THE UNIVERSITY OF CHICAGO

Gender Identity and Representation in the Context of Economic Development in India

By

Bhavya Pandey

April 2024

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science with a concentration in Economics

Faculty Advisor: Professor Christina Brown, PhD
Secondary Reader: Professor Fulya Ersoy, PhD

Abstract

Does legislators' gender identity affect representation, indicated by their participation in legislative activity? I ask whether their gender impacts the extent to which elected representatives present and relay issues to seek government action on. Are females inherently different as legislators in the democracy? This project uses a natural language processing approach along with a dynamic difference-in-differences strategy to investigate. Using a zero-shot classification technique, I create five outcome "scores" to evaluate the Indian Parliament's textual Question Hour data on: development, corruption, accountability & transparency, programmatic representation, and clientilistic representation. The analysis reveals that female representation has minimal to no discernible normalized average total effect per treatment unit on development scores. Slight negative effects are observed in the cases of programmatic and clientelistic representation scores. However, a modest negative effect on accountability and transparency scores contradicts expectations, suggesting nuanced interactions that may not universally improve governance quality through increased female representation. In addition, the only positive effect, though small, on corruption scores reveals the priorities of female representation over time periods. These findings underscore the complexity of gender effects across different governance dimensions and temporal contexts, highlighting the importance of individual temporal effects and their association with specific electoral cycles in informing the dynamism in representation in developing contexts.

Keywords: Political economy; development; dynamic difference-in-differences, natural language processing, classification

1 Introduction & Literature Review

On 19 September 2023, the Constitution Bill (One Hundred and Twenty-Eighth Amendment) Bill 2023 was introduced in the Lok Sabha - the House of the People or the lower house of India's bicameral Parliament - which seeks to reserve one third of all seats for women in the Lok Sabha and the State Legislative Assemblies. The proposal of this legislation comes at a crucial time for India, which seeks to achieve the status of a 'developed' country by the year 2047. Looking at the history of women representation in Lok Sabha, the percentage remained about a mere 5 until the 1970s, and it was only in 2009 that it could reach a double-digit figure. The status quo warrants a question about representation and women's representation in particular, within the context of economic development, in India.

On the one hand, in recent decades, developing nations such as India have experienced a reduction in gender bias when it comes to political participation (Kapoor and Ravi, 2014) through the tool of universal adult franchise. While the share of male voters has been fairly stagnant, that of female voters has been on the rise. However, despite this increased participation, the limited representation and success of women candidates in parliamentary and state elections raise doubts about whether there has been a significant representation of nearly half the population of the country at different levels of governance. Political economists have long maintained that underrepresentation of women in leadership positions threatens democracy and hence negatively impacts development (Wittman, 1989).

Some literature suggests that women like other minorities and historically marginalised groups tend to be better representatives of their group or community because of a shared sense of culture, experience, and traditions (Phillips, 1995). On the contrary, literature also supports that women legislators prefer to let themselves not be seen only as representatives of women and thrust emphasis on general participation in legislative activities (Diamond, 1977). Regardless, the multitude of reasons for pushing for a greater representation as well as participation of women in the public sphere through political activity are clear and evidenced by scholars – this enables all kinds of women to have a better voice in decision-making, combats the issue of women being relegated to the private sphere, and somewhere or the other, helps accommodate their perspectives and interests through a channel that can truly affect policies. The milieu of these ideas eventually affects the manner in which a government ends up responsible to the diverse particular groups that it serves.

Against the backdrop of India's aspiration for development by 2047, the historical underrepresentation of women in legislative bodies prompts a critical examination. Political

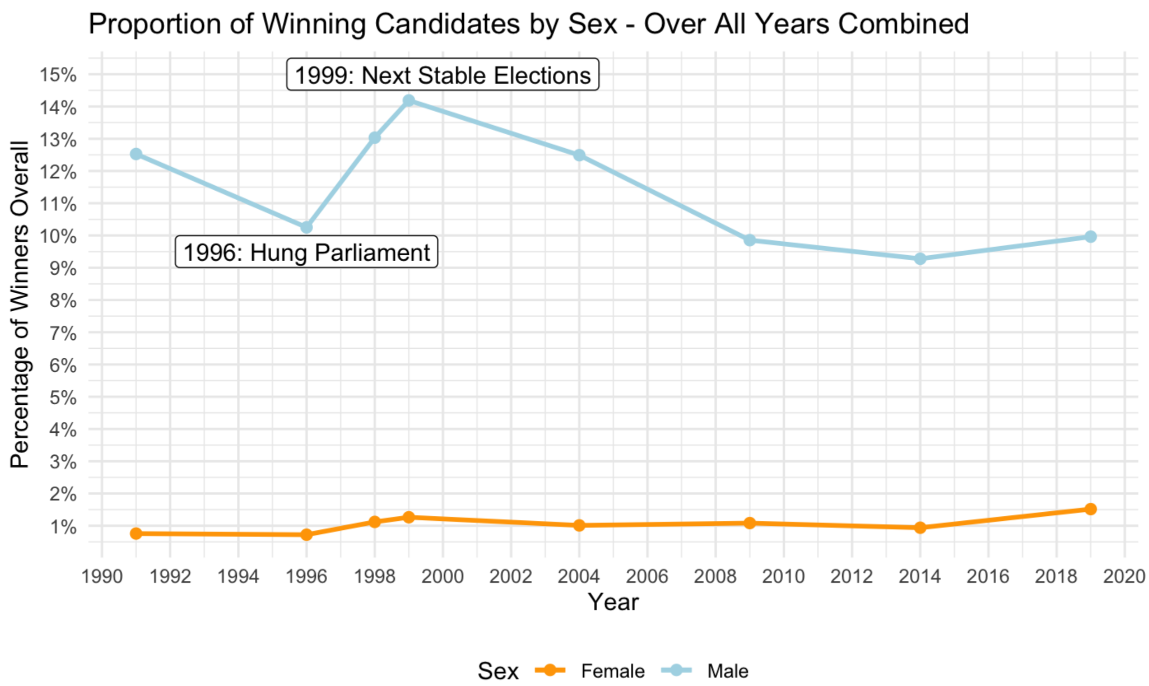

Figure 1: *Proportion of Winners in General Elections, by Gender*

Share of winning candidates amongst all candidates over the years, by the gender of the candidate. The year 1996 saw a hung parliament followed by two elections in quick succession in 1998 and 1999.

scientists and sociologists have theoretically reflected on some of the possible reasons that could explain the small number of female politicians in India (Hussain, 2022):

- There exist systemic electoral disadvantages to women in the electoral system
- Male candidates are more likely to have an incumbency as well as remarkably greater fiscal edge in elections
- In societies deep-rooted in patriarchy such as India, individual attitudes are difficult to disentangle from political outcomes
- As a result, male leaders tend to be characterized as rationality and resilience – fit for roles in political leadership
- Whereas women are considered to be characterized by docility and emotionality – unsuitable traits for leaders.

Besides the discussion on the participation of women in legislative bodies, there is much less evidence about their performance or behavior as representatives. While several studies on parliamentary questions in democracies examine the motivations behind asking a large number of questions on specific topics, including gender, caste, direct electoral links, experience, seniority, and constituency demographics (Bird, 2005, Ayyangar and Jacob, 2015, Bandiera et al., 2020), the extant literature on parliamentary questions in India is focused on quantifying involvement of individual MPs, and the analysis of the question content from a political economy of development angle remains largely unexplored (Sen et al., 2019).

Parliamentary questions can therefore be studied more effectively by grouping data according to the area of interest that is being covered. Questions' policy foci serve as a stand-in for knowledge of the House Members' true priorities and concerns. In the specific context of gendering parliamentary questions for example, evidence from Britain (Bird, 2005) has shown that women members of parliament were more likely than their male colleagues to refer to 'women' and 'gender' in both written and oral questions. The male members were more inclined to refer to 'men than their female colleagues. Although the questions address a wide range of concerns, the members of parliament shared a common understanding of which issues should be linked to 'women'. Representations of traditional manhood or womanhood, however, upheld conservative gender roles and risked essentialising sexual categories.

This project seeks to contribute to the growing body of political economy literature examining the role of legislators in developing democracies by leveraging data from the Indian Parliament and Election Commission, and a dynamic difference-in-differences identification strategy. The research design for this project can be defined in two stages – the creation of the outcome variable with a causal interpretation using the Question Hour text as data, followed by the execution of a dynamic differences-in-difference strategy in multiple time periods.

After trying a series of computational methods including word embeddings (using the common *word2vec* technique) and topic modelling (using Latent Dirichlet Allocation (LDA)), a zero-shot classification approach was adopted to appropriately use the question 'text as data'.

Since I wanted to ascertain the effects as a result of the election of a male versus female legislator from a constituency, over multiple elections, upon my outcome variables, I turn to a dynamic or generalized difference-in-differences estimator which accommodates irregular adoption periods. In other words, different units (constituencies) may enter into treatment at different times. This is often the case in practice, where some policy is rolled out to aggregate level entities (i.e., districts, counties, states, etc.) over a long time horizon. The generalized or dynamic differences-in-differences estimator can accommodate periods of treatment withdrawal, and even multiple treatment histories which is particularly relevant in my setting as I are bound to have constituencies switching between male and female legislators over multiple elections.

1.1 Representation

As the world's largest democracy, India is often faced with the challenge of representation. In her pioneering work on representation, Hannah Pitkin put forth that 'representation is acting in the interests of those being represented' (Pitkin, 1972: 209) and only this be deemed as *substantive representation.* It refers to all those activities of the representative performed on the terms, and absolute translation of will of those being represented. The performance of the representative is thus evaluated by the effectiveness and the extent to which the representative advances the policy outcomes serving the best interests of whom they represent (Hussain, 2022). In this vein, scholars partly shirk the concept of *descriptive representation*, the simple increase in numbers of women legislators in legislative bodies or positions of authority, and partly assert that descriptive representation will lead to substantive representation.

