



OPEN

# The Norwegian Parliamentary Debates Dataset

DATA DESCRIPTOR

Jon H. Fiva<sup>1,4</sup>, Oda Nedregård<sup>2,4</sup>✉ & Henning Øien<sup>3,4</sup>

Recent advancements in computing power and machine learning techniques have facilitated the digitization of new corpora, as well as new methods for studying high-dimensional data. This has enabled empirical investigations of fundamental questions in the social sciences that were previously restricted by technical limitations or data availability. In this note, we introduce a new dataset covering debates in the Norwegian Parliament in the 1945–2024 period. This dataset, which covers close to one million speeches, includes information about speeches (full text, date of speech, and chamber), speakers' status (parliamentary president, member of parliament, deputy member of parliament, or cabinet minister), as well as speaker background characteristics (party affiliation, committee membership, district affiliation, rank on electoral lists, gender, and birth year). This dataset will enable extensive research into political representation in a party-centered electoral framework. More broadly, this dataset serves as a vital resource for interdisciplinary research, enabling studies on the evolution of language, rhetoric, and the broader socio-economic factors influencing legislative behavior.

## Background & Summary

Parliamentary debates are a cornerstone of democratic governance, providing a platform for elected representatives to discuss and contest policy matters. In debates, Members of Parliament (MPs) affect the legislative discourse by advancing political perspectives and interests, as well as signaling policy stands to colleagues and electoral stakeholders such as the media, party elites, and constituents. Legislators always engage in discourse before voting on bills<sup>1</sup> but also after, where the plenary is used to justify how they voted<sup>2</sup>. Until recently, large-scale studies of legislative debates have been constrained, largely due to the complexities involved in compiling data into a usable format and the lack of appropriate natural language processing tools<sup>3,4</sup>. Advancements in computing power and machine learning techniques have facilitated the digitization of new corpora, as well as new methods for understanding and measuring key concepts within political economy, such as representation<sup>5,6</sup>, polarization<sup>7,8</sup>, campaigning<sup>9,10</sup>, and media bias<sup>11,12</sup>.

In this note, we introduce a new dataset covering legislative debates in Norway, an archetypal party-centered parliamentary democracy<sup>13</sup>. Spanning from 1945 to 2024, this dataset constitutes an important addition to the literature, enabling expanded research into contexts beyond the candidate-centered frameworks that have dominated scholarly attention. It allows researchers to shed light on parties' objectives and internal organization, which has been referred to as the 'black box' or 'secret garden' of politics<sup>14,15</sup>.

In a series of companion papers, we have used parts of this data to address key questions on political behavior and policymaking in party-centered environments. Using data from the 1981–2021 period, Fiva, Nedregård and Øien (2024)<sup>16</sup> compare legislators from the same political party and policy committee and document substantial differences in speech patterns between MPs with opposing background characteristics (e.g., gender and social background). This suggests that politicians' social ties and group identities matter for policymaking even in party-centered environments – where parties are omnipotent and there are no incentives to cater to voters based on individual characteristics. Using a failed-renomination design, Fiva and Nedregård (2024)<sup>17</sup> study how electoral incentives affect legislator behavior using data from the 1993–2021 period. They find that legislators reduce their efforts after learning that they cannot be reelected. However, there is no evidence that this affects their adherence to the party line. Nedregård (2024)<sup>18</sup> use data from the 1995–2021 period to study legislators' coverage of unemployment issues when their electoral districts are experiencing economic distress. Her results suggest that political parties strategically allocate such speeches to party members with comparatively strong

<sup>1</sup>BI Norwegian Business School, Department of Economics, Oslo, 0484, Norway. <sup>2</sup>Sciences Po, Department of Economics, Paris, 75007, France. <sup>3</sup>Norwegian Institute of Public Health and Department of Health Management and Health Economics, University of Oslo, Oslo, 0316, Norway. <sup>4</sup>These authors contributed equally: Jon H. Fiva, Oda Nedregård, Henning Øien. ✉e-mail: [oda.nedregard@sciencespo.fr](mailto:oda.nedregard@sciencespo.fr)

incentives to cater to party elites at the expense of their localities. By collecting and cleaning data for the entire period after World War II, we hope to open new avenues for understanding how party-centered democracies function.

A significant advantage of using speeches to determine the position of political actors is that speeches are provided regularly (not just prior to elections) and by a range of different political actors<sup>19</sup>. The high dimensionality of speeches allows for substantial flexibility, enabling researchers to capture nuances that would not be captured by traditional measures (e.g., roll-call votes or surveys). By following speeches over time, we can track changes in policy positions of individuals and parties at a high frequency, and observe how these positions change in response to major events. Additionally, speeches are generally less subject to the constraints of agenda-setting and party leadership discipline, offering a richer insight into legislators' policy positions compared to analyses of their voting behavior in the legislature<sup>20</sup>.

This note outlines the data collection process, describes the institutional setting, and provides an overview of the dataset. To validate the dataset, we analyze how key characteristics—specifically gender and party affiliation—affect speech-making. Furthermore, we confirm that the trends observed in the parliamentary speech data correspond with broader societal changes in Norway.

## Methods

**Data collection.** Our dataset covers all debates in the Norwegian Parliament from December 1945 to June 2024. Parliamentary debates from December 1945 to June 1998 are digitized using Optical Character Recognition (OCR) readings of the online historical archives of the Norwegian Parliamentary Proceedings (*Stortingsforhandling*) of the *Norwegian National Library* (<https://www.nb.no/>). Speeches from the remaining period were sourced directly from the Storting's online API (<https://data.stortinget.no/>).

