Local Candidate Dataset[*]

Jon H. Fiva[†]   Rune J. Sørensen[‡]   Reidar Vøllo[§]

January 10, 2024

Abstract

We introduce a dataset at the candidate level that spans local (municipal and county) elections in Norway from 1971 to 2023. Using information from publicly available electoral lists, we create variables that encompass candidates' background characteristics such as age, gender, occupation, and place of residence. Additionally, we include variables capturing electoral outcomes, including individual-level win margins and information about leadership positions. The information contained in the dataset enables researchers to trace the career progression of candidates from local to national political offices. This note provides a comprehensive description of each variable incorporated in the dataset.

When using the data, please cite as follows:

Fiva, Jon H., Rune J. Sørensen and Reidar Vøllo (2024): Local Candidate Dataset. Available at `www.jon.fiva.no/data.htm`.

[†]BI Norwegian Business School. E-mail: jon.h.fiva@bi.no
[‡]BI Norwegian Business School. E-mail: rune.sorensen@bi.no
[§]BI Norwegian Business School. E-mail: reidar@voello.com

1 Introduction

Norway is a unitary state with three governmental tiers. Elections for the two sub-central tiers, i.e., municipalities (local governments; *kommuner*) and counties (regional governments; *fylkeskommuner*), are held simultaneously every fourth year in September. Our data cover candidates participating in these elections during the 1971–2023 period.

The number of municipalities and counties have remained relatively stable over our sample period, as illustrated in Figure 1. However, before the 2019 local elections the number of municipalities was reduced from 428 to 356. Similarly, the number of counties was reduced from 18 to 10.[1] From January 1, 2024, the number of county governments increased again to 14 (excluding Oslo) and Ålesund was split into two new municipalities (Ålesund and Haram).[2] The dataset uses official municipality (knr) and county (cnr) codes. With consolidations or splits, the affected municipalities and counties receive new codes. This also happens to some unaffected municipalities.[3]

Figure 1: Number of municipalities and counties by election year

---
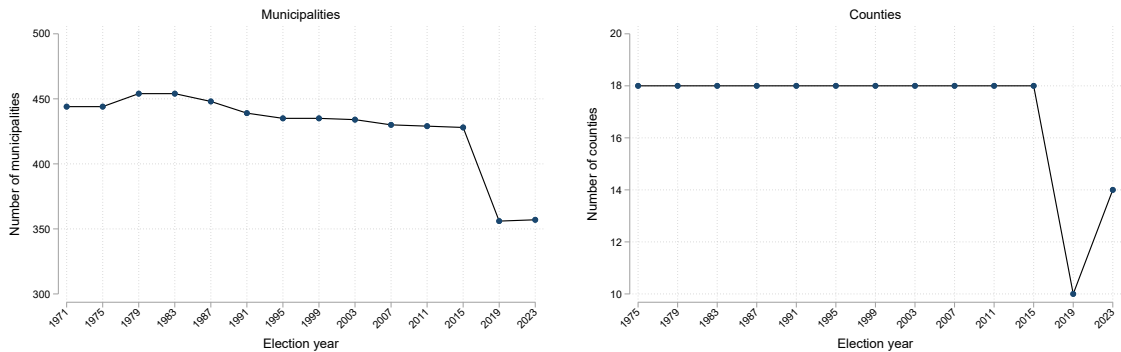
[1]The capital, Oslo, has no regional government. The local government is responsible for both local and regional public services.

[2]"Viken" was split into three new counties (Østfold, Akershus, and Buskerud), "Vestfold and Telemark" was divided into two new counties (Vestfold and Telemark), and "Troms and Finnmark" was divided into two new counties (Troms and Finnmark). Until 2020, the municipalities Lunner was part of Oppland county and Jevnaker was part of Buskerud. These municipalities became part of Viken from January 1, 2020 and part of Akershus from 2024. Akershus also includes the preexisting municipalities of Røyken and Hurum as consequence of these municipalities being merged with Asker in 2020.

[3]For instance, Bærum municipality's code changed from 219 (1971-2015) to 3034 in the 2019 election, then to 3201 in the 2023 election. This is due to Bærum being affected by changes at the county level. It was first located in Akershus county ($cnr = 2$) in the election years 1971-2015, then to the Viken county ($cnr = 30$) in 2019, and subsequently to a modified Akershus county ($cnr = 32$).