This leads us to further question the manner in which identity and representation intersect to affect the behavior of women legislators in the Parliament – the notion that women represent women's issues better and prove to be better performing representatives for their interests. Some literature suggests that women like other minorities and historically marginalised groups tend to be better representatives of their group or community because of a shared sense of culture, experience, and traditions (Phillips, 1995). On the contrary, literature also supports that women legislators prefer to let themselves not be seen only as representatives of women and thrust emphasis on general participation in legislative activities (Diamond, 1977).

Regardless, the multitude of reasons for pushing for a greater representation as well as participation of women in the public sphere through political activity are clear and evidenced by scholars – this enables all kinds of women to have a better voice in decision-making, combats the issue of women being relegated to the private sphere, and somewhere or the other, helps accommodate their perspectives and interests through a channel that can truly affect policies. The milieu of these ideas eventually affects the manner in which a government ends up responsible to the diverse particular groups that it serves.

All this boils down to motivate my research question: does legislators' identity affect descriptive representation reflected through their participation or performance in legislative activity? This paper seek to estimate the causal effect of the gender of the legislator on their propensity to participate in the Parliament in a certain way – raising particular issues relevant to their gender, conforming to gender-norms while participating in debates, or taking interest in identity-conforming ideas through their politics. Essentially, I ask whether gender impacts the extent to which elected representatives translate and relay issues to seek government action. In other words: "How does the representation in the Indian Parliament influence the legislative focus on policies and issues?"

1.2 Socio-economic Significance

Two key papers form the basis of the socio-economic significance of investigating this question. Primarily, Chattopadhyay and Duflo, 2004 have evidenced through a randomized control trial that greater representation of women in local government bodies helps people better through differential allocation of public goods by female politicians. Specifically, leaders invest more in infrastructure that is directly relevant to the needs of their own genders. The study finds significant evidence that the reservation of Village Council heads for women leads to policy choices more reflective of women's preferences, particularly in areas directly relevant to women's needs. This suggests that women leaders prioritize different public goods than their male counterparts, such as infrastructure that benefits women directly (e.g., water facilities). The paper builds on the "citizen-candidate" model to explain the observed policy shifts under women's leadership. The authors propose that women, when elected, bring distinct preferences to the policy-making process, which are not merely a reflection of the electorate's median preferences but are also shaped by gender-specific priorities.

Continuing the exploration from this paper, Jacob, 2014 finds that female legislators ask 24% fewer parliamentary questions, and there are also significant differences in question content, even after controlling for important covariates. Using the intuition that the gender of winners in close elections between women and men is quasi-randomly determined, this article employs a regression discontinuity design at the margin of victory to reestimate the gender effect. The results indicate that gender has zero causal effect on the volume of questions. However, this analysis paves the way for a deeper analysis into the question content as well as mechanisms which may describe the differences between male and female legislators due to unobserved attributes that affect legislative activity – this is where my study picks up the analysis from – making use of more textual data, leveraging computational tools, and finding effect sizes over time using the dynamic differences-in-difference approach.

Secondly, Baskaran et al., 2018 find substantial evidence indicating a notable increase in economic activity within constituencies that choose to elect women, and there is no indication of adverse effects spreading to nearby constituencies led by men. This observation aligns with an overall positive trend in growth. Leveraging a regression discontinuity design, they assert that female legislators tend to exhibit lower tendencies toward criminal behavior and corruption, demonstrate greater effectiveness, and are less susceptible to political opportunism. While this suggests that women legislators are effective to democratic principles in policy implementation, it remains unclear whether the participation of women in decision-making reflects a similar trend at the apex level, and to which extent their participation is affected by their male counterparts in the Parliament. Further, it is also estimable whether female legislators gravitate toward passive 'softer issues' (Hussain, 2022), targeted interventions for geographically localised issues, or active identity-based issues through their legislative participation, forming the foundational basis for this study.

Essentially, this project seeks to assess the causal effect of legislator's gender on participation and issue advocacy within the Parliament. In doing so, it aims to contribute to contemporary political-economic ideas by elucidating how gender dynamics influence legislative behavior and its implications for governance outcomes thereby impacting development progress. Ultimately, it envisions to deconstruct the multifaceted role of gender identity in shaping political representation and its ramifications for achieving inclusive and responsive governance, adding an essential layer to the ongoing discourse on women's participation in democratic processes by connecting disparate threads, and shedding light on the complex relationship between gender identity, representation, and legislative performance.

The paper is organized as follows: this section discussed the contours of relevant literature and motivations, Section 2 lays down the background and setting, Section 3 delves into the data and computational methods, Section 4 and Section 5 get into the model and empirical strategy in-depth, respectively, Section 6 shares results, while Section 7 and Section 8 conclude with a discussion into the contributions of the paper as well as scope for furthering this research.

2 Background

India has a parliamentary system, as defined by its constitution, with power divided between the central government and the states. India's democracy is the largest in the world. General elections for Members of Parliament (elected to the Lok Sabha – also known as the House of Commons or Lower House in the bicameral parlimentary system) are conducted every five years. The Prime Minister of India, who is the leader of the party or political alliance having a majority in the general elections to the Lok Sabha is the leader of the executive branch of the Government of India.

The Lok Sabha has 543 constituencies across the nation, seats to which are filled by election using first-past-the-post voting, plus two seats to which representatives from the Anglo-Indian community are nominated by the President of India. The Election Commission of India, an autonomous statutory body with powers established by the Constitution, conducts the entire electoral exercise. Last in 2019, it was overseeing an election in which there are 815 million registered voters of which over 100 million are new voters. The mind-boggling logistics include the deployment of 11 million poll and security personnel at 830,000 polling stations. India has a multi-party system with the 7 registered parties at the national level. The three largest parties in India are the Indian National Congress (INC), the Bharatiya Janata Party (BJP), and the Communist Party of India (CPI).

Members of the Lok Sabha (House of the People), or the lower house of India's Parliament, are elected by a vote of all adult Indian citizens over the age of 18 from a list of candidates running in their constituencies. Every adult Indian citizen can only vote in their constituency. Candidates who win Lok Sabha elections are referred to as 'Members of Parliament' and serve for five years, or until the body is dissolved by the President on the advice of the Council of Ministers. The house meets in the Lok Sabha Chambers of the Sansad Bhavan in New Delhi to discuss the creation of new laws, as well as the repeal or improvement of existing laws affecting all Indian citizens.

In a democracy like India, elections serve as a test of citizen's preferences. Once elected, the Parliament becomes an arena for the winning candidates to perform this duty. This typically involves attending parliamentary sessions, serving as members of committees, voting on legislation, participating in debates and other proceedings in the parliament as a legislator, in addition to serving their constituents as a public representative.

The first hour of a day's proceedings in both the Upper and Lower House of the Indian Parliament is devoted to the 'Question Hour'. During this time, Members of Parliament or MPs across parties raise questions on every aspect of administration and government activity. Each question is posed to a relevant ministry, and it is the prerogative of the Ministers associated with their respective ministries to provide adequate information and clarifications on matters of public concern. The questions presented by representatives provide valuable insights into the active participation, priorities, and issues that concern the elected representatives in the Parliament.

Question Hour is the first hour of a sitting session of the Lok Sabha devoted to questions that Members of Parliament raise about any aspect of administrative activity. The

> (a) the State-wise details of the Skill Development centres established in the country especially for women; (b) the number of women imparted training in these centres and the number of women who got employment; (c) the location-wise number of skill development centres being run in the country; (d) whether the Government is considering to increase the number of skill development centres to provide self employment to the unemployed youth with a view to end unemployment in the country; and (e) if so, the number of proposals and places where such centres are proposed to be set up in the country and if so, the details thereof and if not, the reasons therefor?

Figure 2: *Sample Question Text*

An example query raised by a female legislator to the then Minister of Skill Development.

concerned Minister is obliged to answer to the Parliament, either orally or in writing, depending on the type of question raised and the text of answers as well as question is to be submitted in writing to the Secretariat. Records have been maintained historically, and this textual data has been made available by the Lok Sabha Secretariat for public use. Figure 2 displays a sample question.

There are four key types of questions possible: starred, non-starred, short notice questions, and questions to private members:

1. Starred questions are those for which an oral answer is expected within 15 days. Answer to such question may be followed by supplementary questions by member. These questions tabled for a specific day are printed in green colour and are marked with asterisk sign, in order to distinguish from other questions.

2. Non-starred questions are those for which a written reply is expected. After the reply has been provided, no supplementary question can be asked. A notice period is to be given to the minister to reply to a question (which usually varies from 10 to 21 days). These questions are printed in white colour and not more than 230 questions can be listed for a day in the Lok Sabha.

3. Short notice questions are those which are asked on matters of urgent public importance and thus, can be asked on a shorter notice i.e. less than 10 days. These questions can be answered orally and supplementary questions can be asked. These questions are printed in light pink colour.

4. Questions to private members are those which are asked to members who are not

ministers. These questions are related to private member's bills, parliamentary committees, private member resolutions. These questions are printed in yellow colour.

5. If a member seeks to ask a question urgently and cannot wait for the duration of the notice period, then the member can do so provided it is accepted by the Speaker. Such questions are called supplementary questions.

