MKKP, Együtt and Momentum) have voted tactically in almost all districts. Együtt voters were less likely to vote tactically where they fielded stronger candidates (such as in Budapest 6 district) and the same goes for one MKKP candidate in Győr 4 district (their candidate was featured in a TV debate few weeks before the election). Momentum had large swings between almost no tactical voting to almost 4000 votes in Budapest 1 and Pest 2 districts where LMP candidates stood against Fidesz, and really needed to make extra gains to win the seat (Budapest 1 was successfully gained).

Among larger parties, Jobbik benefited the most in terms of votes gained as in many districts, they gained an extra 5000 to 9000 votes from left-wing opposition parties. Still, it was not enough to win seats against Fidesz (except for Fejér 4). MSZP and DK candidates lost a total of 10000 votes overall (because of their voters voting for Jobbik in rural areas), but they gained 10 seats in Budapest and 1 in Csongrád (Szeged) thanks to the extra votes from all opposition party supporters (including Jobbik). So in Budapest, tactical voting worked very well and with slightly higher share of tactical voting, left-wing parties would have gained 5 more seats.

There are some interesting districts too where LMP or other parties performed much better than expected. The early mentioned districts where Együtt candidates got more votes than the party itself, LMP performed well in 4 OEVKs, namely Bács-Kiskun 1, Borsod 5, Pest, 6 and Nógrád 1. In Bács-Kiskun 1, Borsod and Nógrád 1, the party had three strong candidates (a former MP and two mayors), but the Pest 6 was a district where some websites advocated tactical voting for LMP, while my website did not. Still, some people felt that they should support LMP here and in the end, it pushed their expected vote share up by almost 15%.

In some districts, there was very limited number of people who switched votes mostly in rural seats (such as in Hajdú-Bihar 6 or or Győr 4). Sometimes it was even unclear who the stronger candidate is and both MSZP and Jobbik candidates gained votes (but they still lost by more than 15000 votes against Fidesz). Overall, tactical voting was successful in Budapest and in some major cities, but otherwise, Fidesz was just too strong in many areas and therefore tactical voting was not enough to turn around more seats.

**6 Pattern in 2018 results, prediction errors and new prediction models**

**6.1 Overview**

The first step to understand patterns in the 2018 results (without comparing them to 2014) was to see how people voted in different areas. For that, I devised a set of predictors such as the OEVK profile type (depending on the number of voters living in rural areas in a district) and looked at the town size categories.

To predict the election results and find where parties overperformed and underperformed compared to original expectations, I ran three different prediction models against the final results:

* Pre-election predictions (UNS with candidate strength)
* GLM prediction (using 2014 voting data aggregated on town level and data from the Hungarian Central Statistical Agency (KSH) on demographics and other predictors)
* UNS prediction (using the 2014 OEVK-level data)

The following sections will refer to sections in the main project summary file[[6]](#footnote-6) created in Markdown.

**6.2 Patterns in the 2018 results (Part 2 in Markdown)**

After merging and dividing the election data in Part 1 of the markdown file, I looked at the following predictors on the OEVK level:

* OEVK profile (0 to 6)

The OEVK profiles were the following:

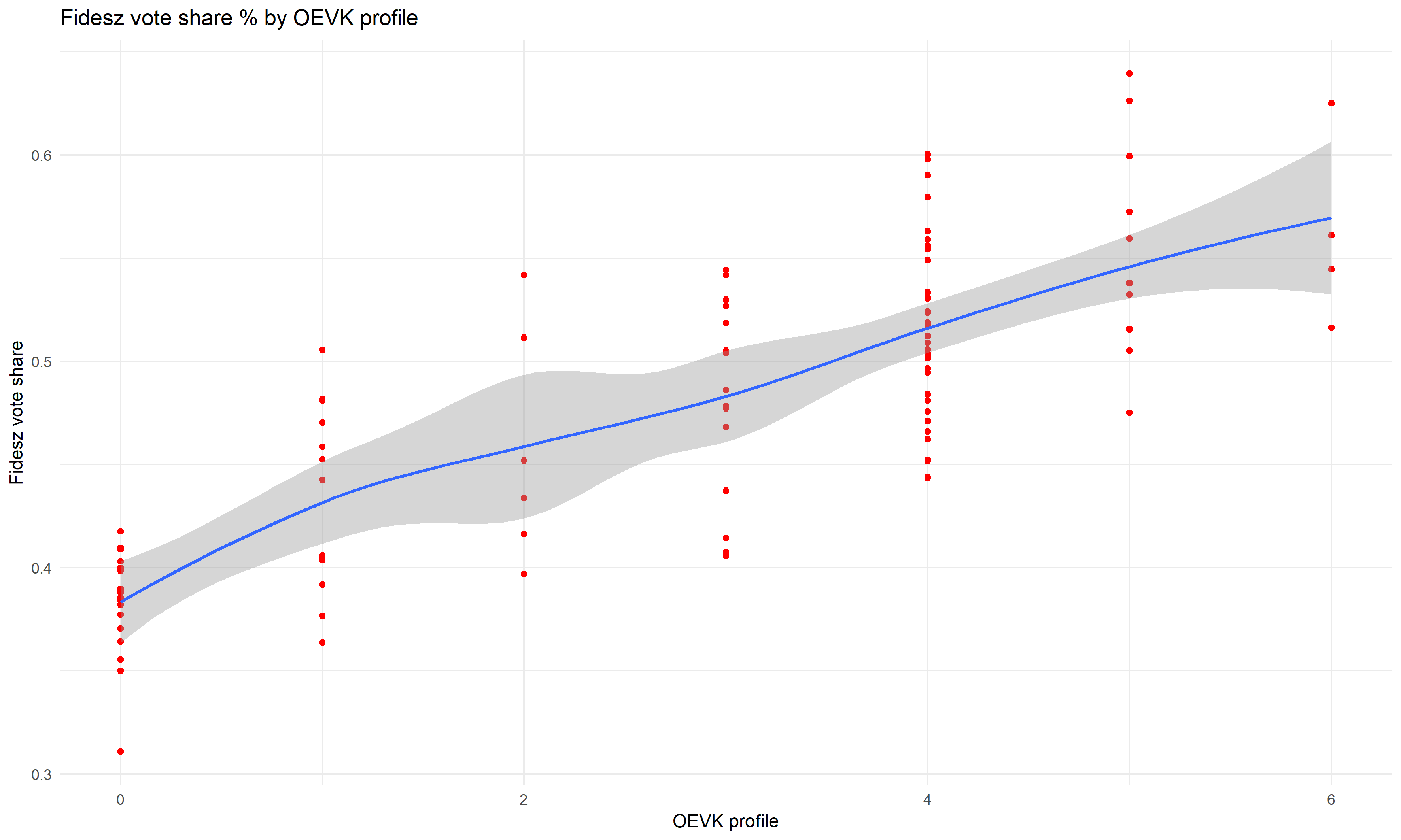
|  |  |  |
| --- | --- | --- |
| Profile ID | Description | No. of OEVKs |
| 0 | Budapest | 52 |
| 1 | County capital with few or no smaller towns |
| 2 | County capital or major city with 5-15 smaller towns |
| 3 | Major city or cities with 10-25 smaller towns |
| 4 | Some smaller cities combined with 30-40 villages | 54 |
| 5 | One or two smaller town with 50-80 villages |
| 6 | One smaller town with 100+ villages |