To obtain metadata on speakers, we match speaker names with the 2024 version of the *Norwegian Parliamentary Elections* dataset<sup>21</sup> (available at <https://www.jon.fiva.no/data.htm>). We incorporate information regarding individuals' party affiliation and parliamentary status (MP, deputy MP, and/or cabinet member). For individuals who are promoted to cabinet without being elected to Parliament (not covered by<sup>21</sup>), we consult *Politikerarkivet* (<https://polsys.sikt.no/storting/politikerarkiv>).

The “President” in our dataset refers to the presiding officer (speaker), who is responsible for leading the debate and ensuring that procedural rules are followed. This role is most frequently filled by the elected President of the Parliament, but it can also be occupied by Vice Presidents of the Presidium. All members in the Presidium are MPs. In the dataset, the speeches of presidents are not marked by individual characteristics, as the speeches are purely procedural.

The Norwegian Agency for Shared Services in Education and Research (Sikt) has approved the collection of parliamentary legislative debates. The processing of personal data is deemed lawful based on public interest (*allmennhetens interesse*).

**Institutional Setting.** Norway operates as a unitary constitutional monarchy with a parliamentary system of governance, marked by a high prevalence of minority governments<sup>22</sup>. Over the span of our 80-year study period, Norway experienced minority governments for about 41 years, as detailed in Table 1. A distinctive feature of the Norwegian system is its adherence to negative parliamentarism; while there is no mandatory vote for the confirmation of a new cabinet in the legislature, the legislature retains the power to dismiss the cabinet at any time through a majority vote<sup>23</sup>.

Norwegian national elections have since 1921 used a closed-list proportional representation system with multi-member districts (for more details, see<sup>21</sup>). Since the conclusion of World War II, elections have occurred every fourth year in the fall. An unusual feature of the Norwegian constitution is the lack of any provision for early parliamentary dissolution<sup>24</sup>. Most of the legislative debates since 1945 have been conducted in the plenary chamber *Storting*. In a small fraction of debates involving non-financial legislation, the legislature was divided into two chambers: the *Odelsting* (lower chamber with three-fourths of the MPs) and the *Lagting* (upper chamber with one-fourth of the MPs)<sup>25</sup>. This division operated until 2009, despite the legislature being elected as a single body. Bills required approval from both chambers. In practice, the Lagting rarely dissented and primarily served as a rubber stamp for the Odelsting's decisions.

The closed-list proportional representation system used in Norway gives parties strong tools to discipline their elected officials since a candidate's probability of being elected depends largely on their ballot rank. This gives party organizations a central role in Norwegian politics.

Table 2 lists the seven political parties that have dominated Norwegian politics since 1945. Although some parties have changed names during our sample period, we use the same party acronyms in our dataset for consistency.

The importance of parties is reflected in the parliamentary discourse. According to the Storting's *Rules of Procedure*, the standing committee from which the debate originated proposes the distribution of speaking time between the parties, while the political parties may submit their lists of proposed speakers. In significant regular debates – such as the throne speech (*trontaledebatten*), government declaration (*regjeringserklæringen*) and main budget debate (*finansdebatten*) – the president consults directly with the leaders of the party groups to determine the conduct of the debate. Based on this information, the president creates a list outlining the order of speakers and the allocation of speaking time<sup>26</sup>.

While the standing committees meet continuously throughout the parliamentary period, temporary committees are given specific mandates to focus on particular issues or tasks. One example of a temporary committee is *Valgkomiteen* (“the Election Committee”), whose mandate is to select candidates for the standing committees. We collect information on committee membership at the time of the speech from <https://www.stortinget.no/no/Representanter-og-komiteer/>.

Time period	Prime minister	Parties	Parliamentary basis	Appointment reason
Jun 1945–Nov 1945	Einar Gerhardsen (A)	A, NKP, H, V	—	<i>Unification Cabinet</i>
Nov 1945–Nov 1951	Einar Gerhardsen (A)	A	Majority	General elections
Nov 1951–Jan 1955	Oscar Torp (A)	A	Majority	Change prime minister
Jan 1955–Aug 1963	Einar Gerhardsen (A)	A	Majority	Change prime minister
Aug 1963–Sep 1963	John Lyng (H)	H, V, Sp, KrF	Minority	Government crisis
Sep 1963–Oct 1965	Einar Gerhardsen (A)	A	Minority	Government crisis
Oct 1965–Mar 1971	Per Borten (Sp)	Sp, H, V, KrF	Majority	General elections
Mar 1971–Oct 1972	Trygve Bratteli (A)	A	Minority	Government crisis
Oct 1972–Oct 1973	Lars Korvald (KrF)	KrF, V, Sp	Minority	Government crisis
Oct 1973–Jan 1976	Trygve Bratteli (A)	A	Minority	General election
Jan 1976–Feb 1981	Odvar Nordli (A)	A	Minority	Change prime minister
Feb 1981–Oct 1981	Gro H. Brundtland (A)	A	Minority	Change prime minister
Oct 1981–Jun 1983	Kåre Willoch (H)	H	Minority	General elections
Jun 1983–Sep 1985	Kåre Willoch (H)	H, KrF, Sp	Majority	Government expansion
Sep 1985–May 1986	Kåre Willoch (H)	H, KrF, Sp	Minority	—
May 1986–Oct 1989	Gro H. Brundtland (A)	A	Minority	Government crisis
Oct 1989–Nov 1990	Jan P. Syse (H)	H, KrF, Sp	Minority	General elections
Nov 1990–Oct 1996	Gro H. Brundtland (A)	A	Minority	Government crisis
Oct 1996–Oct 1997	Thorbjørn Jagland (A)	A	Minority	Change prime minister
Oct 1997–Mar 2000	Kjell M. Bondevik (KrF)	KrF, Sp, V	Minority	General elections
Mar 2000–Oct 2001	Jens Stoltenberg (A)	A	Minority	Government crisis
Oct 2001–Oct 2005	Kjell M. Bondevik (KrF)	KrF, H, V	Minority	General elections
Oct 2005–Oct 2013	Jens Stoltenberg (A)	A, SV, Sp	Majority	General elections
Oct 2013–Jan 2018	Erna Solberg (H)	H, FrP	Minority	General elections
Jan 2018–Jan 2019	Erna Solberg (H)	H, FrP, V	Minority	Government expansion
Jan 2019–Jan 2020	Erna Solberg (H)	H, FrP, V, KrF	Majority	Government expansion
Jan 2020–Oct 2021	Erna Solberg (H)	H, V, KrF	Minority	Government reduction
Oct 2021 –	Jonas Gahr Store (A)	A, Sp	Minority	General elections