The website of the Lok Sabha lays down the purpose of parliamentary questions as follows:

> *Through the Question Hour the Government is able to quickly feel the pulse of the nation and adapt its policies and actions accordingly. It is through parliamentary questions that the Government remains in touch with the people in as much as members are enabled thereby to ventilate the grievances of the public in administrative matters. Questions enable Ministries to gauge the popular reaction to their policy and administration. Questions bring to the notice of the Ministers many loopholes which otherwise would have gone unnoticed. Sometimes questions may lead to the appointment of a Commission, a Court of Enquiry or even Legislation when matters raised by Members are grave enough to agitate the public mind and are of wide public importance.*

3 Data & Methods

3.1 Natural Language Processing

A key component of the following research design and the usage of natural language processing techniques is inspired by multiple papers making use of machine learning techniques for economic research questions – Algaba et al., 2020 (overviews semantic analytics methods that use textual, audio, and visual data in economic science to survey this emerging field that they term *sentometrics*), Hansen and McMahon, 2016 (uses the text-based approach for quantification of natural language data), Gentzkow et al., 2019 (focuses purely on text data and review appropriate statistical methods to address it), and Bandiera et al., 2020 (focuses on reducing the high-dimensionality of data by constructing an index employing LDA). The particular context of using methods like LDA for Question Hour data has been discussed by political scientists (Sen et al., 2019), however, the usage of such techniques to create data for establishing causality in such contexts is yet to be explored in literature, baring the gap that this paper seeks to fill. Finally, an important source of methodological as well as theoretical inspiration has been Gill and Hall, 2015 which uses natural language processing and random assignment to study the impact of judicial identity on the text of legal rulings.

Prior work has studied gender and legislative behavior in the Indian Parliament by using parliamentary questions (Jacob, 2014). In this analysis of questions (raised between 1980 and 2009), the author finds that female parliamentarians are overall less likely to raise questions than male parliamentarians in the Lok Sabha. Additionally, for a ten-year period, he employs a hand-coding method to categorize parliamentary questions using the "subject line" (typically a summary of less than five words) provided by the Lok Sabha Secretariat. Given the availability of novel computational tools, hand coding is a painstaking process and falls short of the efficiency (and arguably, the objectivity) provided by using computational tools.

3.2 Data

This study uses a publicly available raw dataset (Bhogale, 2019) of the text of the question-answer pairs (N = 298,293) recorded in the Question Hour in the Lower House of the Indian Parliament from 1999 to 2019 spanning four terms of the House. This data also contains detailed information about the legislator tabling the question, as well as which Ministry it is addressed to. This data is supplemented by the electoral metadata about electoral candidates from the same time period (N = 35,671) made available by the Election Commission of India (Asopa et al., 2023) to match the demographic information (such as name, gender, party, constituency, position during election etc.) of the legislator raising the question, to the content of the question or issue they raise during the Question Hour.

Election Year	**No. of Female Candidates**	**No. of Female Winners**
1999	302	54
2004	385	52
2009	567	62
2014	673	66
2019	727	79

Table 1: Number of Constituencies with Female Winners by Election Year

In Table 1, the number of female winners represents the constituencies with female candidates, juxtaposed with the total number of female candidates who participated in the elections for the corresponding year.

3.2.1 Question Hour Data Source

The first dataset contains information on the Parliamentary Questions raised in the Lok Sabha between 1999 and 2019 (i.e between the 13th and the 17th Lok Sabha). The data has been extracted from the Lok Sabha's official website and treated in order to add information about the Members (constituency name, state, gender etc.) from the data shared by the Election Commission of India in its Statistical Reports.

3.2.2 Elections Data Source

The second data repository contains data of the results of all Indian election results held since 1962 to the most recent election, that in 2019. The data comes from the Statistical Reports published by the Election Commission of India (ECI). Additional variables, mostly derived from ECI data, have been added. Information from the Association for Democratic Reforms is also included. Table 2 summarizes the key variables in this dataset.

Variable	**Mean**	**Std Dev**	**Min**	**Max**
Valid Votes	900099.11	253843.92	29860.00	1763757.00
Electors	1456417.48	338537.89	37619.00	3368399.00
Turnout Percent	62.12	12.00	8.94	91.67
Vote Share Percent	7.64	15.27	0.01	78.8
Margin	31498.48	81132.55	0.00	712215.00
Margin Percent	3.61	8.81	0.00	70.1
Year	2009	6.68	2014	2019
Male or Female	0.07	0.26	0.0	1.0

Table 2: Summary Statistics for Election Data (1999-2019, N = 35,671)

In Table 2, Valid Votes refers to the sum of all votes to all candidates in a particular election for a particular constituency, Electors refers to the total number of registered electors in the constituency as per ECI, Turnout Percent refers to the percent turnout in the constituency, Vote Share Percent refers to the vote share received by the candidate in a given constituency, Margin refers to the difference in votes between a candidate and the next ordered candidate (ordered by position), and Margin Percent refers to the percent margin of a candidate in compared to the next position candidate.

3.3 Methods

Five distinct outcome variables are designed to measure the difference between male and female legislators through different outcomes (computed "scores"):

- ***corruption***: whether the question raised is pertaining to corruption in governance or government functioning,
- ***development***: whether the question raised is pertaining to the development of their constituency, state, or nation,
- ***accountability and transparency***: whether the question raised is pertaining to issues pertaining to the ideals of accountability and transparency in the democracy,
- ***programmatic representation***: whether the question raised is pertaining to programmatic motivations (pertaining to rules, policies, programs, legislations, etc.) or,
- ***clientilistic representation***: whether the question raised is pertaining to clientelistic ideals (pertaining to appeasing the demography, constituency, etc. of the legislator) in terms of representation.

The task of allocating scores to each question in the data, for each outcome above was explored using three distinct automated approaches using natural language processing tools: word embeddings, topic modelling, and zero-shot classification.

First, word embeddings that create an index of vector similarity between words in the dataset were considered. They help find words that are "similar" to the each of the keywords specified by the researcher pertaining to the final outcome criteria. In order to maintain contextual and logical similarity of words, a word-embedding algorithm called *word2vec* (Mikolov et al., 2013) was trained to produce the model using the question hour dataset. Appendix Figure 12 provides a sample of the visualization of the vectorization of words through this method. However, this method relied on employing the continuous-bag-of-words (CBOW) model to read in the context window words and tried to predict the most likely center word which impacted the speed of the algorithm in being able to accurately produce outcome scores for each category, and also was infeasible for the entire corpus to be considered to have scores for.

The second approach considered for score-generation was that of topic modelling using the Latent Dirichlet Allocation (LDA) algorithm. This model maps a set of documents to unobserved topics, which aids in clustering similar documents into topic clusters that can be manually examined and labeled. I also leveraged a propensity score matching (PSM) technique to find out whether each gender group's questions align with the specific topics that LDA yielded and I defined by hand (a view into this analysis is presented by Figure 3). The PSM would have ensured that I can create balance between the two gender groups while I are creating my outcome variable by ensuring that the mapping function ensures a causal interpretation of the outcome. While it can be useful for aspect extraction within each category, this method required for manual encoding of topic labels and very high computational power for successful execution at the scale of the question hour corpus.

Finally, this study uses the zero-shot classification technique to allocate scores pertaining to each classification category as per each question text using a cross-lingual model on top of the pre-trained natural language processing model, XLM RoBERTa, a Cross-Lingual Robustly Optimized Bidirectional Encoder Representations from Transformers Model (Conneau et al., 2020). This model is based on Facebook's RoBERTa model released in 2019. It is a large multi-lingual language model, trained on 2.5TB of filtered CommonCrawl data. It represents some of the most sophisticated pre-trained models prevalently being used in natural language processing for text classification tasks. And while it is computationally intensive to run the model given the large amount of data in my corpus, the use of a Graphical Processing Unit (GPU) optimized this process greatly.

Associating each question-answer pair to a topic in question (corruption/ development/ accountability/ representation type) using this method yields the quantified agreement of each question-answer pair to each topic - this enables us to effectively and accurately create the outcome variable for each of the topics as "scores". This agreement would be called the mapping function.

The underlying model is trained on the task of Natural Language Inference (NLI), which takes in two sequences and determines whether they contradict each other, entail each other, or neither. The model is fine-tuned on XNLI which includes 15 languages, including Hindi and many of the official Indian languages which may be references in questions raised in the Parliament. This can be adapted to the task of zero-shot classification by treating the sequence which I want to classify as one NLI sequence (called the premise) and turning a candidate label into the other (the hypothesis). If the model predicts that the constructed premise entails the hypothesis, then I can take that as a prediction that the label applies to the text. The score is determined as a "multi-label" hence for each of the categories.

As described in Grimmer et al., 2022 I required that the mapping function reduce the dimensionality of the text data and use it to create some kind of a relative quantitative score (based on topic allocation or classification). Grimmer et al., 2022 details steps to follow, in terms of splitting the training data for outcome-generation, validation of the model, etc, which have been undertaken to ensure robustness of the outcome measures particularly in the context of acting as outcome "scores" for a further causal inference analysis. Appendix Figure 11 presents a sample of scores for a series of sample questions and the top labels attributed to them.

