

Institutional Change after Franchise Expansion: Evidence from British India, 1916-1940

Mitchell Bosley* and Htet Thiha Zaw†

July 8, 2023

Abstract

Universal suffrage, the unconditional provision of voting rights for every adult citizen, is a key element of democracy. The transition towards an expanded electoral franchise is expected to transform the policy formation process in legislature. While existing research identifies such a change by evaluating the policy outcomes, our understanding remains limited on how such changes were discussed in the legislature, a key actor in policy formation. We fill this gap with the first-ever text-as-data analysis of the British India legislature over three decades (1916-40), showing the impact of franchise expansion in 1920 and 1935. Contrary to our expectations, there was no evidence to support a change in the legislature's policy priorities between elected and non-elected members. Instead, we observe divergence from the elected members by only some subgroups of the new elected members. The findings provide a new perspective on colonial legislatures and their role in policy formation after franchise expansion.

*PhD Candidate, Department of Political Science, University of Michigan. e-mail: mcbosley@umich.edu

†PhD Candidate, Department of Political Science, University of Michigan. e-mail: htzaw@umich.edu

I Introduction

Can the introduction of more democratic forms of political institutions transform the policy formation process in the legislature? Existing findings on electoral franchise expansion, an essential element for any steps towards more democratic political institutions, highlight significant changes in state-building process, such as government size and fiscal capacity (Aidt, Dutta, and Loukoianova 2006; Cappelli 2016; Husted and Kenny 1997). However, they predominantly focus on policy outcomes, for instance collected revenues or goods and services provision, while the process through which these changes were considered, discussed, and formed remained largely in the black box.

This is particularly the case in the historical emergence of legislative institutions in colonial states in the twentieth century. While we now have an extensive amount of work on the government policies in colonial states and their long-run consequences (Acemoglu, Johnson, and Robinson 2001; Banerjee and Iyer 2005; Dell 2010), our understanding remains limited on how governments craft, discuss, and form these policies. Specifically, until recently, few discuss the role of legislative institutions in policy formation, and most that do focus on settler colonies outside Asia and Africa (Gailmard 2017, 2019; Paine 2019). Therefore, by extending our knowledge on the consequences of electoral franchise expansion to colonial contexts, we can bring to light not only the transformations policy formation process after electoral franchise expansion but also the policy formation process itself in colonial states.

We then focus on the legislative history British India, politically and economically the most important colony in the British Empire, and one that experienced extensive reforms in legislative institutions and electoral franchise during the interwar period in 1919 and 1935. We develop two original data to investigate the consequences of colonial electoral franchise expansion on the legislative institution and the government interest in education between 1920 and 1940. First, we consolidate the list of individual legislators (both elected and nominated) from *The India Office List*, an annual publication covering all the individuals associated with the colonial government, including the legislature. These lists allow us to infer the legislators who came to the position through a popular election and who were not. Second we construct text-as-data of legislative debates, using extensive records from the British India Parliament Digital Library.

While we look at the post-franchise policy formation process across several topics, our discussion focus on education. Existing literature consistently suggest a positive association between the regime change to democracy and the level of government investment in education (Ansell and Lindvall 2013; Harding and Stasavage 2014; Lindert 2004). Explanations on this association conceptualize education as an good with economic benefits to be distributed across the electorate. For instance, education allows governments to remain competitive in a globalizing economy, or fulfill the demands from the electorate through fiscal transfers from high-income to low-income individuals. Despite the positive association, recent research on educational development in non-democratic states challenges the purported origins of education in democratization. It points out that several first instances of state involvement in education emerged under authoritarian governments, preceding by decades the expansion of electoral franchise. Rather than as a policy tool for economic development, or a response to electoral demands, education in these cases served as a weapon for political control as it allowed governments to craft national identities, establish political order, and influence individual loyalty towards the state (Aghion et al. 2018; Paglayan 2020; Testa 2018).

In this paper, we argue that, by moving beyond the prevailing associations between regime type and purpose of education driving state involvement (economic distribution in a democracy and political order under a non-democracy), we can better understand how the governments in power responded to more than electoral and economic incentives in their education policy making during transitions towards more democratic political institutions. Specifically, we propose that, as electoral franchise expanded, the incumbent government were more motivated to encourage participation in political institutions via state-controlled education, especially when the political legitimacy of said institutions affected by franchise expansion remained weak. Driven by the incentive to increase political participation in weakly legitimate institutions, we expect that the government investments and interest in education increased in the places after franchise expansion, and especially in the places with low electoral turnout.

We then study if this "education for participation" theory can help us understand the history of education in colonial states (specifically, British India), many of which saw the introduction of expanded legislative institutions and electoral franchise in the twentieth century; we therefore build on the recent at-

tention towards the development of legislative institutions in colonial states (Gailmard 2017, 2019; Opalo 2021; Paine 2019).

Preliminary results from our structural topic model suggest that there is no evidence franchise expansions in 1920 and 1935 led to a change in the policy priority for education in the legislature. Looking at the results from Structural Topic Models (STMs), we conclude that overall topic prevalence did not change for education after both events of franchise expansion. Similarly, based on the ideal point estimates, we do not observe a significant difference in policy preferences on education between elected and non-elected members.