**Table 1.** Norway’s governments 1945–2024. *Note:* The parties are the Communist Party (NKP), Socialist Left Party (SV), the Labour Party (A), the Centre Party (Sp), the Christian Democratic Party (KrF), the Liberal Party (V), the Conservative Party (H), and the Progress Party (FrP). Source: <https://www.regjeringen.no/no/om-regjeringa/tidligere-regjeringer-og-historie/>.

## Data Records

We provide several variables at the speaker level (Table 3). All variables are available for the full 1945 – 2024 period, except for *time*, which is only available from March 2004 onwards. The data is accessible at <https://doi.org/10.6084/m9.figshare.27118044><sup>13</sup>.

The full transcript of each speech is stored in the variable *text*. There are two official written standards of the Norwegian language: *Bokmål* and *Nynorsk*. Legislators can choose to have their speeches transcribed in either language. We apply a language identification algorithm (<https://github.com/saffsd/langid.py>; last accessed 19.08.2024) to determine the language used in each speech. In our sample, 87% of the speeches are in *Bokmål*.

In the original documents, approximately 1% of speeches begin with a procedural remark enclosed in parentheses, such as “from the floor” (*fra salen*). We extract these parenthetical phrases from the start of *text* into a separate variable, *procedural*. Additionally, we extract other parenthetical statements noted down by the transcriber that appear elsewhere in the *text* variable, such as “the president gavel” (*presidenten klubber*) and “laughter in the chamber” (*latter i salen*). Some parenthetical statements remain in the text because they provide non-procedural content that is useful for understanding the speech.

**Number of speeches.** The dataset covers 982,333 speeches. Figure 1 illustrates the trend in the number of speeches over time by chamber, revealing two noteworthy patterns in the data. Firstly, the majority of speeches occurred in plenary sessions. Specifically, during the 1945–2009 period, Odelsting and Lagting speeches constituted 13% and 3% of the total speeches, respectively. Secondly, there has been a substantial increase in the overall number of speeches over time. In Fig. 1, and all subsequent analyses, parliamentary sessions are defined by the year in which they concluded.

The rise in the number of speeches can be partly attributed to the growth in the number of legislative representatives over time. Initially, the legislature comprised 150 representatives in 1945, increasing to 157 in 1977, adjusting to 155 in 1981, rising to 165 in 1993, and ultimately reaching 169 in 2009.

**Speakers and speech length.** Speeches are delivered by four types of speakers: MPs (50.9% of speeches), Parliamentary Presidents (29.5%), Cabinet Members (15.4%), and Deputy MPs (3.7%). In addition, there was

Party acronym	Party name (election years)	English translation
A	Det norske Arbeiderparti (1945–2009)	Labour Party
A	Arbeiderpartiet (2013–2017)	Labour Party
FrP	Anders Langes parti (1973)	Progress Party
FrP	Fremskrittspartiet (1977–2017)	Progress Party
H	Høyre (1945–2017)	Conservative Party
KrF	Kristelig Folkeparti (1945–2017)	Christian Democratic Party
SV	Sosialistisk Folkeparti (1961–1969)	Socialist Left Party
SV	Sosialistisk Valgforbund (1973)	Socialist Left Party
SV	Sosialistisk Venstreparti (1977–2017)	Socialist Left Party
Sp	Bondepartiet (1945–1957)	Centre Party
Sp	Senterpartiet (1961–2017)	Centre Party
V	Venstre (1945–2017)	Liberal Party

