How Political Term Length Shapes Time Horizon in Policy-Making? Evidence from India's State Building*

Jun Goto† *GRIPS*
Tomoko Matsumoto‡ *Tokyo Univ. of Science*
Shunsuke Tsuda§ *Univ. of Essex*
Junichi Yamasaki¶ *Kyoto Univ.*

December 12, 2025

Abstract

This paper examines how politicians' term lengths shape the temporal orientation of policy discourse during the state-building period. We exploit a natural experiment in post-independence India, where the first members of the Rajya Sabha (upper house) drew lots in November 1952 to receive randomly assigned terms of two, four, or six years. Then, we apply word embedding techniques to parliamentary debates to measure temporal perspectives along two dimensions: past-versus-future orientation and immediate-versus-long-term orientation. We find that legislators with longer terms adopt significantly more future-oriented and long-term perspectives in their policy discussions. Seeded topic modeling reveals that longer-term legislators emphasize foundational institutional reforms requiring sustained coordination—tax systems, property rights—while shorter-term legislators focus on immediate security threats, often justifying positions through personal experiences and historical narratives rather than forward-looking institutional analysis. We also demonstrate that electoral incentives drive these effects: only elected members facing future electoral accountability respond to term length assignments. These findings show that electoral time horizons systematically influence whether legislators prioritize short-term responsiveness or long-term institutional development during critical state-building periods.

Keywords: state building, term length, temporal orientation, natural language processing, electoral incentives

*We acknowledge comments from Takashi Kurosaki, Hiroshi Sato, and Sujata Visaria. We are also grateful for financial support from the Japan Society for the Promotion of Science (19K01653 and 18K12701).

†Associate Professor. E-mail: ju-goto@grips.ac.jp

‡Associate Professor. E-mail: tomoko.matsumoto@rs.tus.ac.jp

§Assistant Professor, Department of Economics. E-mail: shunsuke.tsuda@essex.ac.uk

¶Associate Professor. E-mail: yamasaki@econ.kyoto-u.ac.jp.

1 Introduction

Legislators in their state-building period need to establish institutional foundations that shape a country's long-run socioeconomic development (Acemoglu et al., 2001, 2005). Yet, during such turbulent times, they simultaneously face pressing challenges—such as economic crises, social unrest, or security threats—that demand immediate attention to maintain public support and social stability (Besley and Burgess, 2002; Myerson, 2011).[1] Given limited political and economic resources, legislators must balance competing temporal priorities: whether to address urgent, short-term needs or to invest in building enduring institutional frameworks. This raises an important but underexplored question: what shapes legislators' temporal perspectives especially in state building?

One of the most prominent influences on their temporal perspectives will be the length of politicians' terms. Term length determines how soon the next election arrives, shaping incentives accordingly (Dal Bó and Rossi, 2011). A short term might incentivize them to pursue initiatives to address imminent, urgent issues. Longer terms, by contrast, encourage them to focus on problems defined over longer horizons and to design fundamental institutions that require extensive coordination. However, research on how term length shapes politicians' temporal perspectives in policy formation remains scarce.

This paper asks whether legislators' term lengths affect the time horizon of their policy discussions during state building. Focusing on post-independence India, we exploit an exogenous variation in tenure created by a constitutional rule applicable for the first members of the Rajya Sabha, the upper house. The Constitution set a six-year term for all members of the parliament (MP). However, it required that one-third of seats be renewed every two years. To reconcile these conflicting stipulations, all members elected in the first election drew lots six months after the first session and were randomly assigned terms of two, four, or six years. We use this natural experiment to compare their policy speeches in parliamentary sessions after the lottery and identify the causal effect of term length on the time horizon of policy deliberation.

The key challenge in measuring individual temporal orientation is that politicians rarely articulate their time horizons explicitly in speeches, making direct count-based measurement difficult. Therefore, we quantify their temporal orientation using recently developed machine learning algorithms applied to parliamentary debate transcripts. We decompose their temporal orientation implicitly embedded in their speeches into two complementary dimensions: (i) a past-versus-future axis and (ii) an immediate-versus-long-term axis. First, the past-versus-future axis captures whether legislators adopt a forward-looking or backward-looking perspective when proposing policies. Specifically, we construct a dictionary containing a small set of words that unambiguously denote future or past concepts. Then we assess how strongly each sentence of speeches aligns with the future end relative to the past end based on word

[1] Recent political science literature examines how territorial integration, security consolidation, and the construction of fiscal capacity shape early policy choices (Grieco, 2024; Queralt, 2022). Queralt (2025) reviews this body of work and underscores the trade-off between addressing immediate redistributive and security demands versus making longer-term investments in institutions.

embeddings. Second, we examine an immediate-versus-long-term dimension. This dimension captures whether legislators emphasize urgent, short-term issues or fundamental, long-term institutional development when proposing policies. We again use word embeddings to position each sentence along this axis, anchored by words that clearly represent immediate versus long-term contexts.

We collect complete speech data from the official records of the Rajya Sabha.[2] Main data covers all statements made by the initial 216 members of the chamber between November 29, 1952 (the lottery day), and the end of the 3rd Session in May 1953. We also assemble politician-level data including each member's elected state, birthplace, age, educational attainment, previous occupation, and party affiliation. In total, main data contains 43,551 statements comprising 564,389 sentences. This period spans the closing weeks of the 2nd Session in late 1952 and the entirety of the 3rd Session (Budget Session) in early 1953. In aggregate, the Rajya Sabha held 60 sittings in 1952 and 100 sittings in 1953, situating our dataset within the first two full legislative cycles of the newly established chamber.

We identify the causal effect of term length on politicians' policy discussions by comparing the temporal orientation of politicians across randomly assigned two-, four-, and six-year terms in the post-lottery period. The balance tests across term-length categories on pre-determined characteristics validate our identification strategy: the randomization was implemented rigorously and no arbitrary adjustments were introduced among legislators. This is consistent with historical accounts that the lottery produced an unpredictable and effectively random allocation of tenure.

We obtain two main findings. First, our estimation results show that longer tenure leads to more future-oriented policy discussions. Specifically, members with longer terms adopt a significantly more future-oriented perspective than their short-tenure counterparts. For example, moving from a two-year to a six-year term increases the future-versus-past orientation score by approximately 0.85 standard deviations. This effect is economically meaningful, corresponding to a shift in temporal framing equivalent to placing 2.2 times more weight on discussing future policy needs relative to discussing past historical events.

Second, we find that longer-term members adopt a significantly more long-horizon perspective in their policy discussions. Moving from a two-year to a six-year term increases the long-term-versus-immediate orientation score by approximately 0.78 standard deviations. This effect corresponds to a shift in framing equivalent to placing 2.1 times more weight on discussing long-term institutional development relative to addressing immediate, urgent issues.

To address potential threats to internal validity, we conduct placebo tests using parliamentary debates from before the tenure lottery, when all members expected to serve full six-year terms and no one knew their eventual assignment. Applying our identification strategy to pre-lottery speeches, we find no significant differences in temporal orientation between groups that would later receive different term lengths. This null finding, contrasting with the

[2]Available at `https://sansad.in/rs/debates/officials`. All texts were manually downloaded from this website. We applied optical character recognition (OCR) to extract text from PDF documents and manually verified the digitized output to correct any character corruption or recognition errors.

post-lottery patterns, provides evidence that our results reflect the causal effect of randomly assigned term lengths rather than pre-existing differences between legislators.

One might be concerned that our main findings on linguistic patterns simply reflect differences in the quantity of parliamentary participation rather than genuine changes in discourse content. Prior research has shown that term length can affect legislative effort (Dal Bó and Rossi, 2011). However, in our context, all Rajya Sabha members attended the full session schedule and actively participated in debates. To address this concern empirically, we aggregate the sentence-level data into member-by-sitting-day speech data and test for differences in participation patterns across randomly assigned term-length groups. Comparing the number of statements delivered, total sentences spoken, and total words used across the two-, four-, and six-year term groups, we find no significant variation in any of these participation metrics. This finding indicates that politicians do not alter their speaking frequency or intensity in response to their term length assignment. Therefore, our results reflect genuine changes in the temporal orientation of policy discourse rather than differential levels of effort or parliamentary engagement.

Our main results establish that longer term lengths shift legislators toward more future-oriented and long-term policy discourse. However, these semantic measures do not reveal what kinds of policies legislators emphasize or how they construct their arguments. We therefore examine the substantive content underlying these linguistic patterns.

To examine whether these linguistic patterns correspond to substantive policy differences, we apply seeded Latent Dirichlet Allocation (LDA) to categorize parliamentary speeches into policy-relevant topics. The results reveal a clear divergence. Members with longer terms engage more frequently with topics linked to long-run institutional development—tax reform, property rights, infrastructure, and Five-Year Plans. By contrast, shorter-term members focus on urgent security concerns—border disputes, military conflict, and defense—often invoking historical context to justify immediate responses. These findings reinforce our main results: longer tenure shifts legislators' attention toward policy domains requiring sustained coordination, while shorter tenure draws attention toward pressing security issues.

We then examine how legislators with different term lengths construct their policy arguments. Legislators with longer terms focus on what should be done to prepare for upcoming challenges, citing numerical justifications and objective evidence. In contrast, legislators with shorter terms ground their policy positions in personal experiences and historical narratives. These patterns suggest that longer tenure leads legislators to frame policies based on concrete evidence and forward-looking analysis, whereas shorter tenure is associated with discourse that draws on historical context and past experiences.

Finally, we investigate the mechanism underlying our main findings. We exploit an institutional feature of the Rajya Sabha: the chamber includes both elected members chosen by state legislative assemblies and nominated members appointed directly by the President of India. This dual composition allows us to test whether electoral incentives drive the observed effects. If electoral considerations underlie our findings, only elected members—who face future electoral accountability—should exhibit differences in temporal orientation based

on their assigned term length. Nominated members, who do not face reelection pressures, should be insensitive to term length variation.

To test this hypothesis, we estimate our baseline specification with an interaction term between term length and an indicator for elected members. The results strongly support the electoral incentive mechanism. The coefficient on term length alone is close to zero and statistically insignificant, indicating that nominated members show no relationship between assigned term length and temporal orientation. In contrast, the interaction term yields results nearly identical to our main findings: elected members with longer terms adopt significantly more future-oriented perspectives. This contrast reveals that electoral accountability is the key driver, confirming that the observed effects stem from electoral incentives rather than other factors associated with term length.

Literature and Contributions Our paper contributes to three strands of literature. First, we advance research on state-building and institutional development (Besley and Persson, 2011; Burgess et al., 2015; Charnysh, 2019; Tilly, ed, 1975), including work on the political origins of state capacity (Besley and Persson, 2009). Recent empirical work highlights how specific policies and shared experiences contribute to state-building, including compulsory schooling (Bandiera et al., 2019) and national football success (Depetris-Chauvin et al., 2020). We complement this literature by examining how politicians' term lengths—a previously understudied institutional feature—shape the temporal orientation of policy discourse during state-building periods.

Second, our paper extends the literature on political economy and electoral accountability, particularly research examining how term lengths affect legislative behavior. Previous studies have established that term length influences various dimensions of political performance, including legislative effort (Dal Bó and Rossi, 2011; Titiunik, 2016) and policy outcomes (Amacher and Boyes, 1978; Kalt and Zupan, 1990; Lott and Davis, 1992), as well as accountability and policy cycles shaped by reelection incentives (Ferraz and Finan, 2011; Rogoff, 1990; Alesina and Tabellini, 1990) and legislative turnover and ideology (Ash et al., 2017). We add new evidence to this literature by demonstrating that term lengths causally shape the temporal orientation of policy discourse. Our findings reveal that longer terms incentivize legislators to adopt more future-oriented and long-term perspectives, highlighting how electoral time horizons influence not only legislative effort but also the substantive framing of policy debates.

Finally, we contribute to the growing literature on text-as-data in economics (Ash et al., 2025b; Gentzkow et al., 2019), including applications that construct policy indicators from documents and deliberation transcripts (Baker et al., 2016; Hansen et al., 2018; Ash et al., 2025a). We develop novel methods to measure politicians' temporal orientation—a concept that is theoretically important but difficult to capture through surveys or interviews due to social desirability bias. By combining these methods with a natural experiment, we demonstrate how computational text analysis can reveal political preferences that are otherwise difficult to measure systematically.