2 Historical Background

Since the mid-seventeenth century and before 1858, most of the present-day India's territories were under what was known as the Company rule, under an organization known as the East India Company. Once a loose coalition of British overseas traders, the company eventually grew to become one of the largest commercial organizations at the time, controlling around half of the world's trade (Farrington 2002). Along with its growing economic power, the company gradually took over vast territories across the Indian subcontinent, either putting them under direct company control or administering them through indigenous rulers.

Most territories of present-day India were formally organized as a British crown colony under Government of India Act of 1858, which effectively transferred political administration from the British East India Company to the crown government. Enacted just after a violent rebellion, the law was part of the numerous institutional reforms that followed, one being the establishment of Imperial Legislative Council in 1861. This legislature was mainly composed of nominated members by the governor and elected members from organizations such as Bombay Chamber for Commerce.

We note that such expansion of legislature and electoral franchise was not unique to British India. Instead, it represents a larger wave of transformation where elections for legislative institutions emerged throughout colonial states. Figure 1 shows the number of colonial states that had held its first legislative

election in the first half of the twentieth century.

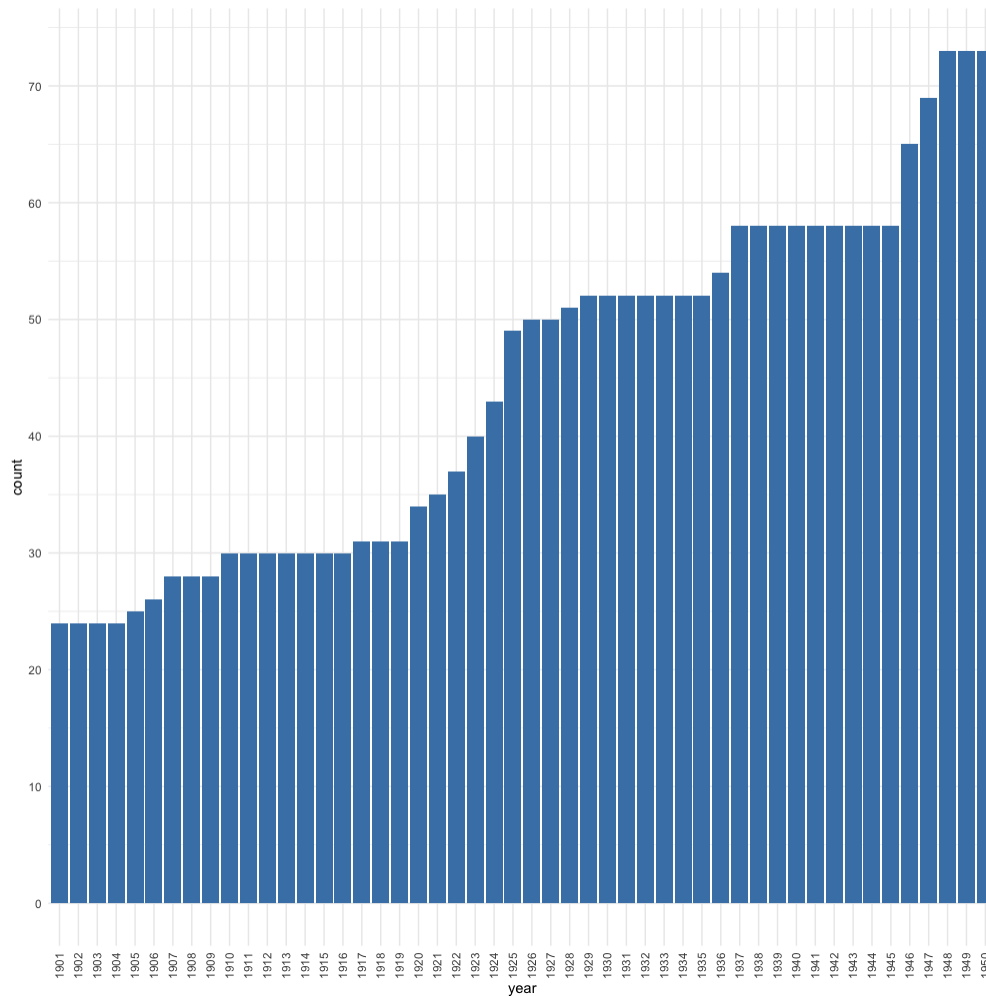


Figure 1: Number of colonial states that had held its first legislative election, 1901-1950. Data from Paine 2019.

Following the First World War, the 1919 Government of India Act introduced two key reforms that fundamentally reshaped the Indian Parliament: the bicameral division of the Indian parliament the Imperial Legislative Assembly and Council of States, and the expansion of the electoral franchise to indigenous peoples of British India (subject to gender and property restrictions). After 1919, then, both houses of the Indian Parliament contained appointed (unelected) and elected members. Notably, many elected members were members of the indigenous peoples who had been granted suffrage. Their seats could either be broadly geographic (such as Madras) or sector-specific (e.g., Muslim landholders). Figure 2 shows that the expansion resulted in a significant increase in legislature size, from roughly 60 members before 1920

to around 200 members by 1921. The growth mainly came from elected members, which increased from around 20 members to around 130 members, now representing roughly two-third of the legislature.