**Table 2.** Party names and their time periods in use.

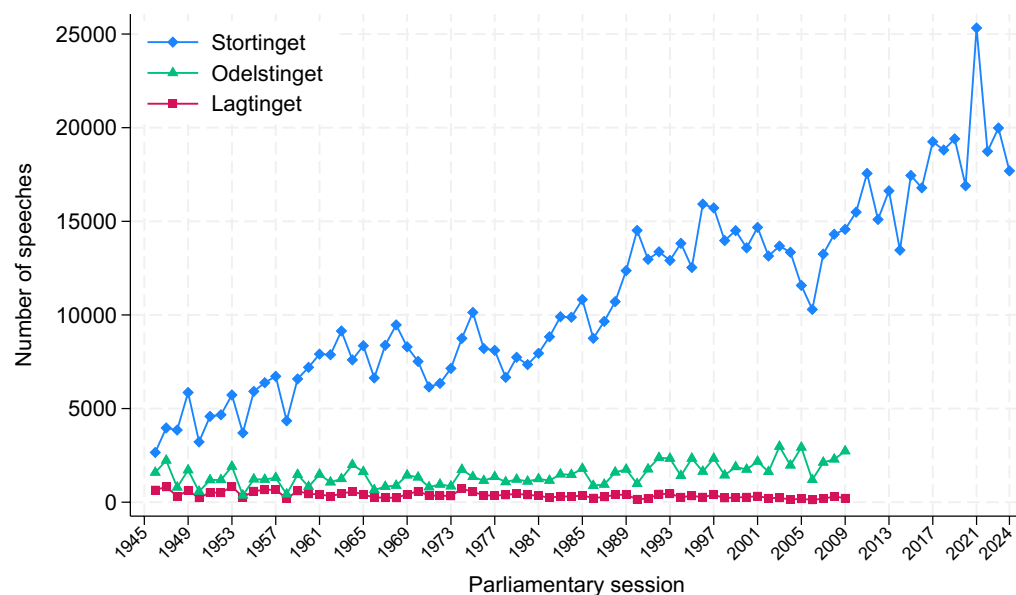
Name	Meaning
date	Date of speech (date-month-year)
time	Time of speech (hour-minute-seconds in 24-hour format)
session	Parliamentary session during which the speech was made
election	Year of the preceding parliamentary election
name	Edited name of speaker
name_original	Speaker's name from original data sources
id_speech	Speech identifier
id_person	Personal identifier
pid_v28	Personal identifier from the 2024 version of the Norwegian Parliamentary Elections dataset <sup>21</sup>
text	Full transcript of the speech
procedural	Parenthetical procedural statements extracted from the original text
reactions	Reactions noted down by the transcriber
language	Language of the speech (either <i>Bokmål</i> or <i>Nynorsk</i> )
pdf_source	PDF source of speech
chamber	Chamber of the Storting where the speech was made
status	Speaker's status at the time of speech
committee	Speaker's committee membership at the time of speech
partyname	Speaker's party name (as listed on ballot) <sup>21</sup>
party	Speaker's party acronym (see Table 2) <sup>21</sup>
elected	Indicates whether the speaker was elected in the last election <sup>21</sup>
deputy	Indicates whether the speaker was elected as deputy in the previous election <sup>21</sup>
cabinet	Indicates whether the speaker is a member of cabinet during the four-year election period <sup>21</sup>
rank	Speaker's ballot position in the last election <sup>21</sup>
district	Speaker's election district in the last election <sup>21</sup>
female	Speaker's gender (female = 1; male = 0) <sup>21</sup>
birthyear	Speaker's year of birth <sup>21</sup>

**Table 3.** Description of variables in the Norwegian Parliamentary Debates dataset.

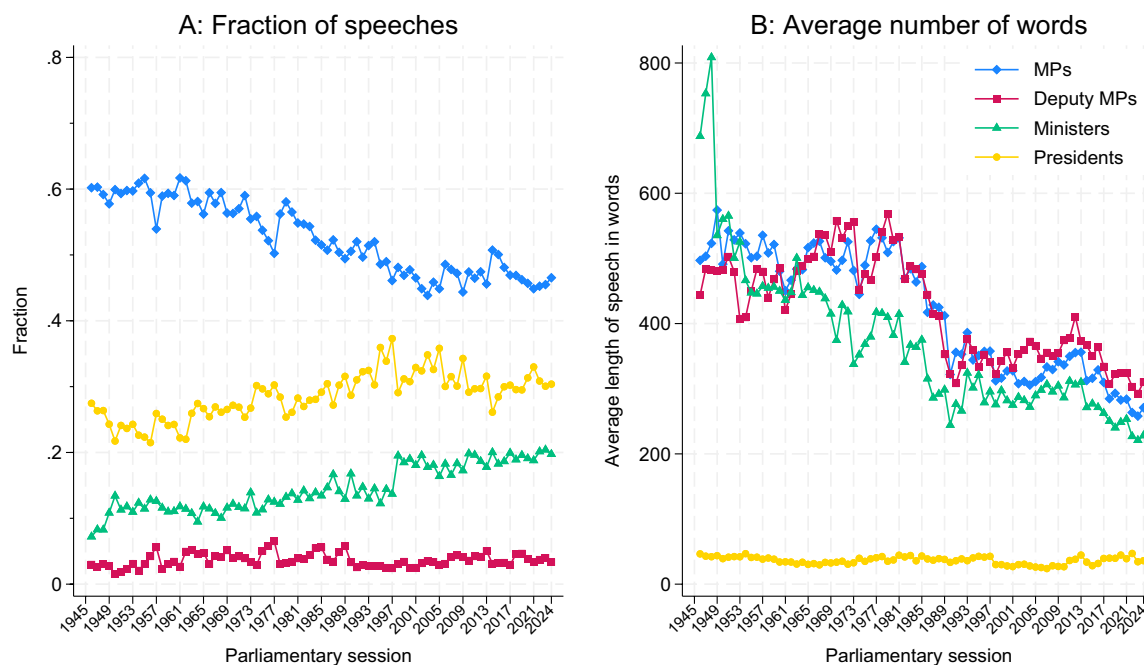
a limited number of other speakers in Parliament during our sample period. These *not matched* speakers, who account for about 0.4% of speeches, are excluded from the subsequent analysis. The 'not matched' category includes instances where speaker identification is ambiguous due to common names (e.g., "Nilsen" could refer to both Nils Alfred Nilsen and Nils Emaus Nilsen, both elected in 1949), speeches from *Parliamentary hearings* where experts or previous cabinet members are invited to the Parliament, and a limited number of cases due to OCR failures.

MPs appointed to cabinet positions are required to resign their seats in the Parliament and their vacant seats are replaced by deputies next in line on the party's list<sup>21</sup>. Figure 2 shows that the fraction of speech given by ministers has increased over time at the expense of regular MPs (Panel A).