2 Context

This section elaborates a policy context and a natural experiment in the Rajya Sabha (Council of States) of India, which was instituted in April 1952, following India's independence from British rule. We provide an overview and context related to the natural experiment.

2.1 Institutional Setting

The Indian Parliament The parliament of India is the supreme legislative body (*The Constitution of India*, 1950). It consists of two chambers: the Rajya Sabha (Council of States) and the Lok Sabha (House of the People). The Rajya Sabha is the upper house of parliament, with members who are elected by the state and territorial legislatures (Article 80). The Rajya Sabha is a permanent body, thus not subject to dissolution (Article 83).[3]

The Role of the Rajya Sabha The Rajya Sabha plays a crucial role in India's legislative process, with authority to enact laws, scrutinize government policies, and participate in debates on national and international issues (Articles 107–111, *The Constitution of India* 1950). Members can raise issues of public concern and participate in policy discussions (*Rules of Procedure and Conduct of Business in the Council of States*, 1952), making it an ideal setting for analyzing how term length affects policy discourse during state-building.

The Election of the Rajya Sabha The first election of the Rajya Sabha was held in 1952, establishing the chamber with 216 members (*Rajya Sabha Debates, Official Report*, 1952–1953). Of these, 204 were elected by state legislatures through proportional representation using the single transferable vote (*Representation of the People Act, 1951*, 1951), while 12 were nominated by the President of India. Under Article 80(3) of the Constitution, nominated members are selected for their distinguished contributions to literature, science, art, or social service—a provision designed to bring professional expertise into parliamentary debates without requiring electoral participation (*The Constitution of India*, 1950). The chamber was first convened on April 3, 1952, with its first sitting on May 13, 1952 (*Rajya Sabha Debates, Official Report*, 1952–1953). Importantly, members whose terms have ended are eligible for re-election without term limits, creating incentives for elected members to demonstrate policy achievements to the state legislators who will vote in future elections.

Legislative Procedures and Speaking Rules The Rajya Sabha operated under a structured session system with specific procedural rules governing parliamentary business (*Rules of Procedure and Conduct of Business in the Council of States*, 1952). Sessions typically commenced at quarter past eight in the morning or quarter to ten, with the Chairman (ex-officio Vice President) presiding over proceedings (Article 89, *The Constitution of India* 1950). The

[3]The Lok Sabha is the lower house of parliament and is made up of 545 members. 543 of these members are elected from the states and union territories, while the President can nominate two members from the Anglo-Indian community if they feel that the community is not adequately represented (Article 81).

chamber followed formal protocols including member oath-taking ceremonies, question periods, and structured debates on legislative matters. Speaking opportunities were governed by established rules: members could raise questions on designated days (following the House of Lords procedure with three starred questions on two days per week), with questions answered in order of receipt (*Rules of Procedure and Conduct of Business in the Council of States*, 1952). The Chairman held authority to manage debate duration and could fix time limits for protracted discussions after taking the sense of the Council. Sessions included various types of business including budget discussions, bill considerations, committee appointments, and messages from the House of the People. All members, whether elected or nominated, had equal speaking rights subject to the Chairman's permission, reflecting the parliamentary culture of the early 1950s when formal speaking restrictions were limited and members could participate freely in policy discussions.

Timeline of Sessions and Analysis Period Our analysis focuses on the period following the term length lottery on November 29, 1952 (*Council of States (Term of Office of Members) Order, 1952*, 1952). The 1st Session began on May 13, 1952, and extended through the year. The 2nd Session continued in late 1952, during which the lottery took place. The 3rd Session (Budget Session) was held in early 1953, concluding in May 1953 (*Rajya Sabha Debates, Official Report*, 1952–1953). Our main analysis uses speeches from the post-lottery period—the closing weeks of the 2nd Session in late 1952 and the entirety of the 3rd Session—when members were aware of their assigned term lengths. For placebo tests, we use pre-lottery speeches from the period when all members expected to serve full six-year terms.

2.2 The Natural Experiment: Random Term Length Assignment

To evaluate the influence of term durations on policy discussions, we leverage a natural experiment created by the random assignment of term lengths to the initial Rajya Sabha members.

Constitutional Mandate The Indian Constitution mandated that Rajya Sabha members serve six-year terms, with one-third of seats renewed every two years to ensure continuity (Article 83, *The Constitution of India* 1950). To establish this staggered retirement system for the inaugural cohort, the Constitution required that initial term lengths be determined by lot (Fourth Schedule).

Lottery Procedure To implement this system, the President issued the Council of States (Term of Office of Members) Order on 30 September 1952 (*Council of States (Term of Office of Members) Order, 1952*, 1952), mandating that members' terms be randomly assigned to three groups with different end dates: April 2, 1954 (two-year terms), April 2, 1956 (four-year terms), and April 2, 1958 (six-year terms). This created effective term lengths of approximately 2, 4, and 6 years respectively. The randomization was implemented through a public lottery held by the Election Commission on 29 November 1952. Members were grouped by

state for the random assignment, ensuring geographic balance across term lengths.[4] Both elected and nominated members were subject to the same randomization process—a feature we exploit in our mechanism analysis to test whether electoral incentives drive the observed effects. The lottery resulted in 72 members assigned six-year terms, 71 members assigned four-year terms, and 71 members assigned two-year terms. The complete assignment was published in the Gazette of India Extraordinary, providing transparent documentation of the randomization outcome.

Unanticipated Announcement of the Lottery Crucially for our identification strategy, members had no advance knowledge of how term lengths would be assigned. While the Constitution mandated staggered retirement, the specific mechanism was left to the President, who was empowered under Section 154 of the Representation of the People Act, 1951, to make special orders "after consulting the Election Commission" to fix member tenures (*Representation of the People Act, 1951*, 1951). Between the chamber's first sitting on May 13, 1952, and the lottery on November 29, 1952, all members operated under the assumption that they would serve full six-year terms. Since the assignment mechanism was unknown until the lottery was conducted, legislators had no basis on which to anticipate their eventual term length, providing strong support for the exogeneity of our treatment.

Balance Tests To validate the randomization, we conduct balance tests comparing pre-determined characteristics—including age, education, occupation, party affiliation, and state of election—across the three term-length groups. Specifically, we regress each pre-determined member attribute on term length (in years), controlling for state and party fixed effects, at both the sentence level and the member level (see Appendix B for estimation details). As shown in Appendix Figure A.1b, the coefficients on term length are consistently small and statistically insignificant across all characteristics. Joint tests fail to reject the null hypothesis of perfect balance (sentence-level: $p = 0.892$; politician-level: $p = 0.756$). These results confirm that the lottery produced balanced groups and support the validity of our identification strategy.

Selective Attrition We also verify that our results are not driven by selective attrition. While some members resigned or passed away before completing their assigned terms, these departures were rare during our study period (late 1952 through May 1953) and unrelated to term length assignment. Our main analysis focuses on speeches delivered during this period, when all original members remained active participants.

[4]Small states were combined into a single group: Bhopal, Bilaspur-cum-Himachal Pradesh, Delhi, and Kutch (*Council of States (Term of Office of Members) Order, 1952*, 1952).

Table 1: Descriptive Statistics

	Two-year term	Four-year term	Six-year term
Basic information			
% of members who made statements before the lottery	89.12	88.57	89.34
% of members who made statements after the lottery	88.87	89.44	88.15
Statement data (before the lottery)			
# of statements per day	3.45	3.01	3.48
# of sentences in each statement	129.52	126.33	127.12

Notes: Our dataset contains 43,551 statements comprising 564,389 sentences from all 216 members who participated during the target period. In aggregate, the Rajya Sabha held 60 sittings in 1952 and 100 sittings in 1953, situating our dataset within the first two full legislative cycles of the newly established chamber.

3 Data

Parliamentary Records We collected complete speech records from the official Rajya Sabha Debates covering the post-lottery period from November 29, 1952, through May 1953.[5] These records contain all statements by the initial 216 members, including speaker identification, session information, and member characteristics such as state and party affiliation. All statements are recorded in English in the official minutes, with English translations provided in appendices for any speeches originally delivered in Hindi or other regional languages. We applied optical character recognition (OCR) to extract text from the original PDF documents and manually verified the digitized output to correct recognition errors. After cleaning and processing, we decomposed each statement into individual sentences for analysis using standard sentence boundary detection. Our final dataset comprises 43,551 statements comprising 564,389 sentences across 160 parliamentary sittings (60 in 1952 and 100 in 1953).

Sample and Unit of Analysis Our main analysis uses sentence-level data. We conceptualize the data-generating process as legislators producing individual sentences, each reflecting a discrete expression of policy discourse. This granular unit is appropriate for two reasons. First, our embedding-based measurement approach operates at the sentence level, where semantic content can be most accurately captured by the machine learning algorithms. Second, the structure of early Rajya Sabha debates makes sentence-level analysis empirically appropriate: parliamentary proceedings in 1952–53 consisted largely of rapid exchanges of short questions and responses rather than extended uninterrupted speeches, with frequent interjections and cut-ins from other members. Aggregating to the statement or speech level would therefore not correspond to natural discourse units in this setting. Aggregating further to the date or session level would deviate substantially from the actual data-generating process in which legislators formulate policy positions sentence by sentence, and would conflate distinct policy topics discussed across different agenda items within the same sitting. We include all statements recorded in the official debates during our study period; no debates or speakers

[5] The official debates are available at `https://sansad.in/rs/debates/officials`. All texts were manually downloaded from this website.

were excluded based on language or transcript completeness, as all records are available in English. Consistent with this design, most utterances are short: roughly two-thirds of statements contain two sentences or fewer (median 2), with a long right tail of rare long speeches. Daily debates also feature rapid exchanges—on high-volume days (e.g., April 23, 1953; December 11, 1952), the median statement remains 1–2 sentences—reinforcing the sentence as the natural unit for embedding-based measurement. (Appendix Tables A.1 and A.2 report the underlying distributional details.)

Descriptive Statistics Table 1 presents descriptive statistics. Participation rates are consistently high across all term-length groups: approximately 89% of members made at least one statement both before and after the lottery, with no significant differences between groups. This reflects the parliamentary culture of the 1950s, when formal speaking rules were limited and members could speak freely with the Chairman's permission. Average daily participation is also balanced across groups, with members making 3.0-3.5 statements per sitting day regardless of assigned term length.

4 Construction of Key Outcomes

4.1 Overview of Temporal Orientation Measures

Our empirical analysis focuses on politicians' time horizons in legislative discourse. Since speakers rarely state their planning horizons explicitly, we must infer temporal orientation from the language they use. We measure temporal orientation along two conceptually distinct dimensions using complementary methodological approaches.

Our main outcomes employ an embedding-based method that positions each sentence along two semantic axes: (i) a *past-versus-future axis* that captures whether legislators adopt forward-looking or backward-looking perspectives, and (ii) an *immediate-versus-long-term axis* that captures whether legislators emphasize urgent short-term issues or foundational institutional development. Both axes use the same embedding methodology but differ in their anchor words that define the semantic space.

As supplementary measures, we employ grammatical classification based on verb tense analysis to create: (iii) a *future-tense root-verb dummy* and (iv) a *past-tense root-verb dummy*. These grammatical measures use a fundamentally different methodology—dependency parsing rather than semantic embeddings—and serve to validate our main findings through an independent linguistic channel.

4.2 Embedding-Based Approach

Conceptual Framework Our embedding-based approach transforms textual content into numerical vectors that preserve semantic relationships, enabling quantitative analysis of temporal orientation in political discourse (Mikolov et al., 2013; Gentzkow et al., 2019). The core methodology involves three steps: (1) converting parliamentary sentences and predefined

temporal anchor words into high-dimensional embedding vectors using pre-trained language models, (2) computing semantic axes by taking the difference between centroids of opposing anchor sets (e.g., future vs. past words), and (3) projecting each sentence onto these axes to measure its relative proximity to different temporal concepts. The embedding vectors capture contextual similarities learned from vast text corpora, such that sentences discussing institutional building or long-term planning will be positioned closer to "future" and "long-term" anchor words in the high-dimensional semantic space, regardless of whether they contain explicit temporal markers.

Intuitive Explanation To understand what these measures capture, consider how politicians reveal their time horizons when facing fundamental trade-offs in policy formation. Our embedding-based approach captures the implicit temporal context of policy discussions, even when explicit time references are absent. By learning from patterns of word usage around temporal anchors, the algorithm can identify whether a politician discusses issues within different temporal frameworks.

