De-Polarizing Politics: Faith Meets Science in the French Third Republic*

Gianandrea Lanzara (University of Bologna)
Mara P. Squicciarini (Bocconi University and CEPR)
Nico Voigtländer (UCLA and NBER)

This version: February 2026

Abstract

We investigate whether individual politicians can help reduce emotionality and polarization in parliamentary debates. Focusing on the French Third Republic (1880–1914), we study Members of Parliament (MPs) educated in Jesuit schools (collèges) – institutions combining religious formation with rigorous scientific, philsophical, and rhetorical training. Using newly digitized data on Jesuit school enrollments and over one million parliamentary speeches, we identify Jesuit-educated MPs (JeMPs) and analyze their linguistic and behavioral patterns through text embeddings and large language models. We find that JeMPs systematically employed less emotional and less polarized rhetoric, emphasizing reason and science in their arguments. They also engaged more directly with preceding speeches, as reflected in greater lexical overlap. We then examine the dynamics within parliamentary debates, where the order of speakers was by and large pre-assigned. Even after controlling for speaker-discussion fixed effects, we find that JeMP interventions reduced subsequent speakers' emotionality and polarization. These effects were particularly pronounced during the contentious 1905 Separation of Church and State Law.

JEL: K42, N43, O10, Z10

Keywords: Schooling, Political Polarization, Mediation, Speech Analysis

*We thank Jérôme Adda, Greg Clark, Tommaso Sonno, Juliane Uhl, and audiences at 1st CEPR Applied Micro-Economic History Workshop, ASREC 2024, Bologna University, CUFE, IIES, LISER, NYU Stern, Southern Denmark University, Tilburg University, Tor Vergata University, University Halle-Wittenberg, University of Bergamo, Utrecht, and Warwick for insightful comments and suggestions. This research was funded by the European Union (ERC) under grant agreement 101042055 - EDIPO. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

"...every good Christian is to be more ready to save his neighbor's proposition than to condemn it. If he cannot save it, let him inquire how he means it; and if he means it badly, let him correct him with charity."

— Ignatius of Loyola (1491–1556), *The Spiritual Exercises*, Annotation 22 (Loyola, 1914).

1 Introduction

Political polarization has become a defining feature of contemporary democracies. Beyond shaping the tone and animosity of political discourse, polarization can undermine legislative effectiveness, weaken institutional checks and balances, and increase policy uncertainty, with negative economic consequences (McCarty, Poole, and Rosenthal, 2006; Baker, Bloom, and Davis, 2016). A large literature documents how to measure polarization in political language and traces its evolution in legislative speech (Gentzkow, Shapiro, and Taddy, 2019), as well as its institutional and electoral underpinnings (Fiorina and Abrams, 2008). Yet much less is known about whether political actors themselves can mitigate polarization within legislative arenas. Can certain types of politicians – through their training, rhetoric, or mode of engagement – act as mediators in polarized environments, moderating debate and reducing the intensity of partisan conflict?

We examine this question in the context of the French Third Republic (1870–1914), a period marked by intense ideological conflict between a progressive republican coalition and its conservative opponents. We focus on members of parliament (MPs) who were educated in Jesuit schools (*collèges*), whom we refer to as Jesuit-Educated MPs (JeMPs). Jesuit education combined rigorous training in theology and humanities, but also in mathematics and the natural sciences, as well as an extensive instruction in rhetoric. Importantly, students in Jesuit schools practiced formal disputations designed to cultivate disciplined reasoning and engagement with opposing viewpoints.[1] We digitize the complete enrollment registries of Jesuit schools in France from the Archives Jésuites, which record student names, surnames, and grade-level information. These records cover approximately 70,000 individual students enrolled between 1850 and 1901. We match these student rosters to the universe of French MPs using detailed biographical information. We identify 284 parliamentarians with a Jesuit education, representing approximately 8% of the total number of MPs in our sample period. We then combine this matched dataset with a newly assembled corpus of all parliamentary

[1]The standardized Jesuit curriculum was codified in the *Ratio Studiorum*, rooted in the pedagogical principles of Ignatius of Loyola (1491–1556), who had founded the Society of Jesus in 1540. From the 16th century onward, Jesuits became one of the most influential teaching orders in Europe, operating a vast network of schools and universities. Historians of education emphasize the central role of Jesuit institutions in shaping early modern European schooling and their reputation for intellectual rigor and disciplined pedagogy (see for example O'Malley (1993) and Grendler (1989)).

debates in France between 1880 and 1914, covering more than one million speeches as well as the full set of roll-call votes cast during these decades.

We document three main findings. First, using Natural Language Processing (NLP) based measures of emotionality and polarization, we show that JeMPs' speeches were systematically less emotional and less polarized than those of other MPs, even within political area and after controlling for topic as well as speaker and speech characteristics. JeMPs also placed greater emphasis on scientific reasoning in their interventions. Second, we document evidence that JeMPs affected the debating style of other legislators. Exploiting institutional features of the Chamber that largely predetermined the order of speakers within a discussion, we estimate event-study specifications around the first JeMP intervention. Following a JeMP speech, subsequent interventions became significantly less emotional and less polarized. These effects persist with speaker-by-discussion fixed effects, ruling out endogenous sorting of less emotional or less polarized MPs into the post-JeMP speaking order. To assess whether the estimated effects reflect distinctive features of Jesuit-educated MPs rather than generic characteristics of certain legislators or districts, we implement placebo exercises that assign "placebo JeMP status" in three ways: i) randomly selected non-Jesuit MPs; ii) the 5% least emotional (or least polarized) non-Jesuit MPs; iii) MPs elected in the 10% least emotional or polarized départements. Our results indicate that the moderating influence we document is not replicated by randomly chosen legislators, by MPs with similarly moderate rhetorical profiles, or by representatives from less polarized electoral environments.

Finally, we examine the 1905 *Law Separating Churches and State*, one of the most polarizing reforms of the Third Republic that made secularism the guiding principle of the country. The 1905 Law ended state recognition and funding for any religious affiliation, required worship to be organized through tightly defined associations, and reallocated many assets of the former worship establishments to public bodies and local authorities. Despite strong material and ideological reasons to oppose the reform, JeMPs employed comparatively neutral and less emotional language in this debate. Moreover, their interventions reduced the incidence of extreme pro-secularization rhetoric among subsequent speakers without symmetrically amplifying extreme opposition. We further provide exploratory evidence that JeMP participation in debates reduced pro-secularization votes among other MPs, consistent with a moderating influence that may extend beyond rhetoric.

Taken together, our findings suggest that depolarization need not rely exclusively on institutional redesign or shifts in mass preferences. It can emerge endogenously through the composition of the political elite: when legislators possess training that emphasizes disciplined reasoning, engagement with opposing viewpoints, and restraint in rhetoric, their participation can moderate deliberation and attenuate conflict. This highlights a complementary channel through which polarization may be shaped—not only by institutions or voters, but by the formative experiences of those who hold

office.

We contribute to several strands of literature. First, we add to the growing body of work measuring political polarization in legislative speech and behavior (McCarty et al., 2006; Bonica, 2014; Gentzkow et al., 2019). A large literature has developed tools to quantify ideological divergence in roll-call voting and high-dimensional text data. Our contribution shifts the focus from aggregate trends to individual-level heterogeneity in rhetorical style and, crucially, to its spillover effects within debates. Second, our findings speak to the political economy literature on the causes and consequences of polarization. Prior work emphasizes institutional and party-level drivers of polarization (Fiorina and Abrams, 2008; Moskowitz, Rogowski, and Jr., 2024; Canen, Kendall, and Trebbi, 2020, 2026). This literature primarily analyzes polarization through shifts in party competition and strategic positioning, rather than through differences in how individual legislators engage in deliberation. We complement this perspective by showing that even in highly polarized environments, the presence of certain legislators can measurably affect the rhetoric of others and alter the dynamics of debate.

Third, our paper contributes to research on political communication and elite behavior in polarized environments. Bäck, Carroll, Renström, and Ryan (2023) show that partisan source cues (that is, attributing an otherwise neutral message to either a voter's most liked or most disliked party) reinforce partisan social identities among voters and heighten intergroup differentiation in the electorate; by contrast, we examine how heterogeneity among elites themselves shapes the tone and dynamics of deliberation within legislative institutions. Henríquez, Sabat, and Sullivan (2023) show that cross-party "likes" among Chilean deputies are positively associated with subsequent cross-coalition voting, interpreting such interactions as costly signals of ideological agreement. Whereas they study strategic signaling between legislators in social media and its relationship to roll-call behavior, we examine how exposure to particular types of legislators shapes the tone and dynamics of deliberation within parliamentary debate itself. Fishkin, Siu, Diamond, and Bradburn (2021) document substantial reductions in both issue-based and affective polarization following structured, balanced cross-partisan discussion in a national deliberative field experiment ("America in One Room"). Their findings suggest that exposure to institutionalized, balanced deliberation can moderate political attitudes. Similarly, Doornbosch, van Vuuren, and de Jong (2025) examine three facilitated bipartisan face-to-face dialogues among U.S. members of Congress and identify communicative strategies – such as embracing multiple frames and foregrounding superordinate identities – that foster convergence across ideological and identity divides. Relatedly, Downey, Kuk, and Nanes (2024) provide experimental evidence that affective polarization is sustained in part by voters' misperceptions about how well their issue positions align with those of their co-partisans. The authors show that correcting these inaccurate beliefs reduces partisan affect, suggesting that polarization can be mitigated when individuals recognize cross-cutting issue misalignment. The results from these three papers are consistent with

the idea that systematic training in rhetoric and awareness/engagement with opposing views – such as that received by Jesuit-educated MPs – may foster moderation. We investigate this effect in a natural legislative setting, where deliberation unfolds naturally within an actual parliament rather than in structured deliberative settings. We identify a specific group of politicians whose background predicts a systematic moderating influence on debate – even during one of the most contentious reforms of the period.