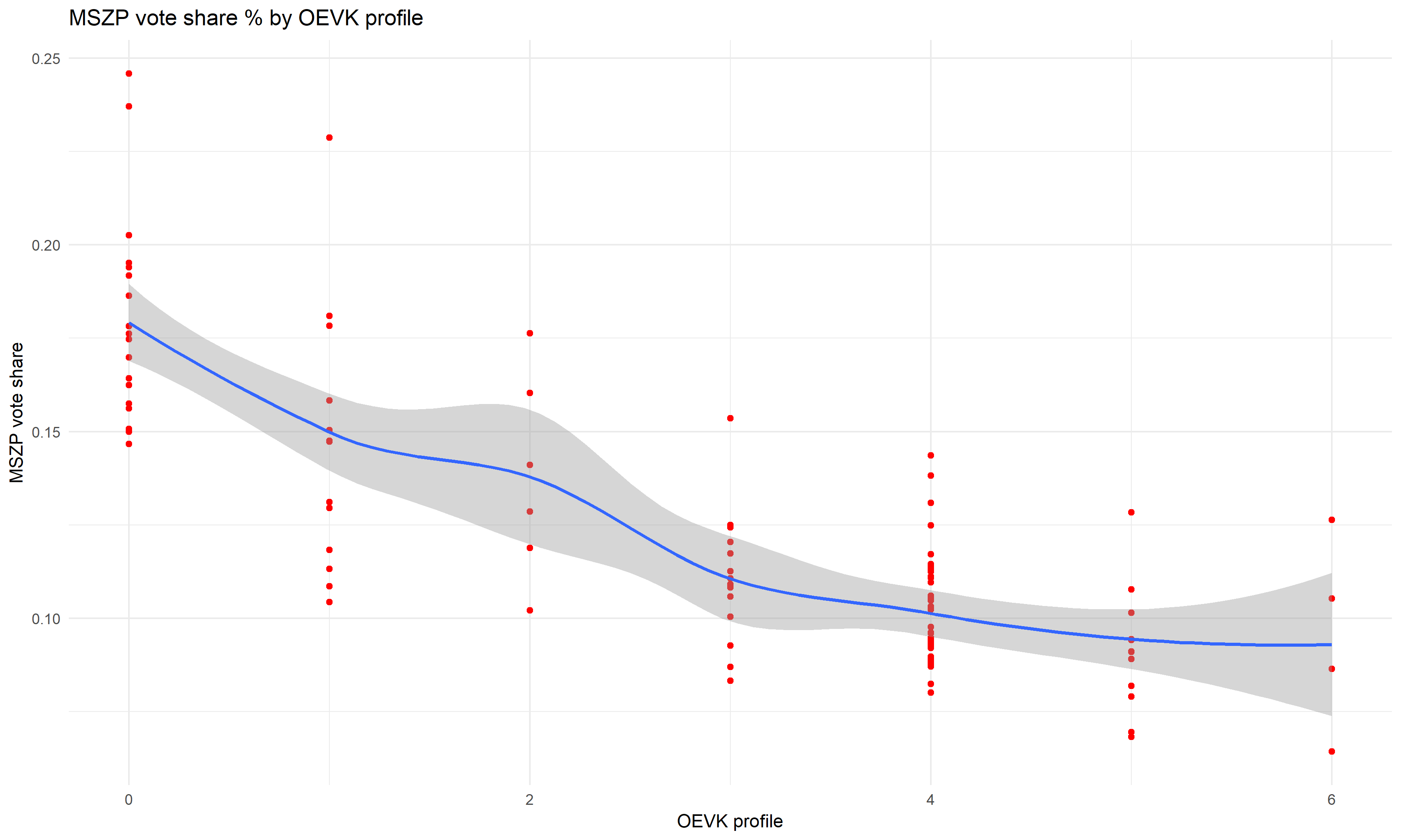
* Town size

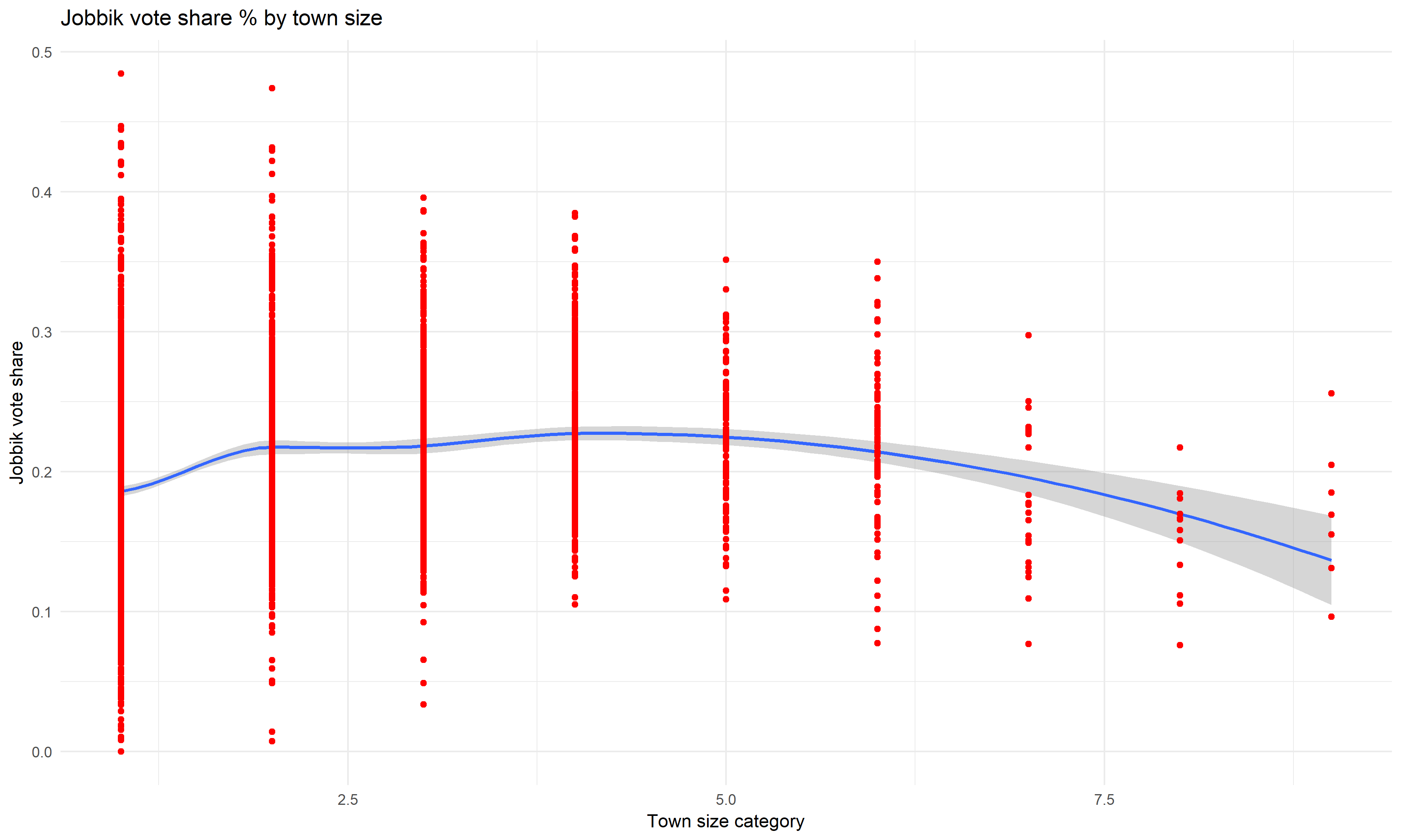
|  |  |  |  |
| --- | --- | --- | --- |
| Town size ID | Number of voters | Number of towns | Total voters in category |
| 1 | 1-500 | 1314 | 223530 |
| 2 | 501-1000 | 676 | 323814 |
| 3 | 1001-2000 | 581 | 551065 |
| 4 | 2001-5000 | 366 | 730765 |
| 5 | 5001-10000 | 112 | 547586 |
| 6 | 10001-30000 | 87 | 1058764 |
| 7 | 30001-60000 | 22 | 764393 |