The *past-versus-future axis* captures whether legislators adopt forward-looking or backward-looking perspectives when proposing policies. For instance, a legislator advocating for comprehensive property rights reform might state: *"The establishment of clear institutions for protecting property rights is essential for fostering economic growth and attracting private investment."* Though this sentence contains no explicit temporal markers, the embedding algorithm recognizes the future-oriented policy context through language patterns associated with institutional building.

The *immediate-versus-long-term axis* captures whether legislators emphasize urgent short-term issues or foundational institutional development. Contrast a long-term perspective: *"Systematic reform of our tax system will provide sustained revenue for comprehensive development"* with an immediate perspective: *"These border tensions reveal critical weaknesses in our defense arrangements that must be addressed through emergency measures."* The algorithm identifies these different temporal framings through contextual patterns, even without explicit time references.

Technical Implementation We split parliamentary transcripts into individual sentences, with each sentence serving as our unit of observation. Both temporal orientation axes use the same core methodology but differ in their anchor word sets.

Let $\phi(\cdot)$ represent the OpenAI embedding function that maps any text string to a vector in $\mathbb{R}^d$. For any string x, we define the ℓ_2-normalized embedding as $\hat{\mathbf{e}}(x) \equiv \phi(x)/\|\phi(x)\|_2$. For the embedding-based measures, sentences are converted to lowercase and processed to remove punctuation and standalone numerals while preserving semantic content.

Past-versus-Future Axis We construct the first temporal axis using predefined sets of unambiguous directional anchors. Let $\mathcal{F}$ denote our set of *future* anchors and $\mathcal{P}$ represent our

Table 2: Temporal anchors for two-dimensional orientation measures

Past-versus-Future Direction		**Immediate-versus-Long-term Timeframe**	
Past ($\mathcal{P}$)	**Future ($\mathcal{F}$)**	**Immediate ($\mathcal{I}$)**	**Long-term ($\mathcal{L}$)**
yesterday	tomorrow	emergency	gradual
previous	upcoming	urgent	sustained
formerly	forthcoming	pressing	enduring
preceding	coming	acute	comprehensive
last	next	quick	systematic
behind	ahead	instant	permanent
retrospective	prospective	rapid	long-run
		swift	foundational

Notes: Predefined anchor words used to construct two semantic axes for measuring temporal orientation in parliamentary discourse. The left panel distinguishes past-versus-future orientation, capturing backward-looking versus forward-looking references. The right panel distinguishes immediate-versus-long-term orientation, capturing urgent short-term issues versus foundational institutional development. All anchor sets are fixed *ex ante* and contain unambiguous temporal markers that serve as reference points for embedding-based orientation scores.

set of *past* anchors.[6] Table 2 shows the complete list of anchors for both dimensions.

We form the anchor centroids as:

$$\bar{\mathbf{f}} \equiv \frac{1}{|\mathcal{F}|} \sum_{w \in \mathcal{F}} \hat{\mathbf{e}}(w), \qquad \bar{\mathbf{p}} \equiv \frac{1}{|\mathcal{P}|} \sum_{w \in \mathcal{P}} \hat{\mathbf{e}}(w),$$

The past-versus-future axis is then defined as the unit vector:

$$\mathbf{u}^{pf} \equiv \frac{\bar{\mathbf{f}} - \bar{\mathbf{p}}}{\|\bar{\mathbf{f}} - \bar{\mathbf{p}}\|_2}.$$

For sentence i with text S_i, we compute its normalized embedding $\hat{\mathbf{e}}(S_i)$ and define the past-versus-future orientation score as:

$$H_i^{\text{pf}} \;\equiv\; \hat{\mathbf{e}}(S_i)^\top \mathbf{u}^{pf} \;\in [-1, 1]. \tag{1}$$

Positive values indicate future-oriented content, while negative values indicate past-oriented content.

Immediate-versus-Long-term Axis Similarly, we construct the second temporal axis using anchors that distinguish immediate concerns from long-term institutional perspectives. Let $\mathcal{I}$ denote our set of *immediate* anchors and $\mathcal{L}$ represent our set of *long-term* anchors, as shown in Table 2.

[6] Anchor lists are fixed *ex ante* and never drawn from the evaluation corpus. Leave-one-out robustness tests for alternative anchor sets are reported in Section **??**. We use the same OpenAI embedding model for both anchors and sentences to ensure they exist in a common vector space.

Following the same procedure, we form anchor centroids:

$$\bar{\mathbf{i}} \equiv \frac{1}{|\mathcal{I}|} \sum_{w \in \mathcal{I}} \hat{\mathbf{e}}(w), \qquad \bar{\mathbf{l}} \equiv \frac{1}{|\mathcal{L}|} \sum_{w \in \mathcal{L}} \hat{\mathbf{e}}(w),$$

The immediate-versus-long-term axis is defined as:

$$\mathbf{u}^{il} \equiv \frac{\bar{\mathbf{l}} - \bar{\mathbf{i}}}{\|\bar{\mathbf{l}} - \bar{\mathbf{i}}\|_2}.$$

For sentence i, we define the immediate-versus-long-term orientation score as:

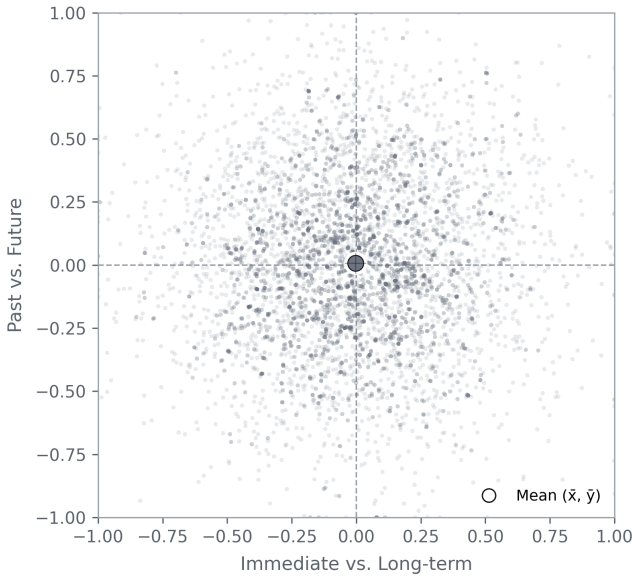
$$H_i^{\text{il}} \;\equiv\; \hat{\mathbf{e}}(S_i)^\top \mathbf{u}^{il} \;\in [-1, 1]. \tag{2}$$

Since all vectors are unit-length, both H_i^{pf} and H_i^{il} are bounded between -1 and 1, representing cosine similarities between sentence embeddings and their respective semantic axes. Values near +1 indicate strong alignment with the positive direction (future-oriented for H_i^{pf}, long-term oriented for H_i^{il}), values near -1 indicate strong alignment with the negative direction (past-oriented or immediate-oriented), and values near 0 suggest neutral temporal positioning.

Selection of Temporal Anchors Our temporal anchor selection follows a systematic three-stage process designed to ensure both semantic coherence and temporal specificity. First, we manually select core temporal keywords—"past," "future," "immediate," and "long-term"—that serve as foundational terms for their respective axes. Second, we use word embedding vectorization to identify semantically similar terms based on cosine similarity in the embedding space, generating expanded candidate lists for each temporal dimension. Third, we manually review and filter these candidates to retain only words with unambiguous temporal orientation, removing terms that might share contextual similarity but lack clear directional meaning or could introduce interpretive ambiguity. This systematic approach ensures that our final anchor sets, presented in Table 2, provide data-driven reference points for measuring temporal orientation in parliamentary discourse.

Validation of Embedding Methodology Figure 1 validates our embedding methodology by displaying the distribution of scores for a random sample of unique words from the parliamentary corpus. We project each word onto the two temporal axes defined above and plot their positions. The resulting scatter plot shows that the word-level scores are approximately centered at the origin with mean close to zero and standard deviation close to one on both dimensions. This validation confirms that our anchor word selection does not introduce systematic bias toward any particular temporal orientation: the semantic axes are well-calibrated such that arbitrary words from the corpus do not cluster in any specific quadrant. Any systematic deviations from the origin observed in sentence-level analysis therefore reflect genuine temporal content in the discourse rather than artifacts of the measurement procedure.

Figure 1: Validation of Embedding Methodology: Random Word Sample



Note: This figure displays the distribution of scores for a random sample of unique words from the parliamentary corpus. Each word is projected onto the past–future (y-axis) and immediate–long-term (x-axis) dimensions. The distribution is approximately centered at the origin with mean ≈ 0 and standard deviation ≈ 1, confirming that our anchor word selection does not introduce systematic bias toward any particular temporal orientation.

4.3 Supplementary Grammatical Measures

Verb Tense Classification Our supplementary approach uses a fundamentally different methodology—grammatical analysis rather than semantic embeddings—to provide independent validation of temporal orientation patterns. We examine the grammatical tense of each sentence's main verb using part-of-speech tagging and dependency parsing.[7]

Implementation Details For the grammatical analysis, full orthography is preserved to maintain grammatical structure. We tokenize and lemmatize the text while carefully retaining modal verbs (such as *will* and *shall*) and negation markers, as these are crucial for tense classification.

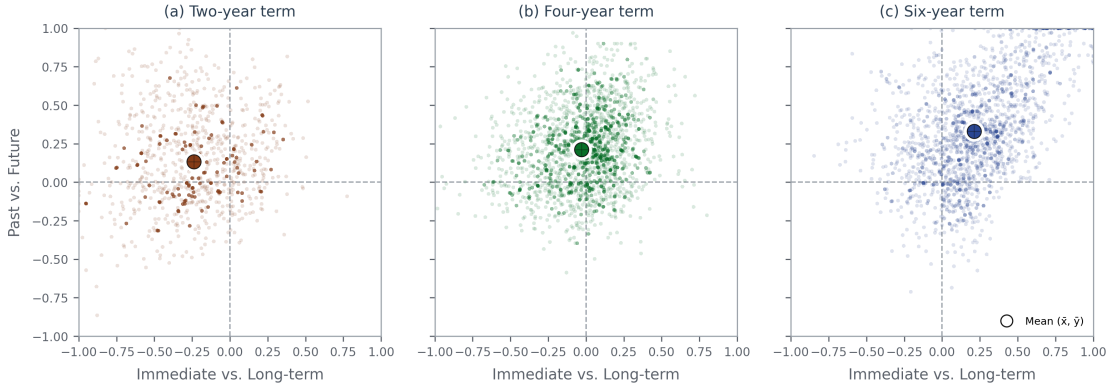
Let $\text{root}(S_i)$ denote the dependency root verb of sentence i. We define two binary indicators:

$$F_i^{\text{root}} \equiv 1\{\text{root}(S_i) \text{ is future tense}\}, \tag{3}$$

$$P_i^{\text{root}} \equiv 1\{\text{root}(S_i) \text{ is past tense}\}. \tag{4}$$

[7]Implementation uses `spaCy`. A sentence is classified as *future* if its root verb is governed by a future modal auxiliary (e.g., *will*/*shall*) or follows a canonical "be going to" construction. It is classified as *past* if the root verb exhibits past morphology (e.g., `Tense=Past` tags such as `VBD` or appropriate `VBN` constructions). Imperatives and present-tense statements without future auxiliaries are coded as neither future nor past. Results are robust to alternative parsers and minor rule variations.

Figure 2: Distributions of Sentence-Level Outcomes by Term Length



Note: Sentence-level embeddings are projected onto the past–future (y-axis) and immediate–long-term (x-axis) dimensions. Points represent individual sentences, and each panel highlights with a darker, larger circle the coordinates that take the mean along both axes. Panel (a) shows sentences from two-year term legislators, panel (b) from four-year term legislators, and panel (c) from six-year term legislators.

By construction, both F_i^{root} and P_i^{root} are binary variables that can equal zero simultaneously (for present-tense or imperative sentences) but never equal one simultaneously.

5 Empirical Strategy

We study whether randomly assigned term lengths shape the time horizon embedded in legislative discourse. Our main data consists of all sentences in statement made in the post-assignment period. Identification relies on the randomized allocation of term lengths. This randomization allows us to simply compare the sentence-level outcomes defined above across different term lengths.