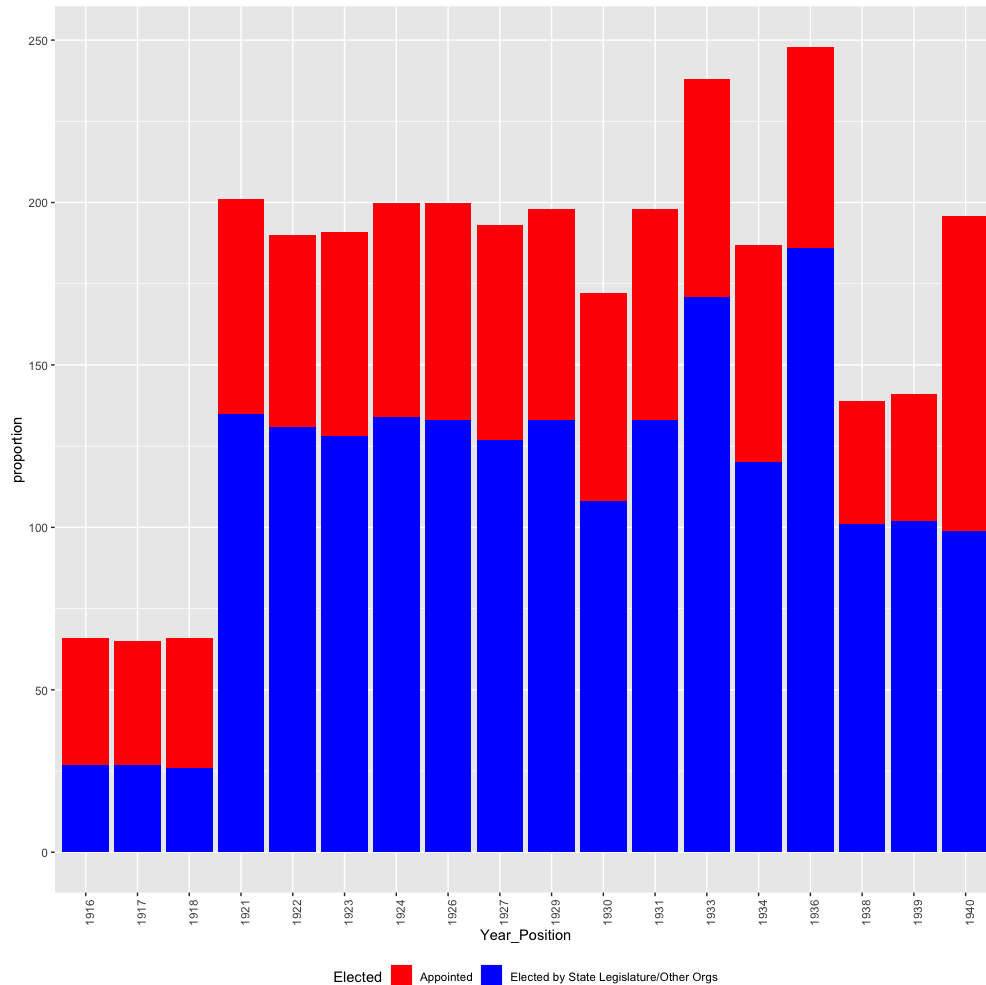


Figure 2: Number of colonial legislative members, 1916-1940. Red represents appointed members and blue represents elected members. Figures after 1920 are from both chambers. Data collected by authors from the British India Office List.

3 Data

To evaluate the effect of suffrage expansion on legislative behavior we create a novel historical data created from the documents covering British India parliamentary debates from 1919 to 1940; this totals nearly 2000 days of legislative speeches and nearly 260000 individual speeches. Figure 3 shows one page from a 1921 debate from the Legislative Assembly. Each of these documents contain daily questions and answers

as well as debates on various topics in Legislative Assembly and Council of States.

To generate these data, we first used a web scraping procedure to collect the scanned PDFs for each day of parliamentary debates from the Parliament of India Digital library¹ We then used the Tesseract Optical Character Recognition (OCR) software to translate each of these PDF files to plain text.

With the help of undergraduate RAs, we have also constructed a dataset of biographical details for members of parliament, which includes information such as whether an MP was elected, what constituency they are elected by (Muslim, non-Muslim, Sikh, Landholders, etc.), and their party affiliation. We are also in progress of constructing datasets of district-level results from elections, as well as of the electoral rules that determined who could and could not vote in each district.

With this, we constructed two datasets. In the first, each row is a page from a legislative day, and the OCR'd text is kept unedited. In the second, we parsed individual speeches using regular expressions to construct a dataset where every row is an individual speech. We then merge this speech-level debate data to biographical level data by legislative body, year, and MP name using a fuzzy matching procedure, which leads to some attenuation of the usable speeches. In total, there are 123,604 rows in the page-level dataset comprised of debates from the upper and lower houses between 1919 and 1940, and 142,856 rows in the speech-level dataset, comprised of debates from the lower house between 1921 and 1940.