Across all speaker types, speeches are, on average, 286 words long. They vary in length from one word (e.g., "President!") to 14,291 words (the opening speech of the 1994–1995 parliamentary session held by the president). Panel B of Fig. 2 shows that speech length has been trending downward throughout our sample period. During our 80-year sample period, the average speech was roughly cut in half for MPs, deputy MPs,

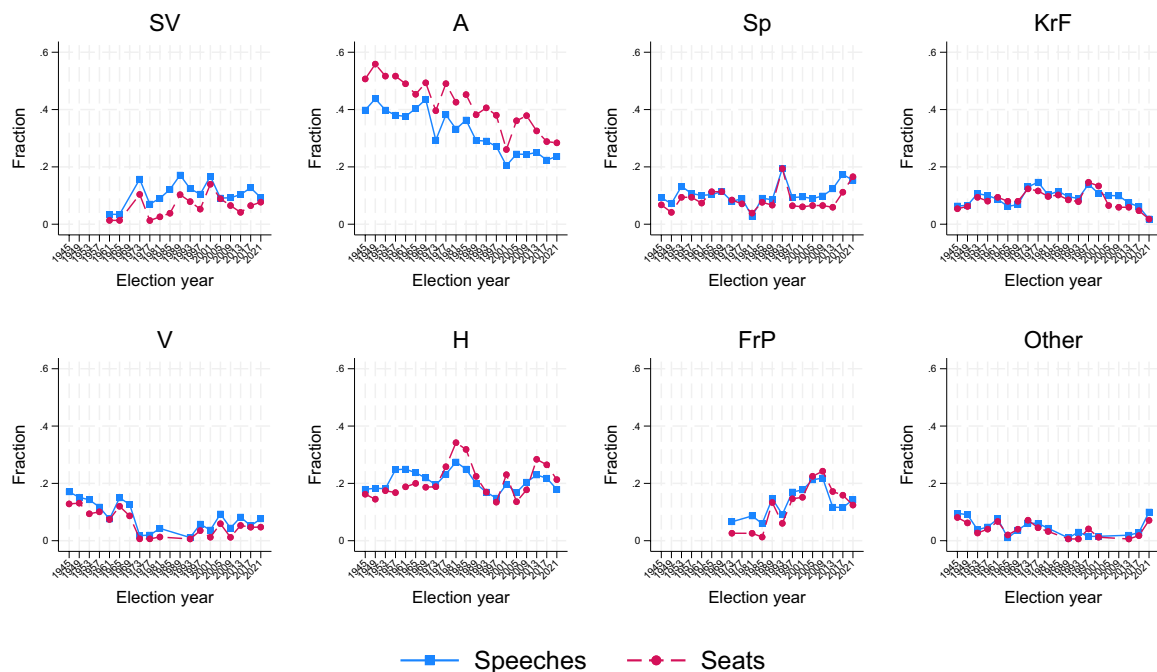


**Fig. 1** Number of speeches by Parliamentary session and chamber. Notes: The Norwegian Parliament, operated until 2009 with an internal division into two chambers – *Odelstinget* and *Lagtinget*.



**Fig. 2** Fraction of speeches and average length of speech by speaker status over time. Notes: Speeches by “non-matched” individuals are excluded.

and ministers (from about 500 words to about 250 words). This trend was accompanied by modifications in the “General Limitations in Speaking Time” as stipulated in the Storting’s *Rules of Procedure*<sup>27</sup>. For example, in 1989, the maximum duration for the initial speech in debates was reduced from one hour to 30 minutes, and this was further reduced to 15 minutes in 2017. These adjustments not only reflect the trend toward shorter speeches but also actively contribute to reducing the average speaking time. While these guidelines establish the norm, the president can approve exceptions for longer speaking times when necessary. Ministers tend to give



**Fig. 3** Party representation over time measured by fraction of seats and fraction of speeches. Notes: The speech share is based on the number of speeches given. Candidates elected on joint lists are classified according to the parliamentary party group they were part of.

slightly shorter speeches than MPs, except for the three first sessions following World War II. Across our sample, speeches by presidents tend to be very short.

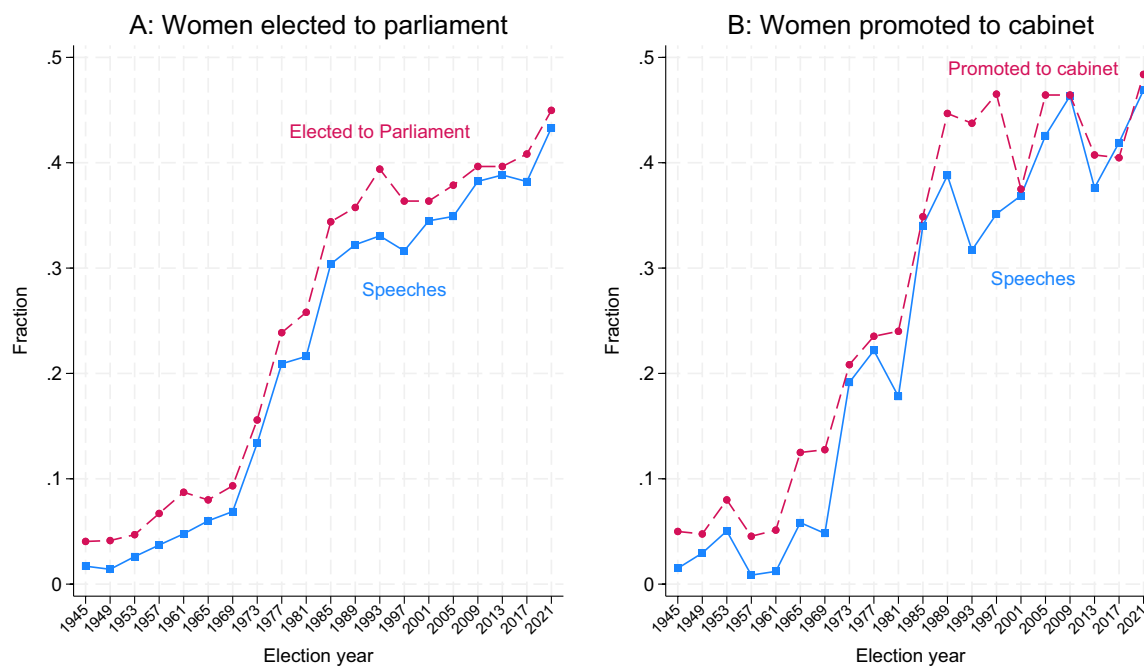
**Political parties.** As previously mentioned, the closed-list proportional representation system used in Norway gives parties a central place in Norwegian politics. The Labor Party (A) has been a dominant force in Norwegian politics since World War II and has held a position in the cabinet for 50 years during our 80-year sample period (Table 1). Figure 3 illustrates that, throughout our study period, members of the Labor Party speak less than their proportionate representation in Parliament suggests. On the contrary, smaller parties, such as the Socialist Left Party (SV) and the Liberal Party (V), tend to speak more than their proportionate share. That parties with higher seat shares speak less than their fair share reflects the importance of committees in the Norwegian plenary<sup>23,28</sup>.