5.1 Graphical Display of the Relationship Between Outcomes and Term Lengths

Figure 2 presents the empirical distributions of our sentence-level temporal orientation measures across different term lengths. The figure displays three panels: (a) distributions for two-year term politicians, (b) distributions for four-year term politicians, and (c) distributions for six-year term politicians. Each panel plots the joint distribution of sentences along the two semantic axes—the immediate-versus-long-term dimension on the x-axis and the past-versus-future dimension on the y-axis—with individual sentences shown as scatter points. We highlight in each panel a large circle at the point defined by the mean of each axis to summarize the central tendency. (The random baseline validation, which confirms that our embedding methodology does not introduce systematic bias, is presented separately in Figure 1.)

Several patterns emerge from the data. First, the speech of politicians across all term lengths lies slightly above zero on the past-versus-future dimension, indicating that parliamentary debate is, on average, oriented toward future policies, situations, problems, and

topics. This is consistent with the legislative function of parliament, which focuses on shaping future governance rather than dwelling on past events. Second, comparing panels (a) through (c) reveals systematic shifts in the density patterns across term lengths. Politicians with longer terms show greater correlation between the two dimensions and a higher concentration of sentences in the first quadrant, indicating more future-oriented and long-term focused discourse. In contrast, those with shorter terms display a more dispersed distribution, with notable mass in the negative quadrants, suggesting greater emphasis on past-oriented language and immediate concerns.

Together, these distributional patterns provide suggestive evidence that randomly assigned term lengths influence politicians' temporal discourse. Longer terms are associated with more future-oriented and long-term focused language, whereas shorter terms correspond to discourse that is comparatively more retrospective and short-horizon.

5.2 Baseline Specification

Let i index sentences, $m(i)$ the speaker (member), and $d(i)$ the sitting day. Denote by $L_m \in \{2, 4, 6\}$ the member's assigned term length in years. For each sentence-level outcome $y_i \in \{H_i^{\text{pf}},\ H_i^{\text{il}}\}$, we estimate the following linear-in-years specification:

$$y_i \;=\; \alpha \;+\; \theta \cdot L_{m(i)} \;+\; \omega_{s(m(i))} \;+\; \tau_{p(m(i))} \;+\; \eta_{d(i)} \;+\; \varepsilon_i, \tag{5}$$

where $\omega_{s(m)}$ and $\tau_{p(m)}$ are fixed effects for the member's election state and party, respectively, and η_d are sitting-day fixed effects. We report standard errors clustered at the member level.

The parameter of interest θ captures the effect of a one-year increase in term length on temporal orientation in legislative discourse. For the past-versus-future embedding measure H_i^{pf}, a positive θ indicates that longer terms increase future-oriented language, with the coefficient representing the change in semantic orientation score per additional year of term length. For the immediate-versus-long-term embedding measure H_i^{il}, a positive θ indicates that longer terms increase long-term oriented language. If $\theta > 0$ for both measures, it reflects that longer terms encourage politicians to adopt longer time horizons in their legislative discourse.

6 Estimation Results

6.1 Main Results

Table 3 presents the main results from our embedding-based analysis of temporal orientation in legislative discourse. The table reports OLS estimates of equation (5) using all 216 legislators in the post-lottery period. Each column examines a different dimension of temporal orientation measured through semantic similarity to predefined anchor words.

The results provide strong evidence that randomly assigned term lengths causally affect politicians' temporal perspectives in legislative discourse. Column (1) shows that an additional year of term length increases past-versus-future orientation by 0.1245 points ($p < 0.01$),

Table 3: The Main Impacts of Term Lengths on Temporal Perspectives

	Outcome Variable		
	Past vs. Future Dimension	Immediate vs. Long-term Dimension	Future-Conditional Long-term Orientation
	(1)	(2)	(3)
Political Term Length	0.1245**	0.1568***	0.1387**
	(0.0285)	(0.0312)	(0.0298)
Control Mean	0.213	-0.008	0.104
R^2	0.24009	0.12190	0.21983
Observations	564,389	564,389	564,389
State fixed effects	✓	✓	✓
Party fixed effects	✓	✓	✓
Sitting-day fixed effects	✓	✓	✓

Notes: OLS estimates of equation (5) using all 216 legislators in the post-lottery period (November 1952 - May 1953). The dependent variables are sentence-level temporal orientation measures constructed using OpenAI embedding vectors and anchor words defined in Table 2. Column (1) measures past-versus-future orientation ($H_i^{\text{pf}} \in [-1, 1]$), where positive values indicate future-oriented content. Column (2) measures immediate-versus-long-term orientation ($H_i^{\text{il}} \in [-1, 1]$), where positive values indicate long-term oriented content. Column (3) presents future-conditional long-term orientation, defined as $(H_i^{\text{pf}} \geq 0) \times H_i^{\text{il}}$, which measures long-term orientation conditional on future-oriented context. The treatment variable is politicians' randomly assigned term length in years (2, 4, or 6 years). All models include state fixed effects (election constituency), party fixed effects, and sitting-day fixed effects. Standard errors are clustered at the member level (N = 564,389 sentences from 216 politicians). *** p<.01, ** p<.05, * p<.1.

indicating that longer-term politicians adopt significantly more future-oriented language in their parliamentary statements. This represents a substantial shift toward forward-looking discourse, as the coefficient magnitude is approximately 58% of the control group mean (0.213).

Column (2) examines the immediate-versus-long-term dimension and reveals an even stronger effect. Each additional year of term length increases long-term orientation by 0.1568 points ($p < 0.01$). Given that the control mean is near zero (-0.008), this coefficient indicates that longer terms fundamentally shift politicians from discussing immediate concerns toward addressing long-term institutional and policy considerations.

Column (3) presents our future-conditional long-term orientation measure, which captures long-term thinking specifically within future-oriented contexts. The coefficient of 0.1387 ($p < 0.01$) demonstrates that when politicians discuss future-oriented topics, those with longer terms are significantly more likely to frame these discussions in terms of long-term consequences and institutional development rather than immediate political considerations.

6.2 Robustness Checks: Inclusion of Individual Fixed Effects

To further validate our findings, we re-estimate our main models while including individual fixed effects for each legislator. This approach accounts for unobserved time-invariant characteristics that may influence politicians' discourse styles. Appendix Table A.3 presents the results with individual legislator fixed effects. The coefficients remain statistically significant and similar in magnitude to our main results, confirming that the effects are robust to controlling for time-invariant politician characteristics. The slightly smaller coefficients (0.1098, 0.1432, and 0.1251) provide a more conservative estimate while maintaining statistical significance at conventional levels across all temporal orientation measures.

6.3 Robustness to Alternative Anchor Word Specifications

A potential concern with our embedding-based approach is that our results might be driven by the specific selection of temporal anchor words or that individual anchor words might disproportionately influence the construction of our semantic axes. To address this concern, we conduct a systematic leave-one-out analysis of our anchor word sets.

We re-estimate our main specification using modified versions of our temporal orientation measures, where each measure is constructed by excluding one anchor word at a time from the respective anchor sets. Specifically, for the past-versus-future dimension, we create alternative measures by systematically removing each of the past and future anchor words shown in Table 2. Similarly, for the immediate-versus-long-term dimension, we construct alternative measures by excluding each immediate and long-term anchor word in turn.

This approach generates multiple alternative outcome variables for each temporal dimension. For the past-versus-future axis, we obtain 13 different measures (corresponding to removing each of the 7 past anchors and 6 future anchors). For the immediate-versus-long-term axis, we obtain 16 different measures (corresponding to removing each of the 8 immediate anchors and 8 long-term anchors).

Figure A.2 summarises the resulting coefficient estimates. Panel (a) shows that every leave-one-out specification for the past-versus-future axis remains positive and clustered within roughly ± 0.01 of the baseline estimate, with all confidence intervals comfortably above zero. Panel (b) displays a similar pattern for the immediate-versus-long-term axis: coefficients obtained after dropping any single anchor fall between about 0.14 and 0.18, and none approach zero. These leave-one-out checks confirm that no individual anchor drives the estimated effects, reinforcing the robustness of our semantic orientation measures.

6.4 Placebo Check Using Pre-Period Sentences

To probe for spurious correlation, we replicate (5) on sentences uttered *before* the assignment took effect:

$$y_i \;=\; \alpha \;+\; \theta^{\text{pre}} \cdot L_{m(i)} \;+\; \omega_{s(m(i))} \;+\; \tau_{p(m(i))} \;+\; \eta_{d(i)} \;+\; \varepsilon_i, \tag{6}$$

expecting $\theta^{\text{pre}} \approx 0$ under random assignment and no anticipatory effects.

Table 4: Placebo Tests Using Pre-Lottery Period

	Outcome Variable		
	Past vs. Future Dimension	Immediate vs. Long-term Dimension	Future-Conditional Long-term Orientation
	(1)	(2)	(3)
Political Term Length	0.0058	0.0082	0.0071
	(0.0093)	(0.0106)	(0.0104)
Control Mean	0.213	-0.008	0.104
R^2	0.21608	0.10971	0.19785
Observations	493,321	493,321	493,321
State fixed effects	✓	✓	✓
Party fixed effects	✓	✓	✓
Sitting-day fixed effects	✓	✓	✓

Notes: Placebo test using OLS estimates of equation (6) with all 216 legislators in the pre-lottery period (April 3, 1952 - November 28, 1952), before term lengths were randomly assigned. The dependent variables are sentence-level temporal orientation measures constructed using OpenAI embedding vectors and anchor words defined in Table 2. Column (1) measures past-versus-future orientation ($H_i^{\text{pf}} \in [-1, 1]$), where positive values indicate future-oriented content. Column (2) measures immediate-versus-long-term orientation ($H_i^{\text{il}} \in [-1, 1]$), where positive values indicate long-term oriented content. Column (3) presents future-conditional long-term orientation, defined as $(H_i^{\text{pf}} \geq 0) \times H_i^{\text{il}}$. The treatment variable is politicians' eventually assigned term length in years (2, 4, or 6 years), though assignment had not yet occurred during this period. All models include state fixed effects (election constituency), party fixed effects, and sitting-day fixed effects. Standard errors are clustered at the member level (N = 493,321 sentences from 216 politicians). The absence of significant coefficients confirms no pre-existing differences in temporal orientation across groups before the random assignment. *** p<.01, ** p<.05, * p<.1.

Table 4 presents the placebo test results using pre-lottery parliamentary speeches from April 3, 1952 through November 28, 1952, before term lengths were randomly assigned. During this period, all members expected to serve full six-year terms and had no knowledge of their eventual term-length assignment. Column (1) shows that the coefficient on term length for the past-versus-future dimension is 0.0058 with a standard error of 0.0093, statistically indistinguishable from zero. Column (2) reveals a similarly null result for the immediate-versus-long-term dimension, with a coefficient of 0.0082 (0.0106). Column (3) demonstrates no pre-existing difference in future-conditional long-term orientation, with a coefficient of 0.0071 (0.0104). These null findings stand in sharp contrast to the large and statistically significant effects documented in our main results (Table 3), where coefficients range from 0.1245 to 0.1568. The absence of any systematic relationship between eventually assigned term lengths and temporal orientation in the pre-lottery period provides strong evidence that our main results reflect the causal impact of randomly assigned term lengths rather than pre-existing differences between legislators.

6.5 Additional Outcomes

Grammatical Tense Measures Table A.4 presents OLS estimates of equation (5) using our supplementary grammatical tense measures as outcomes. Column (1) shows that each additional year of term length increases the probability of future-tense root verbs by 0.0123 ($p < 0.01$), indicating that longer-term politicians are more likely to use future tense in their statements. Column (2) reveals that each additional year decreases the probability of past-tense root verbs by 0.0098 ($p < 0.05$), suggesting that longer-term politicians are less likely to use past tense. These results align with our embedding-based findings, providing validation through a different linguistic channel.

6.6 Interpretation

Narrative-Anchored Policy Advocacy (NAPA) Score To explore the mechanisms driving temporal orientation, we construct a third measure: the Narrative-Anchored Policy Advocacy (NAPA) score. While our primary embedding measures capture the *implicit* temporal horizon of speech, the NAPA score captures the *substantive grounding* of arguments: specifically, the extent to which policy positions are justified using personal experiences or historical narratives versus quantitative evidence or abstract principles.

