

Ádám Sanyó

Predicting the Outcome of the 2018 Hungarian General Elections

Tactical Voting Summary and Election Predictions using UNS and GLM

Central European University

Department of Economics and Business

Supervisor: Dr. Miklós Koren

Budapest, 2018

Acknowledgements

I would like to thank the following people who helped me with different parts of this project during and after the election:

Rudolf Hornig (Software Developer) – web design consultation and template setup

Péter Hajdú (Engineer) – setting up the Perl script to auto-generate data to webpages

Kristóf Menyhért (CEU Student) – data visualization on the website

László Sándor (Economist, Researcher) – cleaning T-Star data sources

Miklós Koren (Associate Professor at CEU) – discussion on prediction models

Andrea Fazekas (Social Media Manager) – helping to run Facebook campaigns

Contents

Acknowledgements	2
List of Tables and Figures	4
1 Election system and tactical voting overview	5
2 Technical solutions and data sources	6
2.1 Running the taktikaiszavas.hu website	6
2.2 Data sources used for predictions	8
3 Pre-election prediction model – UNS with additional factors	9
3.1 Polling data by Common Country Movement	9
3.2 Three steps of the UNS prediction	9
4 Tactical voting predictions	10
4.1 Tactical voting levels by different parties	10
4.2 Formal withdrawals and the effect on tactical voting numbers	11
4.3 The final pre-election prediction model	11
5 Election results and the effect of tactical voting	12
5.1 Election results	12
5.2 Tactical voting numbers	14
5.3 Impact of tactical voting in voting districts	18
5.4 Tactical voting by city type and county	19
6 2018 results, prediction errors, and new prediction models	21
6.1 Overview	21
6.2 Patterns in the 2018 results (Part 2 in Markdown)	22
6.3 Pre-election prediction and deviation (Part 3 in Markdown)	25
6.4 GLM prediction using KSH data on town level (Part 4 in Markdown)	28
6.5 GLM prediction using town size and OEVK profiles on OEVK level (Part 5 in Markdown)	31
6.6 UNS prediction comparing 2014 to 2018 on OEVK level (Part 6 in Markdown)	34
7 Conclusion	37

List of Tables and Figures

Figure 1: Typical view of a webpage showing expected vote shares in one constituency	7
Table 1: Vote shares and predicted seats prior to election day	12
Table 2: Vote shares and deviations from predictions	12
Table 3: List of districts by Fidesz party vote share (ascending)	13
Table 4: Party vote shares in 23 districts of Budapest (Fidesz descending order)	14
Table 5: Party vote shares in 23 major cities (Fidesz descending order)	14
Table 6: Difference in party list votes and candidate votes for each party	15
Figure 2: Difference in party list vote and candidate vote of smaller parties	16
Figure 3: Difference in LMP and MSZP-P/DK and Jobbik party list and candidate vote shares (MSZP-P/DK gains in descending order, MSZP-P/DK vs. Jobbik on joint axis)	17
Figure 4: Votes gained or lost due to tactical voting by town size (ascending order)	20
Table 7: Vote gains and losses in Budapest districts	21
Table 8: OEVK profiles based on different town composition	22
Table 9: List of town sizes IDs	22
Figure 5: Fidesz vote share by OEVK profile in each OEVK	23
Figure 6: MSZP vote share by OEVK profile in each OEVK	24
Figure 7: Jobbik vote share by town size category in each town	24
Figure 8: LMP vote share by town size category in each town	25
Figure 9: Deviation of Fidesz candidate's vote share from pre-election prediction	26
Figure 10: Deviation of Jobbik candidate's vote share from pre-election prediction	27
Figure 11: Deviation of MSZP candidate's vote share from pre-election prediction	28
Figure 12: GLM prediction of Fidesz candidate's vote share in each town	30
Figure 13: GLM prediction of Jobbik candidate's vote share in each town	31
Figure 14: GLM prediction of Fidesz candidate's vote share in each OEVK	33
Figure 15: GLM prediction of Jobbik candidate's vote share in each OEVK	34
Table 10: Input values and UNS multiplier values for each party	35
Figure 16: UNS prediction error of Fidesz candidate's vote share in each OEVK	36
Figure 17: UNS prediction error of Fidesz candidate's vote share in each OEVK	37

1 Election system and tactical voting overview

In 2012, a new election system was 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 was added to this system where the winning candidates in OEVKs carry over those votes that are not needed to win the seat (their votes minus the second place candidate’s 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 in this system than in the former, more proportional election system.

Because of this change, opposition parties had to coordinate their efforts in voting districts to win as many seats as possible. This coordination was quite limited with only 43 candidates withdrawing from the race in 29 seats¹. 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 (“whom should we vote for?”), for which a small sample (500 respondents) of telephone polling was carried out in all districts. This site also provided recommendations on the strongest candidates.

The election predictions on taktikaiszavazas.hu was based on a mixture of constituency-level polling in 20 OEVKs, which was carried out by Medián and Závecz Research, two polling agencies, and the 2014 election data, on which a uniform national swing (UNS) model was applied. The polling was ordered by Közös Ország Mozgalom (Common Country Movement or KOM), an NGO 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

¹ Collection of all opposition withdrawals:
https://index.hu/belfold/2018/valasztas/2018/04/07/ol_maradt_talpon_egyetlen_ellenzeki_eselyes_a_vissza_lepesek_utan_osszefoglalo/ (Hungarian)

possible to identify 40-45 districts (about 40% of all seats) where opposition parties had a 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 subsequent sections, I will explain the following aspects of my election prediction project:

- Different data sources and solutions used to run the website, and create the predictions
- 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, taktikaiszavas.hu, was set up as a static website running as a GitHub page. The reason for this approach was to reduce the impact of DDoS (denial of service) attacks. Each visitor is redirected to Cloudflare as well 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 were predefined in the page template fed from two different sources.

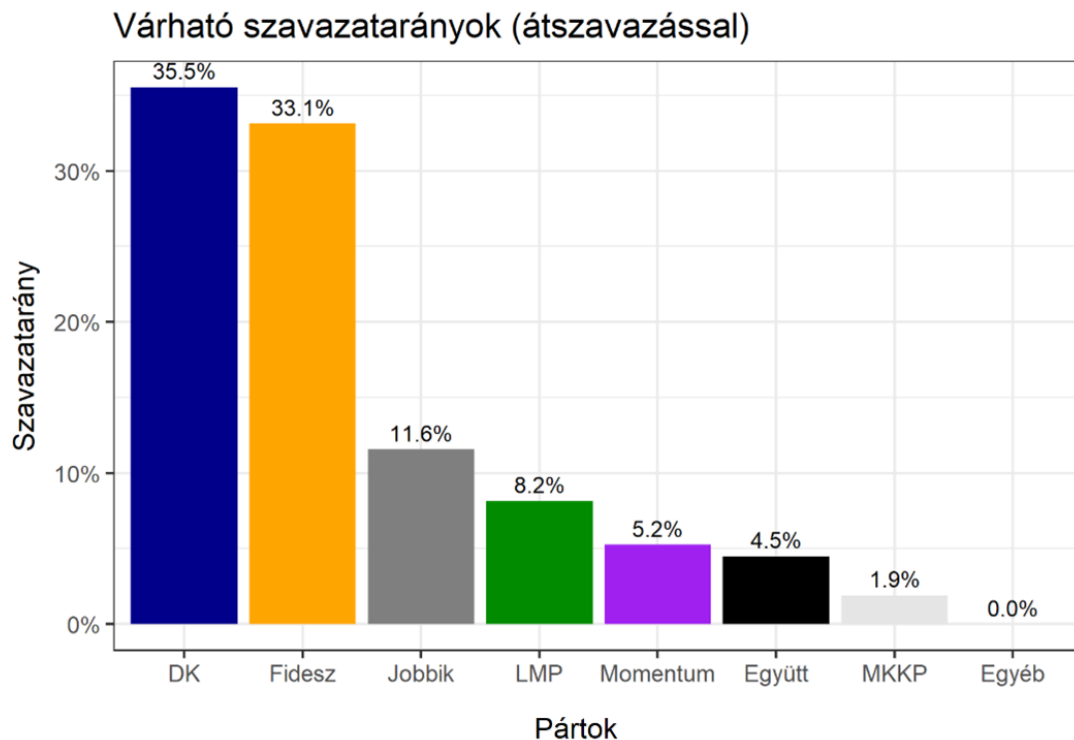
The tables are pulled from a CSV file with a Perl script. The script automatically amends values and generates new single static pages (one for each single seat race). It was relatively easy to update the website every week when new, adjusted predictions were published.

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

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.

2018. április 8-i becslések (átszavazással):










Becslés megbízhatósága a legesélyesebb jelölről: A baloldali jelölt az egyértelmű esélyes

Legesélyesebb ellenzéki jelölt: Ara-Kovács Attila (DK)

Márki-Zay Péter ajánlása: Ara-Kovács Attila (DK)

A V18 ajánlása: Ara-Kovács Attila (DK)

Jelöltek	Várható szavazatarány egyéniben (zárójelben az arányok átszavazás esetén)	
Kocsis Máté - Fidesz-KDNP 	32.1%	(33.1%)
Dűró Dóra - Jobbik 	14.9%	(11.6%)
Ara-Kovács Attila - DK 	23.5%	(35.5%)
Jakabfy Tamás - LMP 	12.1%	(8.2%)
Dr. Cseh Katalin - Momentum 	7.8%	(5.2%)
Baranyi Krisztina - Együtt 	7.2%	(4.5%)
Racsó Gábor - MKKP 	2.3%	(1.9%)

Várható győztes: Ara-Kovács Attila (DK)

Figure 1: 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 election (2014) 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 stood in the encoding of constituency and town names as those with special accents 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 to be used for analysis. These files are the following:

Initial sources (CSV)²

- vote_counts_precincts_2b (2014 data)
- 2018_egyeni_listas_29May (2018 data)
- 6 additional files from KSH data (T-Star data)

Generated and cleaned output files (CSV)

- 2018_UNs_baseline (used as the prediction base)
- 2018_by_oevk_clean (2018 results for each OEVK)
- oevek2014_clean_new (2014 results for each OEVK)
- oevek2014_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)

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

² You can find these files in the document package in the “capstone_adat” library

In addition to the 2014 data and other population data from KSH, the summary of 20 constituency-level polls was 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 Pre-election prediction model – 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 late March - early April, 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. The results were used to calculate both tactical voting, considering 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 figures I used for the UNS predictions.