Fourth, a growing literature examines the formation and selection of political elites, highlighting how educational and career pathways shape who enters politics and how they govern once in office. Bó, Finan, Folke, Persson, and Rickne (2017) show that individuals with higher ability and stronger earnings potential are more likely to enter politics when the returns to office increase. At the same time, education shapes civic engagement and political behavior more broadly (Milligan, Moretti, and Oreopoulos, 2004), and the content of schooling influences long-run ideological beliefs (Cantoni, Chen, Yang, Yuchtman, and Zhang, 2017). State leaders' educational backgrounds affect governance outcomes and reform behavior (Besley, Montalvo, and Reynal-Querol, 2011; Dreher, Lamla, Lein, and Somogyi, 2009). This literature suggests that education shapes how politicians behave in office, suggesting that distinctive training environments (in our case, Jesuit schools) can influence the style and tone of legislative deliberation.

Finally, our historical setting contributes to the literature on the development of democratic institutions and legislative deliberation. The French Third Republic provides a highly polarized environment characterized by deep conflicts over secularization, church–state relations, and republican reform. By assembling a comprehensive corpus of parliamentary speeches and roll-call votes and linking them to individual-level educational backgrounds, we provide new large-scale evidence on deliberation in a foundational democratic period. Our analysis complements work emphasizing the importance of elite norms and informal institutions in sustaining democratic stability (Levitsky and Ziblatt, 2018) and the role of communicative practices in shaping democratic legitimacy (Habermas, 1996). It also connects to empirical studies of legislative speech showing that linguistic patterns predict legislators' ideological positions, highlighting the informational content of parliamentary rhetoric (Diermeier, Godbout, Yu, and Kaufmann, 2012). More broadly, the paper suggests that the structure of democratic debate is shaped not only by formal rules and electoral incentives, but also by the intellectual traditions and formative training of those who hold office.

2 Historical Background

2.1 Jesuits in France

At the beginning of the 18th century, the Society of Jesus was highly influential in France (van Kley, 1975). Jesuits ran a large network of schools (*collèges*) across the country and provided a high-quality education for "future leaders of French society" (Thompson, 2018). This was based on rigorous training in humanities, rhetoric, theology, but also in sciences and mathematics. The Jesuits' contribution to sciences is well-known, especially during the Enlightenment era. A prominent example is the Jesuit mathematician and physicist Louis-Bertrand Castel (1688–1757). Castel, who entered the Jesuit order in 1703, was a prominent Paris-based intellectual and served as science editor of the Jesuit *Journal de Trévoux*.[2] He wrote on physics and universal gravity, proposed the famous "ocular harpsichord," while engaging in a critique of Newton's spectral description of prismatic color.[3]

The Jesuit school network was influential not only because it shaped elite formation, providing its students with a humanistic and scientific training, but also because it created a nationally-connected community that many contemporaries often perceived as a rival source of authority to the crown, or later to the republican state. This perception partly contributed to a strong anti-Jesuit campaign in the early 1760s and to the later ban of the order in 1764.[4] The expulsion of Jesuits from the country lasted until 1814, when the Society of Jesus was permitted to return to France.

After a fragile rebuilding of the Jesuit school network, which suffered an additional closure in 1828, it was the 1850 Falloux Law that legitimated and expanded the space for non-state schooling and triggered the emergence of a longer-term and flourishing community of Jesuit educational institutions.Among these were the famous *Collège de l'Immaculée-Conception* (commonly called Vaugirard) and the *Ecole Sainte-Geneviève de la rue des Postes* (Langdon, 1978).[5]

A new attack against the Society of Jesus came with the Republican government. After the Franco-Prussians wars and a short period dominated by conservative and monarchist forces (1870–1876), from 1877 onward, the political equilibrium shifted toward the progressive reforms implemented by

[2]The *Journal de Trévoux (Mémoires de Trévoux)* was a scientific journal, founded in 1701 and run largely by Jesuit editors. It published critical reviews of contemporary books across "sciences & beaux-arts" (and broader learned culture). See https://artfl-project.uchicago.edu/journal-de-trevoux for an overview.

[3]Another example of a prominent Jesuit scientist is Antoine Gaubil (1689–1759), a leading astronomer who bridged between the Chinese and European scientific traditions, representing an important example of the Jesuit role in transcontinental scientific communication.

[4]The precipitating event was a scandal involving a Jesuit missionary in the French Caribbean, Father Lavalette. In 1761–62, the magistrates, under the strong influence of the Jansenists, claimed that the Jesuit principles were incompatible with the laws and traditions of the kingdom, thus suppressing the Jesuit collèges and requiring members of the order either to abjure their vows or to depart from French territory. Unable to effectively counter the decisions of the magistrates, Louis XV ultimately ratified the ban in 1764 (Vogel, 2010).

[5]The *Collège de l'Immaculée-Conception* was founded in Paris in 1852. Its nickname comes from its location on the rue de Vaugirard. It soon represented the flagship of the Jesuit collèges as the first school being established in Paris after 1850. The *Ecole Sainte-Geneviève de la rue des Postes* was particularly well-known for providing a post-baccalauréat instruction and its students had an extremely high admission rate at the *grandes écoles*.

the Republicans. One of the main objectives was the redefinition of the role of religion in public life toward building a secular state.[6] Secularization laws were poised to grant the State full control over education. This was intended as a means to ensure cohesion and unity of the citizens beyond any past ideological divisions (Mayeur, 1981). However, to achieve this goal, Republicans had to weaken and dismantle their main competitors in the training of the French youth, i.e., the entire network of Catholic schools. Following the exemplary 1882 Jules Ferry Education Act, most secularization reforms addressed all Catholic institutions.[7] However, in some cases, laws were specifically directed against the Jesuits. For instance, the Ferry administration, orchestrated another attempt to finally expel all Jesuits from France and to close their *collèges*.[8] What drove this anti-Jesuit sentiment, despite their large contribution to the training of the upper-tail human capital of the country? Not too differently from the anti-Jesuit spirit at the beginning of the 19th century, Jesuits were criticized "for the tendency to mobilize youth to their goal of state domination" (Doyle, 2017) and for training the Catholic elite of the time to cover important governmental positions, so that "the whole state bureaucracy, [...] was under siege by Jesuit affiliates" (Bush, 1975, p. 127).

The attempt to close all Jesuit collèges was not as effective as initially planned by the Republicans. Many schools were reorganized or reconstituted through alumni support and lay intermediaries.[9] By the early 1890s, the vast majority of Jesuit collèges became again fully operational, managing to recruit a large set of students (Bush, 1975). The definitive closure took place with the 1901-1904 *Lois Anti-congreganistes*, when thousands of religious schools were closed – and the Jesuit collèges shared the same fate.[10] The secularization policies culminated in the 1905 Law "Concerning the Separation of Churches and State," which firmly enshrined the principle of laïcité in France.[11]

[6]This was accompanied by a broader package of "Republican laws," which covered, for instance, freedom of the press, and freedom to form trade unions.

[7]The 1882 Jules Ferry Education Act is considered a milestone of the secularization program promoted by the Republican government that made primary education mandatory, free, and secular (see Squicciarini, 2020).

[8]As Jules Ferry himself puts it, in one of his speeches at the Chamber of Deputies, in 1879: "In education it's these [Jesuits] who give the tone; it's these who furnish the example for others; these who enjoy the success...who serve as model for all the ecclesiastical establishments," justifying the particularly strong opposition against the order.

[9]For the Vaugirard school, "under the supervision of *Monseigneur* D. Forges, Bishop of Ténarie and former auxiliary to Cardinal Brossais-Saint-Marc, Archbishop of Rennes, lay teachers provided the instruction; and a board of administration made up of former pupils and chaired by *Monsieur* Édouard Lefébure managed and oversaw everything relating to financial upkeep. As political passions gradually seem to subside, priests and scholastics of the Society reappear in the teaching posts of the classes and in those of advanced studies" (our translation from Delattre, 1949, pp. 1384–1385).

[10]The Law of July 1901 dissolved religious congregations, and the law of July 1904 prohibited all religious congregations from teaching, requiring congregational schools to close within ten years, and preventing the creation of any new congregational teaching institutions (Harrigan, 2001).

[11]We provide further details on the 1905 Law in Section 4.4.

2.2 Jesuit *Collèges* in the 19th Century: Organization and Pedagogy

Starting with the 1850 Falloux Law, the number of Jesuit collèges increased significantly over time, reaching a total of 30 institutions in operation between 1850 and 1905 (including three in Paris). Figure 1 shows the evolution of Jesuit collèges over our period, and Figure FIG: Map Location Jesuit Schools depicts their distribution across French departments.

Jesuit collèges typically enrolled students aged 8 to 18. Recruitment was addressed to relatively high-social-status families, but tuition cuts or scholarships were often granted to lower-income students. Tuition was comparable to that of secular lycées but generally higher than in non-Jesuit ecclesiastical schools (Bush, 1975).[12] Jesuit students were disproportionately successful in terms of educational achievement. For instance, their success rate in the baccalauréat examination was significantly higher than the national average. Even more strikingly, although they comprised only 5% of total secondary school enrollment, they accounted for 14.4% of students admitted to the grandes écoles between 1892 and 1901 (Bush, 1975).

Pedagogically, the curriculum, principles, and teaching methods of the Jesuit collèges were codified in the 1599 Ratio Studiorum, the comprehensive educational code of the Jesuit order. The Jesuit school curriculum blended intensive study of theology and the humanities with strong preparation in mathematics and the natural sciences, along with thorough training in rhetoric. A key feature of Jesuit education was the practice of formal disputations, through which students developed disciplined reasoning and engagement with opposing viewpoints.

In particular, the first years of Jesuit education were devoted to an immersion in classical letters and humanities. Then, students entered a three-year philosophy course, which, following the 1599 Ratio, included logic, physics, and astronomy together with methaphysics and philosphy. Additionally, mathematics was a core subject, to be studied daily by students in the second year of the philosophy course (Farrell, 1970). This structure and emphasis on sciences is evident also in the school curricula of French collèges in the 19th century. As an example, we provide the curriculum of the Jesuit school in Vannes in Figure 3. A second distinctive feature of Jesuit education was the emphasis on rhetoric. Its study was based on canonical authors – especially Cicero and Aristotle – combined with structured reading lists and stylistic training, including tropes, figures, and rules of style (Schwickerath, 1904).