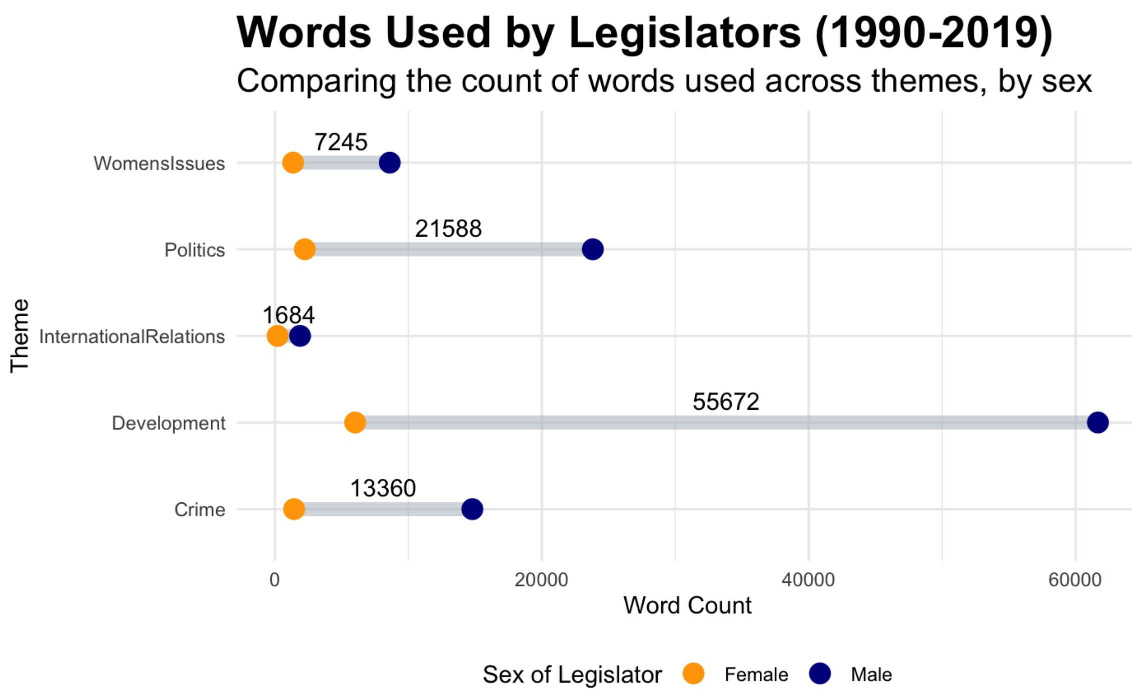

Figure 3: *Normalized Word Count Comparison, by Gender, Leveraging LDA and PSM*

Word counts per theme by gender extracted using Latent Latent Dirichlet Allocation (LDA) and normalized for comparison by using Propensity Score Matching (PSM). The normalization partly adjusts the raw word counts to ensure comparability. These observations align with those put forward in literature (Jacob, 2014), however no causal interpretation can be attributed to these.

Next, I deploy a dynamic difference-in-differences strategy with multiple time periods and an on/off treatment assignment at the constituency-level between the time periods, where each time period represents an election. I are considering the 'gender of the representative of a particular constituency at a given time' as treatment, and measuring its impact on the consequent type/topic of questions raised in the Parliament (based on the outcome variable(s) generated in the steps detailed here).

4 Model

4.1 Identification Approach

Callaway and Sant'Anna, 2021, De Chaisemartin et al., 2022, and De Chaisemartin et al., 2022 are the three landmark papers which have shed light on developing and tuning the dynamic difference-in-differences method this setting uses. While the former's focus was on a staggered treatment ("rollout" design) assignment approach, the latter focused on dynamic effects such that it was possible for units to enter, leave, and re-enter treatment. Various studies have shone a light on the potential biases that can result from applying a standard, twoway fixed-effect (TWFE) regression estimator on such a staggered setup, including Sun and Abraham, 2021, Callaway and Sant'Anna, 2021), and Goodman-Bacon, 2021.

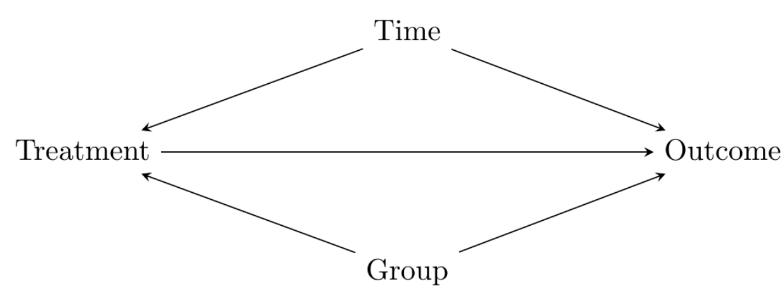

Figure 4: Causal Diagram for Difference-in-differences (Huntington-Klein, 2022)

Instead, differences-in-difference studies in such time-varying contexts take a different approach, as succinctly described in Huntington-Klein, 2022:

> *Seems like we've made things worse by introducing the control group. The key is this, though: now that I have that untreated group, even though we've added a new back door, I can now close both back doors. How is this possible?*
>
> 1. *Isolate the within variation for both the treated group and untreated group. Because I have isolated within variation, I are controlling for group differences and closing the back door through* Group *(the "differences")*
> 2. *Compare the within variation in the treated group to the within variation in the untreated group. Because the within variation in the untreated group is affected by time, doing this comparison controls for time differences and closes the back door through* Time *(the "difference" in those differences)*

4.2 Theory

Exploring the potential differences in legislative behavior between male and female representatives requires an understanding rooted in both political economy and microeconomic theories. This exploration is guided by several pertinent questions:

Why do women bring up different topics than men in legislative discussions? Are these differences attributable to inherent preferences between male and female legislators? Could differing constraints or costs in political engagement influence their legislative priorities? How might institutional settings within democracies differentially impact male and female representatives?

4.2.1 Political Economy of Gender Representation

Building upon the foundational concepts from (Chattopadhyay and Duflo, 2004), this study seeks to integrate the economic theories of political representation and decision-making to scrutinize the impact of gendered leadership. A key theoretical framework employed is the 'citizen-candidate' model, which proposes that political candidates emerge from the populace and inherently possess policy preferences that reflect their personal convictions rather than acting solely as delegates of their constituents. There are a few key features of this model relevant to our discussion.

In contrast to traditional models where candidates can commit to specific policies, the citizen-candidate model suggests that candidates, once elected, are inclined to advocate for policies that align with their personal preferences, as they cannot bind themselves to predetermined policy platforms.

The model also highlights the disparity in candidacy costs, which vary significantly between genders. These costs encompass financial, social, and personal barriers that might discourage potential candidates from running for office, particularly affecting female candidates.

It is posited that policy implementation may inherently favor the preferences of those in power, predominantly men. This bias can potentially be corrected through policies ensuring women's representation, thus allowing for a more equitable reflection of societal needs.

4.2.2 Microeconomic Perspectives on Legislative Behavior

From a microeconomic standpoint, the behavior of legislators can be analyzed through the lens of utility maximization, where representatives weigh the costs and benefits of various policy positions. Studies such as those by (Ranehill and Weber, 2022) and (Portmann and Stadelmann, 2017) invoke the 'median voter theorem', suggesting that under certain conditions, policy outcomes in a democracy tend to reflect the preferences of the median voter. The integration of women into these roles, particularly in contexts where legislative seats are reserved for women, could shift the policy equilibrium. This shift would potentially bring legislative outcomes closer to representing women's preferences, which might deviate from those of the traditional median voter, especially in highly gendered contexts.

Quantifying, testing, and examining the impacts of these ideas is challenging, however, motivates the application of computational approaches to address. Section 7.2 furthers this theoretical discussion in relation with the findings of this study.

5 Empirical Strategy

In terms of the empirical strategy, I consider the estimation of the effect using panel data where different groups (constituencies) are exposed to the treatment at different times. I focus on parameters aggregating instantaneous and dynamic treatment effects, with a clear welfare interpretation. Authors De Chaisemartin and D'Haultfœuille, 2020 have shown that under partial common trends conditions, these parameters can be unbiasedly estimated by weighted sums of differences-in-differences, provided that at least one group is always untreated, and another group is always treated. These estimators are valid if the treatment effect is heterogeneous, contrary to the commonly-used event-study regression. I also compute placebo estimators to test my identifying assumptions.

5.1 Specification

Roughly, are measuring the effects as per the following specification:

$$outcome_{ijt} = \beta_1 + \beta_2 treatment_{ijt} + TPFE_t + CFE_j + u$$

Here,

- $outcome_{ijt}$: the outcome of interest for legislator i at election t within constituency j
- $treatment_{ijt}$: whether the legislator i of constituency j at a given election t is female
- $TPFE_t$: time period fixed effects
- CFE_j: constituency fixed effects
- u: error term
- DID_ℓ estimates β_2 across j for each election t

Further, a few key insights regarding such models, to keep in mind when it comes to interpreting the results, emerge:

- They are designed to evaluate the statistical significance of coefficients for periods preceding the intervention, expecting these to be non-significant if the model is correctly specified.
- Demonstrating that control and treatment groups are statistically equivalent prior to the intervention supports the DID's assumption of parallel trends, although it is not definitive evidence.
- To circumvent the issue of perfect multicollinearity, commonly found in fixed-effects models, one time period is excluded. The period immediately before the intervention, denoted as -1, is often selected as the reference period to be omitted.

5.2 Identification

As I set out to ascertain the causal effect of the gender of the legislator upon the outcomes of interest – their participation in legislation measured through the context of economic development, I consider the theoretical underpinnings of this discussion. Section 4 goes into this discourse further. The key consideration is that there is no such thing as the causal effect of 'gender'. A proximal manner I proceed with thinking about this is: what would be effect of the fact that a constituency is *won* by a man versus a woman. This would be something that is assigned and identifiable and denotes what I proceed with.

The setup for determining the DID_ℓ estimator as described in De Chaisemartin and D'Haultfœuille, 2020 and De Chaisemartin et al., 2022 goes as follows. The assumptions for ascertaining the validity of this estimator along with a discussion about its generalized form can be found in Section 9.