The dataset includes 219,993 unique candidates, totaling 413,152 candidate-years. A majority of 195,700 candidates contend for municipal council elections. There are 51,562 unique individuals who have run for county council elections.[4] The data coverage is most complete for elections in the 2000s, as illustrated in Figure 2.

Figure 2: Data coverage over time by political office
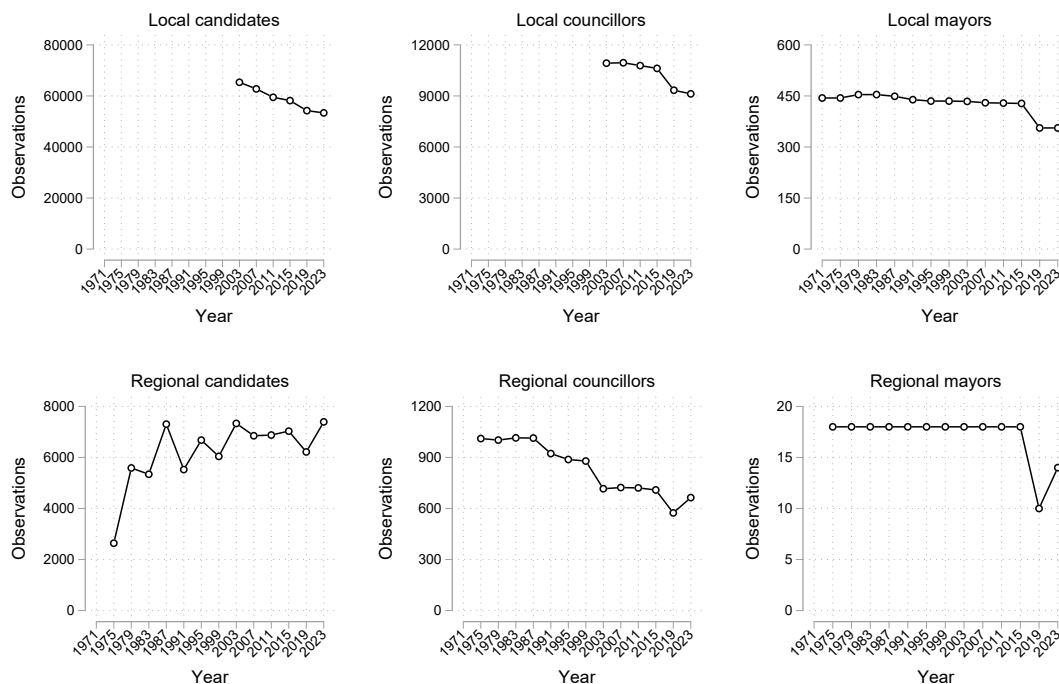
Note: This figure displays the number of observations for candidates, councillors, and mayors, over time. The top-panel provides data for local (municipal) elections. The bottom panel provides data for regional (county) elections. Local candidate data from Oslo is included in the data set for the 1975–1999 period, but is omitted from the figure.

In Figure 3, we display information on number of runs and wins for candidates running in the years 2003-2021. In this period, candidates can run and/or win a seat in up to six elections. Approximately 30% of municipal council election candidates participate in a single election, while around 50% run in either one or two elections. About 17% of municipal council candidates secure a council membership in at least one election. In county council elections, around 45% of candidates compete in only one election, with 66% running in up to two. The average success rate for attaining a seat in county council

---

[4]The total number of county and municipal council candidates appears smaller than the actual total number of candidates because the dataset includes only one personal identifier (*pid*) for each candidate that runs for the municipal and county council elections in the same election year.

elections is 10%, reflecting the more competitive nature of these positions.