The strong emphasis placed on disputation and rhetorical training in Jesuit education can be traced back to Ignatius of Loyola (1491–1556), the founder of the Society of Jesus. In The "Spiritual Exercises," Ignatius articulated a pedagogical and spiritual method grounded in disciplined argumentation. Our opening quote of Ignatius' Paragraph 22 (on top of Section 1) captures the key characteristics of Jesuit pedagogy and rhetoric: the defense of truth and the insistence that such defense be conducted

[12] Tuition ranged from 1000 to 1600 francs for boarding students in Paris and from 800 to 1100 francs in the provinces (Bush, 1975). This corresponds roughly to the average annual income at the time.

with intellectual honesty and respect. This approach was formally institutionalized in the Ratio, which considers disputation among students as a central pedagogical practice, especially for the classes of theology and philosophy. The Ratio prescribed a specific ethos for argumentative engagement. Another illustrative quote is: "In arguing debatable questions, he should defend his view with such modesty and courtesy as to show respect for the contrary view...When it is possible to reconcile diverse views, an attempt should be made to do so" (Farrell, 1970, p. 26). Disputation was, thus, conceived as a disciplined exercise in reasoning, persuasion, and in the reconciliation of differing perspectives. The impact of this disputation tradition is evident in other Jesuit initiatives, such as the Conférence Olivaint. Founded in Paris in 1874, the Conférence Olivaint was a Jesuit student association, dedicated to political discussion and debates with public figures. Meetings were held weekly, and it represented a natural extension of the pedagogical model into the political sphere. Interestingly, the association continues to exist today (www.conferenceolivaint.fr) and, despite not being run by Jesuits, it continues to engage with contemporary political and intellectual life.

3 Data

Our analysis relies on three main sources of the data. First, we digitize the enrollment registries of Jesuit schools in France from 1850 to 1901. Second, we digitize the universe of parliamentary speeches for the French Third Republic from 1881 to 1942. Third, we use individual level data on all MPs during the Third Republic with detailed biographical information. In the remainder of this section, we describe each of these sources, and how we link them together. The final output is a dataset of parliamentary speeches where each speaker, whenever elected as a member of parliament is associated with key personal and political characteristics, among which, importantly, whether they were educated in a Jesuit school. More details on the data construction are provided in Appendix A.

3.1 Jesuit students

Data on students enrolled in Jesuit schools in France are available at the *Archives Jésuites* in Vanves, near Paris[13], which preserve enrollment registries for 30 of the 32 Jesuit schools that operated in France between 1850 and 1901.[14] The student registers are preserved in 142 separate documents dating from 1850 to 1901. For 37 of these documents, only the handwritten (hereafter HW) version has survived. Figure A.1 in Appendix A.1.1 illustrates the student registry for school year 1885-86 at the *Collège de l'Immaculée-Conception*, located in *Rue de Vaugirard* in Paris. As in this example, registers from each Jesuit school typically report the list of students enrolled by school year along with additional information on each student's grade, section, division (a group of students spanning

[13] See https://www.jesuites.com/contact/bureaux-archives-jesuites for an overview.

[14] The archives lack documentation for the schools in Montauban and Mende, part of the Province of Toulouse, as well as for those in Algiers and Oran, located in colonial Algeria and associated with the Province of Lyon.

multiple grades who shared the same boarding facilities), and boarding status. As an exception to the standard format, eight documents report only the name and year of entry for students.

For printed school registries, our digitization process consists of four main steps: (1) We manually scan all front pages, legends, and pages reporting the lists of students in alphabetical order; (2) We perform Optical Character Recognition (OCR) on each document using Google Document AI.[15] (3) We parse and standardize the OCR output to harmonize column types (e.g., name, student number, class, year of entry) despite substantial variation in formatting. (4) With the help of research assistants, we manually reviewed all final Excel files to identify and correct potential OCR errors. For handwritten (HW) documents, we enter the data manually for the entire document. At the end of this process, we obtain a dataset with 268,216 student records. Some of these records belong to the same student observed in multiple school years, while others refer to students who appear only once in the dataset along with their year of entry at their school.

We then reconstruct the school histories of Jesuit-school students based on their names and the years during which they appear in the records. The main objective is to determine the earliest and latest years during which each student attended a Jesuit school. This information, in turn, helps us estimate the student's age, which is ultimately needed to match student records to parliamentarians. Concretely, we begin by constructing the variables $y_r^{\min}$ and $y_r^{\max}$ for each *record* r as follows: (i) if the document reports the school year in standard format, e.g., 18XX–18(XX+1), we assign 18XX to $y_r^{\min}$ and 18(XX+1) to $y_r^{\max}$; (ii) if the document only reports the year of entry, we assign this value to both $y_r^{\min}$ and $y_r^{\max}$. We then implement an iterative procedure that merges records into student-level clusters and updates implied enrollment windows accordingly. In each round, we link records whose current enrollment intervals overlap within a four-year window and whose names match under a conservative fuzzy criterion. The procedure iterates until no further links are formed [16]. As a final step, we use class information reported in the original documents, whenever available, to estimate the age of each student. As a first step, we harmonize class codes across schools and years using the legends provided in the school registries (see Figure A.2). At Jesuit schools, following the French school system of the time, grades ran in reverse order, with *Rhétorique* corresponding to the final year of the standard curriculum (coded as grade 1, and corresponding to grade 12 in modern terminology). Entry-level grades at Jesuit schools varied: while some schools began at grade 7 (i.e., grade 6 in today's upward-counting terminology for school years), others offered preparatory courses starting as early as grade 9 (corresponding to modern grade 4), approximately at age 8. Accordingly, we then estimate the implied age for each record r using the formula: $Age_r = 8 + (9 - Grade_r)$.

[15] A key advantage of Google Document AI is its ability to detect and preserve the tabular structure of the pages.

[16] Further details on the linkage procedure are provided in Appendix A.1.2

3.2 Parliamentary Speeches

The transcripts of parliamentary debates of the French Third Republic were published in the official gazette (*Journal officiel de la République française*) for both the Chamber of Deputies (*Chambre des députés*) and the Senate (*Sénat*). These publications are available online in PDF format on Gallica, the digital repository of the French National Library[17]. While the *Journal officiel* issues had a three-column structure, some pages embed numerical tables that disrupt the three-column layout. Because OCR software may fail to reliably detect column breaks, we use Detectron2 image detection algorithms to segment each page into text columns/blocks and tables, and then apply Google Cloud Vision for optical character recognition on each detected text region.[18]

Next, we split the digitized transcripts into parliamentary sittings, agenda items, and individual speeches. First, we separate distinct sittings within each transcript document (when Parliament met multiple times on the same date, the corresponding record typically reports all sittings). Second, we partition each sitting into consecutive sections of the debate—hereafter, *discussions*—each centered on a specific issue or bill that Parliament was called upon to examine. Finally, we split each discussion into individual speeches. We leverage procedural and stenographic conventions that recur throughout the study period to construct regular-expression patterns that identify sitting boundaries and discussion headers, and extract individual speeches along with the reported name of the speaker.[19]

4 Text Analysis and Empirical Results

In this section, we first describe how we construct our main measures of emotionality and polarization in parliamentary discourse. We then examine how Jesuit-educated MPs differed in their speech patterns and how they affected other speakers during parliamentary debates.

4.1 Text Analysis of Parliamentary Speeches

This section describes how we derive emotionality and polarization scores from speeches in the French Parliament. First, we preprocess the text to harmonize case and accents and retain only informative word stems. Second, we train a Word2Vec model and perform a Latent Dirichlet Allocation (LDA) topic analysis. Finally, we use these techniques to assign emotionality and polarization scores

[17] See https://gallica.bnf.fr. While Gallica provides an OCR-generated text version, we found that OCR errors were too frequent to allow the reliable parsing of individual speeches; we therefore re-digitize the documents from scratch to obtain a higher-quality text extraction.

[18] See Appendix Section A.3 for more details. Detectron2 is an image detection library developed by Facebook AI, see https://github.com/facebookresearch/detectron2. For information on Google Cloud Vision, see https://docs.cloud.google.com/vision/docs/ocr.

[19] For example, the President of the Chamber announces the start of proceedings with the formula *la séance est ouverte*. Each agenda item is introduced by an uppercase header. Each intervention then starts on a new line—typically after a period or an uppercase header—and the speaker is introduced by "M." followed by a surname and a period. See Figure A.4 and Appendix Section A.3 for more details.

to each speech.

Preprocessing. We preprocess parliamentary speeches to reduce dimensionality (Gentzkow, Kelly, and Taddy, 2019). First, we assign part-of-speech tags and retain only adjectives, adverbs, nouns, and verbs. We then normalize tokens by removing accents, converting text to lowercase, stripping punctuation and digits, and keeping only alphabetic strings. Finally, we apply the Porter2 stemming algorithm (`SnowballC`) to reduce words to their morphological roots (Porter, 2009). We further remove uninformative language by dropping short tokens (fewer than three characters), filtering out rare words (appearing in five or fewer documents or ten or fewer times in the corpus), and removing French stopwords using a custom list.[20] Because parliamentary debates contain a large amount of procedural discourse, we also remove procedural words following Gentzkow et al. (2019). To do so, we digitize the Rules of Procedure (*Règlements*) of the Chamber and Senate during the Third Republic (Bonnard, 1926), use an auxiliary Word2Vec model to embed these terms, construct a frequency-weighted "procedural" centroid, and drop the bottom 1% of words in the speech corpus ranked by cosine distance to this centroid.[21] After these steps, we retain 1,468,790 speeches with nonempty text.

Word Embeddings and Topic Analysis. We train a Word2Vec model on the resulting corpus using `gensim` (Mikolov, Sutskever, Chen, Corrado, and Dean, 2013), with 300 dimensions and a context window of 8 words, yielding vector representations for all tokens. We then compute a document embedding for each speech as an inverse-frequency-weighted average of its word vectors. For speech $\mathcal{D}_i$, the document vector is

$$\vec{d_i} \;=\; \frac{1}{|\mathcal{D}_i|} \sum_{w \in \mathcal{D}_i} \frac{\alpha}{\alpha + p(w)} \vec{w}, \tag{1}$$

where $\vec{w}$ is the vector representation of word w, $p(w)$ is its relative frequency in the corpus, and $\alpha = 0.001$ (Gennaro and Ash, 2022).