3.2 Three steps of the UNS prediction

The UNS prediction was constructed the following way:

1. I computed 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. I increased the expected vote share by candidate strength, which is calculated individually for each party candidate. The calculation gave an additional 1-3% 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. I calculated 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 be stronger.

The expected difference between candidates was used in 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 Research³ 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

³ https://index.hu/belfold/2018/valasztas/2018/03/06/zavec_kutatas_kormanyvaltas_szandek (Hungarian)

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.

4.2 Formal withdrawals and the effect on tactical voting numbers

In 29 seats out of 106, 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 where there was a withdrawal from the race, voters could not choose their own party's candidate anymore. In these cases, 80% of voters were reallocated to the strongest candidate. This shows that a clear withdrawal from the race pushes 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 re-allocated between candidates in all seats, and further increased where a candidate had formally withdrawn from the race

For the final prediction, I used the following vote shares for each party:

Party	Vote share % ⁴	Expected number of seats
Fidesz	42%	82
Jobbik	22%	9
MSZP-Párbeszéd	12.5%	8
LMP	8.5%	2
DK	7.5%	4
Momentum	3%	0

⁴ Vote shares and seat predictions are taken from <https://taktikaiszavazas.hu/mandatumbecsles> (Hungarian), this calculation was done expecting 65% turnout (5.2 million voters)

Együtt	2%	1
MKKP	1%	0
Independents	-	0

Table 1: Vote shares and predicted seats prior to election day

Based on this calculation, I expected that roughly 360,000 voters would vote tactically, and this was the realistic scenario. A more optimistic scenario was expecting 550,000 tactical voters with a 70% turnout. In this case, an additional 18 seats would have been taken by the opposition parties. However, 24 wins out of 106 were 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-Párbeszéd	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%	-

Table 2: Vote shares and deviations from predictions

As the table shows, Fidesz was massively underpredicted, while Jobbik, LMP, DK, and Együtt all performed worse than expected. MKKP took on more votes 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 districts⁵. It is important to highlight that three-quarter of 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 in *Section 5.4*.

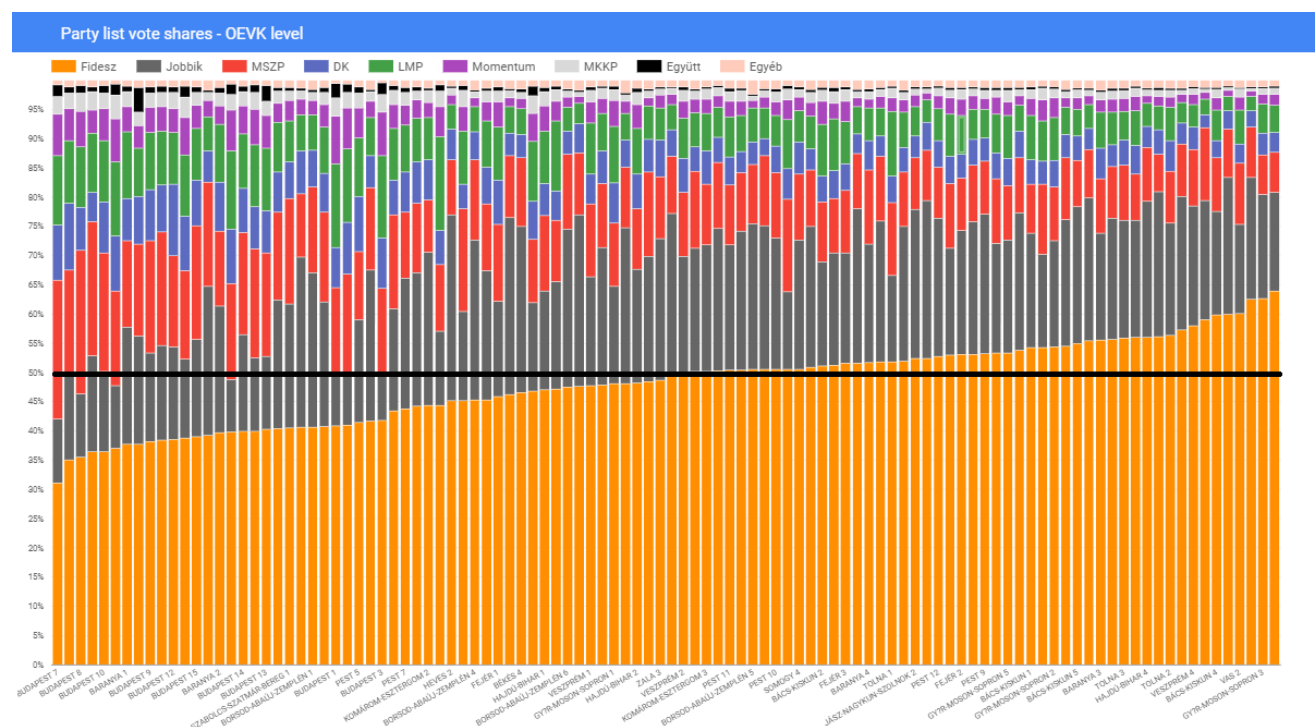


Table 3: List of districts by Fidesz party vote share (ascending)

It was useful 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 underperformed in the 23 major cities of Hungary by getting only 43.35%. Opposition parties had accomplished more in the following areas:

- Budapest district 13, 14, and Szeged for MSZP-P
- Budapest district 1 and 2 for LMP

⁵ Full tables are available here: <https://datastudio.google.com/open/1lfsGiKJsNm1N1sd9cCB2s2nJk0wVHxBC>
https://datastudio.google.com/open/1lkYaL_au-1TtnmQtrni1ljT1_LdkZgLy

- Dunaújváros and Miskolc for Jobbik

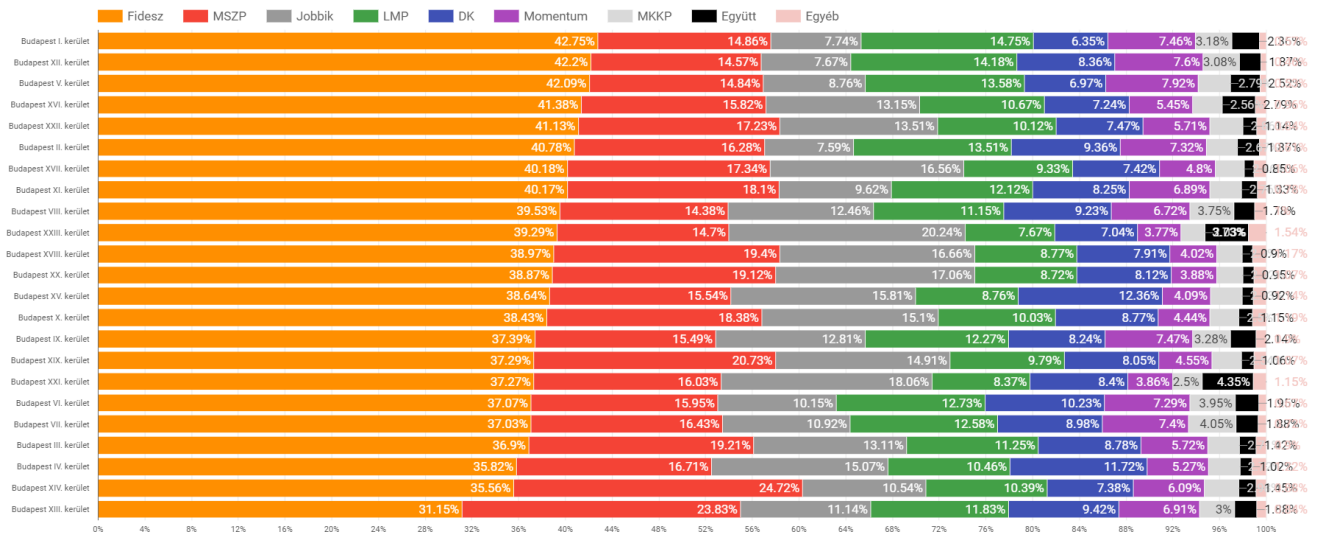


Table 4: Party vote shares in 23 districts of Budapest (Fidesz descending order)