| 8 | 60001-100000 | 12 | 701510 |
| 9 | Over 100000 | 7 | 620037 |

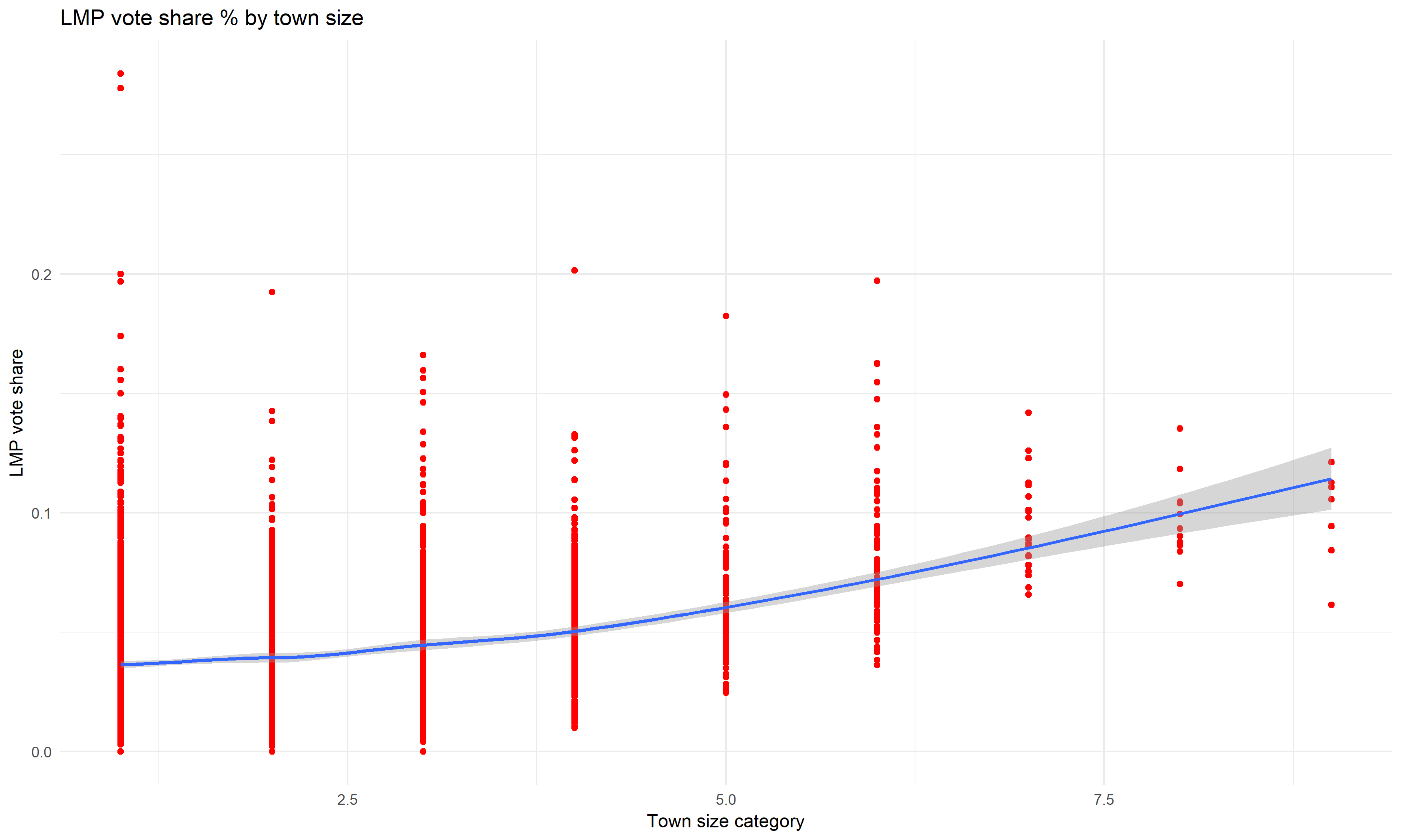
The number of voters in both the OEVK profiles and town sizes were mostly evenly distributed. When looking at the voting records, it was clear that Fidesz and Jobbik performed well in high OEVK profiles (districts with rural areas) and in small towns while left-wing opposition parties were exactly the opposite.