Interestingly, individuals representing parties in government speak less often, and this pattern is particularly noticeable for the Conservatives (H) during the Willoch era (1981–1986) and the Solberg era (2013–2021). Using data from the *The Talk of Norway* dataset<sup>29</sup>, which covers parliamentary speech in the 1998–2016 period, Søyland and Høyland (2022)<sup>23</sup>, document that government MPs participate in fewer debates, but their speeches are longer.

**Gender.** Studying gender differences in parliamentary debates helps to understand how representation and participation have evolved over time, providing insights into the evolution of gender equality in political discourse. Panel A of Fig. 4 contrasts the fraction of speech given by women MPs (blue squares) to the fraction of women elected to Parliament (red circles). Using either metric, women's representation in Parliament has gradually increased over our sample period. However, in almost every four-year election period covered by our data, the fraction of female speech is below the fraction of female MPs, indicating that women speak less than their 'fair share', as in many other European parliaments<sup>20</sup>.

Panel B of Fig. 4 similarly contrasts the fraction of speeches given by women cabinet members (blue squares) with the fraction of women promoted to the cabinet in the same election period (red circles). We observe once again that female speech participation is less than proportionate, particularly prior to the 1970s.

A caveat here is that the number of women promoted to cabinet in the first part of our sample period is very small. In 1945, Kirsten Hansteen (Communist Party) became the first woman promoted to the cabinet in Norway. She was part of the *Unification Cabinet*, which was in office between June and November 1945, in the aftermath of the Second World War. This period is not part of our sample. Aaslaug Aasland (Labor) became in 1948 the second woman promoted to cabinet. She was the minister of social affairs and also acted as prime minister for Oscar Torp in 1953. Panel B of Fig. 4 does not differentiate between individuals who serve the entire election period and those who are cabinet ministers for shorter periods.



**Fig. 4** Gender and speech making. Notes: Panel A is limited to individuals elected to parliament as regular MPs. Panel B is limited to individuals promoted to cabinet during an election period (without taking the duration of tenure as cabinet ministers into account).

### Technical Validation

To validate our data, we have compared the sample size and the names of speakers in the speech records to those in the Norwegian Parliamentary Elections dataset<sup>21</sup>. For those MPs who were nominated to the cabinet without running for election (and hence do not occur in<sup>21</sup>), we validate the speakers against data on Norwegian cabinet ministers from *Politikerarkivet*. We have also conducted some manual validation by comparing speech content and speaker names to PDF samples, as well as speeches listed on the website of the Storting.

To validate the temporal consistency of the dataset, we have plotted the data over time (see Fig. 1). All observed structural breaks correspond to identifiable events, such as the COVID-19 pandemic.