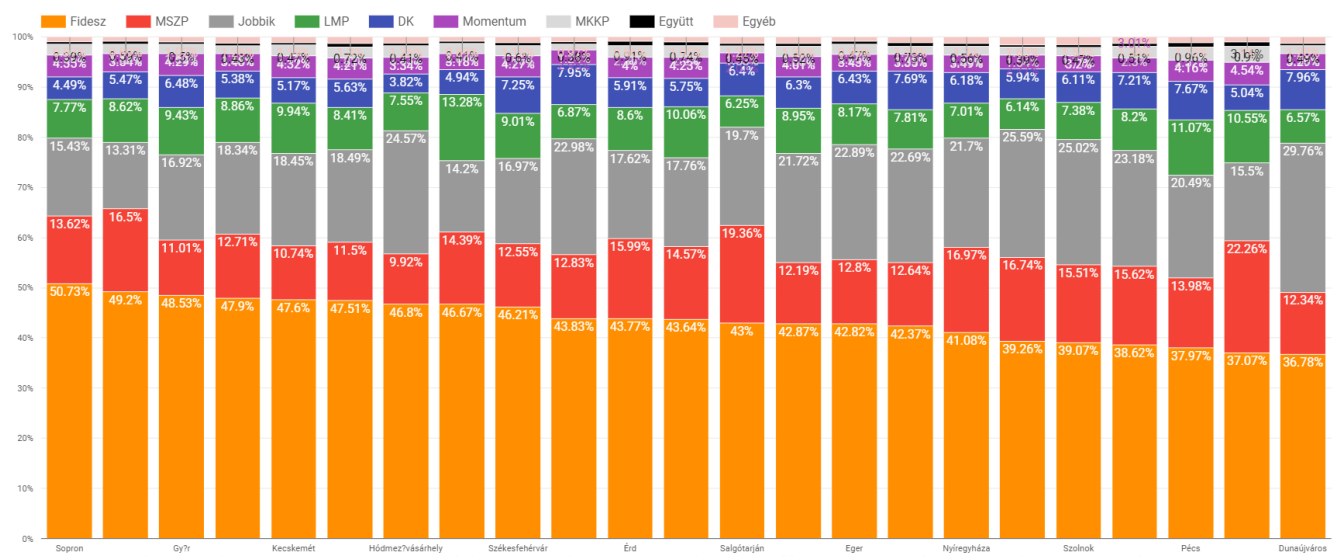


Table 5: Party vote shares in 23 major cities (Fidesz descending order)

5.2 Tactical voting numbers

It was a good starting point to see the difference between the party list and party candidate votes.

The effect of tactical voting was shown instantly in the following figures:

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

Table 6: Difference in party list votes and candidate votes for each party

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-Párbeszéd (MSZP-P) and DK lost 18000 votes, but they gained crucial votes from Jobbik and other smaller parties in key Budapest districts.

The following visualizations⁶ highlight how smaller and larger opposition parties voted in each OEVK. Bars going to the left indicate the 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.

⁶ The full visualization is available as PDF in the document package: *oevk_final_PDF.pdf*

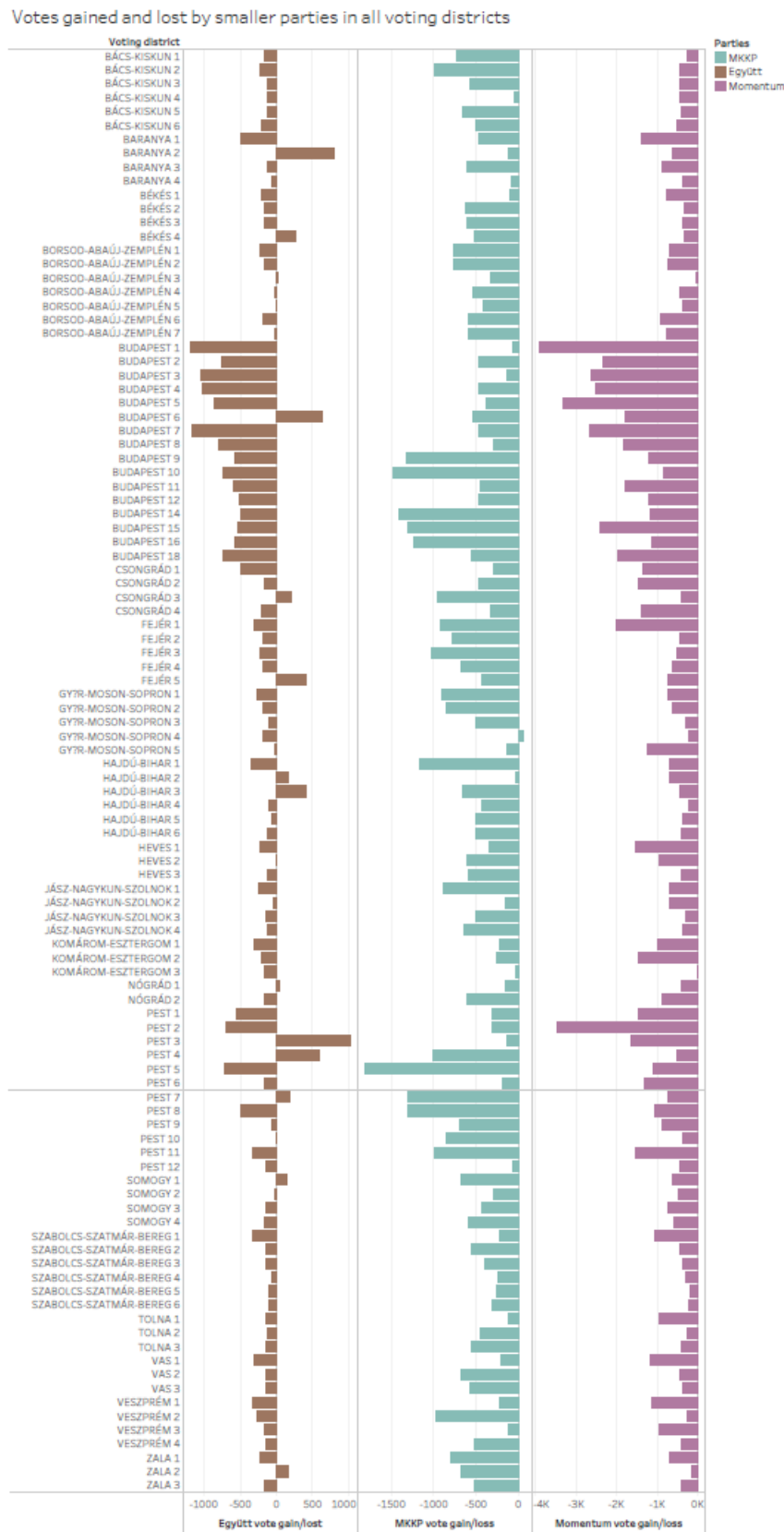


Figure 2: Difference in party list vote and candidate vote of smaller parties

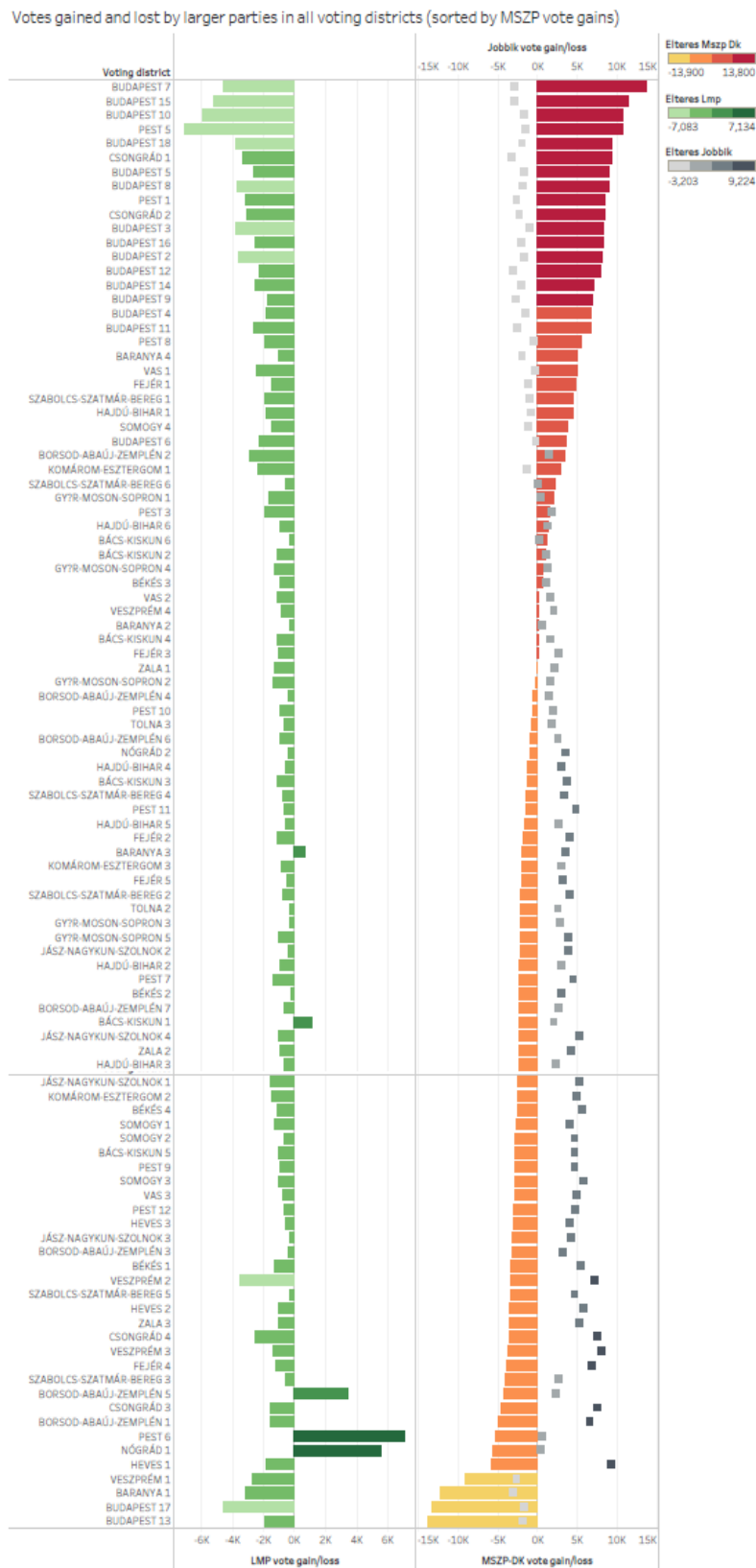


Figure 3: Difference in LMP and MSZP-P/DK and Jobbik party list and candidate vote shares (MSZP-P/DK gains in descending order, MSZP-P/DK vs. Jobbik on joint axis)

5.3 Impact of tactical voting in voting districts

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. 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. In these districts, 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 because in many districts, they received 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 OEVK. MSZP-P 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 a slightly higher share of tactical voting, left-wing parties would have gained 5 more seats.