- When treatment assigned at (g, t) level: $D_{g,t}$ = treatment of each unit (g) at time period t. In this setting, this is denoted by *constituencies g = {1,.., 2727} at election times t = {1, 2, 3, 4, 5} representing election years 1999, 2004, 2009, 2014, 2019.*
- $D_g = (D_{g,1}, \ldots, D_{g,T})$ represents the vector stacking g's treatments from periods 1 to 5, and $\mathbf{D} = (D_1, \ldots, D_G)$ represents the vector stacking the presence of female legislators of all groups at every period. $\mathbf{D}$ is the representation of the study's design.
- We let $\mathcal{D}$ denote the set of values $\mathbf{D_g}$ can take.

- For all $(d_1, \ldots, d_5) \in \mathbf{D}$, we let $Y_{g,t}(d_1, \ldots, d_4)$ denote potential outcome of g at t if $(D_{g,1}, \ldots, D_{g,5}) = (d_1, \ldots, d_5)$, and let $Y_{g,t} = Y_{g,t}(D_g)$ denote observed outcome of g at t.
- $F_g = \min\{t : t \geq 2, D_{g,t} \neq D_{g,t-1}\}$ and denotes the first treatment change.
- This method computes two key estimators – the non-normalized actual-versus-status-quo (AVSQ) effects and the normalized AVSQ effects – where the latter adjusts the effect by the total increments in treatment from the beginning of the period to now.

The design of the measure, hence, is construed as follows:

- Let
$$N_t^g = \#\{g' : D_{g',1} = D_{g,1}, F_{g'} > t\}$$
be the number of groups g' with the same period-one treatment as g, and that have kept the same treatment from period 1 to t.
- To estimate $\delta_{g,\ell}$, the authors construct the following estimator:
$$\mathrm{DID}_{g,\ell} = Y_{g,F_g-1+\ell} - Y_{g,F_g-1} - \frac{1}{N^g_{F_g-1+\ell}} \sum_{\substack{g' : D_{g',1} = D_{g,1}, \\ F_{g'} > F_g - 1 + \ell}} (Y_{g',F_g-1+\ell} - Y_{g',F_g-1})$$
- This is the DiD estimator comparing $F_g - 1$ to $F_g - 1 + \ell$ outcome evolution of g to groups with the same baseline treatment, and that have kept that treatment from period 1 to $F_g - 1 + \ell$.
- For every g,
$$T_g = \max_{g' : D_{g',1} = D_{g,1}} F_{g'}$$
denote last period where there is still a group with the same period-one treatment as g and whose treatment has not changed since the start of the panel.
- For any g, such that $F_g \leq T_g$, and for any $\ell \in \{1, \ldots, T_g - F_g + 1\}$, I know that
$$\delta_{g,\ell} = \mathbb{E}\left[Y_{g,F_g-1+\ell} - Y_{g,F_g-1+\ell}(D_{g,1}, \ldots, D_{g,1}) \mid \mathbf{D}\right]$$
be the expected difference between g's actual outcome at $F_g - 1 + \ell$ and its counterfactual "status quo" outcome i.e. AVSQ effect of g at $F_g - 1 + \ell$.

5.3 Implementation

DID_ℓ, the estimator of the ℓ^{th} dynamic effect, compares first-time switchers' and not-yet switchers' outcome evolution, from the last period before first-time switchers' treatment changes to the ℓ^{th} period after that change.

I organize the dataset into a panel data structure with a binary treatment indicator that identifies whether a group (each constituency in my case) is treated in a given time period. This indicator is created such that there is a switch from 0 to 1 at the time when the treatment is first applied to the group – i.e. the gender of the representative from that constituency at that time period. I use the `did_multiplegt_dyn` STATA package created by the authors De Chaisemartin and D'Haultfœuille, 2020 and De Chaisemartin et al., 2022 to compute the estimator. For each ℓ, the workflow carries out the estimation of the treatment effect by calculating the difference in outcomes between the first-time switchers and not-yet switchers, before and after the treatment. Formally, as discussed in section 5.2, this estimator denotes the average, across all switchers, of DID estimators comparing the F_{g-1} to $F_{g-1+\ell}$ outcome evolution of g to that of groups with the same period-one treatment as g but whose treatment has not changed yet at $F_{g-1+\ell}$. I also compute the robust standard errors (clustered at the constituency level) at each effect time period. For the purpose of my estimation, I compute the normalized DID_ℓ estimator by normalizing by the average total incremental treatment dose received by switchers from F_{g-1} to $F_{g-1+\ell}$ with respect to their period-one treatment. This normalization ensures that the estimator estimates a weighted average of the effects of the current treatment and of its $\ell - 1$ first lags on the outcome.

6 Results

For the development outcome score (Figure 5) the normalized average total effect per treatment unit suggests a negative average total effect after the treatment with a point estimate of -0.0754 and standard errors reported at 0.0439 (Table 4). All estimated effects from the first to the fifth period show negative values, indicating a decline in the development scores following the treatment. The magnitude of these effects varies, with the first and fifth periods (1999 and 2019 respectively) showing larger declines compared to the middle periods (Table 3). In general, these trends seem to indicate that female representation for a constituency unit actually has a negative impact on the development score, corresponding with a lower priority attributed to development-related questions.

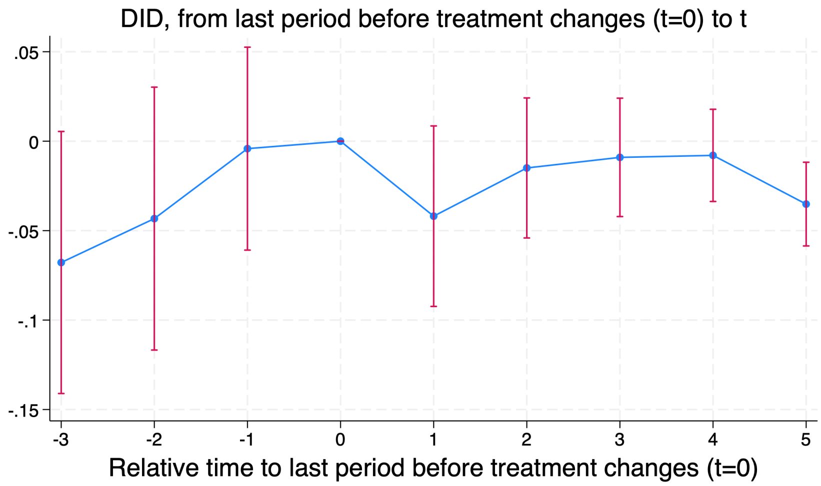

Figure 5: Impact of Treatment on **Development Score**

The corruption score estimates although positive hover around zero over the individual-period estimates over the time periods, whereas the average total effect across all treatment units shows a positive estimate of 0.04256 with a standard error reporting of 0.0256, suggesting a moderate overall positive impact of the treatment on reducing corruption scores (Table 6). This broader effect over the entire treatment period underscores a potentially sustained, albeit moderate, positive impact of the interventions (Figure 6). The immediate effects are stronger and gradually taper off, but remain positive, indicating the larger positive effect in earlier years such as 1999 rather than current or recent times (Table 5).

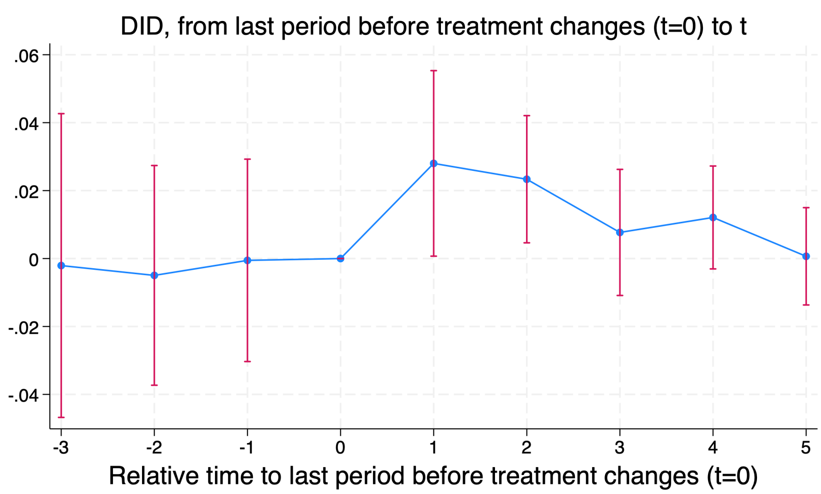

Figure 6: Impact of Treatment on **Corruption Score**

For the accountability and transparency score, all the effects from the first to the fifth period post-treatment show negative values (Table 7), suggesting a decrease in accountability and transparency scores following the change in female representation. This trend is observed over several periods, indicating a sustained impact (Figure 7). The average total effect per treatment unit is -0.0623, with a standard error of 0.0254 (Table 8). This substantial negative average total effect is notable, as it signifies that, on average, increases in female representation are associated with a decrease in accountability and transparency, contrary to some observations in past literature.