¹See <https://eparlib.nic.in>. For an example of the scanned legislative debates, see https://eparlib.nic.in/bitstream/123456789/782704/1/clad_04_03_02-04-1932.pdf.

to every Honourable Member to give reasons why he opposes the amendment and wishes the House to pass the original motion. On the present occasion the Leader of the House has explained his position, that he is willing to accept the dilatory motion in order to expedite the business before the House. The House is entitled to decide, apart from the question of expediting the work before the Assembly, whether they wish to discuss this in full in all its aspects. If there is a general feeling in the House that the suggestion of the Leader of the House should be accepted then the best course would be to deal in the first instance with the amendment only. I should like to know what the general feeling in the House is in that respect. (*Some Honourable Members*: "No postponement", "No postponement"; *Some other Honourable Members*: "Postpone", "Postpone.")

(At this stage Mr. Arthur Moore rose to his feet.)

Mr. President: Do you wish to say anything on this aspect of the question, Mr. Moore?

Mr. Arthur Moore (Bengal: European): I was desiring, Sir, to say that we would like to see your suggestion adopted that we should immediately decide the amendment before the House.

Mr. President (The Honourable Sir Ibrahim Rahimtoola): As the view of the House is not clear, I should like to ascertain exactly what the feeling is. Will those Honourable Members who are in favour of postponing the discussion of the Bill till September please rise in their seats. (*Some Members rose.*) Those against this view will now rise in their seats. (*Some Members rose.*) As the division is about equal, I will allow both the original motion and the amendment to be discussed together.

Kunwar Raghbir Singh (Agra Division: Non-Muhammadan Rural): Sir, I wish to congratulate the Government in general and the Honourable the Leader of the House in particular for bringing this Bill before the House in the interests of the sugar industry in the country. Sir, the Government generally, as we know, are slow to proceed. When there was a question of putting a duty on the import of wheat, they did not agree to it, but when we defeated the Government, they were forced to levy that duty. So, Sir, as you have yourself said this amendment has been described as a dilatory motion and I wish to oppose it, because our country is essentially an agricultural country and the interests of the agriculturists should always be kept in view. When I oppose the amendment and support the original motion, it is on two grounds. In the first place, I support the original motion in the interests of the cultivator himself. Formerly, when protection was proposed, the burden of that protection used to fall on the poor people, but now the case is otherwise. In this connection, I would quote the example of steel, paper, cement and *chuddar*. The protection on these articles fell on the poor people, but this protection which has now been proposed will not fall on the poor people because they generally use *gur* and not white sugar. So, Sir, as I said the benefit of the protection will be for the cultivator and will not fall on the poor people. Secondly, my province, the United Provinces of Agra and Oudh, is a large sugar-producing province. Half the area of the whole of India which is under sugar cultivation is in the United Provinces. So,

Figure 3: One Page from the September 30th 1921 debate of the Legislative Assembly.

4 Methods

4.1 Topic Modeling

To evaluate the effect of suffrage expansion on legislative speech, we perform two analyses. In the first, we fit topic models for each of the page- and speech-level datasets described above using the STM package for the statistical programming language *R*.² A topic model is an unsupervised machine learning model that clusters documents into a pre-specified number of topics. For both models, we set the number of topics K to be 60, and use the default inference algorithm. A fitted topic model generates two parameters for further evaluation: a matrix β , where each element is the probability that a *word* is generated by a given topic, and a matrix γ , where each element is the probability that a *document* is generated by a given topic. We identify topics of interest by inspecting the top words associated with each topic in the β matrix, and then extract the values of the γ matrix associated with those topics. We then merge these estimates of topic prevalence to available metadata covariates such as date and legislative body. For the speech-level topic model, we also merge the topic prevalence results to member-level covariates.

Using the estimates of topic prevalence from the page- and speech-level dataset, we estimate the effect of 1920 and 1935 suffrage reforms on topic prevalence. More precisely, for the debate page level dataset we estimate linear models of the form

$$y_{tm} = \alpha_t + \beta_t \cdot \text{Inst. Period} + \gamma_t \cdot \text{Month} + \psi_t \cdot \text{Month} \times \text{Inst. Period} + e_{tm}, \quad (1)$$

where y_{tm} is the monthly average of topic prevalence for topic t , α_t is the intercept, β_t is the effect of institutional period, γ_t is the effect of an increase in month, ψ_t is the effect of the interaction of institutional period and month, and the error e_{tm} is clustered by topic and month. For the speech-level corpus, we leverage the fact that we know which speakers were elected vs. unelected, and so we estimate models

²While the STM package allows the user to use metadata covariates to help fit the model, we used only text features for both models.

of the form

$$y_{tme} = \alpha_{te} + \beta_{te} \cdot \text{Inst. Period} + \gamma_{te} \cdot \text{Month} + \psi_{te} \cdot \text{Month} \times \text{Inst. Period} + \epsilon_{tme} \quad (2)$$

where the difference is that we fit models where the outcome variable is y_{tme} , the average topic prevalence of topic t in month m amongst speakers e , where e is a dummy variable that takes a value of 1 if the speakers are elected, and 0 if they are not. In each model the *psi* coefficient can be interpreted as the difference in intercepts at the point of the change between the institutional periods (i.e, before and after suffrage expansion).