We extract latent policy domains via unsupervised topic modeling. Specifically, we apply Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan, 2003) and assign each speech to its highest-probability topic. Our preferred specification uses 20 topics, which map closely to major policy domains (e.g., schooling, labor, military affairs; Appendix Figure A.7).[22] Second, we construct two speech-level measures: emotionality and polarization.

[20] We start from the French stopword list in the `stopwords` package and supplement it with days of the week, months, names of French *départements*, and additional low-information terms (e.g., *faire*, *parl*, *temps*, *nombre*).

[21] Operationally: (i) we preprocess the *Règlements* using the same pipeline (except POS tagging); (ii) we train an auxiliary Word2Vec model on the full speech corpus; (iii) we compute a frequency-weighted centroid of procedural terms; (iv) we compute cosine distances between each vocabulary word and this centroid; and (v) we remove the 1% most similar words. Appendix Figure A.6 illustrates the resulting procedural lexicon.

[22] We estimate the LDA using the `topicmodels` package with Gibbs sampling, following standard choices for priors and convergence diagnostics; see Appendix Section B.4 for full details.

Emotionality and Polarization. To measure emotionality, we follow Gennaro and Ash (2022) and use the French LIWC dictionaries (2007 release) to define lists of *affective* and *cognitive* terms (Tausczik and Pennebaker, 2010). We refine these dictionaries by expanding wildcards (via WordNet lemmas), removing overlaps across categories, trimming semantically distant terms, applying the same normalization and stemming as for the corpus, and removing stopwords. We then compute inverse-frequency-weighted centroid vectors for the emotional and cognitive dictionaries, denoted $\mathbf{E}$ and $\mathbf{C}$, and define the emotionality score of speech i as

$$escore_i \;=\; \frac{\mathrm{sim}(\vec{d_i}, \mathbf{E}) + 1}{\mathrm{sim}(\vec{d_i}, \mathbf{C}) + 1}, \tag{2}$$

where $\mathrm{sim}(\cdot, \cdot)$ denotes cosine similarity. Higher values indicate language that is relatively closer to affective content than to cognitive content (Appendix Figure A.8).

To measure polarization, we adapt the methodology of Gennaro and Ash (2022) to capture partisanship in political speech during the French Third Republic. Because no exogenous dictionaries of ideologically charged language exist for this context, we construct left- and right-wing lexicons from our own data using speeches by MPs classified as left- or right-aligned. Moreover, since polarized language likely varies across policy domains and over time, we allow these lexicons to be year- and topic-specific. For each year-topic group g, we compute inverse-frequency-weighted centroid vectors of left- and right-wing discourse, $\mathbf{L}_g$ and $\mathbf{R}_g$, excluding the president of the chamber and members of government. For a speech i in group g, we compute a right-wing alignment score

$$rscore_i \;=\; \frac{\mathrm{sim}(\vec{d_i}, \mathbf{R}_g) + 1}{\mathrm{sim}(\vec{d_i}, \mathbf{L}_g) + 1}, \tag{3}$$

and define the polarization score as the absolute deviation of this alignment from the group median:

$$pscore_i \;=\; \big|\, rscore_i - \mathrm{median}(rscore) \,\big|. \tag{4}$$

This measure captures how far a speech departs from the median rhetorical position in that year-topic debate, whether toward the left or the right (Appendix Figure A.9).

Support for Secularization Policies. In our empirical analysis, we provide a case study on the role of JeMPs in the parliamentary debates on the Law of Separation of churches and the State, widely regarded as one of the most contentious reforms of the French Third Republic and an ideal laboratory to investigate the role of moderators. These debates comprise 12,010 speeches delivered from February to July 1905 in around 60 distinct parliamentary discussions.

One key advantage of focusing on a single piece of legislation is that roll-call votes allow us to observe each MP's stance on the bill. We leverage this information to construct novel text-based measures of support for, and opposition to, secularization. First, we identify MPs who voted consistently in favor of the Secularization Law and MPs who voted consistently against it in both the first and final roll-call votes, thereby excluding MPs who changed position (and may have shifted their language) during the debate. We then build two dictionaries from the speeches of these pro-secularization and anti-secularization MPs and represent each dictionary by its centroid in the word-embedding space, denoted $\mathbf{P}$ (pro-law) and $\mathbf{A}$ (anti-law). For each speech $\mathcal{D}_i$ with embedding $\mathbf{d}_i$, we compute our *pro-secularization score* as

$$\text{proscore}_i \;=\; \frac{\text{sim}(\mathbf{d}_i, \mathbf{P}) + 1}{\text{sim}(\mathbf{d}_i, \mathbf{A}) + 1},$$

where sim denotes cosine similarity. In analogy with the emotionality and polarizations scores, this score captures how closely a speech aligns with the rhetoric of MPs supporting the law relative to the rhetoric of MPs opposing it.

In the context of this case study, the sample size also allows us to complement this measure with a large language model (LLM) assessment of each speech's pro-secularization. Specifically, we ask chatGPT to score the intensity of support for the Separation law on a five-point scale, with values -2 (strongly against), - 1 (moderately against), 0 (neutral), 1 (moderately in favor), and 2 (strongly in favor), allowing for missing values in case of purely procedural or irrelevant interventions. To help the model interpret nuances and implicit references, we include in the prompt, alongside the target speech, up to four immediately preceding speeches from the same discussion, as well as small set of illustrative examples. We iteratively refined the prompt using a random sample of manually classified speeches, and then evaluated its performance on a separate random validation sample, achieving an accuracy rate of 85%.

4.2 Cross-Sectional Results: Jesuit-Educated MPs Debate Differently

We begin by examining differences in rhetorical style between JeMPs and other legislators. We use the following specification:

$$y_{ist} = \beta_0 + \beta_1 \text{JeMP}_s + \gamma X_{ist} + \delta_t + \varepsilon_{ist}, \tag{5}$$

where y_{ist} is our measure of emotionality or polarization for speech i delivered by speaker s at date t, and JeMP_s indicates whether the speaker was educated in a Jesuit school. The vector X_{ist} includes controls for the speaker's institutional role, speech characteristics and individual characteristics that may shape speech content, δ_t are year fixed effects to absorb common shifts in rhetoric over time, and ε_{ist} is an error term. Hence, the coefficient of interest, β_1, reflects the average gap in emotionality or

polarization between JeMPs and other legislators, conditional on a rich set of observables.

Emotionality. Table 1 presents the results for the (standardized) emotionality score and shows that JeMPs use systematically *less emotional* language than other MPs. Across all specifications, the coefficient β_1 is negative and precisely estimated. In the baseline specification with speaker-role controls and year fixed effects (col. 1), JeMP speeches are about 2% of a standard deviation less emotional. This gap *widens* to roughly 5% of a standard deviation when we add an indicator for discussions that feature at least one JeMP speaker (col. 2), alleviating the concern that JeMPs sort into inherently low-emotional discussions. The estimated gap is essentially unchanged when we additionally control for speech length (col. 3), and it persists, though slightly attenuated, even after adding topic fixed effects (col. 4), fixed effects for the speaker's age and occupation (col. 5), and fixed effects for political affiliation and department of election (col. 6).[23]

Polarization. Table 2 reports the analogous specifications for the polarization score and shows that JeMPs use systematically *less polarized* language than other MPs. Again, the JeMP coefficient is negative and precisely estimated across all columns. In the baseline specification with speaker-role controls and year fixed effects (col. 1), JeMP speeches are about 2% of a standard deviation less polarized - similar to the magnitude we find for emotionality. This pattern remains robust even in the most demanding specification with additional controls for discussion type, speech length, discussion topic, age and occupation fixed effects, as well as political affiliation and departement-of-election fixed effects.

Overall, Tables 1 and 2 establish a robust pattern: JeMPs adopt a systematically less emotional and less polarized rhetorical register, even after netting out the role of topic choice, age, occupation, department of election, and political affiliation. This evidence aligns with the distinctive environment to which JeMPs were exposed during their formative years. As discussed in the historical background (Section 2), Jesuit education combined rigorous training in philosophy and rhetoric with structured disputation – explicitly training students to engage opposing viewpoints constructively – and emphasized the integration of theology and scientific inquiry. To further corroborate this mechanism, we examine whether JeMP speeches are differentially framed around religion and science. Specifically, we use our word embeddings to measure the proximity of each speech to the semantic fields associated with the concepts "religion and "science", and we use these proximity measures as dependent variables in equation (5).

Proximity to Religion and Science. Table 3 examines the extent to which JeMPs referred to religion

[23]The number of observations drops from column 5 to column 6 in Tables 1, 2, and 3 because some speakers could not be matched to an elected member of parliament and therefore have missing biographical information. The sample shrinks further in column 6 because some government members who speak in the Chamber were elected to the Senate; for these speakers, we observe age and occupation but lack our detailed coding of political affiliation.

(cols 1-3) and science (cols 4-6) in their speeches. The dependent variable captures the proximity of each speech to the concepts *religion* and *science* in our main word-embedding space. Figure 4 displays word clouds of the terms most closely associated with these concepts in our corpus. For religion, we find ambiguous results: JeMPs refer slightly more to religion when we only include our standard set of controls (col 1). However, this appears to be driven by their greater participation in debates in which religious issues were particularly salient, as the coefficient becomes zero when we control for topic fixed effects in col 2. When we add further speaker-specific controls on age, occupation, origin, and party (col 3), the JeMP coefficient becomes significantly negative. This suggests that within their political party, JeMPs referred less to religion than their fellow party members of similar background and age. In contrast, we find a coherent pattern of JeMPs using more scientific language across all specifications. We conclude that, among MPs with similar personal and political backgrounds, and conditional on the topic of a debate, JeMPs used *less* religious and more scientific language. One interpretation is that Jesuit-educated MPs deliberately avoided explicit religious appeals in parliamentary debate, instead framing arguments in scientific and rational terms that carried broader legitimacy in an increasingly secular and polarized political environment. Rather than mobilizing confessional rhetoric, they may have translated religious or moral commitments into universalistic, evidence-based arguments. This rhetorical strategy would be consistent with their moderating influence in debate, as secular and scientifically framed language may have lowered the risk of partisan escalation and facilitated engagement across ideological lines.