In about a dozen OEVKs, LMP or other parties performed much better than expected. In districts such as Budapest 6, Pest 3, or Baranya 2, 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, LMP had three strong candidates (a former MP and two mayors), but in Pest 6, the website rendszervaltass2018.hu 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 many districts, there was a very limited number of people who switched votes mostly in rural seats, such as in Hajdú-Bihar 6 or Győr 4. Sometimes it was unclear who the stronger candidate is, and both MSZP-P 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 like Szeged or Érd, but otherwise, Fidesz was too strong in many areas, and therefore tactical voting was not enough to turn around more seats.

5.4 Tactical voting by city type and county

Prior to the election, it was expected that tactical voting would work the same way in every district; the stronger candidate will be supported by 20-25% of people regardless of where they live or what their OEVK profile is. This assumption turned out to be a bit simplistic as people treated tactical voting differently in cities than in smaller villages.

In the context of city type level data⁷, towns are divided into 9 different size categories, ranging from villages with less than 500 voters to cities or Budapest districts with over 100000 voters. Looking at this data, the following patterns are emerging:

- Jobbik managed to capture most of the extra votes in smaller cities, but Jobbik voters supported the left in Budapest and bigger cities.
- Roughly 25% of MSZP/DK voters chose Jobbik in smaller towns, while they benefited from extra votes in Budapest and in cities.
- LMP voters acted in an unusual way: in smaller towns, they did not vote tactically, but in big cities, they lent a massive vote share (about 50%) to MSZP-P or DK candidates. With this choice, they helped these parties to take seats.
- Momentum and MKKP voters were quite good with tactical voting because more than half of their supporters helped the stronger candidate even when their own candidates stood well in a given district.

⁷ Full tabs are available here: <https://datastudio.google.com/open/1UIGEkfA7h5P416TmdElsI5xm0dTKrwsK>

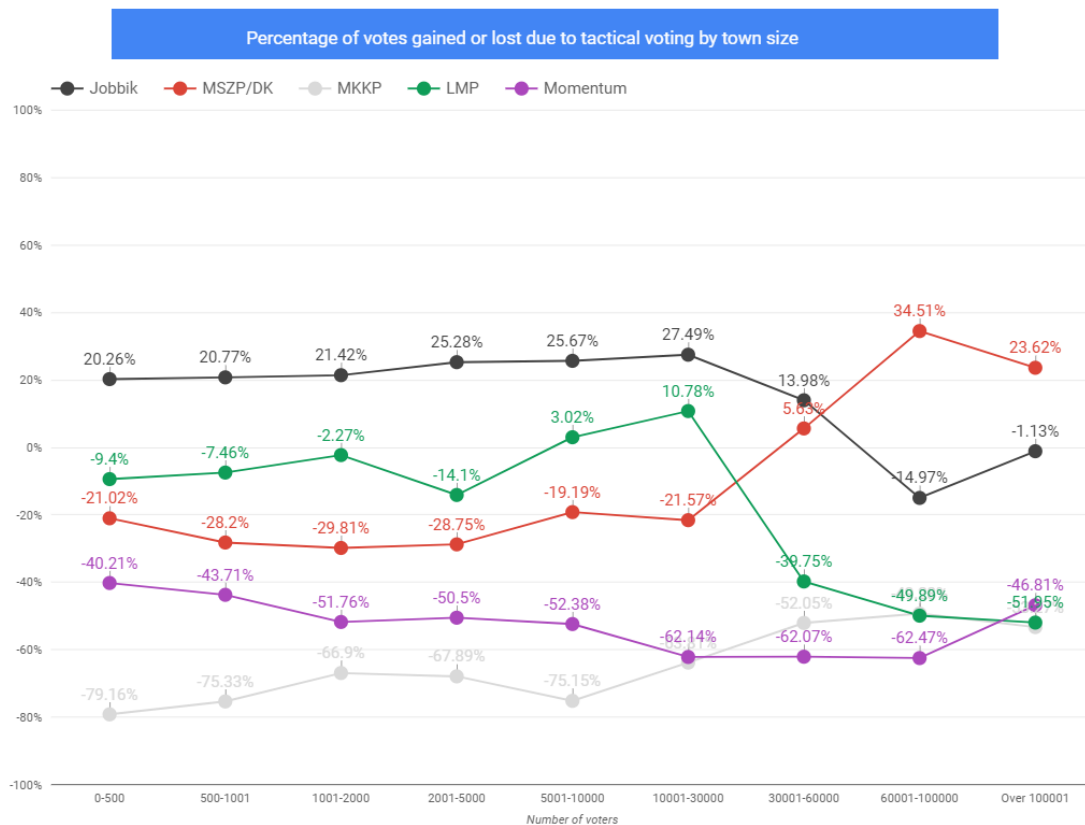


Figure 4: Votes gained or lost due to tactical voting by town size (ascending order)

When it comes to results in Budapest, we can see that people voted in an orderly way as there was only 3 out of 23 districts where more than one candidate benefited from tactical voting. The numbers here are also high ranging from 27% from Jobbik voters to 66% of Momentum voters, with LMP and MKKP in between. MSZP-P and DK got an extra 34% of votes in the capital, thanks to these votes.

Percentage of votes gained or lost due to tactical voting in each district of Budapest

Budapest kerületei	MSZP/DK	LMP	Jobbik	Momentum	MKGP	Együtt
Budapest I. kerület	-100%	229.61%	-35.7%	-100%	-7.94%	-100%
Budapest II. kerület	50.44%	-32.6%	-30.29%	-63.78%	-26.38%	-100%
Budapest III. kerület	60.09%	-74.52%	-22.92%	-36.53%	-78.09%	-100%
Budapest IV. kerület	41.54%	-43.81%	-30.4%	-58.57%	-31.98%	-100%
Budapest IX. kerület	-18.44%	90.47%	-9.22%	-79.63%	-19.78%	42.38%
Budapest V. kerület	-100%	250.09%	-29.04%	-100%	1.94%	-100%
Budapest VI. kerület	77.55%	-47.18%	-32.66%	-100%	-23.11%	-100%
Budapest VII. kerület	79.04%	-45.56%	-34.08%	-100%	-20.59%	-100%
Budapest VIII. kerület	17.6%	3.57%	-6.28%	-66.88%	-31.86%	-1.15%
Budapest X. kerület	42.96%	-30.88%	-30.6%	-48.3%	-100%	-100%
Budapest XI. kerület	58.9%	-53.86%	-33.91%	-58.43%	-33.3%	-100%
Budapest XII. kerület	68.78%	-52.01%	-24.15%	-65.41%	-7.87%	-100%
Budapest XIII. kerület	65.06%	-61.46%	-41.75%	-61.88%	-23.86%	-100%
Budapest XIV. kerület	25.15%	-50.73%	-25.56%	-57.87%	-26.17%	328.29%
Budapest XIX. kerület	56.36%	-42.87%	-32.1%	-54.55%	-100%	-100%
Budapest XV. kerület	54.61%	-50.84%	-36.78%	-54.66%	-32.44%	-100%
Budapest XVI. kerület	-100%	-39.63%	-30.32%	-100%	-100%	1,315.22%
Budapest XVII. kerület	50.53%	-48.77%	-21.5%	-42.75%	-100%	-100%
Budapest XVIII. kerület	69.87%	-100%	-30.16%	-100%	-100%	-100%
Budapest XX. kerület	56.42%	-55.03%	-22.91%	-48.94%	-100%	-100%
Budapest XXI. kerület	-100%	-100%	-20.2%	-100%	-3.12%	899.74%
Budapest XXII. kerület	58.19%	-59.93%	-22.35%	-53.38%	-27.31%	-100%
Budapest XXIII. kerület	-100%	-100%	-1.96%	-100%	29.8%	942.57%
Grand total	33.91%	-36.57%	-26.99%	-66.49%	-45.27%	192.98%

Table 7: Vote gains and losses in Budapest districts

6 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 (generalized linear model) 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 file⁸ 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)
- Town size

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	One or two cities combined with 30-40 villages	54
5	One or two smaller towns with 50-80 villages	
6	One smaller town with 100+ villages	

Table 8: OEVK profiles based on different town composition

Town size ID	Number of voters	Number of towns ⁹	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

Table 9: List of town sizes IDs

⁸ See 106\capstone_adat\project_summary_markdown.md

⁹ The 23 districts of Budapest are counted as separate towns in this categorization.

The number of voters in both the OEVK profiles and town sizes was 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 because since 2010, all of these parties performed like this in general elections, so their voter base has not changed much. It was clear 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).