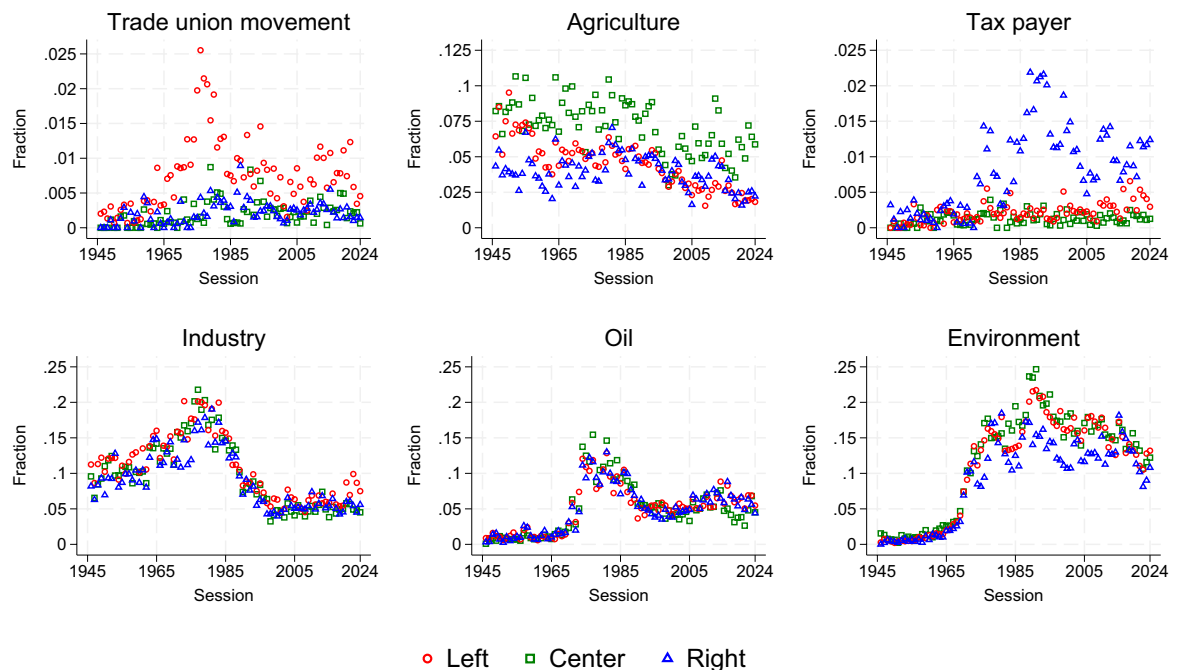
**Topics.** In this section, we present descriptive patterns as a validity check to confirm that the trends observed in the data reflect the broader political trends in Norway. These descriptive patterns also illustrate potential applications of the dataset. Of course, these types of analyses only scratch the surface of what is possible using the rapidly expanding natural language processing methods<sup>3</sup>.

Figure 5 plots the proportion of speeches mentioning the terms listed in the panel headings by party bloc over time. The top-row displays terms we expect to be associated with the left-wing, center and right-wing parties, respectively. Indeed, we observe that politicians from left-wing parties, who traditionally have upheld strong links to the labor unions, mention ‘trade union movement’ (*fagbevegelse*) more than others. While politicians from the center bloc, which includes the Center Party (formerly Farmer’s Party), frequently mention ‘agriculture’ (*landbruk*), right-leaning politicians, including members from the Progress Party which originated as an anti-tax protest movement, frequently mention ‘taxpayers’ (*skattebetaler*) in parliamentary debates.

Over the eight decades covered by our sample, political debates reflect changes in Norwegian society. For instance, the number of individuals employed in industry increased in the first decades post-World War II, peaking at 22% in 1971, before declining to approximately 8% at the end of our sample period (<https://www.ssb.no/virksomheter-foretak-og-regnskap/virksomheter-og-foretak/artikler/slik-jobber-norge>). The bottom-left panel of Fig. 5 illustrates that the frequency of speeches mentioning ‘industry’ (*industri*) mirrors this trajectory. Similarly, the discovery of oil in the North Sea in 1969 and ensuing debates on wealth management triggered a significant rise in speeches referencing ‘oil’ (*olje*). Discussions in Parliament related to environmental conservation have exhibited a comparable pattern, as indicated by the mentions of the environment (*miljø*) shown in the bottom panel of Fig. 5, coinciding with Norway’s establishment of its Ministry of Environmental Protection in 1972.

### Usage Notes

The dataset, which covers almost one million speeches over eight decades, allows researchers to examine a wide range of research questions within political economics. Our data can elucidate how and why legislators behave in a specific manner when participating in legislative debates<sup>20</sup>. In addition to studying representational aspects, the longitudinality of our data allows researchers to show how these have evolved over time and responded to



**Fig. 5** Fraction of speeches mentioning selected terms over time by party bloc. Notes: This figure displays the proportion of speeches that mention the term specified in the panel headings, by party bloc over time. The left-wing parties (Labor and Socialist Left) are marked with red triangles, the center parties (Center Party, Liberals and Christian Democrats) are marked with green squares, and the right-wing parties (Conservatives and Progress Party) are marked with blue triangles. The terms are translated to English from Norwegian. The Norwegian words are (from top left to right): *fagbevegelse*, *landbruk*, *skattebetaler*, *industri*, *olje*, and *miljø*.

changes in the electoral environment. For example Gennaro and Ash (2022)<sup>30</sup>, study the use of emotion and reason in political discourse using speeches from the United States Congress. Their new measurement approach can be directly applied to other empirical contexts such as Norwegian parliamentary debates.

Moreover, the dataset should be useful across the social sciences. Specifically, in political science, it enables the study of topics such as political cleavages, coalition formation, and the evolution of policy debates, among others. In addition, the unique availability of data in Norway, such as detailed records of politicians' backgrounds and comprehensive registry data capturing education, income, social benefits, and more, significantly expands the scope of research questions that can be explored. We include a unique personal identifier for each speaker, which allows the user to merge the speech records with<sup>21</sup>, which contains rich data on all candidates running for election in Norway since 1906. The dataset can also be helpful outside of social sciences. For example, for linguists interested in the evolution of language and rhetoric within political speech.

### Code availability

Figures 1–5 have been created using Stata (version 18). A replication package is available at [www.jon.fiva.no/data/NPD\\_replication.zip](http://www.jon.fiva.no/data/NPD_replication.zip).