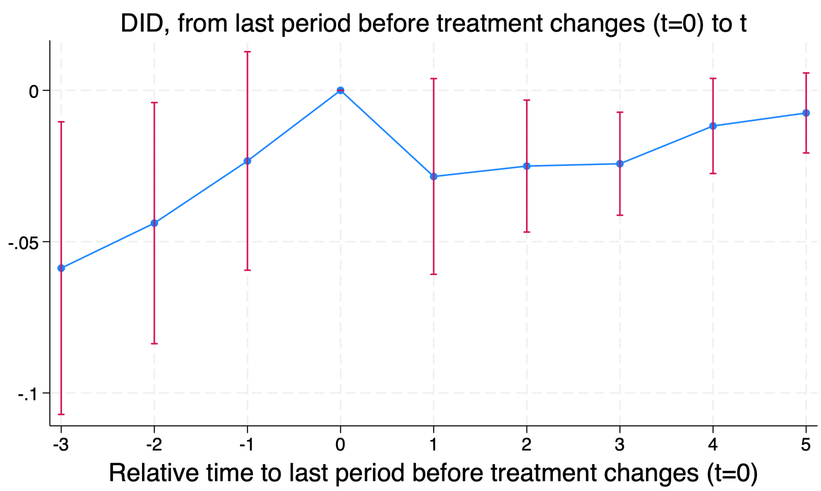

Figure 7: Impact of Treatment on **Accountability and Transparency Score**

In case of the programmatic representation score (Figure 8) the average total effect per treatment unit is -0.0253 with a standard error of 0.0345 (Table 10). This result implies that overall, across all periods and considering the cumulative effects of the treatment, there is a slight decrease in programmatic representation scores associated with changes in female representation. The negative sign indicates a reduction, but given the magnitude and the standard error, this effect is relatively small and not statistically significant at conventional levels, with the effects fluctuating around zero and becoming slightly more negative but still minimal by the fifth period (Table 9). This implies that female representation is associated with a slight decrease in programmatic representation scores. This suggests that constituencies with a greater female representation might see a lesser engagement in programmatic issues through questions raised by legislators, according to the measure used.

For the clientilistic representation score estimates, all five time periods show negative estimates, with the effects progressively getting smaller but remaining negative through the last time period (Table 11). The average total effect per treatment unit is also negative

(-0.0432), with a standard error of 0.0256 (Table 12), which indicates a significant overall decline in clientelistic representation across the study population depicted in Figure 9. This indicates female representation potentially leading to a shift away from clientelism over time.

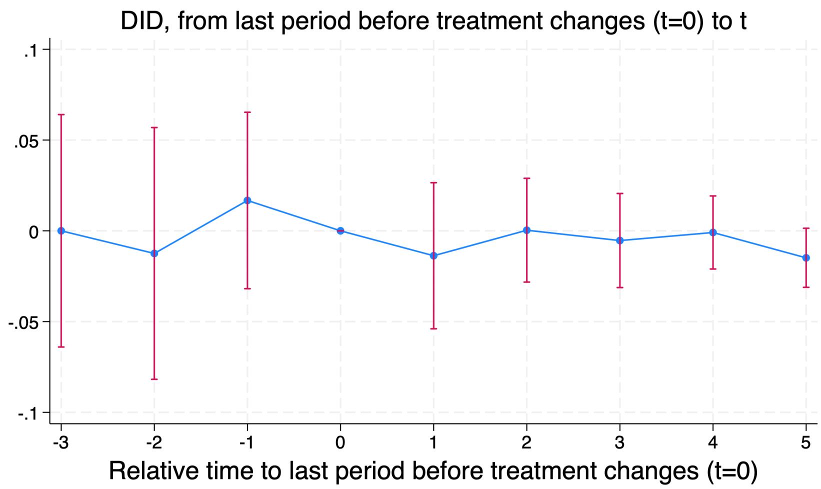

Figure 8: Impact of Treatment on **Programmatic Representation Score**

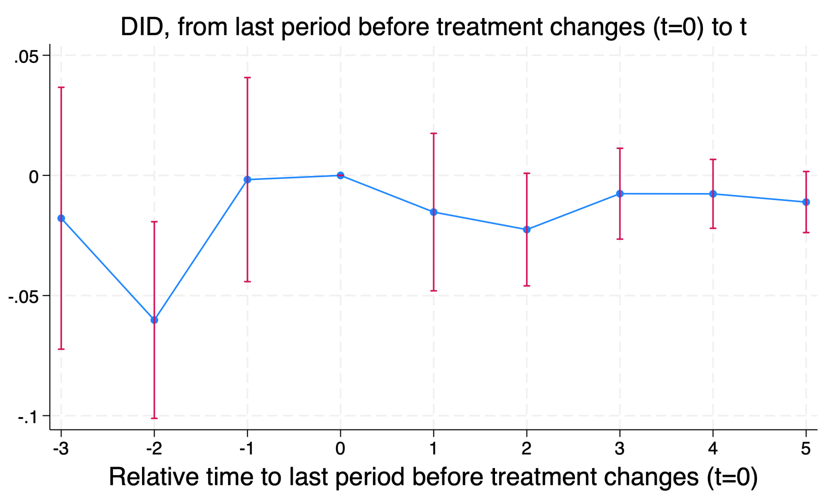

Figure 9: Impact of Treatment on **Clientilistic Representation Score**

Representation involves making citizens' voices, opinions, and perspectives present in the public policy-making process. The finding that gender identity can have varying effects on different legislative priorities is aligned with the idea that descriptive representation (the idea that elected bodies should mirror the gender composition of the general population) does not always translate into substantive representation (policies that advance the interests of that population). This reflects the ideal that empowerment and actual representation must go beyond simply increasing numbers and be directed towards truly shifting participation priorities in favor of those that need it the most.

These findings highlight the complex relationship between female representation and various dimensions of accountability and transparency in governance. On the one hand, they indicate that increases in female representation may have a subtly negative impact on development, transparency, and representation scores. And on the other hand, there seems to be a small impact on corruption score, indicating the priorities of female representation over time. However, it is important to note that the effect sizes are relatively small and not statistically significant at conventional levels – this has been discussed in the context of the dynamic difference-in-differences approach, in detail, in the following section.

7 Discussion

7.1 Methodological Considerations

Three key points pertaining to the inference of results from the dynamic difference-in-differences approach warrant discussion.

Firstly, regular difference-in-differences takes advantage of all the data in the entire “after” period to estimate the effect. However, each period’s effect estimate in the dynamic treatment effects approach relies mostly on data from that one period which results in much less precision in estimates. Generally, it has been observed that the overall average effect is statistically significant however the individual-period effects aren’t. Huntington-Klein, 2022 warns of this challenge in his discussion on dynamic studies, and draws attention to the importance of the more accurate interpretation of individual-period effects.

Secondly, when interpreting the results, all observed effects stand relative to the omitted period-zero effect. As always when we have a categorical variable, everything is relative to the omitted group. So the β_2 coefficient, for example in this context, would imply that the effect two periods after treatment is β_2 higher than the effect in the last period before treatment (Huntington-Klein, 2022). This provides a helpful barometer to contextualize my results in. The variations in effects over different time periods underscore the importance of considering the temporal context in which these changes occur, as different election years bring forth unique challenges and opportunities that shape the observed effects.

Thirdly, since my outcome variables are a set of relative “scores” varying between 0 and 1, this also bears some impact onto the manner in which I can interpret the effects (Appendix Figure 11 for a sample scoring).

Outside of staggered designs, such as that of this study, some initially untreated groups switch to being treated and may remain treated thereafter, while other groups may revert to being untreated just after that first switch, other groups may alternate between treated and untreated, etc. Then, DID_ℓ cannot be directly converted into an effect per unit of treatment received. Instead, its construction gets modified slightly to reflect the notion of the normalized AVSQ – I compute the weighted average estimates – the average number of treatment units received by switchers after their first switch, relative to the counterfactual where they would have kept their period-one treatment all along. Finally, I construct the ratio of these two estimators. This ratio estimates the effect of first switches on the outcome, and scales it by the "first-stage" effect of first switches on the treatments received thereafter. Accordingly, it estimates some average of the change in outcome created by a one-unit change in treatment. This weighted normalization is an aspect which can be leveraged in further extensions of the study to reduce the size of the confidence intervals to make the estimates stronger.

I control for party, size of electorate (in terms of number of electors), and number of votes cast for each constituency. Formally, estimators with controls are similar to those without controls, except that the first-difference of the outcome is replaced by residuals from regressions of the first-difference of the outcome on the first-differences of the controls and time fixed effects. Those regressions are estimated in the sample of control *(g,t)*s: i.e. *(g,t)*s such that group g's treatment has not changed yet at t. Those regressions are also estimated separately for each value of the period-one treatment. Estimators with controls are unbiased even if groups experience differential trends, provided such differential trends can be fully explained by a linear model in covariates changes.

To check for the parallel trends assumption, I compute placebo estimates – the ℓ^{th} placebo compares first-time switchers' and not-yet switchers' outcome evolution, from the $(\ell+1)^{th}$ to the ℓ^{th} period before first-time switchers' treatment changes. Thus, the ℓ^{th} placebo assesses if parallel trends holds over 3 consecutive periods, ℓ periods before switchers switch (which is indicated in the result figures as the time periods before $t = 0$). Such first-difference placebos may be useful to test the no-anticipation assumption. Placebos compare the outcome evolution of switchers and of their controls, before switchers' treatment changes for the first time. Under the parallel trends and no-anticipation assumptions underlying the event-study estimators computed by `did_multiplegt_dyn`, the expectation of the placebos is equal to zero. Thus, placebos can be used to test those assumptions, by testing the null that all placebos are equal to zero. I find no statistically significant results in the tests of joint nullity of the placebos for each of the outcome variables.

The DID_ℓ estimators with controls are similar to those without controls, except that the first-difference of the outcome is replaced by residuals from regressions of the first-difference of the outcome on the first-differences of the controls and time fixed effects. It is also important to note that these related DID estimators of the instantaneous treatment effect as well as that of the dynamic treatment effects, and of placebo tests of the parallel trends assumption are robust to heterogeneous effects.

7.2 Theoretical Considerations

The following discussion aims to relate some of the findings from the dynamic difference-in-differences analysis to the theoretical conjecture discussed in Section 4.2.

Firstly, our empirical findings reveal nuanced effects of female representation across different governance outcomes. Notably, the limited impact on development could be interpreted through the lens of the citizen-candidate model, where the non-commitment nature of political candidates allows them to pursue policies that align with their personal preferences once elected. This suggests that issues raised may perhaps reflect a consensus that transcends gender differences in policy priorities.