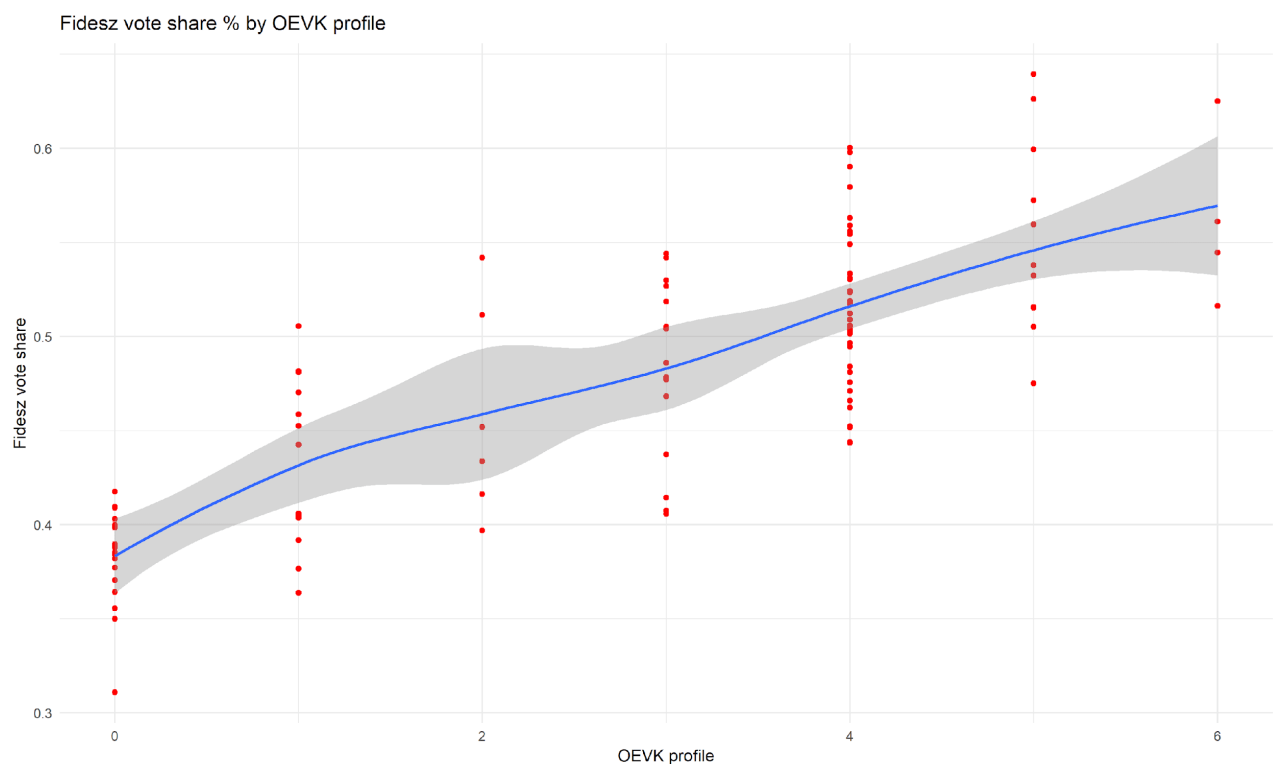


Figure 5: Fidesz vote share by OEVK profile in each OEVK

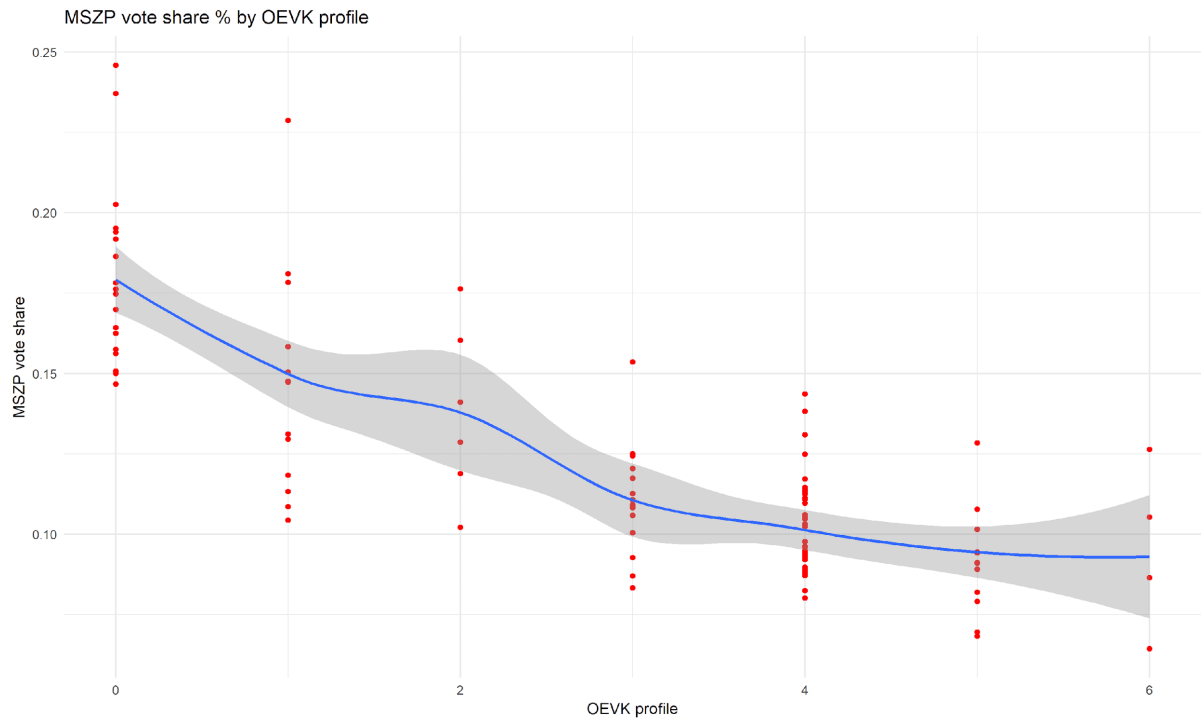


Figure 6: MSZP vote share by OEVK profile in each OEVK

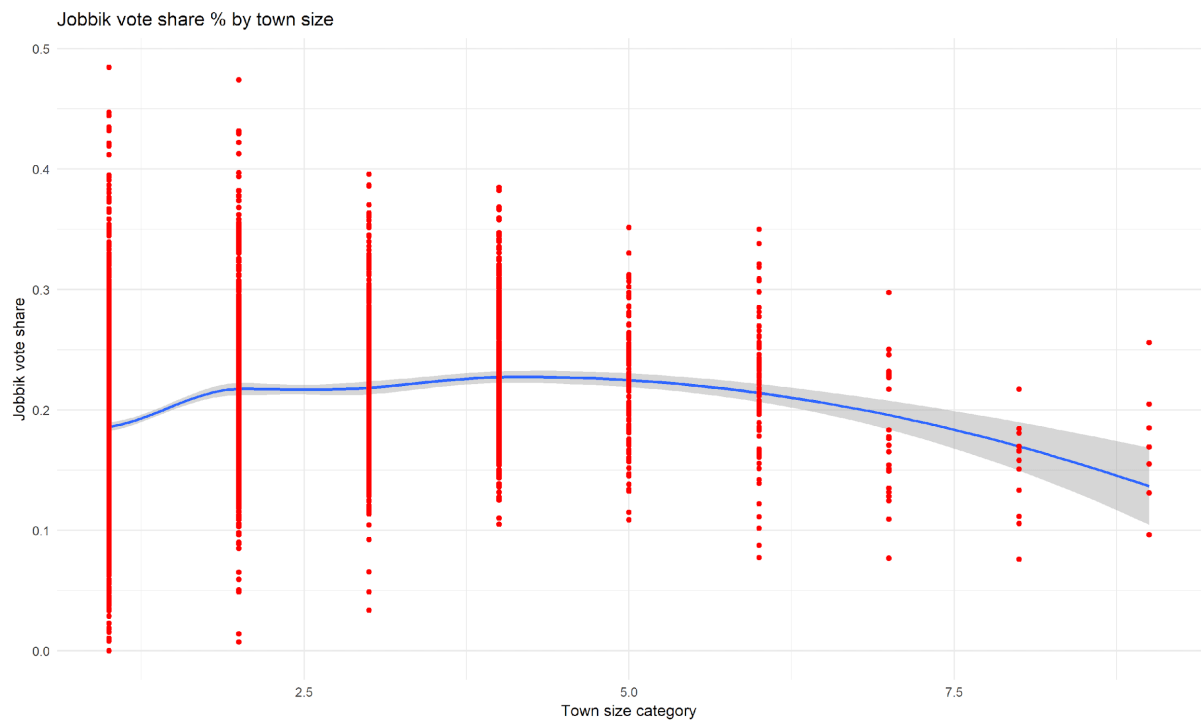


Figure 7: Jobbik vote share by town size category in each town

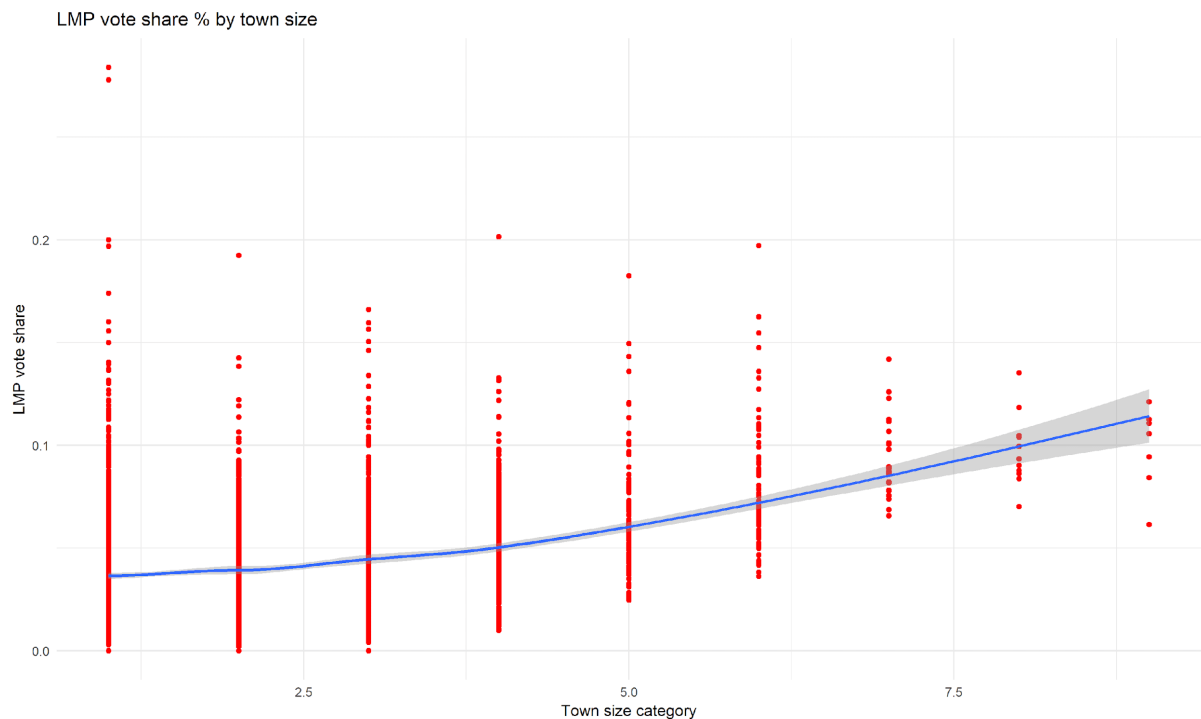


Figure 8: LMP vote share by town size category in each town

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. Their strongest performance was in Jobbik target seats, such as Heves and Borsod counties