Engaging with Preceding Speakers. Finally, Table 4 examines whether JeMPs differed from other MPs in the extent to which they engaged with preceding speakers. The dependent variable is the Jaccard index, a measure of lexical overlap, between each speech and the preceding speech.[24] Across all three specifications – baseline controls, topic fixed effects, and additional speaker-level controls – we find that JeMPs were significantly more likely to refer to and respond directly to the immediately preceding intervention. This pattern holds within parties and conditional on debate topics, indicating that JeMPs were systematically more engaged in ongoing exchanges than their fellow legislators. By explicitly acknowledging and addressing prior arguments, JeMPs may have transformed debates from parallel monologues into structured interaction, thereby limiting rhetorical escalation. We therefore turn next to examine whether JeMP interventions led to measurable reductions in emotionality and polarization among subsequent speakers.

[24]More generally, the Jaccard index measures the similarity between two sets and is defined as the size of their intersection divided by the size of their union. In our context, it equals the number of unique words shared by two speeches divided by the total number of unique words across both speeches. Because the measure is mechanically normalized by the combined vocabulary size of the two speeches, it is independent of speech length; we therefore omit speech length as a control in Table 4.

4.3 Event Studies: Jesuit-Educated MPs Influenced Political Debates

Having established that JeMPs adopt a more rational, neutral, and science-oriented rhetorical style, we next ask whether this distinctive register can shape the dynamics of parliamentary debates, reducing their emotional tone and intensity of partisan conflict. To address this question, we implement an event-study design that exploits the timing of the first JeMP intervention within a discussion. Our main specification includes speaker $\times$ discussion fixed effects, which absorb all determinants of a speech rhetorical style that are fixed for a given speaker within a given discussion. As a result, identification of the event-time coefficients comes from within-speaker changes in emotionality or polarization within the same discussion, comparing speeches delivered before versus after a JeMP takes the floor. More specifically, we use the following specification:

$$y_{isd} = \sum_{\substack{k \in \{-5 \ldots +5\} \\ k \neq -1}} \beta_k D_{i,t-k} + \gamma X_{isd} + \alpha_{sd} + \varepsilon_{isd} \tag{6}$$

where y_{isd} is the emotionality or polarization score for speech i delivered by speaker s in discussion d, and $D_{i,t-k}$ is an indicator that speech i occurs k speech turns relative to t, the first JeMP intervention in discussion d. The control vector X_{isd} includes quartiles of speech length, an indicator for procedural speeches, as well as fixed effects for the speaker's speech count within a discussion (i.e., whether the speech is the speaker's 1st, 2nd, 3rd,...intervention), to flexibly absorb systematic patterns in rhetoric over the course of a debate. Finally, the term α_{sd} denotes speaker $\times$ discussion fixed effects, and ε_{isd} is an error term. The key identification assumption is that the error term – i.e., the unobserved factors that affect the speech tone – are not systematically correlated with the timing of the JeMP entry.

The main objects of interest are the event-time coefficients β_k. The post-JeMP coefficients ($k \geq 0$) trace out the dynamic response of emotionality or polarization following JeMP entry. The pre-JeMP coefficients ($k < 0$) provide a falsification test, allowing us to assess the presence of pre-trends in rhetorical style closely preceding the JeMP intervention. We omit the dummy for $k = -1$ (the speech immediately preceding JeMP entry), so all β_k are interpreted relative to that baseline.

Figure 5 reports the event-study estimates for the standardized emotionality score. The x-axis indexes speech order relative to the first JeMP intervention ($t = 0$), and the y-axis plots the estimated β_k coefficients from equation (6), along with 90% confidence intervals, relative to the omitted category $k = -1$. Two main patterns emerge from the figure. First, following the JeMP speech the emotional tone of the parliamentary debate declines markedly for five speaking turns, reaching a trough of about 6.5 percent of a standard deviation at $t + 4$. While emotionality becomes to fall already at $t + 1$, the post-JeMP coefficients are statistically significant at the 90% confidence level from $t + 2$ through $t + 4$. Second, all the pre-JeMP coefficients are statistically indistinguishable from zero, indicating

that JeMPs do not systematically intervene at particular stretches of parliamentary debate when the emotional intensity of the discussion is already trending upward or downward.

Figure 6 reports the analogous results for the polarization score. We again find a marked decline immediately after the JeMP intervention. The depolarizing effect is sizeable, with estimates ranging from 5 percent of a standard deviation at $t + 1$ to 9 percent of a standard deviation at $t + 5$. As in Figure 6, none of the pre-JeMP coefficients are statistically different from zero.

Reassuringly, there are no pre-trends in Figures 5 and 6, suggesting that JeMPs did not enter the debate when it was already on a trajectory to become less emotional or polarized. In addition, we view the concern that JeMPs may have strategically timed their interventions unlikely for several reasons. First, speaking order was to a large extent predetermined by parliamentary procedure, limiting the scope for fine-grained strategic timing. Second, even if speakers had substantial discretion, it is not clear how JeMPs would have systematically forecast sudden changes in rhetorical tone at such short horizons, especially given the flat pre-trends and the rich set of controls included in the specification. In particular, note that our use of speaker-discussion fixed effects renders it impossible that JeMPs could have forecast the waxing and waning of debating styles based on the next speaker on the agenda. Finally, even if JeMPs could time their interventions around sudden declines in rhetorical heat and had incentives to do so, rationalizing our results would still require that such timing ability or incentives be systematically stronger for JeMPs than for other MPs.

Placebo Exercises. Another concern is that the estimated decline in emotionality and polarization following JeMPs interventions may not reflect a uniquely JeMP effect, but could rather arise mechanically from generic properties of low-emotional speakers. To address this concern, we implement three placebo exercises. First, we randomly sample 5% of MPs (excluding JeMPs) and use them as a treatment group in the event-study. We repeat this procedure 1,000 times and show the distribution of the event-time coefficients, for periods 0 to 5, in Figure 7 for emotionality and in Figure 8 for polarization. These figures show that the observed pattern of JeMP interventions is indeed unique, as our event study coefficients typically lie outside the distributions of coefficients obtain from the placebo exercises. Second, we estimate year-specific MP fixed effects from regressions of emotionality and polarization (excluding JeMPs). We then identify the 5% of MPs with the lowest fixed effects (i.e., those who are systematically the least emotional or least polarized speakers within that year) and treat them as placebo JeMPs in the event-study framework. The results show that interventions by these "calmest" MPs do not reproduce the magnitude or persistence of the JeMP effect. In other words, being individually less emotional was not sufficient to shift the tone of subsequent speakers. This exercise suggests that JeMPs possessed attributes – credibility, authority, cross-ideological legitimacy, or institutional standing – that allowed them to influence the rhetorical equilibrium of the debate. In the same vein, we perform a third placebo exercise to test whether department-level characteristics

may driving the results. We estimate department-of-election fixed effects for emotionality and polarization (excluding JeMPs). From those, we identify the 10% of departments with the lowest fixed effects and use the MPs from these departments as placebo JeMPs. To match the number of actual JeMPs in the sample, we randomly extract MPs from the selected departments and repeat the placebo exercise in 1,000 different simulations. Again, we do not find comparable post-intervention declines in this placebo, suggesting that the JeMP effect is not simply attributable to constituency-level moderation or local political culture.

4.4 Case Study: The 1905 Separation Law

The 1905 Separation Law is widely regarded as one of the most significant—and polarizing—pieces of legislation in modern French history. Building on earlier republican efforts to reorganize church–state relations, it made secularism the guiding constitutional principle of public life. The law guaranteed freedom of conscience and the free exercise of religion, while also establishing that the state would no longer recognize, fund, or support any religion. It further imposed strict rules on religious activity through the framework of the 1901 Law on Associations, requiring religious groups to operate as regulated associations. Finally, implementation involved a mandatory inventory of church property to distinguish assets deemed necessary for worship from those that could be confiscated. Importantly, the 1905 Separation Law was the final building block in an anti-clerical agenda that had already contributed to the closure of Jesuit schools through the 1901 Law on Associations and the 1904 "Loi Combes", which prohibited teaching by members of religious congregations and organized liquidation of their assets. These features make the Separation Law an ideal setting to study the role of Jesuit-educated MPs in a highly polarized political environment.

To conduct this analysis, we focus on the discussions in the Chamber of Deputies specifically devoted to the Separation Law, which we identify using section titles and verify with a manual check. This sample comprises 12,010 speeches delivered in the Chamber of Deputies in 60 distinct discussions between February 2 and July 3, 1905, the date of the final vote on the full text of the bill. Around 60% of JeMPs voted against the law, as opposed to 40% of other MPs. Furthermore, 80% of the JeMP speeches in this sample come from JeMPs who voted against the law, while this percentage is close to 50% for other MPs. This is not surprising: these parliamentarians had been educated in institutions that the incumbent left coalition had recently dismantled, giving them strong personal and political reasons to oppose the reform.

Cross-sectional Results. Against this background, one might expect JeMPs to deliver their arguments in a more inflammatory tone. Table 5, however, shows the opposite. Even on this highly-sensitive policy issue, JeMPs use relatively neutral language: in the pooled cross-section, their speeches are not statistically different from those of the average MP in terms of pro-secularization slant (col. 1),

partisan polarization (col. 3), or religious content (col. 4). By contrast, JeMPs' interventions are less emotional (col. 2) and more science-oriented (col. 6).