Secondly, the candidacy costs - both social and economic - that differ significantly between genders could be influencing the legislative focus of female representatives. For example, the slight negative associations observed in clientelistic representation scores might indicate that female legislators, facing higher social scrutiny, may opt to steer away from controversial or highly clientelistic issues that could potentially attract negative attention or backlash, thus affecting their re-election prospects or standing within the legislature.

Similarly, since legislators are seen as rational actors who maximize their utility under given constraints, the strategic focus on corruption-related issues by female legislators could be interpreted as an effort to maximize their utility in terms of building trust and credibility among constituents while being within the constraints of re-election prospects. This is particularly relevant considering the higher candidacy costs they face, where establishing a strong ethical and accountable image could be crucial for overcoming barriers to entry and re-election.

While speculative, these considerations are important to be critically examined in this vein of research in order to delineate the underlying mechanisms behind the observations this study has discerned. One of the ways in which this can be approached by future research is through conducting a heterogeneity analysis at the margin of victory to identify patterns or groups that can further an understanding of these theoretical considerations in practice.

8 Conclusion

This study delineates a complex terrain where the effects of changes in female representation manifest diversely across various governance dimensions and temporal frameworks. While the overall impacts on development scores and accountability, transparency, and representation metrics remain subdued, indicating a restrained direct influence of gender alterations, contrasting positive variations are discernible in the case of corruption scores. The effects, particularly in clientilistic or programmatic representation, reflect a slight decrease which contradicts some expectations, suggesting that increases in female representation do not consistently translate to improved governance outcomes within these specific areas. Moreover, the temporal analysis underscores the pivotal role of timing in interpreting the influence of gender dynamics on governance. These findings emphasize the necessity for a nuanced approach in future research, aiming to unravel these temporal intricacies and their implications for policy-making. This exploration is crucial for crafting strategies that effectively harness gender representation in political leadership, rooted in a comprehensive understanding of these multifaceted impacts. Future work aims to delve into these patterns more thoroughly, as discussed in Section 7.

Data and Code Availability Statement

Data from *"TPCD-IPD: TCPD Indian Parliament Dataset (Question Hour) 1.0", Trivedi Centre for Political Data, Ashoka University, 2019* and *"TCPD Indian Elections Data v2.0", Trivedi Centre for Political Data, Ashoka University, 2021* is used in this paper. The full dataset and documentation can be downloaded from here.

Code for data cleaning and analysis is provided as part of this replication package. It is available here.

References

Algaba, A., Ardia, D., Bluteau, K., Borms, S., & Boudt, K. (2020). Econometrics meets sentiment: An overview of methodology and applications [Citations: 55]. *Journal of Economic Surveys*, *First published*. https://doi.org/10.1111/joes.12370

Asopa, N., Gangwani, S., Sunaad, A., & Verniers, G. (2023). TCPD Indian Elections dataset (TCPD-IED), 1951-1962 Codebook 1.0 [Trivedi Centre for Political Data, Ashoka University].

Ayyangar, S., & Jacob, S. (2015). Question hour activity and party behaviour in india. *The Journal of Legislative Studies*, *21*(2), 232–249.

Bandiera, O., Prat, A., Hansen, S., & Sadun, R. (2020). Ceo behavior and firm performance. *Journal of Political Economy*, *128*(4).

Baskaran, T., Bhalotra, S., Min, B., & Uppal, Y. (2018, May). *Women legislators and economic performance* (tech. rep. No. 47). Helsinki, Finland, UNU-WIDER. https://doi.org/https://doi.org/10.35188/UNU-WIDER/2018/489-6

Bhogale, S. (2019). TPCD-IPD: TCPD Indian Parliament Dataset (Question Hour) 1.0 [Trivedi Centre for Political Data, Ashoka University].

Bird, K. (2005). Gendering parliamentary questions. *The British Journal of Politics and International Relations*, *7*(3), 353–370.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, *72*(5), 1409–1443.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale.

De Chaisemartin, C., et al. (2022). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

De Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96. https://doi.org/10.1257/aer.20181169

Diamond, I. (1977). *Sex roles in the state house*. Yale University Press.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535–574. https://doi.org/10.1257/jel.20181020

Gill, M., & Hall, A. (2015). How judicial identity changes the text of legal rulings. https://doi.org/10.2139/ssrn.2620781

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/https://doi.org/10.1016/j.jeconom.2021.03.014

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.

Hansen, S., & McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication [38th Annual NBER International Seminar on Macroeconomics]. *Journal of International Economics*, *99*, S114–S133. https://doi.org/https://doi.org/10.1016/j.jinteco.2015.12.008

Huntington-Klein, N. (2022). *The effect: An introduction to research design and causality* (1st ed.) [162 B/W Illustrations]. Chapman & Hall. https://www.routledge.com/The-Effect-An-Introduction-to-Research-Design-and-Causality/Huntington-Klein/p/book/9781032125787

Hussain, S. (2022). Performance of women in parliament: A quantitative study of the questions asked by women members in lok sabha (1999–2019). *Economic and Political Weekly*, *57*(31), 52–59.

Jacob, S. (2014). Gender and legislative performance in india. *Politics and Gender*, *10*, 236–264.

Kapoor, M., & Ravi, S. (2014). Women voters in indian democracy: A silent revolution. *Economic and Political Weekly*, *49*(12), 63–67.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space.

Phillips, A. (1995). *The politics of presence: Democracy and the group representation*. Oxford University Press.

Pitkin, H. F. (1972). *The concept of representation*. University of California Press.

Portmann, M., & Stadelmann, D. (2017). Testing the median voter model and moving beyond its limits. *Social Science Quarterly*, *98*(5), 1264–1276. https://www.jstor.org/stable/26612477

Ranehill, E., & Weber, R. A. (2022). Gender preference gaps and voting for redistribution. *Experimental Economics*, *25*(3), 845–875. https://doi.org/10.1007/s10683-021-09741-8

Sen, A., Ghatak, D., Kumar, K., Khanuja, G., Bansal, D., Gupta, M., Rekha, K., Bhogale, S., Trivedi, P., & Seth, A. (2019). Studying the discourse on economic policies in india using mass media, social media, and the parliamentary question hour data. *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies*, 234–247. https://doi.org/10.1145/3314344.3332489

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/https://doi.org/10.1016/j.jeconom.2020.09.006

Wittman, D. (1989). Why democracies produce efficient results. *Journal of Political Economy*, *97*(6), 1395–1424. https://www.jstor.org/stable/1833245

9 Appendix

9.1 Discussion of the Assumptions for Computing Difference-in-differences Estimator

De Chaisemartin et al., 2022 make the following assumptions for their proposed estimators to be estimable, consistent, and unbiased. I discuss the confirmation of these assumptions in my setting:

- there have to exist units with the same period-one treatment
  - I confirm this with the presence of constituencies with male (female) legislators at t = 1
- there have to exist at least one set of units sharing the same period-one treatment within which there is heterogeneity in the date at which groups change treatment for the first time
  - I confirm presence of constituencies with male (female) legislators at t = 1 that changed to female (male) at different times t = {2, 3, 4}
- unit's current outcome does not depend on future treatments
  - I confirm that the current election outcome does not depend on future outcomes
  - the vice-versa may be possible – such that future election outcomes are dependent on past outcome – this notion is accounted for by the authors
- if two units have the same period-one treatment, then they have the same expected evolution of their status-quo outcome
  - male (female) legislators are a similar to each other regardless of constituencies they represent by virtue of the fact that both are equally likely to be winners or losers
  - further, the authors note that this assumption can be relaxed if treatment is binary, which is the case in my setting

9.2 Generalized Dynamic Difference-in-differences Estimator

The generalized regression model utilized in difference-in-differences event studies is formalized as follows:

$$Y_{gt} = \alpha + \sum_{k=T_0}^{-2} \beta_k \times treat_{gk} + \sum_{k=0}^{T_1} \beta_k \times treat_{gk} + X_g'\Gamma + \phi_s + \gamma_t + \epsilon_{gt} \tag{1}$$

In this framework:

- $treat_{gk}$ is a binary indicator set to 1 if the observation is within the initial treated period for group g relative to period k, and 0 otherwise, including for groups never receiving treatment.
- T_0 and T_1 denote the range of periods before and after the start of the treatment used in the analysis.
- X' represents a vector of control variables.
- ϕ and γ encapsulate the fixed effects for state and time, respectively.
- Regression estimates are derived with standard errors that are clustered at the group level.