- 26



- MSZP vote collapsed in western part of Hungary such as Vas, Veszprém, and Győr counties, meaning that these areas had become solid Fidesz districts. Also, MSZP-P 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.

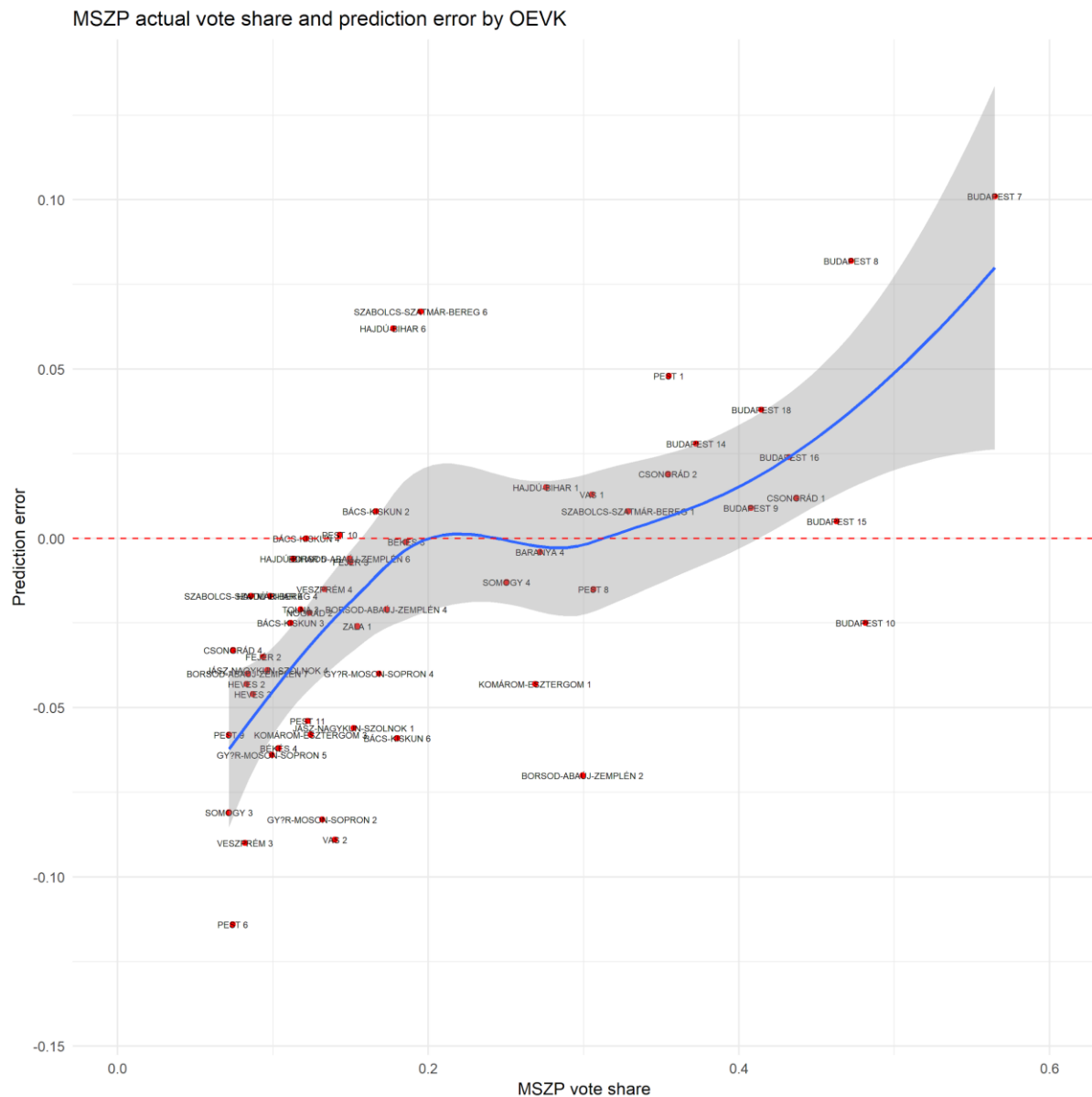


Figure 11: Deviation of MSZP candidate's vote share from pre-election prediction

6.4 GLM prediction using KSH data on town level (Part 4 in Markdown)

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 + 22 predictors such as gender ratio, average income per person, distance from county capital and Budapest, and so on.

Results of Budapest with more than 1 million voters counted as one data point versus the 1400+ micro-towns with less than 500 voters counted as 1400+ data points. Since the weighing of these data points was left out in the data cleansing, the additional predictors told very little about why parties got more votes, or fewer votes than expected. Because of that, none of the 22 predictors turned out to be significant, and they did not reduce the 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 the following counties:

- Baranya (2)
- Borsod (5)
- Somogy (15)
- Veszprém (19)
- Zala (20)

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 (20), but the picture was mixed in Borsod (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.