Figure 3: Political Candidacy: Number of Runs and Successful Elections
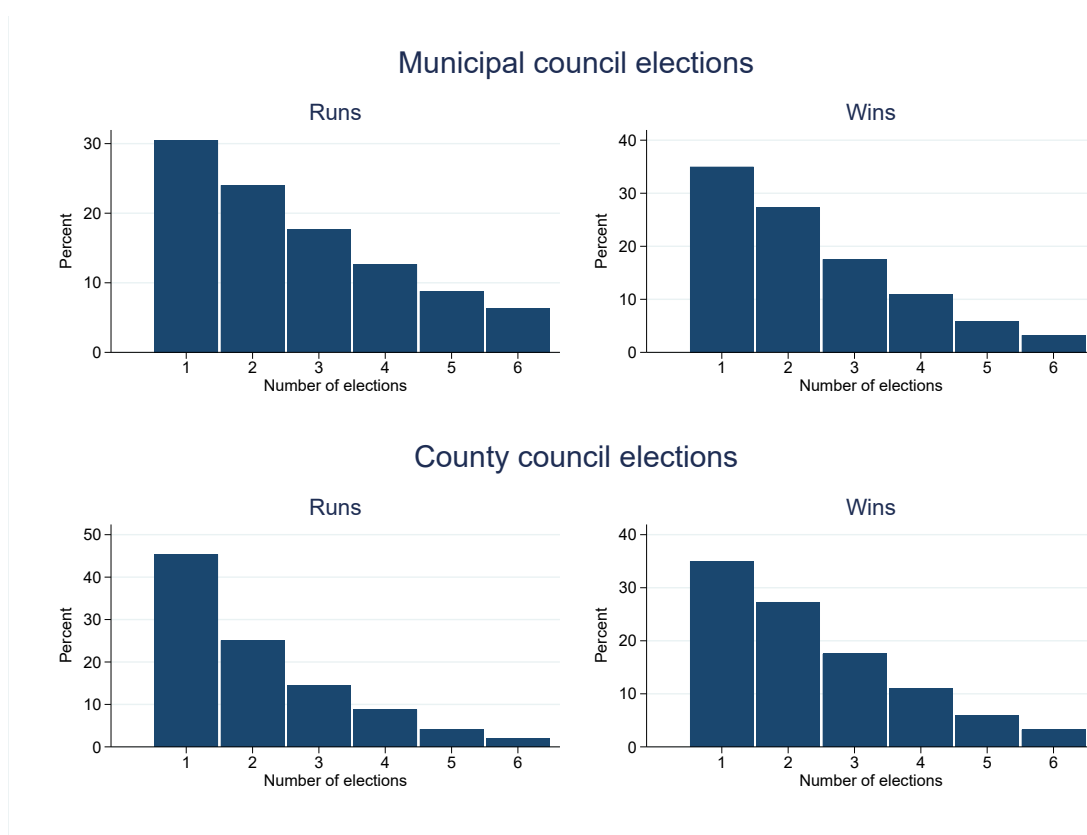
Please note:

1. The dataset includes information of candidates only in the election years where they are running and/or elected. A balanced panel dataset can be created by expanding the dataset with candidate-years that include years when candidates are not running or elected.

2. The dataset includes complete information on both candidates and representatives in the municipal and county council elections starting in the election year 2003 (see description below).

3. The dataset provides a detailed coding of candidates' party affiliations, using the official (list) names of parties and joint lists. Any changes in official party names or

joint lists necessitate additional coding for analyzing shifts in individual-level party affiliations over time or across government levels (municipal, county, national).

4. The dataset contains data on the occupations of county council candidates, as sourced from election lists dating from 2003 onwards. The dataset does not include occupational information for municipal council candidates. Note that many of the election lists at the local and regional levels lack occupational details for the candidates.

2 Municipal council elections

Data for municipal council elections are primarily taken from Fiva and Røhr (2018) and cover the 2003–2015 period (245,791 candidate-year observations). In addition, we collect data for the 2019 and 2023 elections (58,182 and 54,244 observations), data for Oslo for the 1975–1999 period (2,327 candidate-year observations), all mayors covering the 1971–1999 period (3,553 candidate-year observations), and all deputy mayors covering the 1987–1999 period (1,757 observations). Cirone et al. (2021) study political careers by merging these data with the Fiva and Smith (2017) data set.[5]

At the local level, the electoral system is open-list ("flexible-list") proportional representation (PR) where party seats are (since 2003) distributed using the Modified Sainte-Laguë method. Citizens affect the election outcome by voting for a party list and by casting personal votes for particular candidates. Parties can opt to give certain candidates an increased share of the poll (25% of the total number of votes received by the party). Candidates with such a pre-advantage are listed at the top of the ballot in boldface (for further details, see Fiva and Røhr, 2018).