Within-Discussion Results. Next, we examine whether JeMPs may have influenced the pro-secularization stance of other MPs through their engagement in the debate. In this setting, the limited number of discussions prevents us from implementing the rich event-study specification in equation (6), which relies on interacted speaker and discussion fixed effects and narrow windows around the first JeMP intervention. Instead, in a difference-in-differences spirit, we compare speeches delivered before and after the first JeMP intervention within a discussion, restricting the post-intervention window to the next 30 speaking turns. To focus on JeMPs' potential role as moderators, we construct outcomes that capture *extreme* pro-secularization rhetoric using both our Word2Vec embedding-based and LLM-based measures. Specifically, we define (i) a *very pro-secularization* indicator equal to one if the embedding-based pro score lies above the 90th percentile, and (ii) a *very pro-secularization* indicator equal to one if the LLM classifies the speech as *strongly in favor* of secularization.[25]

For each speech i delivered by speaker s on date t in discussion d, we estimate

$$y_{isd} = \beta_0 + \beta_1 \,\text{PostJeMP}_{i,d} + \gamma X_{isd} + \alpha_s + \delta_t + \varepsilon_{isd} \tag{7}$$

where y_{isd} is one of these measures of stance. The key regressor, $\text{PostJeMP}_{i,d}$ equals one for speeches delivered after the first JeMP intervention in discussion d and zero otherwise. The vector X_{isd} includes indicators for quartiles of speech length and a dummy for procedural speeches. We also include speaker fixed effects α_s to absorb all time-invariant determinants of a speaker's rhetorical style and ideological position, and date fixed effects δ_t to net out shocks common to all speeches on a given day; ε_{isd} is the error term. The coefficient of interest, β_1, measures the average change in pro-secularization stance within a discussion after the first JeMP intervention.

Columns 1 and 3 in Table 6 show that JeMPs interventions reduce fervent support for the Law of Separation, both measured with word-embeddings (col. 1) and a large language model (col. 3). Following a JeMP intervention, speakers are 2-3% points less likely to express strong views in favor of the Law of Separation. At the same time, the coefficients in columns (2) and (3) are small and statistically insignificant, suggesting that JeMPs do not seem to fuel fervent opposition to the law. Together with the results in Table 5, these findings highlight the role of JeMPs as mediators even in a highly-polarized political context in which their main political objective was to oppose the reform.

Voting Behavior. Finally, we ask whether the effect of JeMP interventions on the rhetorical style of subsequent speakers is accompanied by a shift in voting behavior. To provide some exploratory evidence on this question, we use the digitized roll-call votes on the 1905 Law of Separation. These

[25] See section **??** for details on the LLM measure.

roll-call votes include votes on procedural decisions, amendments, single articles, as well as on the full text of the bill. For each of these acts, we coded the exact point of the parliamentary discussion where a vote on the bill occurred. This allows us to partition each discussion into segments defined by consecutive roll-call votes. We then assign to each roll call all speeches in the preceding segment.

We focus on the subset of acts for which left-wing parliamentarians were, on average, at least 25% more likely than right-wing parliamentarians to vote in favor. This restriction helps us ensure that a recorded "for" vote reflects *support* for the Separation Law.[26] For each discussion segment,we compute the share of speeches delivered by JeMPs and examine how this measure of JeMP engagement correlates with the subsequent voting behavior of other MPs. Column 1 of Table 7 shows that segments with greater JeMP engagement are associated with lower support for the Law of Separation. This pattern is robust to including speaker fixed effects (col 2) and is not driven by the presence of JeMPs among the voters themselves, as results are similar when excluding JeMPs from the voting sample (cols 3 and 4). While exploratory in nature, these findings suggest that JeMPs may have influenced not only rhetoric but also the voting behavior of other MPs.

5 Conclusions

We examined whether certain types of political elites can moderate polarization within legislative arenas. Using newly digitized archival records on Jesuit student enrollment – linking them to the universe of French parliamentary speeches and roll-call votes during the Third Republic – we showed that Jesuit-educated MPs employed systematically less emotional and less polarized rhetoric and influenced subsequent speakers to adopt more moderate language. These results persisted under stringent fixed effects specifications, while they were absent in a range of placebo exercises, suggesting that our findings reflect distinctive characteristics of JeMPs rather than generic features of particular legislators or electoral environments. Even during the highly contentious debates surrounding the 1905 separation of Church and State, JeMPs used comparatively neutral rhetoric and dampened extreme discourse among subsequent speakers.

Our findings contribute to the literature on political polarization by shifting attention from institutional or party-level drivers to heterogeneity within the political elite. Much of the existing literature explains polarization through electoral competition, party sorting, or institutional incentives. While these forces undoubtedly shaped aggregate trends, our results suggested that polarization within legislatures was not solely an equilibrium outcome of institutional constraints. Instead, it could be moderated from within, through the presence and participation of individuals with distinctive backgrounds

[26] We adopt this procedure because the political objective of a roll call is often difficult to infer from its title alone (e.g., "On the additional article by M. Allard"), and the full text of the proposed act is not consistently available in the parliamentary records.

and interaction styles.

Beyond these contributions, our findings point to two broader implications. First, they suggest that polarization within legislatures may be a dynamic and path-dependent process. The spillover effects that we document indicate that the tone of debate can shift depending on who intervenes and when. If certain legislators systematically de-escalate rhetoric, early interventions by such actors may move discussions onto lower-temperature trajectories, whereas the absence of moderating voices may allow emotional escalation to persist. This perspective highlights the importance of agenda control, speaking order, and procedural rules in shaping polarization dynamics, and suggests that relatively small changes in elite composition or participation could generate disproportionate effects on deliberative outcomes.

Second, our results raise the possibility that rhetorical moderation functions as a form of elite self-restraint that generates positive externalities within democratic institutions. By tempering the tone of debate, moderating legislators may improve the quality of deliberation not only through their own speeches but also by influencing others. This perspective resonates with arguments emphasizing the importance of informal "norms of forbearance" and mutual toleration for democratic stability (Levitsky and Ziblatt, 2018), as well as with theories that view democracy as sustained by self-enforcing limits among political actors (Weingast, 1997). Yet, the incentives to supply such restraint may be limited, particularly in polarized environments where partisan signaling and mobilization are electorally rewarded. If rhetorical moderation constitutes a form of institutional forbearance, then its provision may depend critically on elite composition and formation, suggesting that the selection and training of political leaders have first-order implications for the resilience of democratic institutions.

References

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring Economic Policy Uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011). Do Educated Leaders Matter? *The Economic Journal 121*(554), F205–F227.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research 3*, 993–1022.

Bó, E. D., F. Finan, O. Folke, T. Persson, and J. Rickne (2017). Who Becomes a Politician? *Quarterly Journal of Economics 132*(4), 1877–1914.

Bonica, A. (2014). Mapping the Ideological Marketplace. *American Journal of Political Science 58*(2), 367–386.

Bonnard, R. (1926). *Les règlements des assemblées législatives de la France depuis 1789: notices historiques et textes*. Paris: Société anonyme du Recueil Sirey.

Bush, J. W. (1975). Education and Social Status: The Jesuit College in the Early Third Republic. *French Historical Studies 9*, 125–140.

Bäck, H., R. Carroll, E. Renström, and A. Ryan (2023). Elite Communication and Affective Polarization Among Voters. *Electoral Studies 84*, 102639.

Canen, N., C. Kendall, and F. Trebbi (2020). Unbundling Polarization. *Econometrica 88*(3), 1197–1233.

Canen, N., C. Kendall, and F. Trebbi (2026). Political Parties as Drivers of U.S. Polarization: 1927–2018. *Journal of Political Economy*. Forthcoming.

Cantoni, D., Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2017). Curriculum and Ideology. *Journal of Political Economy 125*(2), 338–392.

Delattre, P. (Ed.) (1949). *Les Établissements des Jésuites en France depuis quatre siècles: Répertoire topo-bibliographique publié à l'occasion du quatrième centenaire de la fondation de la Compagnie de Jésus, 1540–1940*, Volume 1. Enghien and Wetteren, Belgium: Institut supérieur de Théologie and Imprimerie De Meester Frères.

Diermeier, D., J.-F. Godbout, B. Yu, and S. Kaufmann (2012). Language and Ideology in Congress. *British Journal of Political Science 42*(1), 31–55.

Doornbosch, L. M., M. van Vuuren, and M. D. de Jong (2025). Moving Beyond Us-Versus-Them Polarization Towards Constructive Conversations. *Democratization 32*(1), 96–122.

Downey, M., J. Kuk, and M. Nanes (2024). (Not) Feelin' the (Cross) Pressure: Experimental Evidence on Issue Misalignment and Polarization. Working paper.

Doyle, A. (2017). Catholic Church and State Relations in French Education in the Nineteenth Century: The Struggle between Laïcité and Religion. *International Studies in Catholic Education 9*(1), 108–122.

Dreher, A., M. J. Lamla, S. M. Lein, and F. Somogyi (2009). The Impact of Political Leaders' Profession and Education on Reforms. *Journal of Comparative Economics 37*(1), 169–193.

Farrell, A. P. (1970). *The Jesuit Ratio Studiorum of 1599: Translated into English, with an Introduction and Explanatory Notes*. Washington, D.C.: Conference of Major Superiors of Jesuits.

Fiorina, M. P. and S. J. Abrams (2008). Political Polarization in the American Public. *Annual Review of*

Political Science 11, 563–588.

Fishkin, J. S., A. Siu, L. Diamond, and N. Bradburn (2021). Is Deliberation an Antidote to Extreme Partisan Polarization? Reflections on "America in One Room". *American Political Science Review 115*(4), 1464–1481.

Gennaro, G. and E. Ash (2022). Emotion and Reason in Political Language. *The Economic Journal 132*(643), 1037–1059.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as Data. *Journal of Economic Literature 57*(3), 535–574.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica 87*(4), 1307–1340.

Grendler, P. F. (1989). *Schooling in Renaissance Italy: Literacy and Learning, 1300–1600*. Baltimore, MD: Johns Hopkins University Press.

Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge, MA: MIT Press.

Harrigan, P. J. (2001). Church, State, and Education in France from the Falloux to the Ferry Laws: A Reassessment. *Canadian Journal of History 36*, 51–83.