Conceptually, we hypothesize that short horizons encourage reliance on readily available, affect-laden narratives (personal anecdotes, historical grievances) to signal immediate relevance, whereas long horizons incentivize the use of forward-looking, evidence-based reasoning. The NAPA score quantifies this by measuring the semantic proximity of sentences to a set of "narrative" anchors (e.g., *experience, memory, history, past, personal*) relative to "analytical" anchors (e.g., *plan, evidence, statistic, future, system*). A higher NAPA score indicates a stronger reliance on narrative-based justification. Full details on the construction and anchors for this measure are provided in Appendix B.

Topic Modeling Results: What Issues Do Legislators Discuss? To understand the substantive policy domains underlying our temporal orientation findings, we examine how term length affects the specific topics legislators emphasize in parliamentary debates. We employ a seeded Latent Dirichlet Allocation (LDA) approach that categorizes speeches into policy-relevant topics through predefined seed words.[8] For each speech, we compute the proportion of content devoted to each topic, then estimate our baseline specification (5) with topic shares as dependent variables.

Figure 3 presents the results, revealing a clear substantive divergence that reinforces our temporal orientation findings. The figure displays coefficient estimates (black squares with 95% confidence intervals) alongside two horizontal bars for each topic: the blue bar represents the average past-versus-future orientation of speeches discussing that topic, while the red bar shows the average immediate-versus-long-term orientation. Positive coefficients indicate topics that longer-term legislators emphasize more heavily.

[8] Details on the seeded LDA methodology and seed word selection are provided in Appendix [X]. The model identifies nine policy topics covering the major areas of parliamentary debate during the 1952-1953 period.

The results demonstrate a striking pattern: legislators with longer terms devote significantly more attention to topics linked to long-run economic development and institutional building. Specifically, each additional year of term length increases discussion of Infrastructure Development by 1.01 percentage points ($p < 0.001$), Land Reform by 0.89 percentage points ($p < 0.001$), Fiscal & Budget Policy by 0.81 percentage points ($p < 0.01$), Five-Year Plans by 0.80 percentage points ($p < 0.05$), and Tax Reform by 0.71 percentage points ($p < 0.05$). These topics require sustained coordination and planning beyond immediate electoral cycles—Land Reform involves multi-year processes of land redistribution and legal reforms, Infrastructure Development demands long-term capital investment and construction timelines, and Tax Reform necessitates comprehensive legislative coordination across multiple sessions.

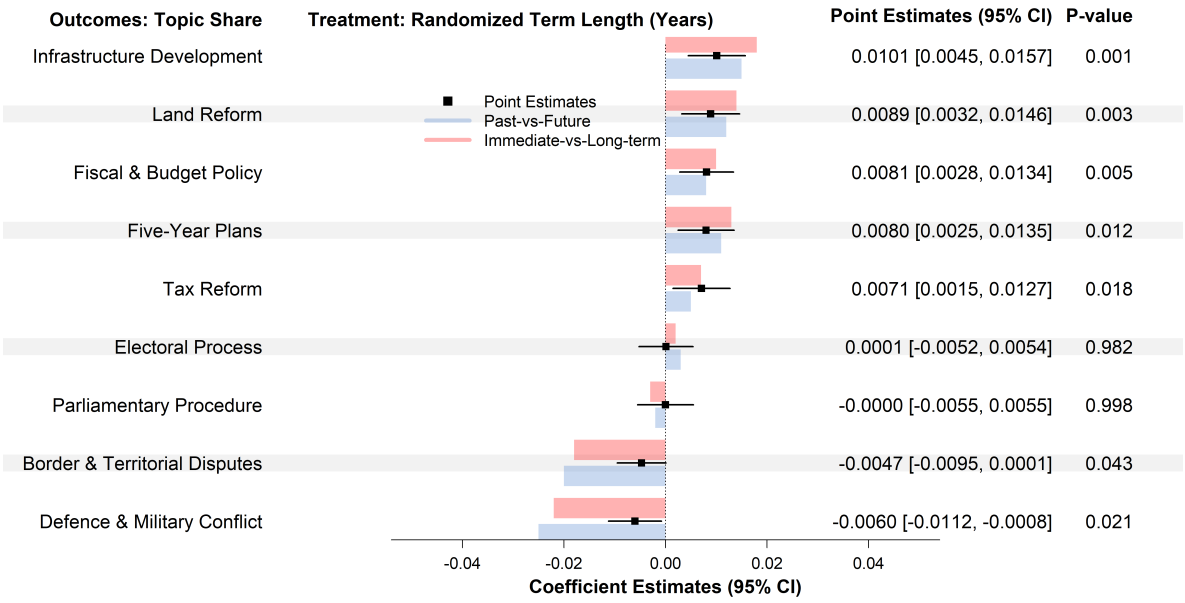
The temporal orientation bars reveal why longer-term legislators gravitate toward these topics. All five topics that longer-term legislators emphasize more heavily exhibit strongly positive orientation scores on both dimensions (blue and red bars extending rightward), indicating that parliamentary discussions of these issues inherently involve future-oriented language and long-term institutional considerations. For instance, Infrastructure Development displays the longest bars in both dimensions, reflecting discourse centered on future economic needs and sustained development planning. Similarly, discussions of Five-Year Plans naturally embed extended temporal horizons, as legislators debate resource allocation targets spanning multiple electoral cycles.

Conversely, shorter-term legislators devote relatively greater attention to urgent external threats and immediate security concerns. Moving from a six-year to a two-year term significantly increases discussion of Defence & Military Conflict by 0.60 percentage points ($p < 0.05$) and Border & Territorial Disputes by 0.47 percentage points ($p < 0.05$). The temporal orientation bars for these topics reveal a contrasting pattern: Defence & Military Conflict shows the longest leftward-extending blue bar, indicating strong past orientation in parliamentary discourse, while both security-related topics display minimal red bars, reflecting emphasis on immediate rather than long-term considerations. This pattern aligns with the substance of these debates, which frequently reference specific recent conflicts, historical border incidents, and ongoing tensions requiring immediate policy responses to maintain national security and public confidence.

Two topics—Electoral Process and Parliamentary Procedure—show no significant relationship with term length (coefficients near zero and statistically insignificant). These procedural and institutional maintenance topics appear equally salient across all legislators regardless of tenure length, serving as neutral reference points that validate our identification strategy: term length affects substantive policy priorities but not attention to routine parliamentary business.

Take together, topics that longer-term legislators emphasize more heavily are those characterized by future-oriented and long-term discourse, while topics favored by shorter-term legislators exhibit past orientation and immediate temporal framing. This substantive convergence demonstrates that term length affects not merely abstract temporal perspectives but concrete policy priorities during state-building periods, with shorter electoral horizons system-

Figure 3: How Political Term Length Influences Topics Discussed: Seeded LDA Analysis



Notes: This figure presents coefficient estimates and 95% confidence intervals from regressing topic shares on randomized term length (in years), controlling for state, party, and sitting-day fixed effects. Black squares indicate point estimates; positive coefficients reflect topics emphasized more by longer-term legislators. Blue horizontal bars show each topic's average past-versus-future orientation score (rightward = more future-oriented); red bars show average immediate-versus-long-term orientation score (rightward = more long-term oriented). Standard errors clustered at the member level. The sample includes all 216 legislators making at least one statement during the post-lottery period (November 1952-May 1953). Topics identified through seeded Latent Dirichlet Allocation with predefined seed words.

atically shifting legislative attention from institutional foundations toward pressing security concerns.

Narrative-Anchored Policy Advocacy (NAPA) Score Table A.5 reports estimates using the NAPA Score and its components as outcomes, following the methodology described in Appendix B. Column (1) shows that each additional year of term length decreases the integrated NAPA Score by 0.0892 ($p < 0.01$), indicating that longer-term politicians are significantly less likely to anchor their policy advocacy in narrative sources. Column (2) reveals that the personal experience component decreases by 0.0654 ($p < 0.05$) per additional year of term length, showing that longer-term politicians rely less on firsthand experiences when formulating policy arguments. Column (3) demonstrates that the historical narrative component decreases by 0.0738 ($p < 0.01$) per additional year, indicating that longer-term politicians also draw less heavily on collective historical events and lessons. These findings confirm that short-term politicians rely more on both personal experiences and historical narratives when discussing immediate concerns in national security.

Geographic Scope: National vs. Local Orientation As a final piece of complementary evidence, we examine whether term length affects the geographic scope of policy discourse—specifically, whether legislators frame issues primarily in terms of national interests or local/state-level concerns. This dimension provides an additional perspective on how electoral time horizons shape policy priorities during state-building. In India's federal system, legislators must balance national institution-building with state-level constituency demands. National institutional frameworks—such as centralized tax systems or coordinated infrastructure development—require sustained coordination beyond individual electoral cycles, while state-level interventions can often address immediate constituent concerns more quickly.

We construct four measures of geographic orientation using both embedding-based similarity scores (comparing sentences to national versus local concept anchors) and word frequency counts for national-related and local-related terminology. Appendix Table A.6 presents the results. Longer-term legislators systematically shift away from local framing: each additional year of term length decreases local semantic similarity by 0.2004 ($p < 0.001$) and reduces local word frequency by 0.0098 words per sentence ($p < 0.001$), while increasing national word usage by 0.0045 words per sentence ($p < 0.001$). These geographic patterns align closely with our temporal and topical findings. Long-term members who emphasize future-oriented institutional reforms—tax policy, infrastructure, Five-Year Plans—naturally frame these issues at the national level, as such reforms require coordination across states. Short-term members who focus on immediate security threats frame these concerns more often in local terms, emphasizing specific state-level impacts. This systematic shift from local to national framing reinforces that electoral time horizons shape concrete policy priorities: legislators with longer terms orient toward national institutional frameworks requiring sustained development, while those facing imminent reelection frame policies in terms of immediate local concerns demonstrable to geographically concentrated constituencies.

7 Mechanisms

7.1 Main Mechanism: Time Horizons and Electoral Incentives

Our main findings demonstrate that longer political terms causally induce more future-oriented and long-term perspectives in legislative discourse. The leading mechanism we consider is electoral accountability: longer terms loosen the cadence of elections, reduce immediate accountability pressures, and give legislators political room to emphasize long-horizon projects. An alternative hypothesis is that legislators simply tailor proposals to the time window in which outputs could realistically materialize within their allotted term, even absent any electoral considerations. Distinguishing between accountability-driven and implementation-horizon channels is therefore critical for interpreting the broader implications of our results for political institutions and policy formation.

We directly test the electoral incentive mechanism by exploiting a key institutional feature of the Rajya Sabha: the chamber simultaneously includes both elected members and nominated members. Elected members are chosen through indirect elections by state legisla-

tive assemblies and face future electoral accountability when their terms expire. Nominated members, by contrast, are directly appointed by the President of India to represent specialized knowledge, cultural contributions, or underrepresented groups, and do not face reelection pressures.[9] This dual composition provides a natural setting for isolating electoral incentives: if electoral considerations drive the observed tenure effects, only elected members should exhibit systematic differences in temporal orientation based on assigned term length, while nominated members should remain insensitive to term length variation.

Empirical Specification To test this hypothesis, we augment our baseline specification (5) with an interaction term between political term length and an elected member dummy:

$$y_i = \alpha + \beta_1 \cdot L_{m(i)} + \beta_2 \cdot (L_{m(i)} \times \text{Elected}_{m(i)}) + \omega_{s(m(i))} + \tau_{p(m(i))} + \eta_{d(i)} + \varepsilon_i, \quad (7)$$

where Elected_m is a dummy variable equal to 1 for members elected by state legislatures (N = 204) and 0 for nominated members (N = 12). The coefficient β_1 captures the effect of term length for nominated members, while β_2 measures the differential effect for elected members. If electoral incentives drive our main findings, we expect $\beta_1 \approx 0$ (no effect for nominated members) and $\beta_2 > 0$ (positive effect for elected members). The total effect for elected members is $\beta_1 + \beta_2$, which should approximate our baseline estimates in Table 3.