These results were not surprising as since 2010, all of these parties performed this way in general elections, so their voter base has not changed much. It was interesting to see though that Fidesz had a very strong support in OEVK types 4 to 6 (over 50% on average) meaning that over half of the seats were taken by them with an absolute majority (something that cannot be overtaken by even standing head to head against the Fidesz candidate).





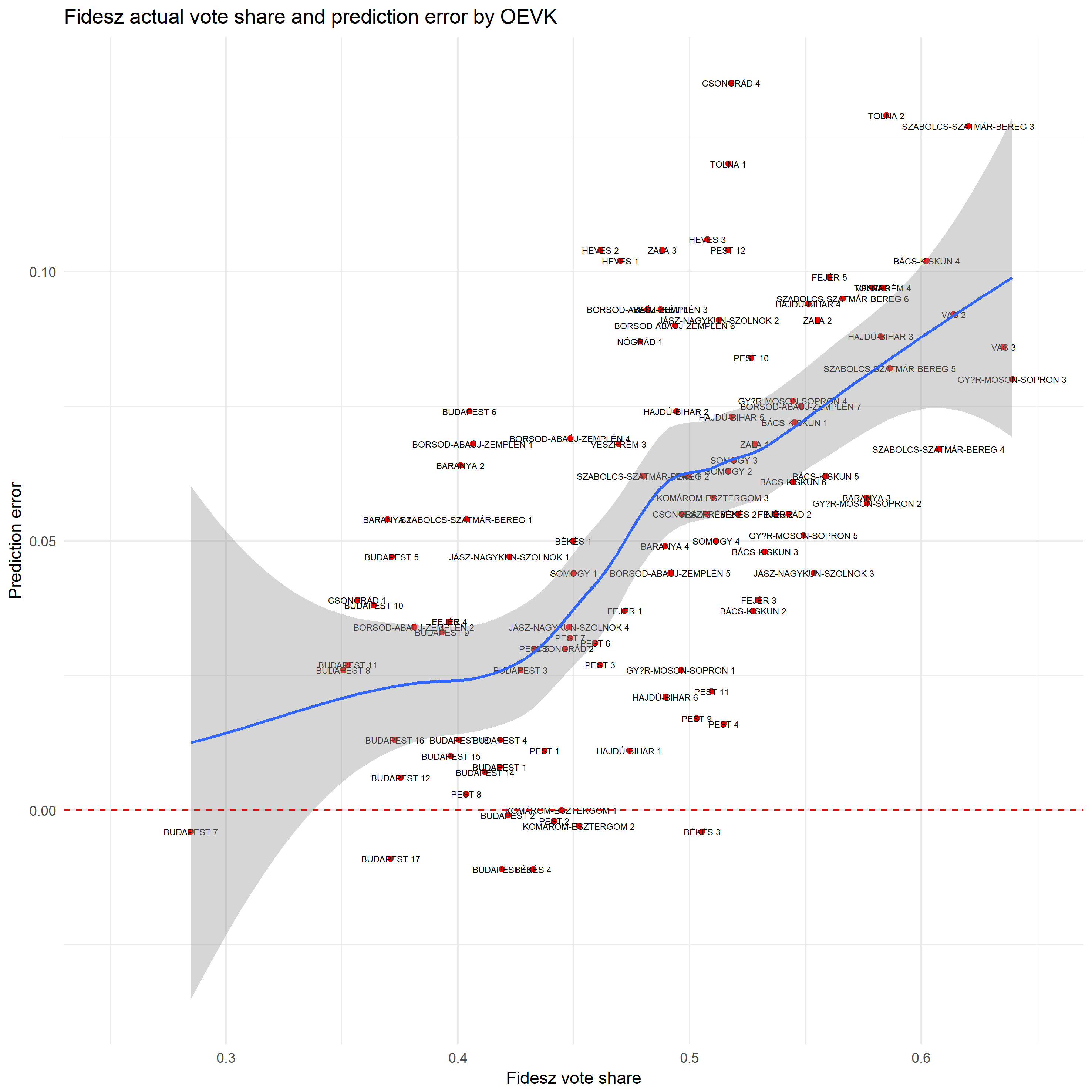




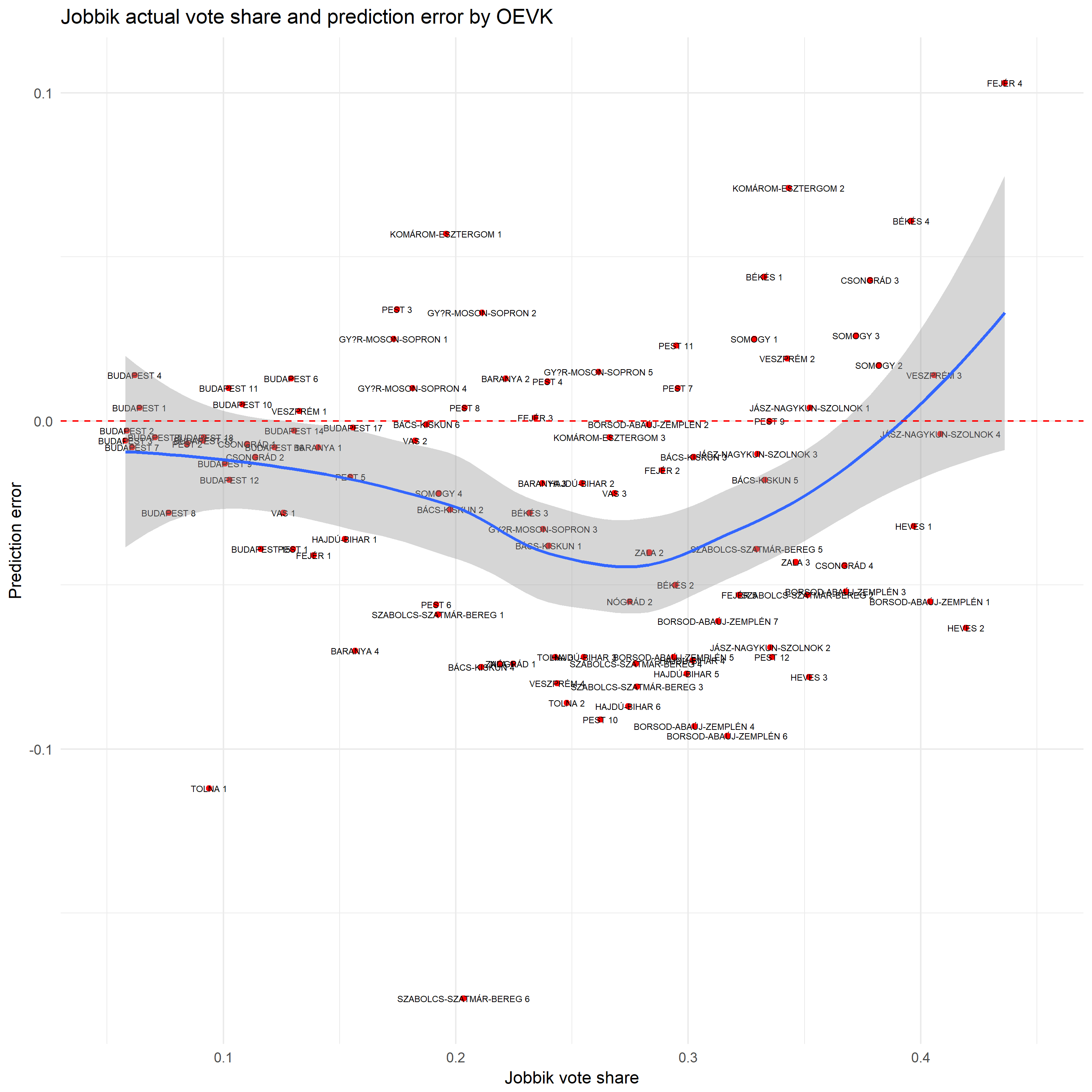
**6.3 Pre-election prediction and deviation (Part 3 in Markdown)**