Henríquez, P. A., J. Sabat, and J. P. Sullivan (2023). Politicians' Willingness to Agree: Evidence from the Interactions in Twitter of Chilean Deputies. *Journal of Information Technology & Politics 20*(1), 92–111.

Langdon, J. W. (1978). The Jesuits and French Education: A Comparative Study of Two Schools, 1852–1913. *History of Education Quarterly 18*(1), 49–60.

Levitsky, S. and D. Ziblatt (2018). *How Democracies Die*. New York: Crown.

Loyola, I. o. (1914). *The Spiritual Exercises*. Christian Classics Ethereal Library. Accessed 27 Feb. 2026.

Mayeur, F. (1981). *Histoire Générale de l'Enseignement et de l'Éducation en France: Tome III. De la Révolution à l'École Républicaine (1789–1930)*. Paris: Nouvelle Librairie de France.

McCarty, N., K. T. Poole, and H. Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean (2013). Distributed Representations of Words and Phrases and Their Compositionality. *arXiv preprint arXiv:1310.4546*.

Milligan, K., E. Moretti, and P. Oreopoulos (2004). Does Education Improve Citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics 88*(9–10), 1667–1695.

Moskowitz, D. J., J. C. Rogowski, and J. M. S. Jr. (2024). Parsing Party Polarization in Congress. *Quarterly Journal of Political Science 19*(4), 357–385.

O'Malley, J. W. (1993). *The First Jesuits*. Cambridge, MA: Harvard University Press.

Porter, M. (2009). The English (Porter2) Stemming Algorithm. http://snowball.tartarus.org/algorithms/english/stemmer.html.

Schwickerath, R. (1904). *Jesuit Education: Its History and Principles Viewed in the Light of Modern Educational Problems* (Second ed.). St. Louis, MO: B. Herder.

Squicciarini, M. P. (2020). Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France. *American Economic Review 110*(11), 3454–3491.

Tausczik, Y. R. and J. W. Pennebaker (2010). The Psychological Meaning of Words: LIWC and Computerized

Text Analysis Methods. *Journal of Language and Social Psychology 29*(1), 24–54.

Thompson, D. G. (2018). The Jesuit Province of France on the Eve of its Destruction in 1762. *Archivum Historicum Societatis Iesu 87*(173).

van Kley, D. K. (1975). *The Jansenists and the Expulsion of the Jesuits from France, 1757–1765*. Yale Historical Publications: Miscellany. New Haven: Yale University Press.

Vogel, C. (2010). The Suppression of the Society of Jesus, 1758–1773. European History Online (EGO).

Weingast, B. R. (1997). The Political Foundations of Democracy and the Rule of Law. *American Political Science Review 91*(2), 245–263.

TABLES

Table 1: Emotionality of JeMP Speeches

Dependent Variable: *Emotionality Score of Parliamentary Speeches*

	(1)	(2)	(3)	(4)	(5)	(6)
JeMP	-0.022***	-0.048***	-0.044***	-0.030***	-0.028***	-0.021***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Speaker Role	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Discussion Type FE		✓	✓	✓	✓	✓
Speech Length			✓	✓	✓	✓
Topic FE				✓	✓	✓
Speaker Age/Occupation FE					✓	✓
Speaker Origin Depart/Party FE						✓
Mean Dep. Var.	0.007	0.007	0.007	0.007	0.039	0.042
R^2	0.05	0.05	0.06	0.29	0.29	0.30
Observations	1,128,866	1,128,866	1,128,866	1,128,866	1,007,382	983,727

Notes: The table shows that Jesuit-educated MPs (JeMPs) had significantly lower levels of emotionality in their parliamentary speeches. Regressions are run at the speech level. *Speaker Role* represents dummies for the President of the Chamber, members of government, rapporteurs, and for procedural speeches. Year FE are fixed effects for the year during which a speech occurred. *Discussion Type FE* is a dummy for discussions with at least one JeMP speaker; *Speech Length* are dummies for quartiles of speech length (measured by total word count); *Topic FE* are dummies where topics are identified using a Latent Dirichlet Allocation algorithm trained with 20 topics; *Speaker Age/Occupation FE* are year-specific dummies for the speaker's age and for 12 occupation categories; and *Speaker Origin Depart/Party FE* includes dummies for the department in which the speaker was elected as well as dummies for the speaker's political area (left, center-left, center, center-right, right, independent). Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Partisan Polarization of JeMP's Speeches

Dependent Variable: Polarization Score

	(1)	(2)	(3)	(4)	(5)	(6)
JeMP	-0.021***	-0.015***	-0.018***	-0.014***	-0.019***	-0.013**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Speaker Role	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Discussion Type FE		✓	✓	✓	✓	✓
Speech Length			✓	✓	✓	✓
Topic FE				✓	✓	✓
Speaker Age/Occupation FE					✓	✓
Speaker Origin Depart/Party FE						✓
Mean Dep. Var.	-0.005	-0.005	-0.005	-0.005	-0.006	-0.004
R^2	0.02	0.02	0.03	0.13	0.13	0.13
Observations	1,128,866	1,128,866	1,128,866	1,128,866	1,007,382	983,727

Notes: The table shows that Jesuit-educated MPs (JeMPs) debated in a less polarized fashion in parliament. Regressions are run at the speech level. See the note to Table 1 for a description of the control variables. Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Proximity to Religion and Science in MP Speeches

Dependent Variable: Proximity to Religion / Science

	Religion			Science		
	(1)	(2)	(3)	(4)	(5)	(6)
JeMP	0.010*	-0.001	-0.012***	0.042***	0.053***	0.056***
	(0.006)	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)
Speaker Role	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Discussion Type FE	✓	✓	✓	✓	✓	✓
Speech Length	✓	✓	✓	✓	✓	✓
Topic FE		✓	✓		✓	✓
Speaker Age/Occupation FE			✓			✓
Speaker Origin Depart/Party FE			✓			✓
Mean Dep. Var.	0.023	0.023	0.053	0.016	0.016	0.037
R^2	0.06	0.48	0.48	0.07	0.35	0.35
Observations	1,128,866	1,128,866	1,128,866	1,128,866	1,007,382	983,727

Notes: The table shows that the parliamentary speeches of JeMPs were closer to science as compared to other MPs, but not necessarily closer to religion. Regressions are run at the speech level. Proximity to religious and scientific language measures the proximity, in a word-embedding space, between each speech and, respectively, the words "religion" and "science." Cols 1-3 reports the results for religion, while cols 4-6 report the results for science. See the note to Table 1 for a description of the controls variables. Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Vocabulary Overlap with the Preceding Speaker

Dep. Var.: Jaccard similarity index with the previous speech

	(1)	(2)	(3)
JeMP	0.012**	0.013**	0.010*
	(0.005)	(0.005)	(0.006)
Speaker Role	✓	✓	✓
Year FE	✓	✓	✓
Discussion Type FE	✓	✓	✓
Topic FE		✓	✓
Speaker Age/Occupation FE			✓
Speaker Origin Depart/Party FE			✓
Mean Dep. Var.	0.011	0.011	0.007
R^2	0.01	0.02	0.02
Observations	1,128,866	1,128,866	983,727

Notes: The table shows that JeMPs engaged more with the arguments of preceding speakers. The dependent variable, the Jaccard index, measures the vocabulary overlap between each speech and the preceding speech. See footnote 24 for detail. Regressions are run at the speech level. See the note to Table 1 for a description of the control variables. Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: The Law of Separation of Church and State: JeMPs' Rhetorical Style

Dependent Variable: as indicated in the table header

	(1)	(2)	(3)	(4)	(5)
	Pro Score	Emotionality Score	Polarization Score	Proximity to Religion	Proximity to Science
JeMP	-0.005	-0.083**	0.028	0.069	0.106***
	(0.036)	(0.034)	(0.022)	(0.048)	(0.037)
Speaker Role	✓	✓	✓	✓	✓
Discussion Type FE	✓	✓	✓	✓	✓
Mean Dep. Var.	-0.000	-0.090	-0.220	0.812	0.127
R^2	0.06	0.06	0.02	0.11	0.09
Observations	10,749	12,010	12,010	12,010	12,010

Notes: The table shows that, when debating the 1905 Law on the Separation of Churches and the State, JeMPs used a relatively neutral language (with a close-to-average Pro Score that measure support for the 1905 Law) and were less emotional and more science-oriented than other MPs. See the note to Table 1 for a description of the controls variables. Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: The Law of Separation of Church and State: Speakers following a JeMP Speech

Dependent Variable: As indicated in the table header

	Word2Vec		LLM	
	(1)	(2)	(3)	(4)
	Very Pro Score	Very Against Score	Very Pro Law	Very Against Law
Post JeMP	-0.034**	0.004	-0.022*	0.014
	(0.016)	(0.018)	(0.012)	(0.017)
Speaker FE	✓	✓	✓	✓
Date FE	✓	✓	✓	✓
Speech Characteristics	✓	✓	✓	✓
Mean Dep. Var.	0.104	0.103	0.076	0.144
R^2	0.24	0.15	0.27	0.24
Observations	2,844	2,844	2,844	2,844

Notes: The table provides suggestive evidence that, during debates on the 1905 Law on the Separation of Churches and the State, JeMP interventions reduced fervent support for the bill without fueling fervent opposition. These results hold both when we code strong support ("Very Pro Score") and strong opposition ("Very against Score") using Word2Vec (cols 1 and 2), and when using LLM (cols 3 and 4, which is feasible in this smaller sample – see Section 4.4 for detail). Regressions are run at the speech level. *Post JeMP* represents an indicator that takes value one after the first JeMP intervention in a discussion. *Speaker FE* represents fixed effects for individual MPs. See the note to Table 1 for a description of other controls variables. Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: The Law of Separation of Church and State: JeMPs Interventions and Voting

Dependent Variable: Dummy for MP Vote in Favor of Secularization

	All MPs		No JeMPs	
	(1)	(2)	(3)	(4)
% JeMP Speeches	-0.023***	-0.023***	-0.024***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
MP FE		✓		✓
Mean Dep. Var.				
R^2	0.03	0.72	0.03	0.72
Observations	10,582	10,582	9,862	9,862