4.2 Text-Based Ideal Points

While the goal of the topic modelling is to determine whether MPs discuss certain topics more after suffrage expansion, and if they are elected or not, it will not tell us whether elected and unelected MPs speak about the same topics in different ways. To estimate the effect of an MP being elected on the way they speak about a particular topic, we used speeches that were most highly associated with the "education" topic from the topic model, and employed a Text-Based Ideal Point (TBIP) estimation model from Vafa, Naidu, and Blei (2020). The TBIP model estimates the parameters of the Poisson factorization model

$$y_{dv} \sim \text{Pois} \left(\sum_k \theta_{dk} \beta_{kv} \exp\{x_{ad} \eta_{kv}\} \right), \quad (3)$$

where y_{dv} is the word count of vocabulary term v in document d , θ_{dk} is a scalar representing the intensity of the non-ideological part of topic k in document d , β_{kv} is a scalar representing the frequency of word v in the non-ideological part of topic k , x_{ad} is the scalar ideal point of author a for document d , and η_{kv} is the scalar representing the frequency of a word v in the ideological part of the topic. The goal is to estimate the parameters of this model, one of which is the ideal point x_{ad} . The high level interpretation of this equation is that for k topics, the data matrix (a document-term matrix where each row is a document and each column is a count of words) is factorized into matrices θ , which contains per-document topic intensities, β , which contains the non-ideological (i.e., common across speakers) dimension of the

topics and η , which contains the ideological (i.e., not common across speakers) topics, and x , which is a scalar ideal point for the author that persists across topics. For a more complete description of the model, see section 2.3 in Vafa, Naidu, and Blei (2020). Because of the intractability of this model, the authors recommend using a variational inference algorithm to estimate the model's parameters. We follow this advice, and run the model on Google Colab, a free cloud-based computing service³.

³For instructions on how to run the model, see <https://github.com/keyonvafa/tbip>

5 Results

5.1 Structural Topic Model (STM) Results

First, we look at the results from STM to uncover the changes in topic prevalence in legislative debates among all legislative bodies after the two events of franchise expansion (Figure 4). Overall, based on the topic prevalence over time, we did not observe a significant rise in education (Topic 20, or Row 1 and Column 4 of the figure) as a topic in legislative debates after the franchise expansions in 1919 and 1935. Instead, while education's prevalence was increasing before 1919, it declined after 1919 and remained relatively stable until 1940.

Next, we narrow down the STM results to include only the debates in Imperial Legislative Assembly (Lower House), which, compared to the Council of State (Upper House), had significantly more elected legislative members (Figure 5). Additionally, we divide the speeches by the electoral status of the speaker (elected or nominated). Similar to the previous results, we find that there were no substantial difference between the two groups in terms of education's topic prevalence. Given this, STM results do not provide any empirical support for a change in legislative focus on education after the two events of franchise expansion.