Candidates' pre-advantage status and the number of personal votes yield the personal poll that forms the basis of the within-party distribution of seats. More specifically, the

---

[5]The Local Candidate Dataset provides only basic information about local candidates who are also running for national office. The Fiva-Smith dataset, which offers more comprehensive details about individuals participating in national elections, can be merged with the current dataset using the *pid_nat* variable.

poll for candidate i running for party l is given by:

$$Poll_{il} = \begin{cases} PersonalVotes_i & \text{if } i \text{ has no pre-advantage} \\ \\ PersonalVotes_i + 0.25 \cdot PartyVotes_l & \text{if } i \text{ has a pre-advantage for list } l \end{cases} \tag{1}$$

The forcing variable in the regression discontinuity (RD) design of Fiva and Røhr (2018) is included in our data set and is given by:

$$WinMargin_{il} = \begin{cases} \frac{Poll_{il} - Poll_l^{S_l+1}}{PartyVotes_l} & \text{if } R_{il} \leqslant S_l \text{ [elected candidates]} \\ \\ \frac{Poll_{il} - Poll_l^{S_l}}{PartyVotes_l} & \text{if } R_{il} > S_l \text{ [non-elected candidates]} \end{cases} \tag{2}$$

where R_{il} denotes candidates' (ex-post) list rank and S_l the number of seats party list l won. Fiva and Røhr (2018) document that candidates who barely win a seat on the local council have around 9 percentage points (43%) greater probability of getting elected in the next election, compared to a candidate who just misses out on a seat on the same party list. In a regression discontinuity design using the same forcing variable, Nowacki (2024) finds that the incumbency advantage is smaller for women than for men.

Other studies have employed $WinMargin_{il}$ to create instrumental variables, aiming at estimating the impact on policy outcomes. For instance, Sørensen (2023) utilized closely contested elections to identify as-good-as random variation in the proportion of council members with higher education. The individual-level election results were transformed into a municipal-level instrumental variable, which was then leveraged to assess its influence on local government efficiency. In a similar vein, Geys et al. (2024) employed close elections to formulate an instrumental variable for the share of public employees in municipal councils. This was then used to evaluate its effect on local government expenditures and staff remunerations.

On the ballot, political parties typically provide candidates' full name, year of birth, and place of residence (as well as rank-position and whether the candidates are given a

pre-advantage). Figure 4 is an example ballot from *Oslo Arbeiderparti* from 2019. The first ten candidates on the ballot are in boldface and hence have a pre-advantage.

Figure 4: Example of Ballot from Oslo Labor Party 2019
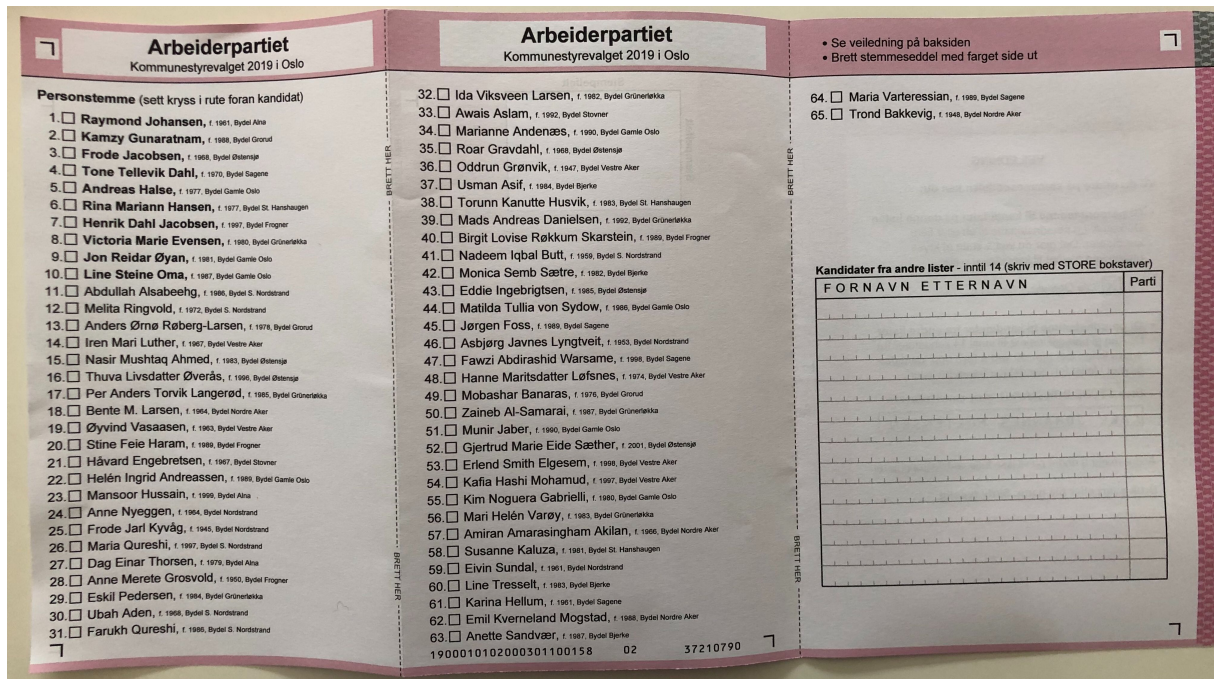
Based on candidates' year of birth and first name we construct variables for candidates' age and gender. Figure 5 displays the age of candidates running for municipal office by election year and gender. The figure provides some evidence that the gender gap in local politics might be closing for the younger cohorts (Fiva and King, 2024).