Received: 28 May 2024; Accepted: 13 November 2024;

Published online: 02 January 2025

### References

1. Proksch, S.-O. & Slapin, J. B. *The Politics of Parliamentary Debate: Parties, Rebels, and Representation* (Cambridge University Press, 2015).
2. Itzkovitch-Malka, R., Mor, G., Oshri, O. & Shaul, S. Talking representation: How legislators re-establish responsiveness in cases of representational deficits. *European Journal of Political Research* (2023).
3. Ash, E. & Hansen, S. Text algorithms in economics. *Annual Review of Economics* **15**, 659–688 (2023).
4. Slapin, J. B. & Proksch, S.-O. Words as data: Content analysis in legislative studies. In Martin, S., Saalfeld, T. & Strøm, K. W. (eds.) *The Oxford Handbook of Legislative Studies* (Oxford University Press, Oxford, 2014).
5. Martin, S. Using parliamentary questions to measure constituency focus: An application to the Irish case. *Political studies* **59**, 472–488 (2011).
6. Lippmann, Q. Gender and lawmaking in times of quotas. *Journal of Public Economics* **207**, 104610 (2022).
7. Gentzkow, M., Shapiro, J. M. & Taddy, M. Measuring group differences in high-dimensional choices: Method and application to congressional speech. *Econometrica* **87**, 1307–1340 (2019).
8. Peterson, A. & Spirling, A. Classification accuracy as a substantive quantity of interest: Measuring polarization in Westminster systems. *Political Analysis* **26**, 120–128 (2018).



9. Cagé, J., Le Pennec, C. & Mougin, E. Firm donations and political rhetoric: Evidence from a national ban. *American Economic Journal: Economic Policy* **16**, 217–56 (2024).
10. Le Pennec, C. Strategic campaign communication: Evidence from 30,000 candidate manifestos. *The Economic Journal* **134**, 785–810 (2024).
11. Gentzkow, M. & Shapiro, J. M. What drives media slant? Evidence from US daily newspapers. *Econometrica* **78**, 35–71 (2010).
12. Groseclose, T. & Milyo, J. A measure of media bias. *The Quarterly Journal of Economics* **120**, 1191–1237 (2005).
13. Fiva, J. H., Nedregård, O. & Øien, H. The Norwegian parliamentary debates dataset. *Figshare* <https://doi.org/10.6084/m9.figshare.27118044> (2025).
14. Marsh, M. et al. *Candidate selection in comparative perspective: The secret garden of politics*, **18** (SAGE Publications Limited, 1988).
15. Hazan, R. Y. & Rahat, G. *Democracy Within Parties: Candidate Selection Methods and Their Political Consequences*. (Oxford University Press, Oxford, UK, 2010).
16. Fiva, J. H., Nedregård, O. & Øien, H. Group identities and parliamentary debates. Working paper available at [www.jon.fiva.no/docs/FivaNedregardOien.pdf](http://www.jon.fiva.no/docs/FivaNedregardOien.pdf) (2024)
17. Fiva, J. H. & Nedregård, O. How does party discipline affect legislative behavior? Evidence from within-term variation in lame-duck status. *Quarterly Journal of Political Science* **19**, 191–216 (2024).
18. Nedregård, O. Staying in line or straying for votes: Party discipline and dyadic representation. Unpublished manuscript, Sciences Po (2024)
19. Klemmensen, R., Hobolt, S. B. & Hansen, M. E. Estimating policy positions using political texts: An evaluation of the wordcores approach. *Electoral Studies* **26**, 746–755 (2007).
20. Back, H., Debus, M. & Fernandes, J. M. *The Politics of Legislative Debates* (Oxford University Press, 2021).
21. Fiva, J. H. & Smith, D. M. Norwegian parliamentary elections, 1906–2013: Representation and turnout across four electoral systems. *West European Politics* **40**, 1373–1391 (2017).
22. Strom, K. W. Norway: A Land of Minority Governments. In *Minority Governments in Comparative Perspective* (Oxford University Press, 2022).
23. Søyland, M. G. & Høyland, B. Committee-membership matters, party loyalty decides. In Back, H., Debus, M. & Fernandes, J. M. (eds.) *The Politics of Legislative Debate*, chap. 31, 634–650 (Oxford University Press, Oxford, 2022).
24. Strøm, K. & Swindle, S. M. Strategic parliamentary dissolution. *American Political Science Review* **96**, 575–591 (2002).
25. Rasch, B. E. Electoral systems, parliamentary committees, and party discipline: the Norwegian Storting in a comparative perspective. In Shaun Bowler, D. M. F. & Katz, R. S. (eds.) *Party Discipline and Parliamentary Government*, 121–140 (Ohio State University Press Columbus, 1999).
26. Stortingets forretningsorden med kommentarer. Available from: <https://www.stortinget.no/globalassets/pdf/forretningsorden/kommentarutgave-forretningsordenen-2024.pdf> Accessed 21. May 2024.
27. Stortingets Presidentskap. Innstilling til stortinget fra stortingets presidentskap (2016–2017). Available from: <https://www.stortinget.no/globalassets/pdf/innstillinger/stortinget/2016-2017/inns-201617-487s.pdf> Accessed 21. May 2024.
28. Nedregård, O. & Søyland, M. Legislative party groups and party cohesion. In Allern, E., Braut-Hegghammer, M., Kolltveit, K. & Rasch, B. E. (eds.) *Oxford Handbook of Norwegian Politics* (Oxford University Press, 2025).
29. Laponi, E., Søyland, M. G., Velldal, E. & Open, S. The talk of Norway: A richly annotated corpus of the Norwegian parliament, 1998–2016. *Language Resources and Evaluation* **52**, 873–893 (2018).
30. Gennaro, G. & Ash, E. Emotion and reason in political language. *The Economic Journal* **132**, 1037–1059 (2022).

## Acknowledgements

We thank Sebastian Hoas for excellent research assistance. Fiva gratefully acknowledge financial support from the Norwegian Research Council (grant no. 314079).

## Author contributions

All authors conceived the overall project. O.N. gathered and cleaned the majority of the raw speech data. J.H.F. and H.Ø. validated the overall data and ensured its correct integration with auxiliary datasets. All authors contributed equally to the drafting and review of the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to O.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025