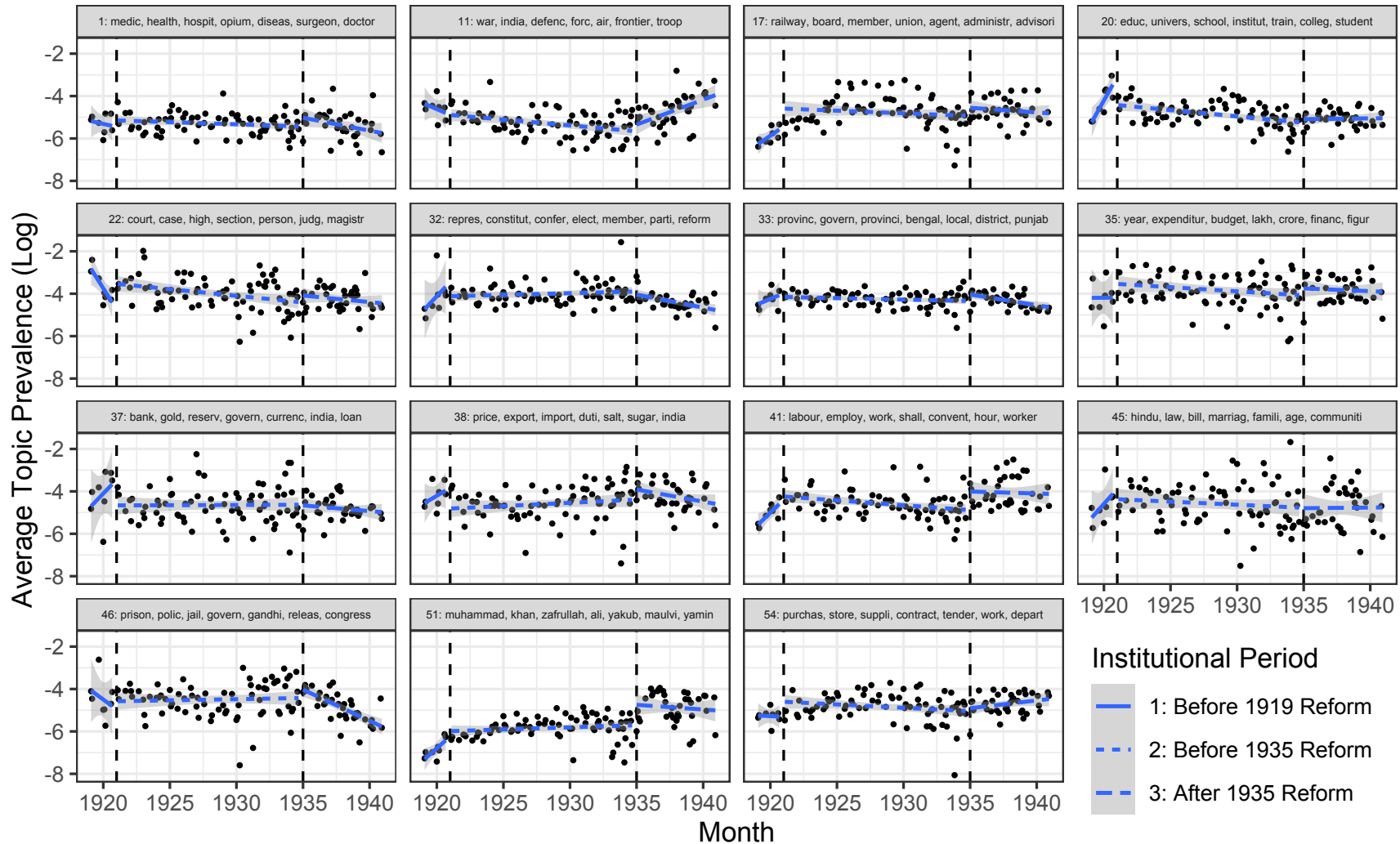


Figure 4: **Estimating the Effect of Suffrage Expansion on Topic Prevalence in all Legislative Bodies, 1919-1940.** This figure shows the distribution of the logged value of average monthly topic prevalence pooled across the pre-1919 reform Legislative Council, and the post-1919 reform Legislative Assembly and Council of States. The text above each panel indicates the topic number, and the top words associated with that topic. The dashed lines indicate lines of best fit for each institutional period, with the shaded areas representing confidence intervals.

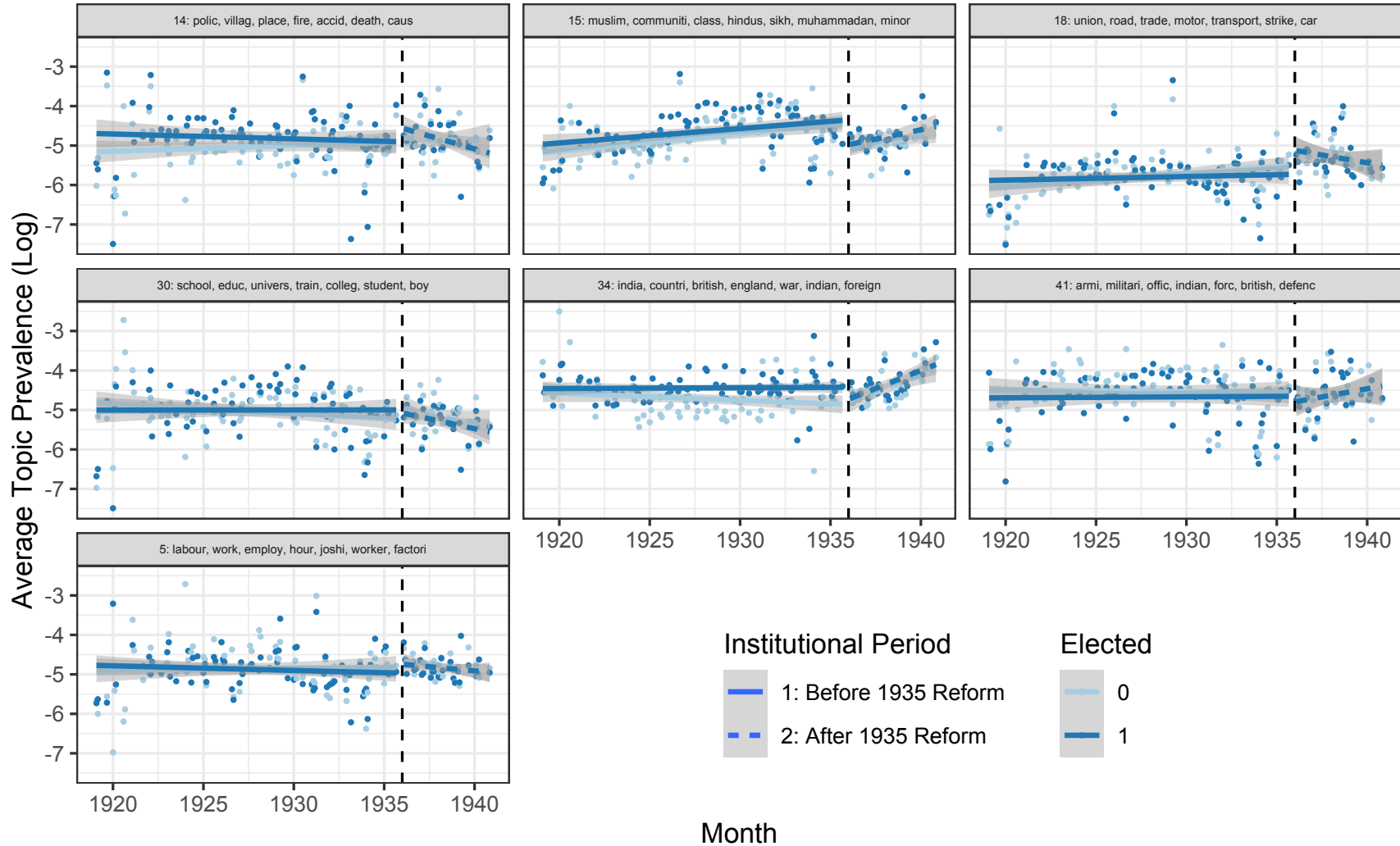


Figure 5: **Estimating the Effect of Suffrage Expansion on Topic Prevalence in the Indian Lower House by year and elected status, 1919-1940.** This figure shows the distribution of the logged value of average monthly topic prevalence pooled across the post 1919 reform Legislative Assembly. The text above each panel indicates the topic number, and the top words associated with that topic. The color of the points indicates whether the logged monthly average of topic prevalence refers to elected or unelected members. The dashed lines indicate lines of best fit for each institutional period, with the shaded areas representing confidence intervals.