Figure 5: Candidates' age in municipal elections by year and gender

3   County council elections

Data for regional elections (county elections) cover the universe of elected candidates 1975–2023, as well as all non-elected candidates in the 2003–2023 period, and about half of non-elected candidates in the 1975–1999 period (80,819 candidate-year observations).[6] These data were originally collected for the Fiva et al. (2021) study.

Elections to the regional office are, like the local elections, decided by the Modified Sainte-Laguë method. Voters have, however, less scope to affect the selection of candidates. Up until 1999, a closed-list electoral system was being used.[7] Since 2003, voters have the opportunity to cast personal votes[8], but the switch to a *flexible list system* appears to have had negligible effect on candidate selection (Christensen et al., 2004). For the 2007–2015 period, we lack information about personal votes and we consequently construct the variable *elected_reg* under the assumption that personal votes did not affect candidate selection. Bergh et al. (2016) report that only eight councillors (1.1%) were elected due to personal votes in 2007.

Like for the municipal elections, we can construct variables capturing candidates age and gender using information provided on the ballot. Figure 6 illustrates that women are still underrepresented in Norwegian politics, in particular for executive offices, such as the mayoral position (*fylkesordfører*) (see also Cirone et al., 2023).

---

[6]Direct elections to the regional level were first held in 1975.

[7]Candidates were assigned seats according to the ranking assigned them by their respective parties, but with one modification: 1/6 of the seats were reserved for candidates from municipalities with no seats once the other 5/6 of the seats had been assigned.

[8]To jump to the top of the list, a candidate needs to receive individual preference votes equivalent to 8% of their party's votes.

Figure 6: Women's representation in regional councils 1975–2023
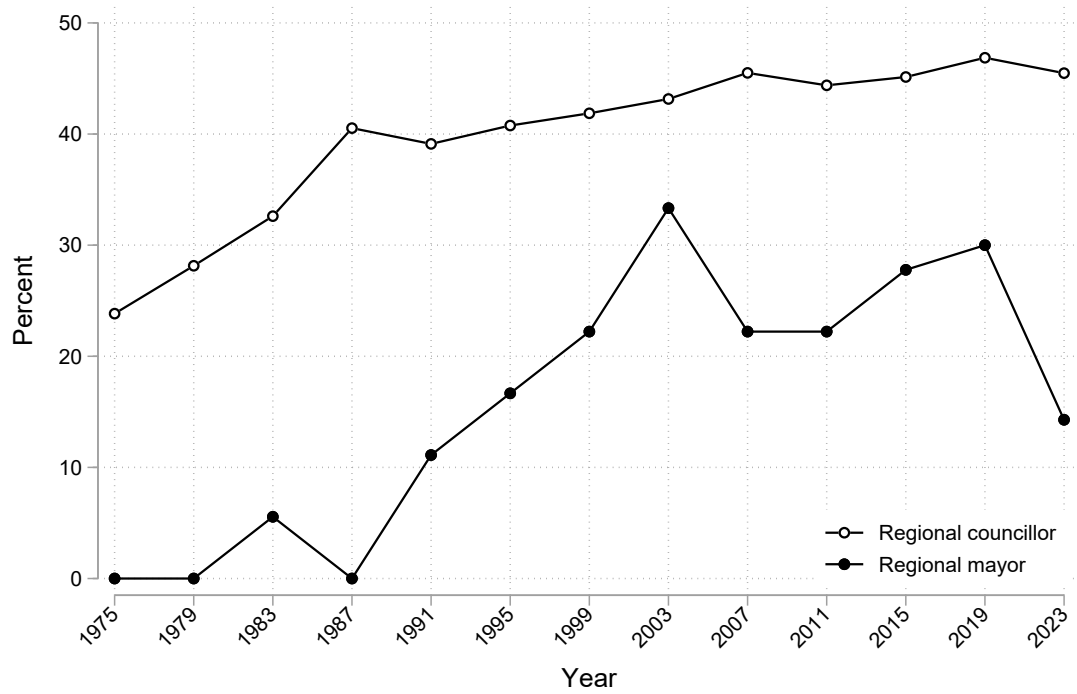
4 Code book