The next step was to look at deviation from pre-election predictions. I collected all predictions into a single spreadsheet to make the comparison easier and at this stage, there were some interesting findings:

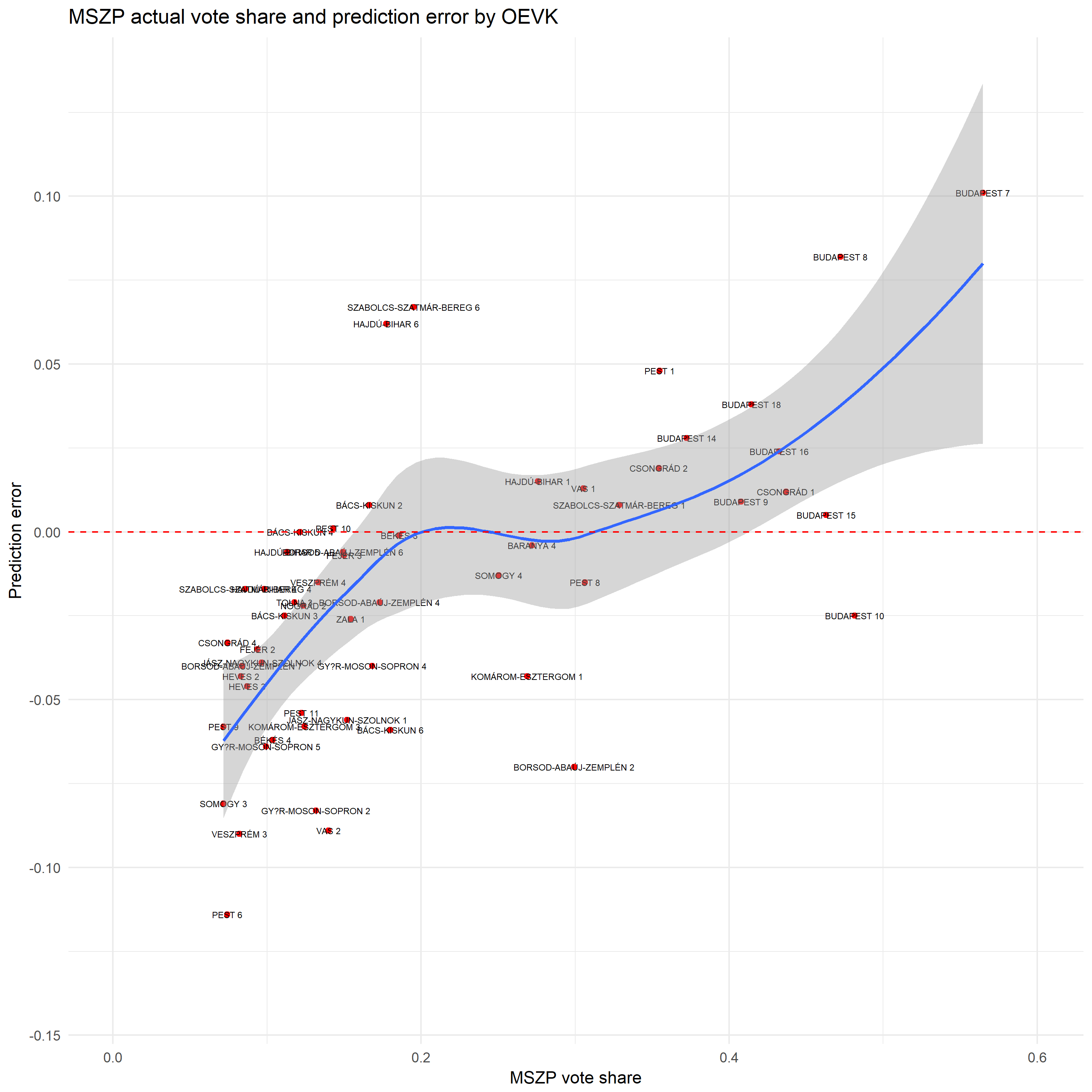
* Since Fidesz got 5.5% more votes than expected, they overperformed in almost all OEVKs compared to the pre-election prediction, but their strongest performance was in Jobbik target seats, such as Heves and Borsod counties



* Jobbik has slightly underperformed in Borsod and Heves meaning that it was likely that Fidesz took their 2014 voter base in key districts. They also surprisingly won Fejér 4, which was the 15th strongest Jobbik district according to pre-election predictions



* MSZP vote has collapsed in western part of Hungary such as Vas, Veszprém and Győr counties meaning that these areas had become rock solid Fidesz districts. Also, MSZP and other left-wing parties performed slightly better in Budapest than expected which led to the expected number of victories (12 out of 18) in the capital.



**6.4 GLM prediction using KSH data on town level (part 4)**

To understand more about the prediction errors, a general linearized model (GLM) was used on aggregated town-level data. The difficulty with this prediction was that first, all data had to be merged with the data from KSH database, and second, there was a need to normalize results. And finally, the weighing of data points would have been needed, which was not done at this stage of the research.