5.2 Ideal Point Estimation Results

We then further investigate if there were different policy preferences between elected versus nominated members, and if these differences, if any, changed over time. To do this, we first look at the distribution of ideal point estimates by electoral status for two periods: 1921-1930 and 1931-40 (Figure 6). In the first period (1921-30), the policy preferences for education remained largely the same between the elected and nominated legislative members. We also observed a noticeable shift in the second period (1931-40), where we observed some differences in ideal point estimates between the two groups. Regardless, from the graphical evidence, we find no strong empirical support for the hypothesis that the policy preferences between elected and nominated members regarding education were significantly different.

Next, we further divide the distribution of ideal point estimates into different sub-groups (Figure 7). Similar to the previous results, while there were some shifts in policy preferences regarding education, we could not observe a substantial difference between different sub-groups of electoral constituencies.

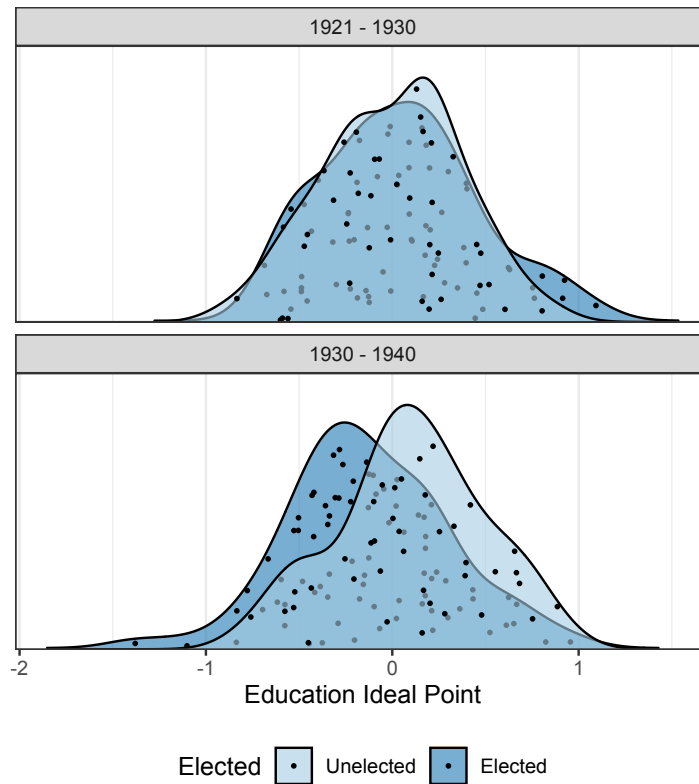


Figure 6: **Inspecting Education Ideal Point Density by Decade and Elected Status, 1921-1940.** This figure breaks down the smoothed density of ideal points over education (as estimated by the Text-Based Ideal Point (TBIP) model) by decade and by whether the MP was elected or not. Individual points under each curve show the actual (i.e., not smoothed) distribution of ideal points.

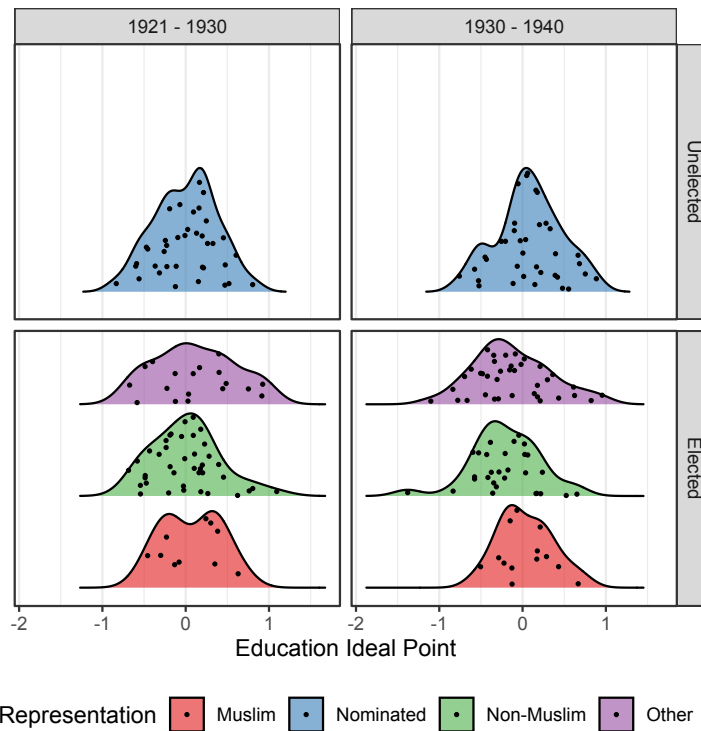


Figure 7: **Inspecting Education Ideal Point Density by Decade, Elected Status, and Representation, 1921-1940.** This figure breaks down the smoothed density of ideal points over education (as estimated by the Text-Based Ideal Point (TBIP) model) by decade, whether the MP was elected or not, and who the MP represented. Individual points under each curve show the actual (i.e., not smoothed) distribution of ideal points.