Candidate variables

- year: Year identifier
- pid_lcd: Personal identifier
- pid_nat: Personal identifier from Fiva and Smith (2017) data set (*pid*).[9]
- candidatename_std: Candidate's name cleaned and standardized
- candidatename_nat: Candidate's name from Fiva and Smith (2017) data set
- firstname: Candidate's first name
- middlename: Candidate's middle name
- lastname: Candidate's last name
- birthyear: Candidate's year of birth
- age: Candidate's age in election year
- female: Indicator variable for gender
- running_loc: Indicator variable for running for municipal election this year
- running_reg: Indicator variable for running for county election this year
- candrun_loc: Number of times candidate is running for municipal office in the dataset
- candwin_loc: Number of times candidate is elected for municipal office in the dataset
- candrun_reg: Number of times candidate is running for county office in the dataset
- candwin_reg: Number of times candidate is elected for county office in the dataset
- first_year: First year candidate is running for local office in the dataset
- last_year: Last year candidate is running for local office in the dataset
- first_year_loc: First year candidate is running for municipal office in the dataset
- last_year_loc: Last year candidate is running for municipal office in the dataset
- first_year_reg: First year candidate is running for county office in the dataset
- last_year_reg: Last year candidate is running for county office in the dataset
- candrun_nat: Number of times candidate is running for national office
- candrun_nat_firstyear: First year candidate is running for national office
- candrun_nat_lastyear: Last year candidate is running for national office
- candwin_nat: Number of terms serving as Member of Parliament (MP)
- candwin_nat_firstyear: First year candidate is winning national office
- candwin_nat_lastyear: Last year candidate is winning national office

---

[9]All variables below capturing information about national elections are from Fiva and Smith (2017). This dataset is available at `www.jon.fiva.no/data.htm`.

- canddeputy_nat_firstyear: First year candidate is elected as deputy MP (*vararepresentant*)

- canddeputy_nat_lastyear: Last year candidate is elected as deputy MP (*vararepresentant*)

- days_nat: Tenure in *Storting* measured in days (as of March 2019)

Variables specific to municipal council elections

- knr_loc: Municipality identifier based on the municipal election

- kname_loc: Municipality name

- cnr_loc: County identifier based on the municipal election

- cname_loc: County name

- party_loc: Candidate's party acronym in the municipal election

- partyname_loc: Candidate's party name in the municipal election

- rank_loc: Candidate's ex ante rank

- rankresult_loc: Candidate's ex post rank (after personal votes has been taken into account)

- candidatename_orig_loc: Candidate's name unedited

- elected_loc: Indicator variable for representation in municipal council

- preadvantage_loc: Indicator variable for being pre-advantaged (*forhåndskummulert*)

- personalvotes_own_loc: Personal votes from voters casting a party vote for candidate's party

- personalvotes_other_loc: Personal votes from voters casting a party vote for other party

- personalvotes_total_loc: Candidate's total personal votes

- preadvnvotes_loc: Additional "votes" from being pre-advantaged (25% of the total number of votes received by the party)

- mayor_loc: Indicator variable for becoming municipal mayor

- depmayor_loc: Indicator variable for becoming municipal deputy mayor

- executiveboard_loc: Indicator variable for becoming elected to municipal executive board

- party_preadvantage_loc: Number of candidates on list with a pre-advantage

- party_votes_loc: Number of party votes (*partistemmer*) for candidate's party

- totalvotes_loc: Total number of party votes (*partistemmer*) across all parties

- party_seats_loc: Number of seats for candidate's party

- sizeofcouncil_loc: Total number of seats in the municipal council

- directmayor_loc: Indicator variable for municipalities with direct mayoral elections