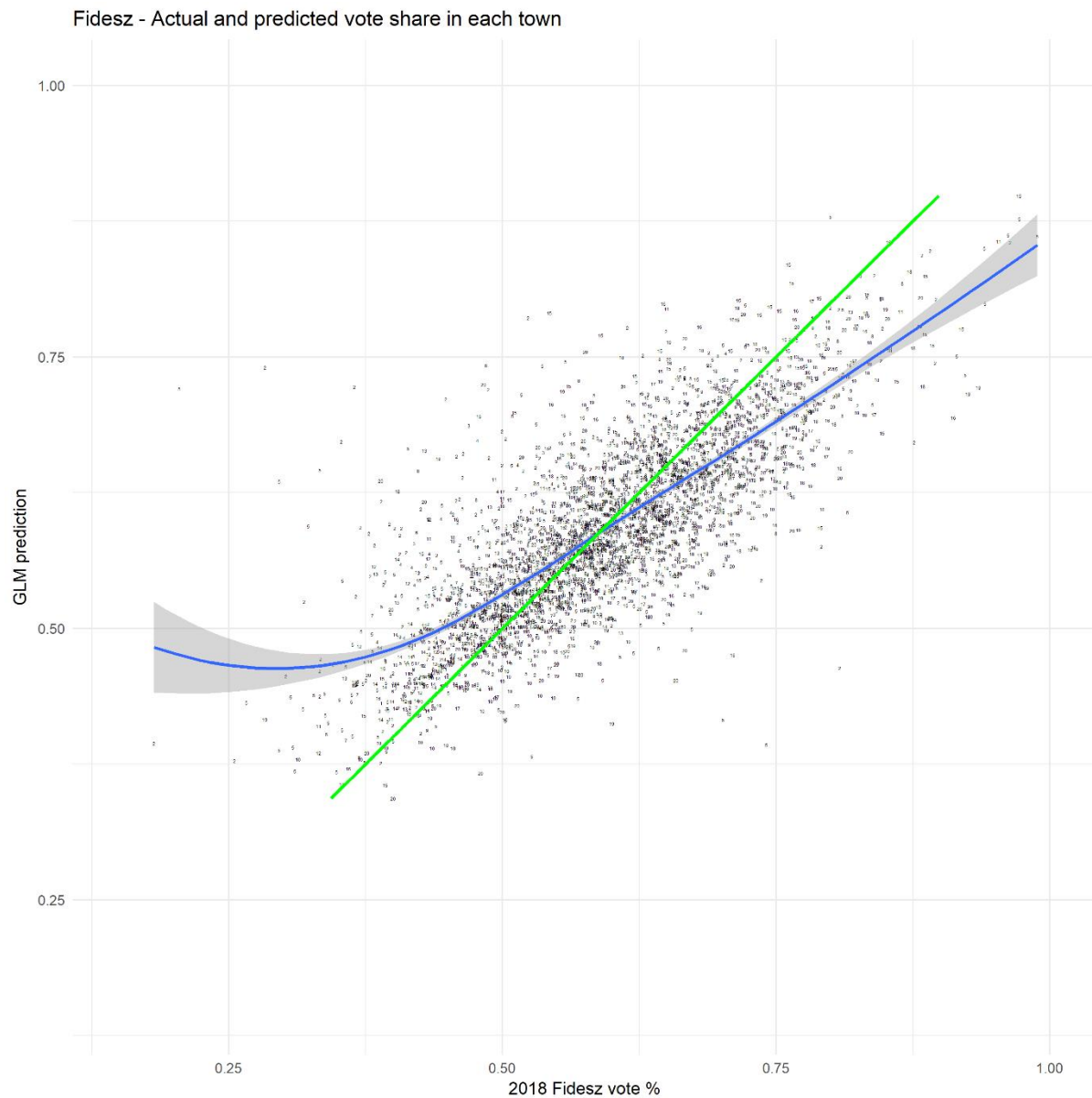


Figure 12: GLM prediction of Fidesz candidate's vote share in each town

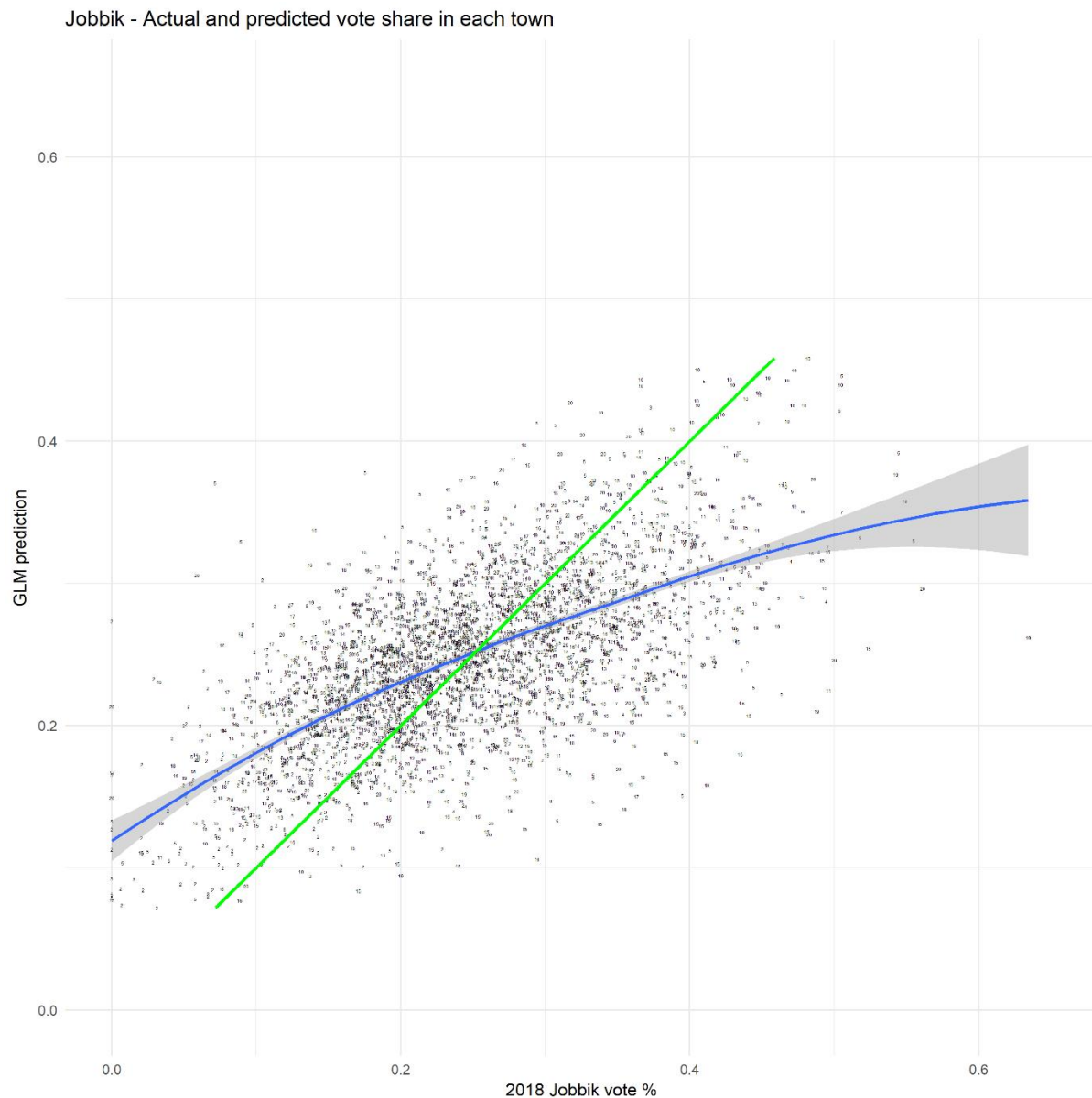


Figure 13: GLM prediction of Jobbik candidate's vote share in each town

6.5 GLM prediction using town size and OEVK profiles on OEVK level (Part 5 in Markdown)

Using GLM on OEVK level was much easier since there are only 106 data points, and there is a 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, 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 of Fidesz and Jobbik, where tactical voting did not influence their final results that heavily, the following 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, and so on. What is interesting 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. 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 partly due to tactical voting.
- 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).