	(1)
(Intercept)	0.045 (0.050)
Elected	-0.046 (0.082)
Decade	-0.174 * (0.068)
Elected x Decade	0.275 * (0.116)
N	224
R ²	0.046
logLik	-116.431
AIC	242.862

*** p < 0.001; ** p < 0.01; * p < 0.05.

Table 1: **Estimating the Effect of Being an Elected MP on Ideal Points Over Education, 1921-1940.** This table describes the results of a simple OLS linear regression where the ideal point of an MP over education policy is regressed against their status as an elected or unelected member, the decade, and the interaction of decade and elected status. The results compliment the relationship shown visually in Figure 6, that the ideal points of the elected and unelected members begin to diverge in the second decade.

6 Discussion and Conclusion

In this paper, we investigate the relationship between government policy interest in education and electoral franchise expansion. By doing so, we bring together the existing literature on electoral franchise expansion, and the literature on the legacies of colonial policies. Contrary to our initial expectations, our preliminary results do not find a substantial change in government interest in education after franchise expansion in British India, either in terms of education's topic prevalence in the legislature or individual legislative member's preferences regarding education. Future research can further explore why, at least in colonial state contexts, the expansion of electoral franchise did not lead to a substantial change in goods and services provision such as education.

References

- Acemoglu, Daron, Simon Johnson, and James A Robinson. 2001. "The colonial origins of comparative development: An empirical investigation." *American Economic Review* 91 (5): 1369–1401.
- Aghion, Philippe, Xavier Jaravel, Torsten Persson, and Dorothée Rouzet. 2018. "Education and military rivalry." *Journal of the European Economic Association* 17 (2): 376–412.
- Aidt, Toke S, Jayasri Dutta, and Elena Loukoianova. 2006. "Democracy comes to Europe: franchise extension and fiscal outcomes 1830–1938." *European Economic Review* 50 (2): 249–283.
- Ansell, Ben, and Johannes Lindvall. 2013. "The Political Origins of Primary Education Systems: Ideology, Institutions, and Interdenominational Conflict in an Era of Nation-building." *American Political Science Review*, 505–522.
- Banerjee, Abhijit, and Lakshmi Iyer. 2005. "History, institutions, and economic performance: The legacy of colonial land tenure systems in India." *American Economic Review* 95 (4): 1190–1213.
- Cappelli, Gabriele. 2016. "One size that didn't fit all? Electoral franchise, fiscal capacity and the rise of mass schooling across Italy's provinces, 1870–1911." *Cliometrica* 10 (3): 311–343.
- Dell, Melissa. 2010. "The persistent effects of Peru's mining mita." *Econometrica* 78 (6): 1863–1903.
- Farrington, Anthony. 2002. *Trading Places: The East Asia Company and Asia, 1600-1834*. The British Library.
- Gailmard, Sean. 2017. "Building a New Imperial State: The Strategic Foundations of Separation of Powers in America" [in en]. *American Political Science Review* 111, no. 4 (November): 668–685. ISSN: 0003-0554, 1537-5943. <https://doi.org/10.1017/S0003055417000235>.
- . 2019. "Imperial Politics, English Law, and the Strategic Foundations of Constitutional Review in America" [in en]. *American Political Science Review* 113, no. 3 (August): 778–795. ISSN: 0003-0554, 1537-5943. <https://doi.org/10.1017/S0003055419000212>.

- Harding, Robin, and David Stasavage. 2014. "What Democracy Does (and Doesn't Do) for Basic Services: School Fees, School Inputs, and African Elections." *The Journal of Politics* 76 (1): 229–245.
- Husted, Thomas A, and Lawrence W Kenny. 1997. "The Effect of the Expansion of the Voting Franchise on the Size of Government." *Journal of Political Economy* 105 (1): 54–82.
- Lindert, Peter H. 2004. *Growing Public: Volume 1, The Story: Social Spending and Economic Growth since the Eighteenth Century*. Vol. 1. Cambridge University Press.
- Opalo, Ken. 2021. "Colonialism and Institutional Persistence: Mixed Legislative Legacies in Ghana and Kenya."
- Paglayan, Agustina S. 2020. "The Non-Democratic Roots of Mass Education: Evidence from 200 Years." *American Political Science Review*, 1–20.
- Paine, Jack. 2019. "Democratic Contradictions in European Settler Colonies." *World Politics*, 1–44.
- Testa, Patrick A. 2018. "Education and propaganda: Tradeoffs to public education provision in non-democracies." *Journal of Public Economics* 160:66–81.
- Vafa, Keyon, Suresh Naidu, and David M. Blei. 2020. "Text-Based Ideal Points." *arXiv:2005.04232 [cs, stat]* (July). arXiv: 2005.04232 [cs, stat].