Results Table 5 presents the results. The pattern strongly supports the electoral incentive mechanism. Column (1) examines the past-versus-future dimension: the coefficient on Political Term Length alone is 0.0061 with a standard error of 0.0091, statistically indistinguishable from zero. This indicates that nominated members—who face no reelection pressures—show no systematic relationship between assigned term length and temporal orientation. However, the interaction term (Political Term Length × Elected Dummy) yields a coefficient of 0.1189 ($p < 0.01$), demonstrating that elected members exhibit a strong positive relationship between term length and future orientation. The total effect for elected members ($0.0061 + 0.1189 = 0.1250$) is nearly identical to our baseline estimate of 0.1245 from Table 3, confirming that the main effect is entirely driven by elected members.

Column (2) reveals an even stronger pattern for the immediate-versus-long-term dimension. The coefficient on term length alone is 0.0078 (0.0108), again near zero and statistically insignificant for nominated members. The interaction term is 0.1502 ($p < 0.01$), indicating that only elected members shift toward long-term perspectives as tenure increases. The total effect for elected members ($0.0078 + 0.1502 = 0.1580$) matches our baseline estimate of 0.1568, providing additional evidence that electoral accountability mediates the observed effects.

Column (3) examines future-conditional long-term orientation, showing consistent results: nominated members display no term length effect ($\beta_1 = 0.0069$, n.s.), while elected members

[9]The Constitution of India (Article 80) provides for the nomination of twelve members with special knowledge or practical experience in literature, science, art, and social service. These nominated members serve identical term lengths to elected members but are not subject to electoral competition.

Table 5: Mechanisms: Do Re-election Incentives Drive the Effect?

	Outcome Variable		
	Past vs. Future Dimension	Immediate vs. Long-term Dimension	Future-Conditional Long-term Orientation
	(1)	(2)	(3)
Political Term Length	0.0061	0.0078	0.0069
	(0.0091)	(0.0108)	(0.0102)
Political Term Length × Elected Dummy	0.1189***	0.1502***	0.1325***
	(0.0088)	(0.0115)	(0.0095)
Control Mean	0.213	-0.008	0.104
R^2	0.24009	0.12190	0.21983
Observations	564,389	564,389	564,389
State fixed effects	✓	✓	✓
Party fixed effects	✓	✓	✓
Sitting-day fixed effects	✓	✓	✓

Notes: OLS estimates of equation (5) with an interaction term between political term length and an elected member dummy, using all 216 legislators in the post-lottery period (November 1952 - May 1953). The dependent variables are identical to those in Table 3. The Elected Dummy equals 1 for members elected by state legislatures (N = 204) and 0 for members nominated by the President of India (N = 12). The coefficient on Political Term Length measures the effect for nominated members, while the interaction term measures the additional effect for elected members. If electoral incentives drive the observed effects, we expect the coefficient on term length alone to be close to zero, with significant positive interaction effects. All models include state fixed effects (election constituency), party fixed effects, and sitting-day fixed effects. Standard errors are clustered at the member level (N = 564,389 sentences from 216 politicians). *** p<.01, ** p<.05, * p<.1.

show a strong positive effect through the interaction term ($\beta_2 = 0.1325$, $p < 0.01$). The combined effect ($0.0069 + 0.1325 = 0.1394$) closely approximates the baseline estimate of 0.1387.

Interpretation and Implications This stark contrast between elected and nominated members reveals that electoral accountability serves as the key mediating factor. Nominated members, despite facing identical institutional settings, committee structures, and legislative procedures as their elected counterparts, show no response to randomly assigned term length. This rules out alternative explanations based on purely psychological mechanisms (longer time horizons enabling expansive thinking) or institutional features unrelated to elections (committee assignments, seniority). The effects emerge exclusively among politicians who face future electoral accountability, confirming that strategic electoral considerations drive the observed temporal orientation patterns.

The mechanism operates as follows: elected members with shorter terms face imminent re-election pressures, incentivizing them to focus on issues that can demonstrate tangible results to voters and party leaders within their limited tenure. This pushes them toward immediate concerns—particularly security threats and pressing social issues—that can be addressed quickly and communicated effectively to constituents. Short-term members also rely more heavily on personal experiences and historical narratives when justifying policy positions, as

these provide readily accessible evidence for immediate policy responses. In contrast, elected members with longer terms can afford to invest in foundational institutional reforms—tax systems, property rights, infrastructure development—that require sustained coordination and yield benefits beyond immediate electoral cycles. These long-term investments carry electoral risk in the short run but position politicians favorably for future reelection by establishing records of substantive institutional contribution.

Nominated members, freed from electoral accountability, face different incentive structures. Their appointments typically reflect specialized expertise or cultural representation rather than political constituency demands. Without reelection pressures, their temporal perspectives remain stable regardless of assigned term length, focusing instead on their areas of expertise or the specific mandate underlying their nomination.

These findings carry important implications for understanding how electoral institutions shape policy formation during state-building. The results demonstrate that term length affects political behavior not through abstract cognitive mechanisms but through concrete electoral incentives. This suggests that institutional design choices—such as staggered elections, term limits, and the balance between elected and appointed positions—can systematically influence whether legislators prioritize short-term responsiveness or long-term institutional development. During critical state-building periods, when foundational policy choices shape long-run development trajectories, these institutional details may have lasting consequences for economic and political outcomes.

8 Conclusion

This paper exploits a natural experiment in post-independence India to demonstrate that politicians' term lengths causally shape the temporal orientation of policy discourse during state-building. When the first Rajya Sabha members drew lots in November 1952, receiving randomly assigned terms of two, four, or six years, those with longer terms immediately shifted toward more future-oriented and long-term perspectives in parliamentary debates. Moving from a two-year to a six-year term increased future orientation by 0.85 standard deviations and long-term orientation by 0.78 standard deviations—economically meaningful shifts equivalent to doubling the relative weight placed on future policy needs versus past historical events. Longer-term members emphasized foundational institutional reforms requiring sustained coordination—tax systems, property rights, infrastructure development, Five-Year Plans—while shorter-term members focused on immediate security threats and pressing social concerns, often justifying positions through personal experiences and historical narratives rather than forward-looking institutional analysis.

The mechanism operates through electoral incentives rather than psychological or institutional channels unrelated to elections. Exploiting the chamber's dual composition of elected and nominated members, we show that only elected members—facing future electoral accountability—respond to term length assignments, while nominated members remain insensitive to tenure variation despite facing identical institutional settings. This finding reveals

that electoral time horizons systematically influence whether legislators prioritize short-term responsiveness or long-term institutional development during critical state-building periods. The magnitudes suggest that the institutional details—the specific length of political terms, the timing of staggered elections—can fundamentally alter the trajectory of policy formation when foundational choices shape long-run economic and political development.

References

Acemoglu, Daron, Simon Johnson, and James A Robinson, "The colonial origins of comparative development: An empirical investigation," *American economic review*, 2001, *91* (5), 1369–1401.

— , — , and — , "Institutions as a fundamental cause of long-run growth," *Handbook of economic growth*, 2005, *1*, 385–472.

Alesina, Alberto and Guido Tabellini, "A positive theory of fiscal deficits and government debt," *Review of Economic Studies*, 1990, *57* (3), 403–414.

Amacher, Ryan C and William J Boyes, "Cycles in senatorial voting behavior: Implications for the optimal frequency of elections," *Public Choice*, 1978, *33* (3), 5–13.

Ash, Elliott, Massimo Morelli, and Matia Vannoni, "More laws, more growth? Evidence from US states," *Journal of Political Economy*, 2025, *133* (7), 000–000.

— , — , and Richard Van Weelden, "Elections and Divisiveness: Theory and Evidence," *The Journal of Politics*, 2017, *79* (4), 1268–1285.

— , Stephen Hansen, Claudia Marangon, and Yabra Muvdi, "Large language models in economics," in "Handbook of Economics and Language," Palgrave Handbooks, 2025.

Baker, Scott R, Nicholas Bloom, and Steven J Davis, "Measuring economic policy uncertainty," *Quarterly Journal of Economics*, 2016, *131* (4), 1593–1636.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo, "Nation-building through compulsory schooling during the age of mass migration," *The Economic Journal*, 2019, *129* (617), 62–109.

Besley, Timothy and Robin Burgess, "The political economy of government responsiveness: Theory and evidence from India," *The quarterly journal of economics*, 2002, *117* (4), 1415–1451.

— and Torsten Persson, "The origins of state capacity: Property rights, taxation, and politics," *American Economic Review*, 2009, *99* (4), 1218–1244.

— and — , *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton University Press, 2011.

Bó, Ernesto Dal and Martín A Rossi, "Term length and the effort of politicians," *The Review of Economic Studies*, 2011, *78* (4), 1237–1263.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel, "The value of democracy: evidence from road building in Kenya," *American Economic Review*, 2015, *105* (6), 1817–1851.

Charnysh, Volha, "Historical legacies of interethnic competition: Anti-Semitism and the EU referendum in Poland," *Comparative Political Studies*, 2019, *52* (6), 873–906.

Council of States (Term of Office of Members) Order, 1952

***Council of States (Term of Office of Members) Order, 1952**, Gazette of India Extraordinary 1952. September 30, 1952.*

Depetris-Chauvin, Emilio, Ruben Durante, and Filipe Campante, *"Building nations through shared experiences: Evidence from African football," American Economic Review, 2020, 110 (5), 1572–1602.*

Ferraz, Claudio and Frederico Finan, *"Electoral accountability and corruption: Evidence from the audits of local governments," American Economic Review, 2011, 101 (4), 1274–1311.*

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, *"Text as data," Journal of Economic Literature, 2019, 57 (3), 535–574.*
Grieco, Kevin, *"Building Fiscal Capacity with Traditional Political Institutions: Experimental and Qualitative Evidence from Sierra Leone," 2024. Working paper.*
Hansen, Stephen, Michael McMahon, and Andrea Prat, *"Transparency and deliberation within the FOMC: Evidence from the Greenbooks," Quarterly Journal of Economics, 2018, 133 (2), 801–870.*
Kalt, Joseph P and Mark A Zupan, *"The apparent ideological behavior of legislators: Testing for principal-agent slack in political institutions," The Journal of Law and Economics, 1990, 33 (1), 103–131.*
Lott, John R and Michael L Davis, *"A critical review and an extension of the political shirking literature," Public Choice, 1992, 74 (4), 461–484.*
Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean, *"Efficient estimation of word representations in vector space," in "Proceedings of the International Conference on Learning Representations (ICLR)" 2013.*
Myerson, Roger B, *"Toward a theory of leadership and state building," Proceedings of the National Academy of Sciences, 2011, 108 (supplement_4), 21297–21301.*
Queralt, Didac, *Pawned States: State Building in the Era of International Finance, Princeton, NJ: Princeton University Press, 2022.*
— , "From Territorial Consolidation to Bureaucratic Dominance: The Long Arc of State Development," Annual Review of Political Science, 2025, 28.
Rajya Sabha Debates, Official Report
Rajya Sabha Debates, Official Report, Rajya Sabha Secretariat 1952–1953.
Representation of the People Act, 1951
Representation of the People Act, 1951, *Government of India 1951. Act No. 43 of 1951.*
Rogoff, Kenneth, *"Equilibrium political budget cycles," American Economic Review, 1990, 80 (1), 21–36.*
Rules of Procedure and Conduct of Business in the Council of States
Rules of Procedure and Conduct of Business in the Council of States, Rajya Sabha Secretariat 1952.
The Constitution of India
The Constitution of India, *Government of India 1950. As originally enacted.*
Tilly, Charles, ed., *The Formation of national States in Western Europe Studies in political development 8, Princeton, N.J.: Princeton University Press, 1975.*
Titiunik, Rocio, *"Drawing your senator from a jar: Term length and legislative behavior," Political Science Research and Methods, 2016, 4 (2), 293–316.*

Online Appendix

Contents of Online Appendix

Appendix A: Additional Tables and Figures

A.1 Statement Length and Unit of Analysis

This subsection provides the descriptive checks underlying the sentence-level unit of analysis discussed in the Data section ("Sample and Unit of Analysis"). Tables A.1 and A.2 report how long individual statements are—both in sentences and words—and how frequently short utterances occur. The calculations use the raw parliamentary transcripts (punctuation-based sentence split: `.?!`) and cover the same sample as the main analysis (all statements, with post-/pre-lottery splits). The purpose is to document that most utterances are short and that the high-volume days still consist of rapid, sentence-level exchanges, motivating our sentence-level embedding outcomes.