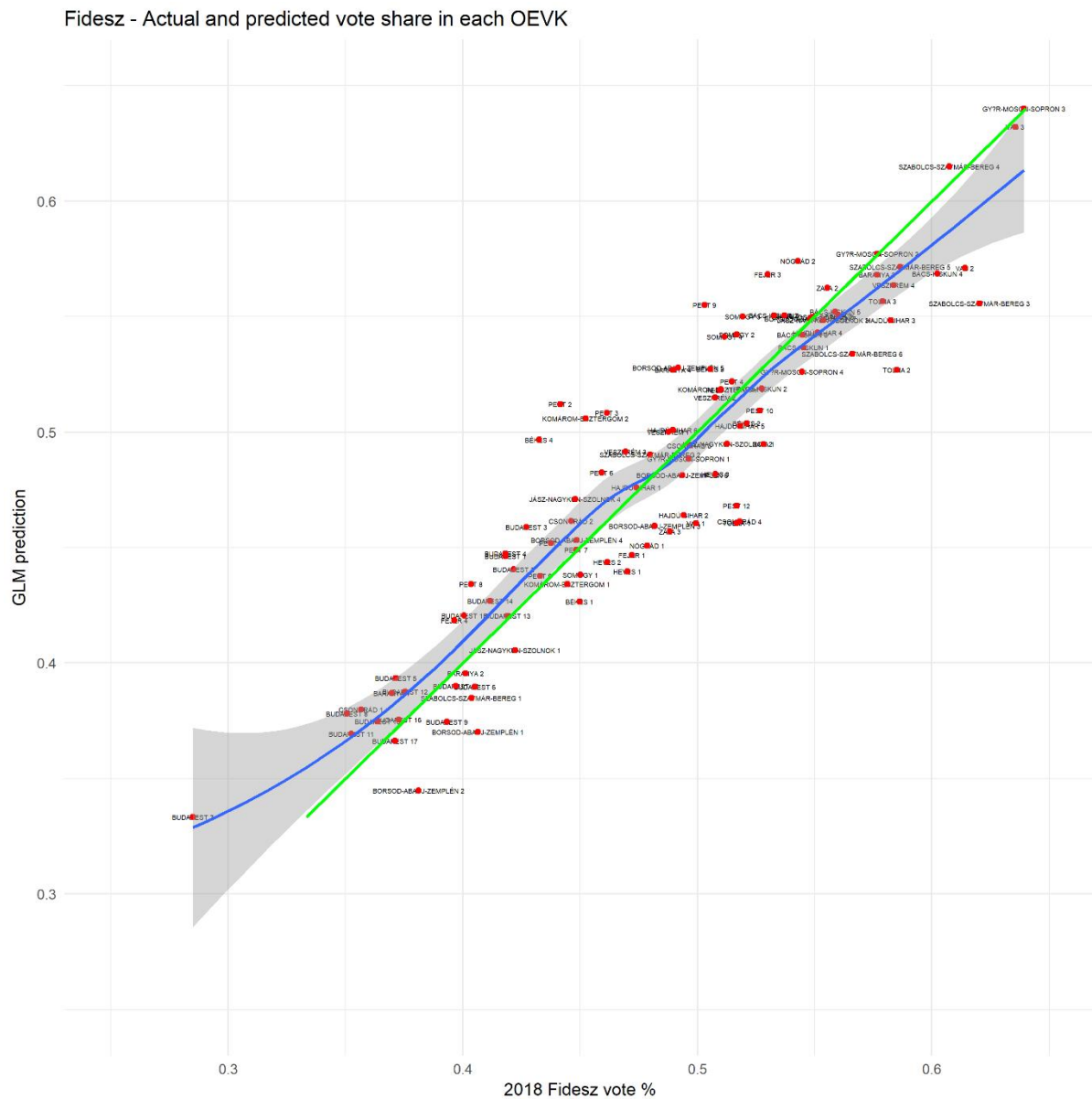


Figure 14: GLM prediction of Fidesz candidate's vote share in each OEVK

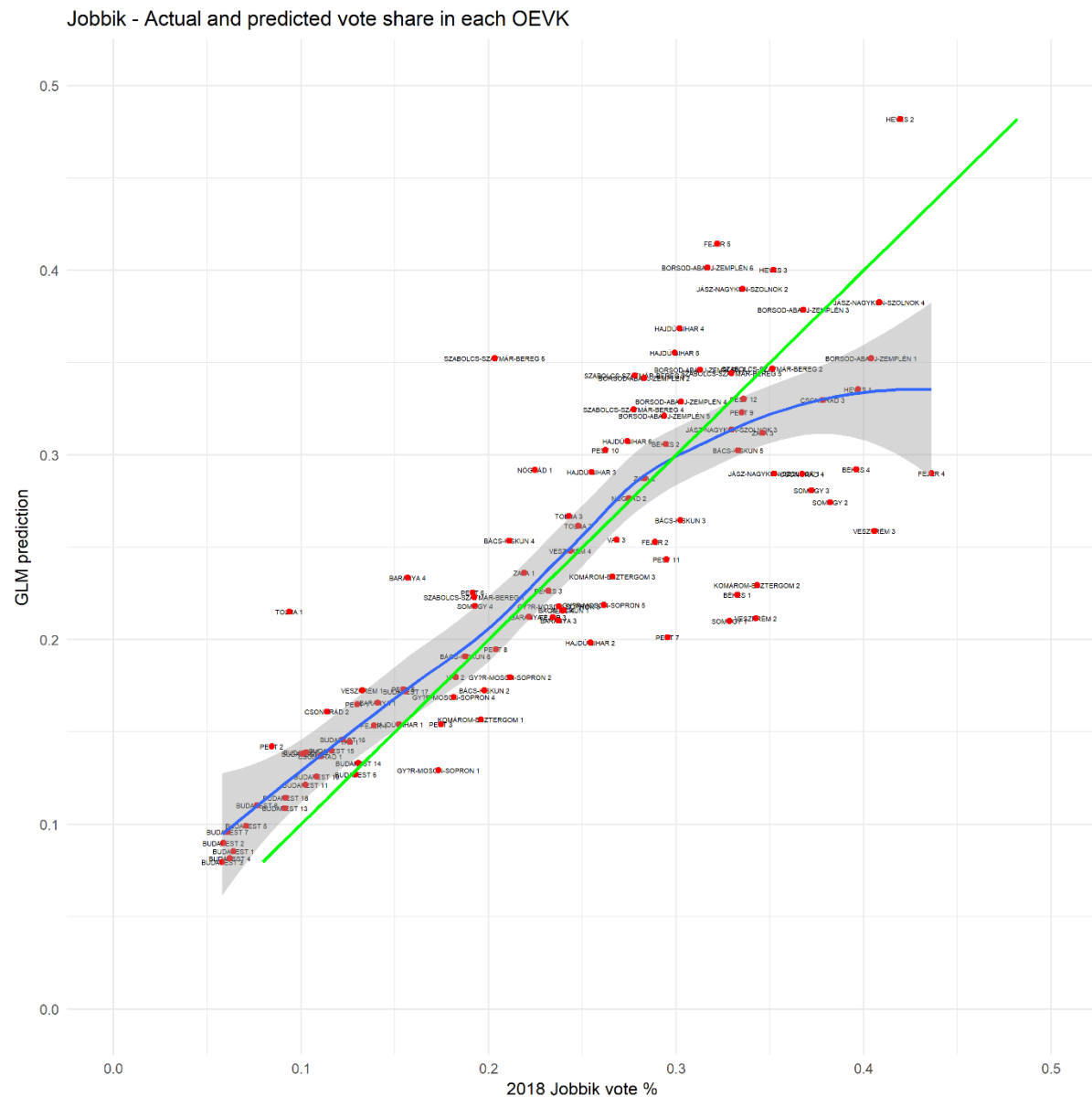


Figure 15: GLM prediction of Jobbik candidate's vote share in each OEVK

6.6 UNS prediction comparing 2014 to 2018 on OEVK level (Part 6 in Markdown)

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. The idea with this prediction was to look at only the UNS changes using the formula explained in *Section 3.2*.

The candidate vote share was calculated using 2014 results x 2018 UNS multiplier.

The following input numbers were used from 2014 and 2018:

Party	2014 vote share	2018 vote share	2018 UNS multiplier
Fidesz	44.13%	47.97%	1.086
Jobbik	20.34%	23.23%	1.142
MSZP/DK	26.86%	17.95%	0.668
LMP	4.98%	5.69%	1.142

Table 10: Input values and UNS multiplier values for each party

The results were similar to what the GLM prediction provided, but there were smaller differences:

- Fidesz performed better in rural OEVKs (with OEVK profile code 3 and above), indicating that they focused more on these areas. Unlike in the GLM prediction, their only strong non-rural area was in Miskolc (Borsod 1 and 2 districts).
- Jobbik performed well thanks to additional tactical votes mostly in cities, but they failed to get extra voters in rural target seats, such as in Jász-Nagykun-Szolnok or Borsod counties.
- UNS does not tell much about LMP or MSZP-P, as their numbers are heavily skewed because of tactical voting. In some areas, both parties over-/underperform by up to 30% of the total vote as they received support from the other party through formal or informal withdrawals.



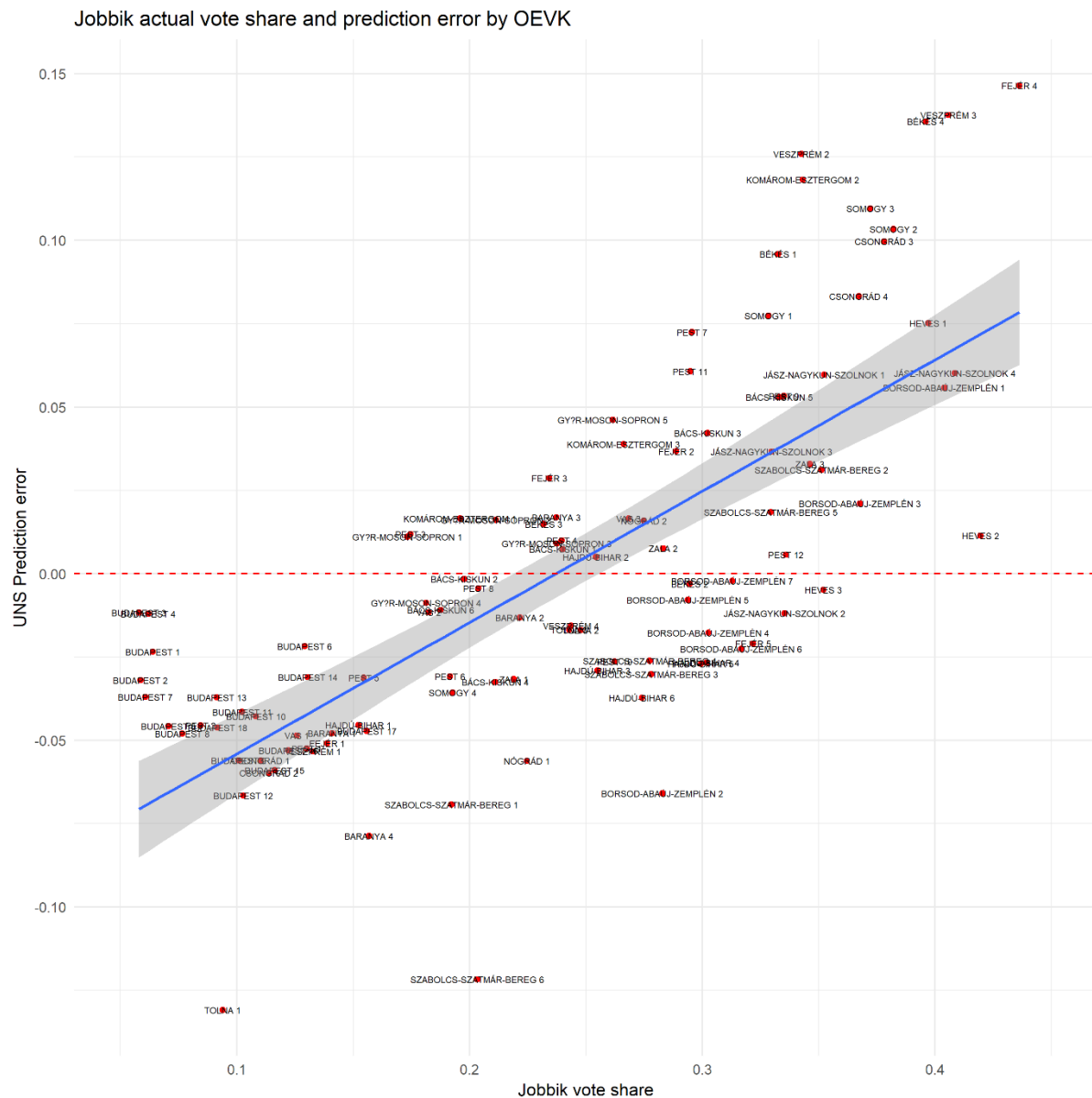


Figure 17: UNS prediction error of Fidesz candidate's vote share in each OEVK

7 Conclusion

As demonstrated in all three types of predictions (pre-election, GLM, and UNS), Fidesz surprisingly took almost all seats outside of Budapest, and the expected breakthrough of Jobbik did not happen. It was also surprising how left-wing parties struggled to gain ground outside of major cities, which foreshadows a situation where there is no credible party against Fidesz.

In terms of tactical voting, it was great to see that people decided to trust opposition party candidates to try and win seats against Fidesz. However, the extent of tactical voting was not enough to defeat a Fidesz candidate who got over 45% of the votes (or sometimes even 40-42% was enough to win the seat). Without tactical voting, however, the results would have been much worse for the opposition, they would have won 1 out of 106 seats.

Since this research tried to capture both the tactical voting and the prediction aspects of this election, there was not enough room to explore either of them in detail. However, the following areas could still be explored to understand more about why Fidesz won this election:

- Analyzing demographics data by city type and normalize results to get a better performing GLM prediction.
- Creating an MRP (Multilevel Regression with Poststratification) analysis using voter demographics and other data to calculate expected vote share of parties.
- Research potential electoral fraud committed in rural areas. There were a handful of cases where it was clear that Fidesz tried to increase the number of voters in key districts, such as in Borsod and Szabolcs county, close to the Ukrainian border. These anomalies are likely to be visible in the data.

For the upcoming 2019 Hungarian local election, all these extra research areas will be explored to provide even better predictions and inform voters about the list of potential opposition candidates. Since the tactical voting initiative had quite an impact on the final results, especially in Budapest, voters may take an informed decision in the upcoming election as well if they know whom to vote for.