9.3 Result Tables

Table 3: Estimation of Treatment Effects: **Development Score**

Effect	Estimate	SE	LB CI	UB CI	N	Switchers
Effect_1	-0.0419	0.0257	-0.0923	0.0085	9114	171
Effect_2	-0.0149	0.0200	-0.0541	0.0242	9017	171
Effect_3	-0.0090	0.0169	-0.0421	0.0240	8939	171
Effect_4	-0.0079	0.0131	-0.0337	0.0178	8888	171
Effect_5	-0.0352	0.0119	-0.0586	-0.0118	8846	171

Table 4: Average Total Effect per Treatment Unit for **Development Score**

Estimate	SE	LB CI	UB CI	N	Switch x Periods
-0.0755	0.0439	-0.1615	0.0106	13973	855

Table 5: Estimation of Treatment Effects: **Corruption Score**

Effect	Estimate	SE	LB CI	UB CI	N	Switchers
Effect_1	0.02801	0.01393	0.0006973	0.0553143	9114	171
Effect_2	0.02335	0.00955	0.0046257	0.042072	9017	171
Effect_3	0.00769	0.00947	-0.01087	0.02625	8939	171
Effect_4	0.01209	0.00773	-0.00305	0.02724	8888	171
Effect_5	0.000657	0.00731	-0.01366	0.01498	8846	171

Table 6: Average Total Effect per Treatment Unit for **Corruption Score**

Estimate	SE	LB CI	UB CI	N	Switch x Periods
0.04257	0.02567	-0.00773	0.09286	13973	855

Table 7: Estimation of Treatment Effects: **Accountability and Transparency Score**

Effect	Estimate	SE	LB CI	UB CI	N	Switchers
Effect_1	-0.02845	0.01649	-0.0607	0.0038	9114	171
Effect_2	-0.02501	0.0111	-0.0468	-0.0032	9017	171
Effect_3	-0.02445	0.0088	-0.0412	-0.0072	8939	171
Effect_4	-0.011	0.0080	-0.0274	0.0039	8888	171
Effect_5	-0.0074	0.00674	-0.0206	0.0057	8846	171

Table 8: Average Total Effect per Treatment Unit for **Accountability and Transparency Score**

Estimate	SE	LB CI	UB CI	N	Switch x Periods
-0.0623315	0.0253859	-0.112087	-0.012576	13973	855

Table 9: Estimation of Treatment Effects: **Programmatic Representation Score**

Effect	Estimate	SE	LB CI	UB CI	N	Switchers
Effect_1	-0.01372	0.02053	-0.05396	0.02652	9114	171
Effect_2	0.00034	0.01458	-0.02824	0.02891	9017	171
Effect_3	-0.00536	0.01322	-0.03127	0.02055	8939	171
Effect_4	-0.00092	0.01021	-0.02105	0.01921	8888	171
Effect_5	-0.01488	0.00830	-0.03114	0.00139	8846	171

Table 10: Average Total Effect Per Treatment Unit for **Programmatic Representation Score**

Estimate	SE	LB CI	UB CI	N	Switch x Periods
-0.02527	0.03454	-0.09297	0.04244	13973	855

Table 11: Estimation of treatment effects: **Clientelistic Representation Score**

Effect	Estimate	SE	LB CI	UB CI	N	Switchers
Effect_1	-0.0153	0.0167	-0.0480	0.0176	9114	171
Effect_2	-0.0225	0.0120	-0.0460	0.0009	9017	171
Effect_3	-0.0076	0.0096	-0.0265	0.0113	8939	171
Effect_4	-0.0077	0.0073	-0.0220	0.0067	8888	171
Effect_5	-0.0111	0.0065	-0.0238	0.0016	8846	171

Table 12: Average Total Effect Per Treatment Unit for **Clientilistic Representation Score**

Estimate	SE	LB CI	UB CI	N	Switch x Periods
-0.0432	0.0256	-0.0934	-0.0071	13973	855

9.4 Additional Figures

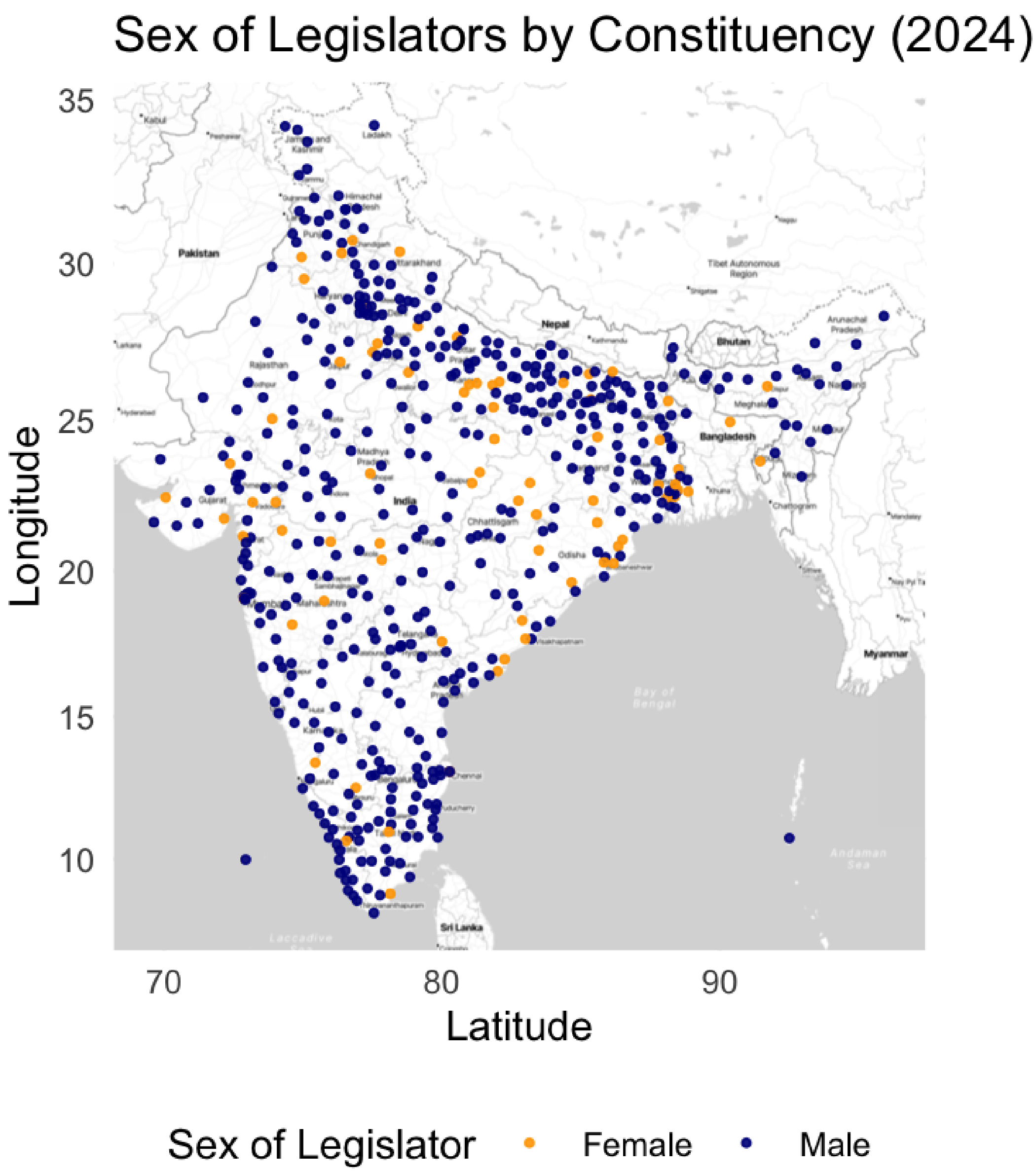

Figure 10: Geographic Distribution of Male and Female Legislators, by Constituency

question_text	top_category	top_category_score	processed_results
(a) whether any scheme is under consideation of the Government to allot waste land and forest land to farmers or landless persons in rural areas for plantation of fruit trees; (b) if so, the details thereof; (c) whether there is any proposal to allocate funds also for this purpose; (d) if so, the details in this regard and details of amount allocated during the last two years, State-wise; and (e) details of the works executed during the last two years?	programmatic representation	0.769578	{'programmatic representation': 0.7695783376693726, 'accountability and transparency': 0.6634148359298706, 'clientelistic representation': 0.2397715002298355, 'development': 0.19319695234298706, 'corruption': 0.018497761338949203}
(a) whether the Government is aware that there is acute scarcity of essential commodities in Andaman and Nicobar Islands; (b) whether the Government has taken steps to distribute/supply all essential commodities including kerosene to the consumers at Fair Price; (c) if so, the details thereof; and (d) if not, the reasons therefor?	accountability and transparency	0.722198	{'accountability and transparency': 0.7221977114677429, 'clientelistic representation': 0.41903725266456604, 'programmatic representation': 0.3838687539100647, 'development': 0.16504906117916107, 'corruption': 0.007125996984541416}
(a) whether the Government has any proposal to develop tourism in Andaman and Nicobar Islands; (b) if so, the details thereof; (c) whether the Government has received any proposal from private organization for setting up of a hotel in the Island; and (d) if so, the action the Government has taken to clear the proposal?	development	0.952534	{'development': 0.9525343775749207, 'accountability and transparency': 0.9019773602485657, 'programmatic representation': 0.6822377443313599, 'clientelistic representation': 0.5837655067443848, 'corruption': 0.007300106342881918}
- (a) the number of Colleges and Research Institutions in Andaman and Nicobar Islands connected with ERNET to provide value added services; (b) the details of its advantages/benefits to the students of rural areas in Andman and Nicobar Islands; and (c) the steps taken by the Government to provide ERNET connectivity to all the colleges in the Islands?	clientelistic representation	0.759415	{'clientelistic representation': 0.7594146728515625, 'programmatic representation': 0.6315813660621643, 'accountability and transparency': 0.616448163986206, 'development': 0.5167775750160217, 'corruption': 0.000335676217218861}
(a) whether the selection of Dealers and allotment of distributorship is made by the Oil marketing companies as per their own guidelines; (b) if so, the manner in which the Government exercises its control over them; (c) the details of the system put in place by the Government to ensure that the lack of its own control may not lead to corruption and acts prejudicial to public interest; and (d) the benefits yielded by the said system to the Government?	corruption	0.725767	{'corruption': 0.7257670760154724, 'programmatic representation': 0.4795593023300171, 'accountability and transparency': 0.26828786730766296, 'clientelistic representation': 0.19287022948265076, 'development': 0.054823681712150574}

Figure 11: Sample Scoring of Question Text

Figure represents a sample of question texts with the corresponding classification (label assigned) for each. The 'processed results' variable depicts the assignment and assigned scores for multi-label assignment of score categories.

Figure 12: Word Cloud curated using Word Embeddings for vectorized questions of Females and Males