- pop_loc: Number of inhabitants in municipality
- poll_loc: Candidate poll in municipal election (see equation 1)
- electionmargin_loc: Distance (in ex ante rank position) to borderline elected/defeated
- pollborderlineelected_loc: Party list borderline elected poll (see equation 2)
- pollborderlinedefeated_loc: Party list borderline defeated poll (see equation 2)
- winmargin_votes_loc: Win margin in raw votes
- winmargin_loc: Standardized win margin (see equation 2)

Variables specific to county council elections

- knr_reg: Municipality identifier based on the county election
- cnr_reg: Election district/county identifier based on the county election
- cname_reg: Election distric/County name
- party_reg: Candidate's party acronym in the county election
- partyname_reg: Candidate's party name in the county election
- rank_reg: Candidate's ex ante rank
- rankresult_reg: Candidate's ex post rank
- candidatename_orig_reg: Candidate's name unedited
- kname_orig_reg: Candidate's hometown as listed on county election ballot
- occupation_reg: Candidate occupation as listed on county election ballot
- elected_reg: Indicator variable for representation in county council
- deputy_candidate_reg: Indicator variable for becoming deputy councilor (*only 2003*)
- mayor_reg: Indicator variable for being elected county mayor
- leveling_seat_reg: Indicator variable for geographical adjustment seat (*utjevnings-mandat*) (*1975-1999*)
- personalvotes_total_reg: Candidate's total personal votes
- party_seats_reg: Number of seats won in county council for candidate's party
- pop_reg: Number of inhabitants in election district/county

Variables describing issues with data

- only_elected_loc: Indicator variable for municipal election lists containing elected candidates only
- only_elected_reg: Indicator variable for county election lists containing elected candidates only
- inconsistency_candidate_loc: Indicator variable identifying inconsistencies between *win_margin_loc* and *elected_loc*

13

- inconsistency_list_loc: Indicator variable identifying inconsistencies between *win_margin_loc* and *elected_loc* for any candidate at the list where the candidate belongs

- inconsistency_list_reg: Indicator variable identifying any problem with the regional election list where candidate belongs

- missing_votes_candidate_loc: Indicator variable for missing personal votes for candidate

- missing_votes_list_loc: Indicator variable for missing personal votes for any candidate at list where candidate belongs

- missing_votes_municipality_loc: Indicator variable for missing personal votes in municipality-year where candidate runs

- missing_votes_candidate_reg: Indicator variable for missing personal votes for candidate

- missing_votes_list_reg: Indicator variable for missing personal votes for any candidate at list where candidate belongs

- missing_votes_county_reg: Indicator variable for missing personal votes in county-year where candidate runs

References

Bergh, J., Christensen, D. A., Folkestad, B., Karlsen, R., and Saglie, J. (2016). Personvalg ved stortingsvalg? *Norsk statsvitenskapelig tidsskrift*, 32(3):217–242.

Christensen, D. A., Midtbø, T., Ringkjøb, H.-E., Svåsand, L., and Aars, J. (2004). Ny personvalgordning og hva så? en analyse av kommune og fylkestingsvalget i 2003. Rapport, Rokkansenteret.

Cirone, A., Cox, G. W., and Fiva, J. H. (2021). Seniority-based nominations and political careers. *American Political Science Review*, 115(1):234–251.

Cirone, A., Cox, G. W., Fiva, J. H., Smith, D. M., and Teele, D. (2023). Gender gaps in political seniority systems. *Available at SSRN*.

Fiva, J. H., Geys, B., Heggedal, T.-R., and Sørensen, R. (2021). Political alignment and bureaucratic pay. *Journal of Public Administration Research and Theory*, 31(3):596–615.

Fiva, J. H. and King, M.-E. M. (2024). Child Penalties in Politics. *The Economic Journal*, forthcoming.

Fiva, J. H. and Røhr, H. L. (2018). Climbing the ranks: incumbency effects in party-list systems. *European Economic Review*, 101:142–156.

Fiva, J. H. and Smith, D. M. (2017). Norwegian parliamentary elections, 1906-2013: Representation and turnout across four electoral systems. *West European Politics*, 40(6):1373–1391.

Geys, B., Murdoch, Z., and Sørensen, R. J. (2024). Public employees as elected politicians: assessing direct and indirect substantive effects of passive representation. *The Journal of Politics*, 86(1):000–000.

Nowacki, T. (2024). The gender gap in political careers under proportional representation. *Journal of Politics*.

Sørensen, R. J. (2023). Educated politicians and government efficiency: Evidence from norwegian local government. *Journal of Economic Behavior & Organization*, 210:163–179.