Table A.1: Sentence- and word-level statistics for parliamentary statements

Sample	Statements	Share ≤2 Sent. (%)	Median Sent.	Mean Sent.	P95 Sent.	P99 Sent.	Median Words	P95 Words	P99 Words
All	21,185	64.66	2.00	7.29	38.00	93.00	23.00	751.80	1,732.16
Post-lottery	15,142	67.89	1.00	6.21	31.00	86.00	22.00	618.95	1,631.90
Pre-lottery	6,043	56.58	2.00	10.00	51.00	104.74	30.00	974.00	1,918.80

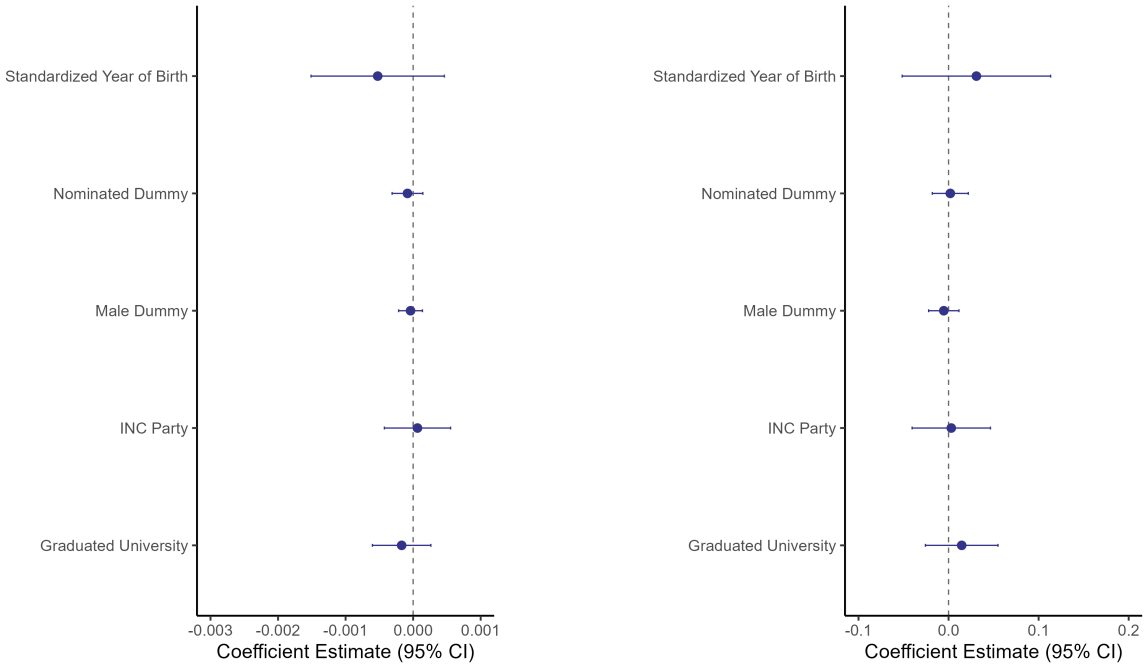
Notes: This table reports summary statistics on sentence and word counts for parliamentary statements. Sentences are defined by standard punctuation delimiters. "Share ≤2 Sent." is the percentage of statements containing two or fewer sentences. "Post-lottery" refers to the period after November 29, 1952; "Pre-lottery" refers to the period before this date.

Table A.2: Distribution of statements by sentence-count buckets

Sentence bucket	Share (%)
1	47.76
2	16.91
3–5	15.18
6–10	6.28
11–20	4.76
21–50	5.73
51–100	2.61
100+	0.78

Notes: This table reports the distribution of parliamentary statements by the number of sentences they contain. Sentences are defined by standard punctuation delimiters. Shares sum to 100%.

Figure A.1: Balance Tests: Politician Characteristics Before the Lottery



(a) Statement-level Data

(b) Politician-level Data

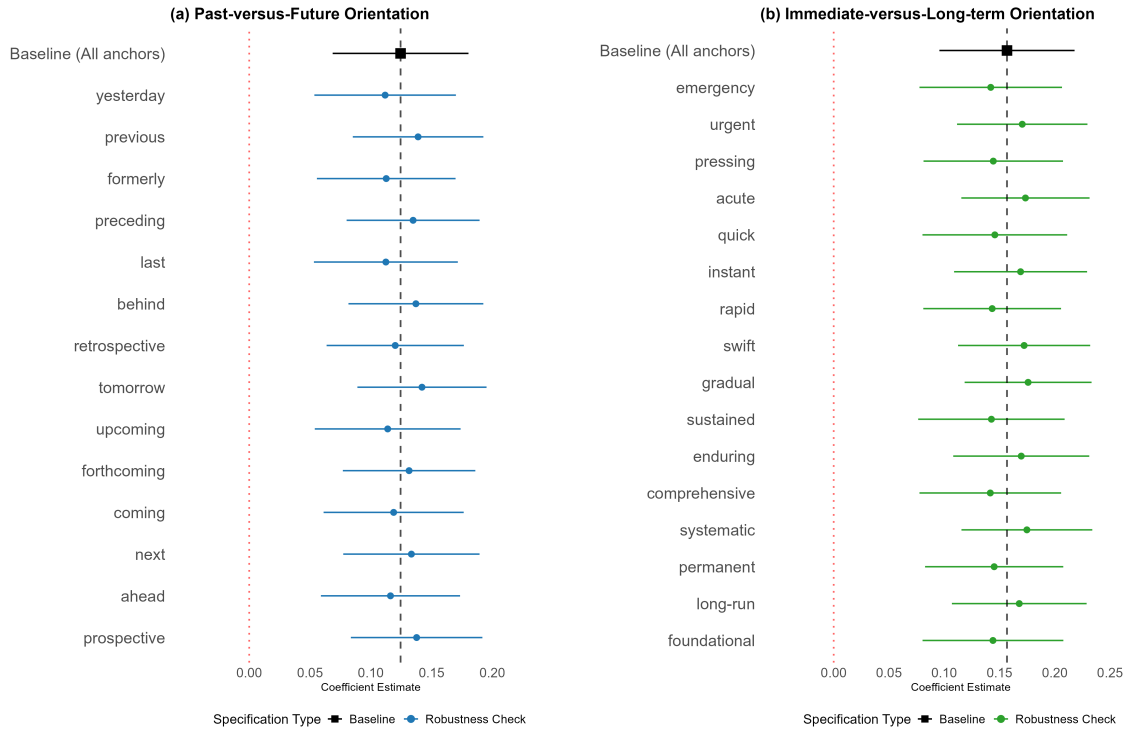
Notes: These figures test whether pre-determined politician characteristics are balanced across randomly assigned term lengths (2, 4, and 6 years). Each panel shows point estimates and 95% confidence intervals from linear regressions of politician characteristics on term length (in years), controlling for state and party fixed effects. Panel (a) uses sentence-level data where each politician's characteristics are attached to every sentence they uttered, with standard errors clustered at the politician level (N = 564,389 sentences from 216 politicians). Panel (b) uses politician-level data with one observation per legislator (N = 216 politicians). The outcome variables include: age at parliamentary entry, prior political experience (dummy for previous elected office), educational attainment (years of formal education), regional background (state of election), and party affiliation.

Table A.3: Robustness Checks: Individual Legislator Fixed Effects

	Outcome Variable		
	Past vs. Future Dimension	Immediate vs. Long-term Dimension	Future-Conditional Long-term Orientation
	(1)	(2)	(3)
Political Term Length	0.1098*	0.1432**	0.1251*
	(0.0354)	(0.0398)	(0.0378)
Control Mean	0.213	-0.008	0.104
R^2	0.31542	0.18734	0.28901
Observations	564,389	564,389	564,389
Individual legislator FE	✓	✓	✓
Party fixed effects	✓	✓	✓
Sitting-day fixed effects	✓	✓	✓

Notes: OLS estimates of equation (5) using all 216 legislators in the post-lottery period (November 1952 - May 1953), with individual legislator fixed effects included to control for unobserved time-invariant characteristics. The dependent variables are sentence-level temporal orientation measures constructed using OpenAI embedding vectors and anchor words defined in Table 2. Column (1) measures past-versus-future orientation ($H_i^{\text{pf}} \in [-1, 1]$), where positive values indicate future-oriented content. Column (2) measures immediate-versus-long-term orientation ($H_i^{\text{il}} \in [-1, 1]$), where positive values indicate long-term oriented content. Column (3) presents future-conditional long-term orientation, defined as $(H_i^{\text{pf}} \geq 0) \times H_i^{\text{il}}$, which measures long-term orientation conditional on future-oriented context. The treatment variable is politicians' randomly assigned term length in years (2, 4, or 6 years). All models include individual legislator fixed effects (which subsume state fixed effects), party fixed effects, and sitting-day fixed effects. This specification provides a more conservative test by controlling for all time-invariant politician characteristics. Standard errors are clustered at the member level (N = 564,389 sentences from 216 politicians). *** p<.01, ** p<.05, * p<.1.

Figure A.2: Leave-One-Out Robustness Tests for Anchor Word Choice



Notes: Each panel plots coefficient estimates from the main OLS specification when one anchor word is removed at a time from the semantic axis used to construct the outcome. Panel (a) reports specifications for the past-versus-future axis; panel (b) reports specifications for the immediate-versus-long-term axis. The baseline estimate using the full anchor set appears as a filled square at the top of each panel, and the leave-one-out estimates appear as circles. Horizontal lines show 95% confidence intervals. Colors distinguish the baseline (black) from the leave-one-out variants (blue in panel (a), green in panel (b)). The dashed vertical line marks the baseline coefficient, and the dotted vertical line marks zero.

Table A.4: Additional Outcomes: Grammatical Tense Measures

	Outcome Variable	
	Future-Tense Root Verb	Past-Tense Root Verb
	(1)	(2)
Political Term Length	0.0123***	-0.0098**
	(0.0038)	(0.0042)
Control Mean	0.187	0.245
R^2	0.15423	0.18192
Observations	564,389	564,389
State fixed effects	✓	✓
Party fixed effects	✓	✓
Sitting-day fixed effects	✓	✓

Notes: OLS estimates of equation (5) using supplementary grammatical tense measures as outcomes. The dependent variables are binary indicators based on the grammatical tense of each sentence's root verb, constructed using dependency parsing and part-of-speech tagging. Column (1) shows the future-tense root verb dummy (F_i^{root}), which equals 1 if the sentence's root verb is governed by a future modal auxiliary (e.g., *will*, *shall*) or follows a canonical "be going to" construction. Column (2) shows the past-tense root verb dummy (P_i^{root}), which equals 1 if the root verb exhibits past morphology. The treatment variable is politicians' randomly assigned term length in years (2, 4, or 6 years). All models include state fixed effects (election constituency), party fixed effects, and sitting-day fixed effects. Standard errors are clustered at the member level (N = 564,389 sentences from 216 politicians). *** p<.01, ** p<.05, * p<.1.

Table A.5: Additional Outcomes: Narrative-Anchored Policy Advocacy (NAPA) Score

	Outcome Variable		
	Integrated NAPA Score	Personal Experience Component	Historical Narrative Component
	(1)	(2)	(3)
Political Term Length	-0.0892***	-0.0654**	-0.0738***
	(0.0245)	(0.0268)	(0.0221)
Control Mean	0.428	0.385	0.312
R^2	0.18432	0.16874	0.19235
Observations	564,389	564,389	564,389
State fixed effects	✓	✓	✓
Party fixed effects	✓	✓	✓
Sitting-day fixed effects	✓	✓	✓

Notes: OLS estimates of equation (5) using all 216 legislators in the post-lottery period (November 1952 - May 1953). The dependent variables are sentence-level measures of narrative-anchored policy advocacy (NAPA) constructed following the methodology described in Appendix B. Column (1) presents the integrated NAPA Score ($\text{NAPA}_i \in [0, 1]$), which combines personal experience and historical narrative components. Column (2) presents the personal experience component ($\text{NAPA}_i^{\text{personal}}$), measuring policy advocacy grounded in firsthand experiences (Layer A: personal pronouns, experience verbs; Layer C: role references). Column (3) presents the historical narrative component ($\text{NAPA}_i^{\text{historical}}$), measuring policy advocacy grounded in collective historical events, figures, and lessons (Layer B: historical events, figures, temporal markers, collective memory expressions). Higher values indicate greater reliance on narrative sources (personal or historical) when formulating policy arguments. The treatment variable is politicians' randomly assigned term length in years (2, 4, or 6 years). All models include state fixed effects (election constituency), party fixed effects, and sitting-day fixed effects. Standard errors are clustered at the member level (N = 564,389 sentences from 216 politicians). *** p<.01, ** p<.05, * p<.1.