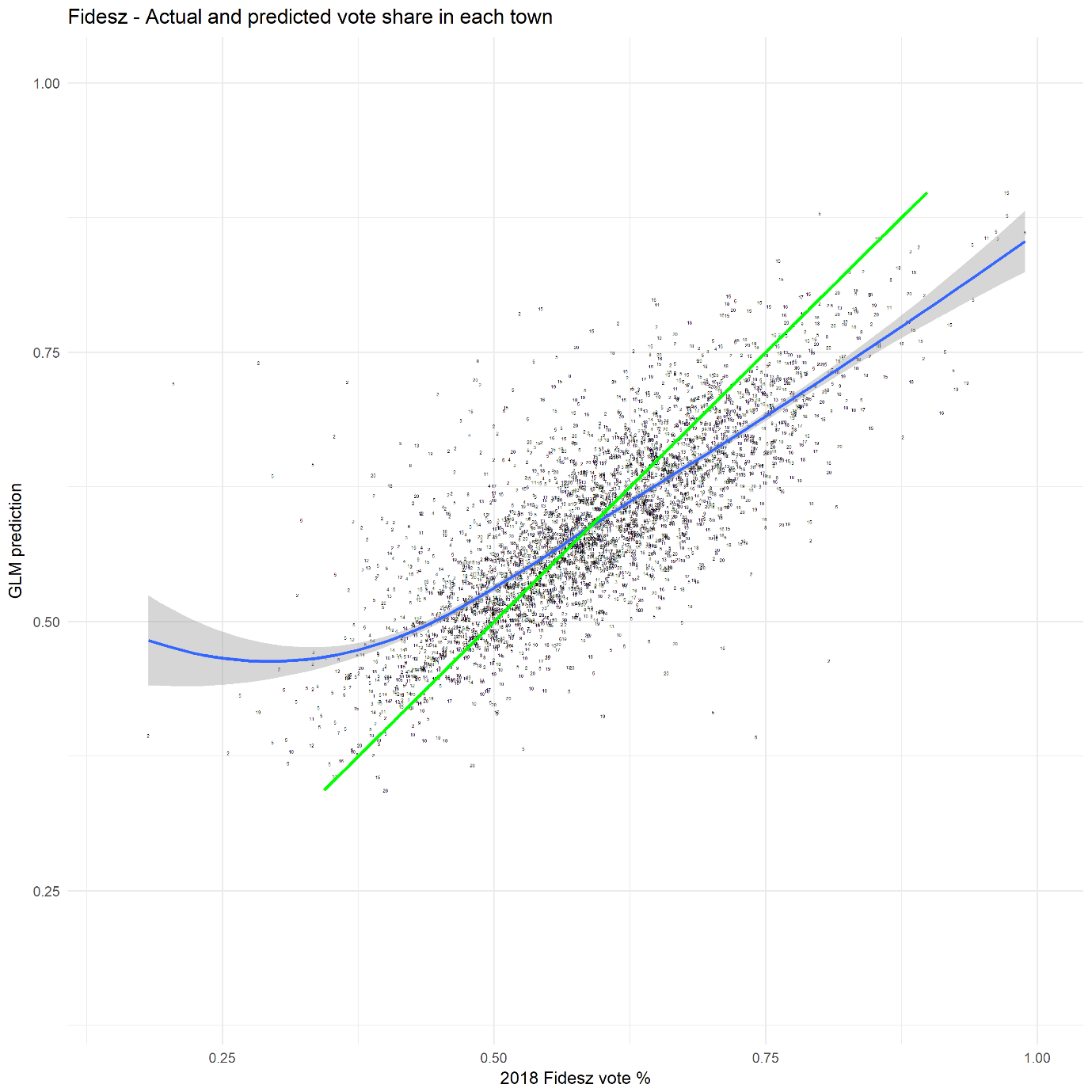
The town level predictions showed all expected vote share of major parties using the following predictors: 2014 results + over 30 predictors such as gender ratio, average income per person, distance from county capital and Budapest, etc.

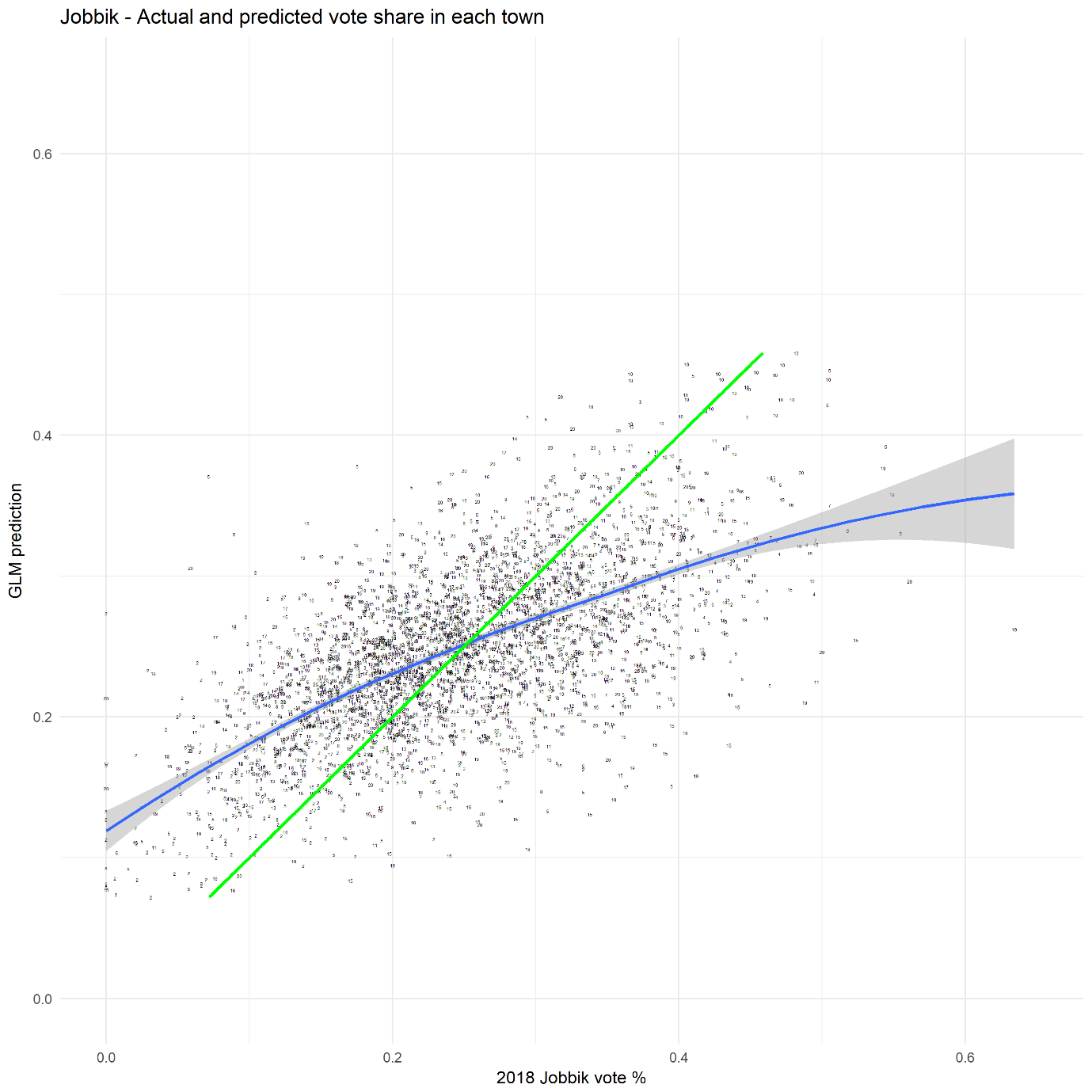
Since the weighing of these data points was left out in the data cleansing (so results of Budapest with close to 1 million voters counted as one data point versus the 1400+ micro-towns with less than 500 voters counted as 1400+ data points), the additional predictors told very little about why parties got more or less votes than expected. Because of that, none of the 30 predictors turned out to be significant, and they did not reduce prediction error in the GLM models.

Still, a pattern in underpredicted towns and cities was visible in this prediction. Looking at the megye\_id (county id) values, we can see that there are dozens of massively overperforming towns in Baranya, Borsod, Somogy, Veszprém, Zala (2,5, 15, 19, 20) counties. These areas all contained at least one opposition target seat meaning that Fidesz could have run a targeted campaign in these areas to increase their vote share.

Jobbik also performed better than expected in Zala (id 20), but the picture was mixed in Borsod (id 5) where they could have taken 3 out of 7 seats.

To have more insight into results on the town level, weighing will need to be done to differentiate small towns from large cities (or the results should be analyzed by town size).





**6.5 GLM prediction using town size and OEVK profiles on OEVK level (part 5)**

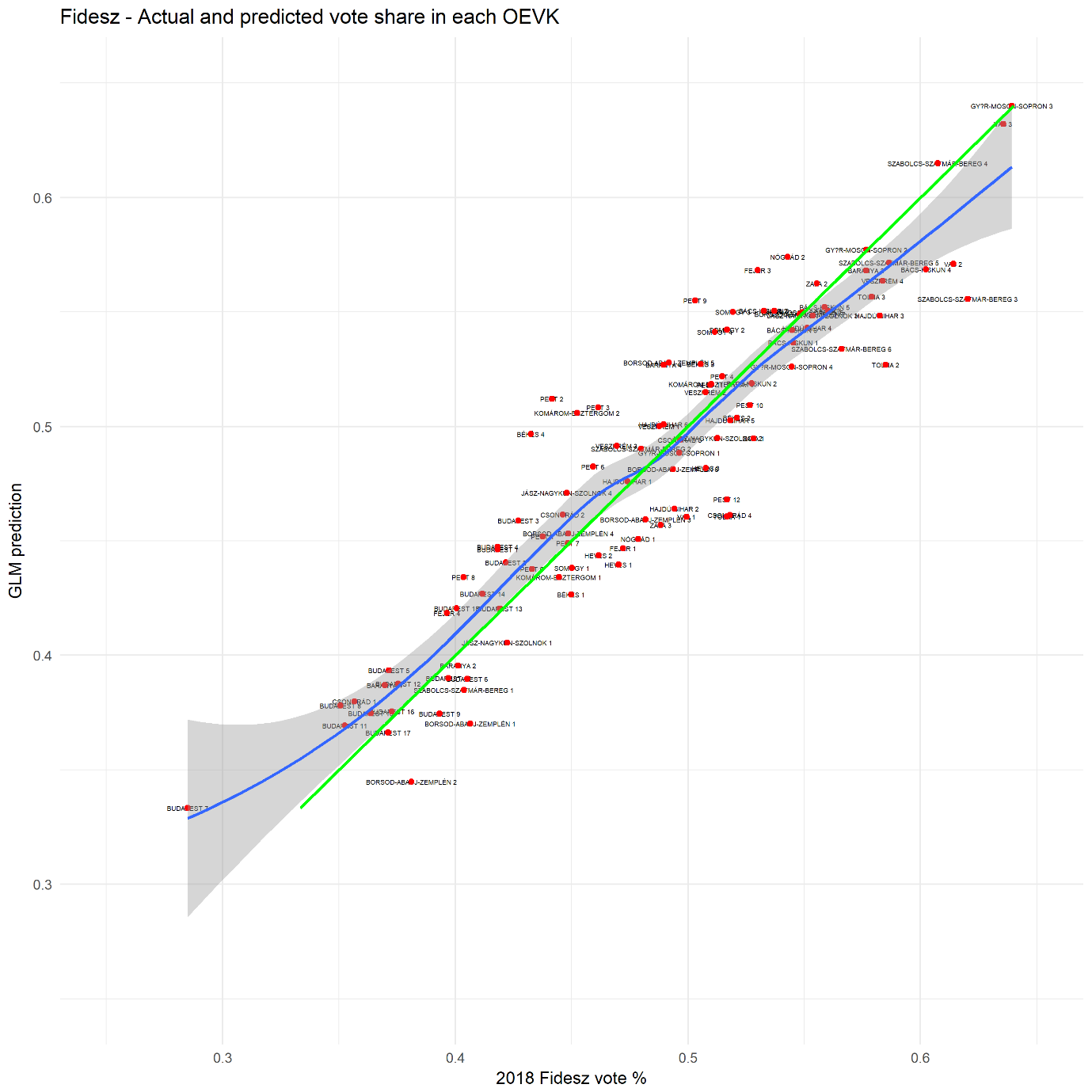
Using GLM on OEVK level was much easier since there are only 106 data points and there is much smaller difference in the size of districts compared to towns. In this model, the 5 predictors were the following:

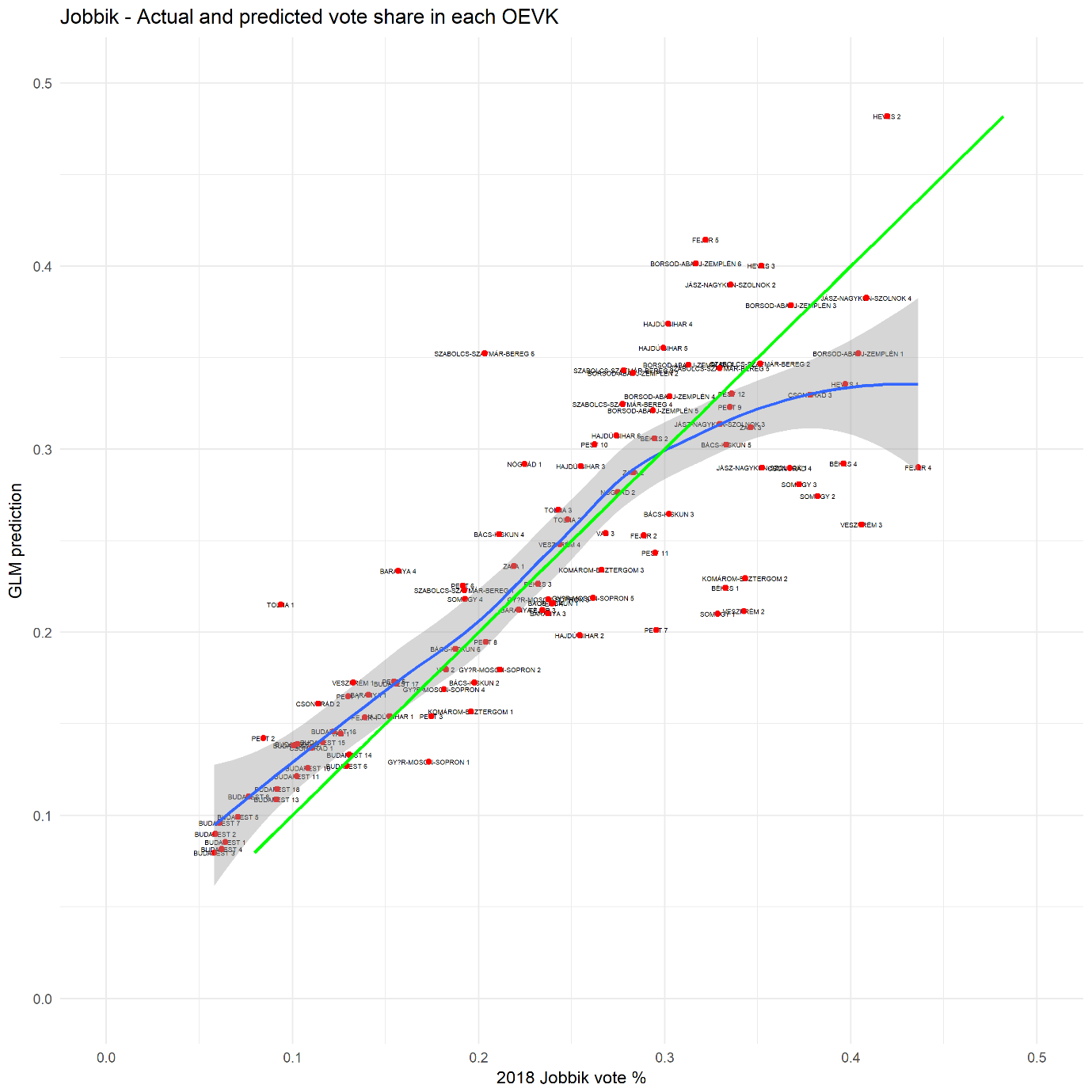
* 2014 vote share
* Number of towns in the OEVK
* Percentage of cities
* OEVK profile
* Number of polling stations

In this model as well, the strongest predictor was the 2014 vote share, and none of the other predictors had a significant impact on the expected support of a party. However, in the case Fidesz and Jobbik, where tactical voting did not influence their final results that heavily, some patterns emerged:

* The overperformance of Fidesz was not that strong, but they did get more votes in almost all target seats (such as Budapest 6, Borsod 1, Heves 2, Pest 12, etc.). What is interesting is that these seats belong to all kinds of OEVK profiles so there is no direct link between OEVK profile and overperformance, something that contradicts the previous findings. So Fidesz was able to mobilize slightly more voters in key districts regardless of the demographics of that OEVK.
* Jobbik did not perform that badly in key districts but they did not manage to go above 45% in key districts, which vote share usually guarantees victory. Even with tactical voting pushing up some key districts like Heves 1 or Pest 12, Jobbik underperformed in rural areas. In non-target seats though (where their vote share is in the 30-35% range), they performed better than expected.
* For the results of MSZP/DK and LMP, it is impossible to use the GLM results reliably, because tactical voting completely changed how these parties received votes in each OEVK.

The key finding here is that the OEVK profile itself does not explain why Fidesz performed better than expected since they managed to get extra voters even in large towns like Miskolc (Borsod 1) or Békéscsaba (Békés 1).





**6.6 UNS prediction comparing 2014 to 2018 on OEVK level (part 6)**

As the last step, I looked at simple UNS data from 2014 to see how a basic model with no other predictors would show the changes from 2014 to 2018.

TBD – similar results as GLM, but Fidesz OEVK 3-4 is better

**7 Conclusion**

Tactical voting results and UNS prediction error highlights

GLM on town level – to be explored

Further research (voter fraud)

1. Collection of all opposition withdrawals: <https://index.hu/belfold/2018/valasztas/2018/04/07/hol_maradt_talpon_egyetlen_ellenzeki_eselyes_a_visszalepesek_utan_osszefoglalo/> (Hungarian) [↑](#footnote-ref-1)
2. <https://index.hu/belfold/2018/valasztas/2018/03/06/zavecz_kutatas_kormanyvaltas_szandek> (Hungarian) [↑](#footnote-ref-2)
3. Vote shares and seat predictions are taken from <https://taktikaiszavazas.hu/mandatumbecsles> (Hungarian) [↑](#footnote-ref-3)
4. Full tables are available here: <https://datastudio.google.com/u/0/reporting/1lfsGiKJsNm1N1sd9cCB2s2nJk0wVHxBC/page/tLPS> <https://datastudio.google.com/u/0/reporting/1IkYaL_au-1TtnmQtrni1ljT1_LdkZgLy/page/9QPS> [↑](#footnote-ref-4)
5. The full visualization is available as PDF in the document package: oevk\_final\_PDF.pdf [↑](#footnote-ref-5)
6. See 106\capstone\_adat\project\_summary\_markdown.md [↑](#footnote-ref-6)