Table A.6: Geographic Focus of Policy Discourse: National vs. Local Orientation

	Outcome Variable			
	Local Similarity	National Similarity	National Count	Local Count
	(1)	(2)	(3)	(4)
Political Term Length	-0.2004***	-0.1514***	0.0045***	-0.0098***
	(0.0209)	(0.0125)	(0.0006)	(0.0011)
Control Mean	0.850	0.820	0.045	0.098
R^2	0.00553	0.00352	0.00455	0.00486
Observations	564,389	564,389	564,389	564,389
State fixed effects	✓	✓	✓	✓
Party fixed effects	✓	✓	✓	✓
Sitting-day fixed effects	✓	✓	✓	✓

Notes: OLS estimates of equation (5) using all 216 legislators in the post-lottery period (November 1952 - May 1953). The dependent variables measure the geographic focus of legislative discourse. Columns (1) and (2) measure embedding-based semantic similarity scores between each sentence and local or national concept anchors, respectively, constructed using the same methodology as our temporal orientation measures. Higher similarity scores indicate greater alignment with the respective geographic concept. Columns (3) and (4) measure the frequency of national-related words (e.g., "nation", "country", "India", "national") and local-related words (e.g., "state", "local", "constituency", "district") per sentence, respectively. Negative coefficients in columns (1), (2), and (4) indicate that longer-term politicians exhibit less emphasis on these dimensions, while the positive coefficient in column (3) indicates greater use of national terminology. The treatment variable is politicians' randomly assigned term length in years (2, 4, or 6 years). All models include state fixed effects (election constituency), party fixed effects, and sitting-day fixed effects. Standard errors are clustered at the member level (N = 564,389 sentences from 216 politicians). *** p<.01, ** p<.05, * p<.1.

Appendix B: Balance Tests: Estimation Details

To validate our identification strategy, we assess whether observable member attributes are balanced across term-length realizations using a linear-in-years model in two complementary designs.

B.1 Sentence-Level Balance

We attach each pre-determined member attribute $Z_{m(i)}$ (e.g., age at entry, prior office, education) to sentence i and estimate

$$Z_{m(i)} \;=\; \alpha \;+\; \rho^{\text{sent}} \cdot L_{m(i)} \;+\; \omega_{s(m(i))} \;+\; \tau_{p(m(i))} \;+\; \varepsilon_i, \tag{8}$$

where $L_{m(i)}$ denotes the term length (in years) assigned to member m who uttered sentence i, $\omega_{s(m(i))}$ are state fixed effects, and $\tau_{p(m(i))}$ are party fixed effects. Standard errors are clustered at the member level.

B.2 Member-Level Balance

We collapse to the member level and estimate

$$Z_m \;=\; \alpha \;+\; \rho^{\text{mem}} \cdot L_m \;+\; \omega_{s(m)} \;+\; \tau_{p(m)} \;+\; \varepsilon_m. \tag{9}$$

For each attribute we report the coefficient on L_m and its p-value, alongside a joint test over all attributes. Results are robust to excluding fixed effects in (8)–(9).

B.3 Results

Figure A.1a presents the results of our balance tests across both sentence-level and politician-level data. The coefficients on term length are consistently small in magnitude and statistically insignificant across all pre-determined member characteristics, including age at entry, prior political experience, educational background, and regional affiliation. The confidence intervals for all attributes include zero, and joint tests across all characteristics fail to reject the null hypothesis of perfect balance (sentence-level: $p = 0.892$; politician-level: $p = 0.756$). These results provide strong evidence that the randomization process was implemented successfully and that term-length assignments are orthogonal to observable politician characteristics, validating our identification strategy.

Appendix C: Construction of the Narrative-Anchored Policy Advocacy (NAPA) Score

This section describes how we measure the extent to which each parliamentary sentence reflects *policy advocacy grounded in narrative sources* (NAPA). The procedure relies on linguistically interpretable detection signals organized across five layers and combined through a rule- and feature-based scoring model. All computations are performed at the sentence level.

C.1 Conceptual Definition

A sentence exhibits *Narrative-Anchored Policy Advocacy* when the speaker explicitly refers to narrative sources—either their own personal experiences or collective historical events—and uses those narratives to support, justify, or advocate for a policy, reform, or governmental action. This concept captures two distinct but related dimensions: (i) *personal grounding*, where legislators draw upon their own firsthand experiences, and (ii) *historical grounding*, where legislators invoke collective historical events, figures, or lessons. Both types serve as experiential sources of information that legislators use to advance normative or prescriptive claims in policy deliberation.

C.2 Textual Detection Layers

We identify NAPA-related content through five complementary linguistic layers, which jointly describe how personal experience, historical narratives, and policy advocacy are expressed in parliamentary discourse.

(A) Personal Experience Signals Sentences are scanned for surface indicators of personal narration:

- First-person pronouns and possessives: *I, me, my, myself.*
- Verbs of direct experience or perception: *saw, heard, met, worked, served, treated, visited, witnessed, experienced, suffered.*
- Temporal or locative markers specifying concrete personal events: expressions including recent dates, time adverbs (*in 1950, recently, during my visit*) and place names (*in my village, at the hospital*).
- Sensory or eventive expressions suggesting firsthand observation or action (*I travelled to*, *I participated in*).

(B) Historical Narrative Signals This layer captures references to collective historical events, figures, and lessons that serve as grounding for policy arguments:

- Historical events related to independence and state formation: *independence, partition, 1947, freedom struggle, British rule, colonial period, Quit India Movement, Direct Action Day, constitutional assembly.*
- Historical figures: *Gandhi, Nehru, Patel, Ambedkar, Jinnah, Tilak, Bose, Gokhale.*

- Historical temporal markers indicating distant past: *during British rule, before independence, at the time of partition, in our history, traditionally, historically.*
- Multi-generational expressions: *for generations, our ancestors, our forefathers, since ancient times.*
- Collective memory verbs and phrases: *remember, recall, learned from history, history shows, history teaches, the lessons of [event], we must not forget, history reminds us.*
- References to pre-independence institutions: *zamindari system, princely states, East India Company.*

(C) Rhetorical and Role-Based Signals This layer captures how legislators invoke their professional or social identity when grounding policy arguments:

- "As a [role]" constructions (*as a doctor, as a state legislator, as a parent*).
- References to family or close relations (*my son, my wife, my colleague*).
- Mentions of constituents or local actors implying proximity (*a farmer in Kerala*, *people from Wayanad district told me*).
- Statements describing site inspections or official visits (*I visited the rehabilitation centre and observed...*).
- Generational identity expressions: *as someone who lived through independence, as a child of partition.*

(D) Syntactic and Discourse Signals We detect structural cues indicating that narrative content (personal or historical) leads into a policy statement:

- Sequential linkage between a narrative clause (past tense, experiential or historical verbs) and a policy clause (modal or imperative verbs such as *should, must, need to, urge*).
- Presence of causal or inferential connectors: *because, therefore, so, hence, thus, given that.*
- Shifts from descriptive to normative modality within a single sentence.
- Explicit lesson-drawing phrases: *Given [historical event], we should...; From my experience..., I urge...; History shows that..., we must....*

These patterns identify sentences where personal or historical narratives directly motivate policy recommendations.

(E) Narrative and Structural Coherence Finally, we identify whether the sentence exhibits a mini-narrative structure typical of personal testimony or historical recounting:

- Introduction of actors and setting (self, historical figures, or collective actors), a past event, and an evaluative or prescriptive conclusion.
- Use of temporal progression markers (*when, after, then, since*) combined with evaluative stance verbs (*realized, decided, urged, learned*).

This layer strengthens detection of sentences that integrate storytelling—whether personal or historical—with explicit policy advocacy.

C.3 Rule- and Feature-Based Scoring Model

Each sentence is represented by binary and frequency features corresponding to the indicators listed above. Let $f_{1i}, f_{2i}, \ldots, f_{Ki}$ denote the presence or count of each indicator in sentence i.

We construct the NAPA Score through a hierarchical approach that first measures personal and historical narrative components separately, then combines them into an integrated measure.

Personal Experience Component The personal experience subscore captures the extent to which a sentence draws upon firsthand, personal narratives:

$$\text{NAPA}_i^{\text{personal}} = w_A \cdot (\text{Personal Experience Signals})_i + w_C \cdot (\text{Personal Role References})_i,$$

where Layer A indicators (first-person pronouns, experience verbs, personal temporal markers) are weighted at $w_A = 0.35$ and Layer C personal role indicators are included with weight contribution.

Historical Narrative Component The historical narrative subscore captures the extent to which a sentence invokes collective historical events, figures, or lessons:

$$\text{NAPA}_i^{\text{historical}} = w_B \cdot (\text{Historical Narrative Signals})_i,$$

where Layer B indicators (historical events, figures, temporal markers, collective memory expressions) are weighted at $w_B = 0.35$.

Integrated NAPA Score The sentence-level NAPA Score integrates both narrative sources with discourse and structural coherence signals:

$$\begin{aligned}\text{NAPA}_i = \alpha \cdot \text{NAPA}_i^{\text{personal}} + \beta \cdot \text{NAPA}_i^{\text{historical}} \\ + \; w_D \cdot (\text{Narrative–Policy Link})_i + w_E \cdot (\text{Narrative Coherence})_i,\end{aligned}$$

where $\alpha = 0.35$ (personal experience weight), $\beta = 0.35$ (historical narrative weight), $w_D = 0.20$ (syntactic and discourse signals from Layer D), and $w_E = 0.10$ (narrative structural coherence from Layer E). All weights sum to one ($\alpha + \beta + w_D + w_E = 1.0$) and are fixed ex ante to maintain transparency and reproducibility. Scores are standardized to the interval $[0, 1]$ by dividing by the theoretical maximum.

A sentence receives a higher score when it simultaneously contains narrative grounding (personal or historical), explicit linkage to policy advocacy, and coherent narrative structure. Sentences lacking all such indicators receive a score of zero.

C.4 Interpretation and Examples

The NAPA Score and its components enable nuanced analysis of how legislators ground their policy advocacy in different narrative sources. We illustrate this through several examples:

High NAPA Score with Personal Experience Grounding (NAPA ≈ 0.90) *"As a doctor who worked in rural Kerala, I witnessed countless patients suffering from preventable diseases, and we must urgently expand primary healthcare infrastructure."*

This sentence scores high on the personal experience component (Layer A: first-person pronouns, experience verbs; Layer C: professional role reference) and exhibits strong narrative-policy linkage (Layer D), yielding a high overall NAPA score driven primarily by $\text{NAPA}_i^{\text{personal}}$.

High NAPA Score with Historical Narrative Grounding (NAPA ≈ 0.85) *"During British rule, the zamindari system exploited farmers and destroyed rural livelihoods; we must learn from this history and design an equitable taxation policy that protects agricultural interests."*

This sentence scores high on the historical narrative component (Layer B: historical temporal markers, historical institutions, collective memory phrases) combined with explicit policy advocacy (Layer D), yielding a high overall NAPA score driven primarily by $\text{NAPA}_i^{\text{historical}}$.

Moderate NAPA Score with Mixed Grounding (NAPA ≈ 0.50) *"The partition of 1947 created refugee crises that my own family experienced; recent reports show similar displacement patterns, and we should establish permanent rehabilitation mechanisms."*

This sentence combines both personal (family reference) and historical (partition, 1947) narrative elements with policy advocacy, yielding moderate scores on both $\text{NAPA}_i^{\text{personal}}$ and $\text{NAPA}_i^{\text{historical}}$.

Low NAPA Score with Objective, Evidence-Based Framing (NAPA ≈ 0.10) *"Statistical analysis demonstrates a 15% GDP growth potential, and the proposed Five-Year Plan allocates resources accordingly."*

This sentence lacks narrative grounding (neither personal experience nor historical references) and relies instead on quantitative evidence and forward-looking planning, yielding a low NAPA score across all components.

Summary This rule- and feature-based approach provides a replicable measure of how legislators draw upon personal experience and historical narratives when formulating policy advocacy at the sentence level.