Notes: The table suggests that, during debates on the 1905 Law on the Separation of Churches and the State, the share of JeMP interventions is negatively related to share of votes in favor of the law. *%JeMP Speeches* measures the share of JeMP speeches (scaled between 0 and 100) in a discussion segment leading up to a roll-call vote. Regressions are run at the MP-discussion-segment level, using all discussion segments in parliament about the Separation Law that were followed by a roll-call vote. Robust standard errors are reported in parentheses. Notation for statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

FIGURES

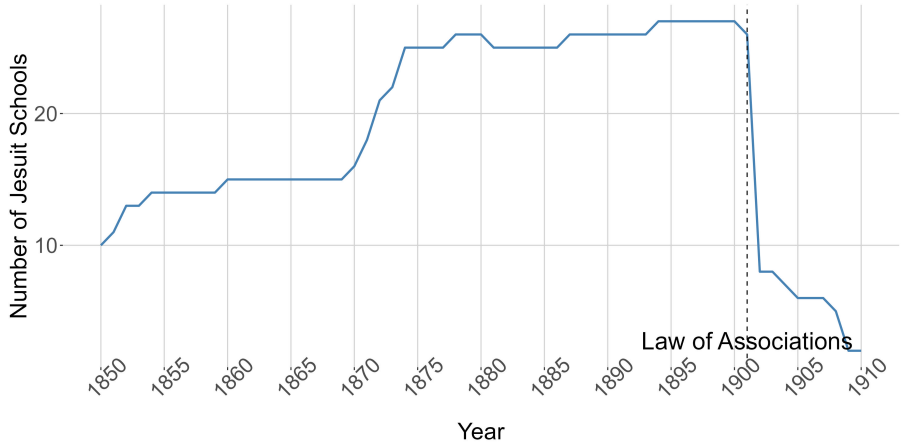


Figure 1: Number of Jesuit Schools in France, 1850-1910

Notes: The figure shows the number of Jesuit schools operating in France between 1850 and 1910 (blue solid line). In 1850, the Falloux Law legitimated non-State schooling and triggered the emergence of Jesuit educational institutions. The increase in 1870 can be explained by the rise of a conservative, Catholic-dominated government after France's defeat in the Franco-Prussian War, which promoted religious schools to restore traditional moral and social order. In 1901, the Law of Association (black dashed line) dissolved religious congregations.

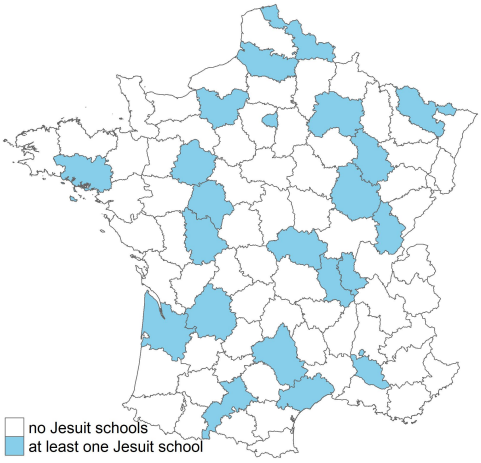
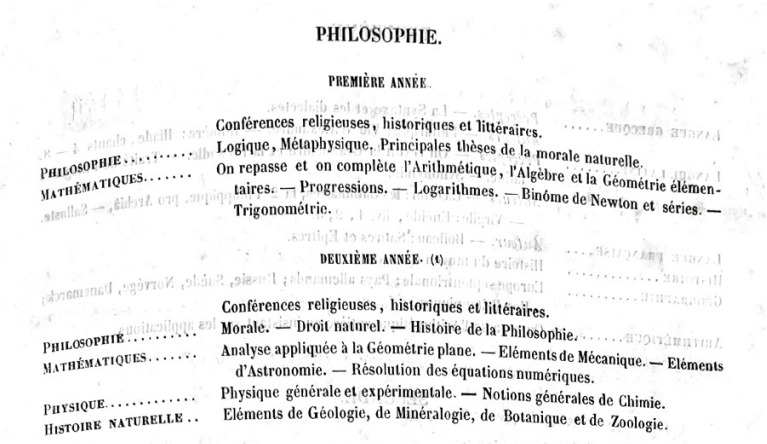


Figure 2: Distribution of Jesuit Schools across French Departments

Notes: The figure shows French departments hosting at least one Jesuit School from 1850 to 1910 (highlighted in light blue).



PHILOSOPHIE.

PREMIÈRE ANNÉE.

PHILOSOPHIE..... Conférences religieuses, historiques et littéraires.
Logique, Métaphysique, Principales thèses de la morale naturelle.
MATHÉMATIQUES..... On repasse et on complète l'Arithmétique, l'Algèbre et la Géométrie élémentaires. — Progressions. — Logarithmes. — Binôme de Newton et séries. — Trigonométrie.

DEUXIÈME ANNÉE. (1)

PHILOSOPHIE..... Conférences religieuses, historiques et littéraires.
Morale. — Droit naturel. — Histoire de la Philosophie.
MATHÉMATIQUES..... Analyse appliquée à la Géométrie plane. — Eléments de Mécanique. — Eléments d'Astronomie. — Résolution des équations numériques.
PHYSIQUE..... Physique générale et expérimentale. — Notions générales de Chimie.
HISTOIRE NATURELLE.. Eléments de Géologie, de Minéralogie, de Botanique et de Zoologie.

Figure 3: Curriculum in the Final Two Years of Instruction at a Jesuit School

Notes: The figure reproduces the curriculum for the final two years of instruction, referred to as "Philosophy" (*Philosophie*), at the Jesuit school in Vannes. Following the 1599 *Ratio Studiorum*, the Curriculum included logic, physics, and astronomy together with methaphysics and philosphy.



(a) Religion



(b) Science

Figure 4: Word Clouds: Religion and Science in Parliamentary Speeches

Notes: The figure shows the word clouds of the 100 words in the corpus with the smallest cosine distance to the words "religion" and "science," computed with our main word-embedding model.

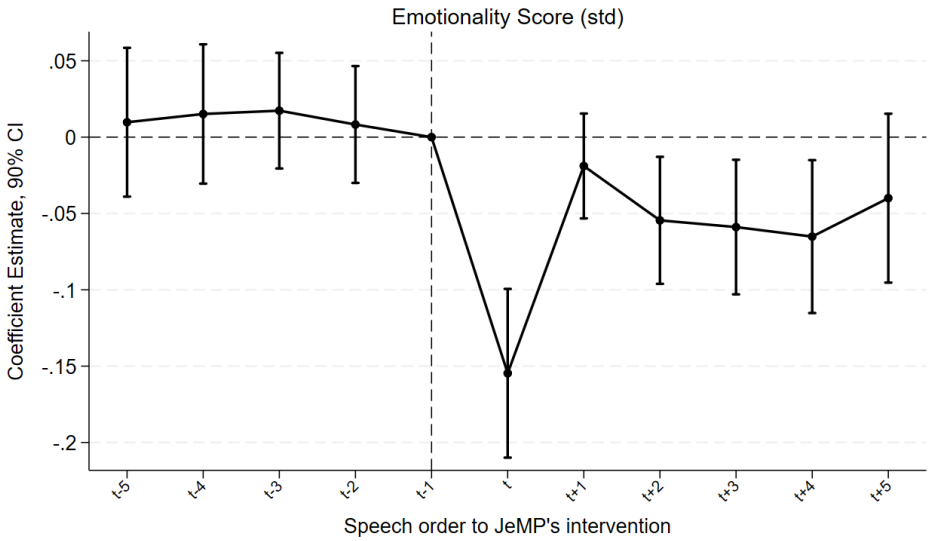


Figure 5: Event Study: JeMPs Reduce the Emotionality of Subsequent Speakers

Notes: The figure shows that following a JeMP speech, the emotional tone of the parliamentary debate declines markedly. The emotionality score measures the relative distance, in a word-embedding space, between each speech and pre-defined dictionaries of affection and cognitive terms. The x-axis indexes speech order relative to the first JeMP intervention ($t = 0$), and the y-axis plots the estimated β_k coefficients from equation (6), along with 90% confidence intervals, relative to the omitted category $k = -1$. The event-study specification includes speaker $\times$ discussion fixed effects, as well as controls for quartiles of speech length, an indicator for procedural speeches, and fixed effects for the speaker's speech order within a discussion. Standard errors are clustered at the discussion level.

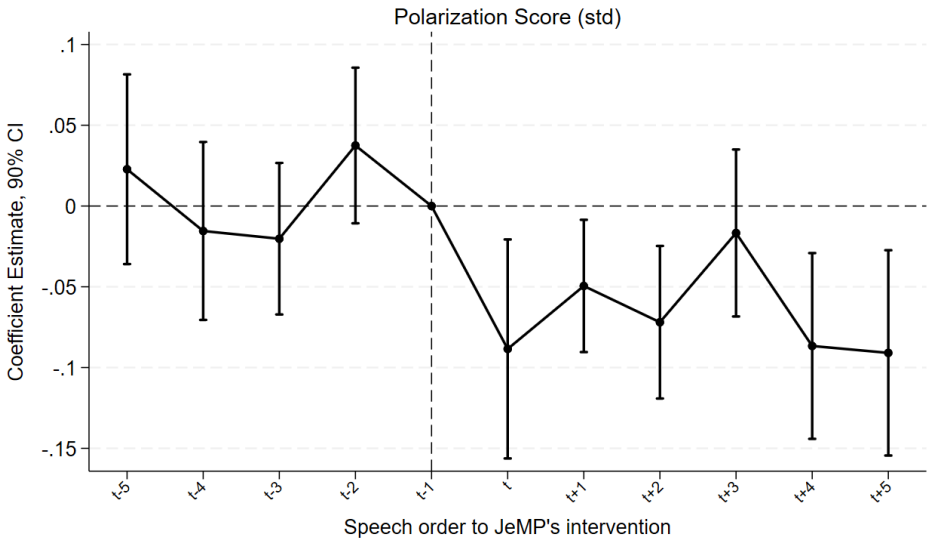
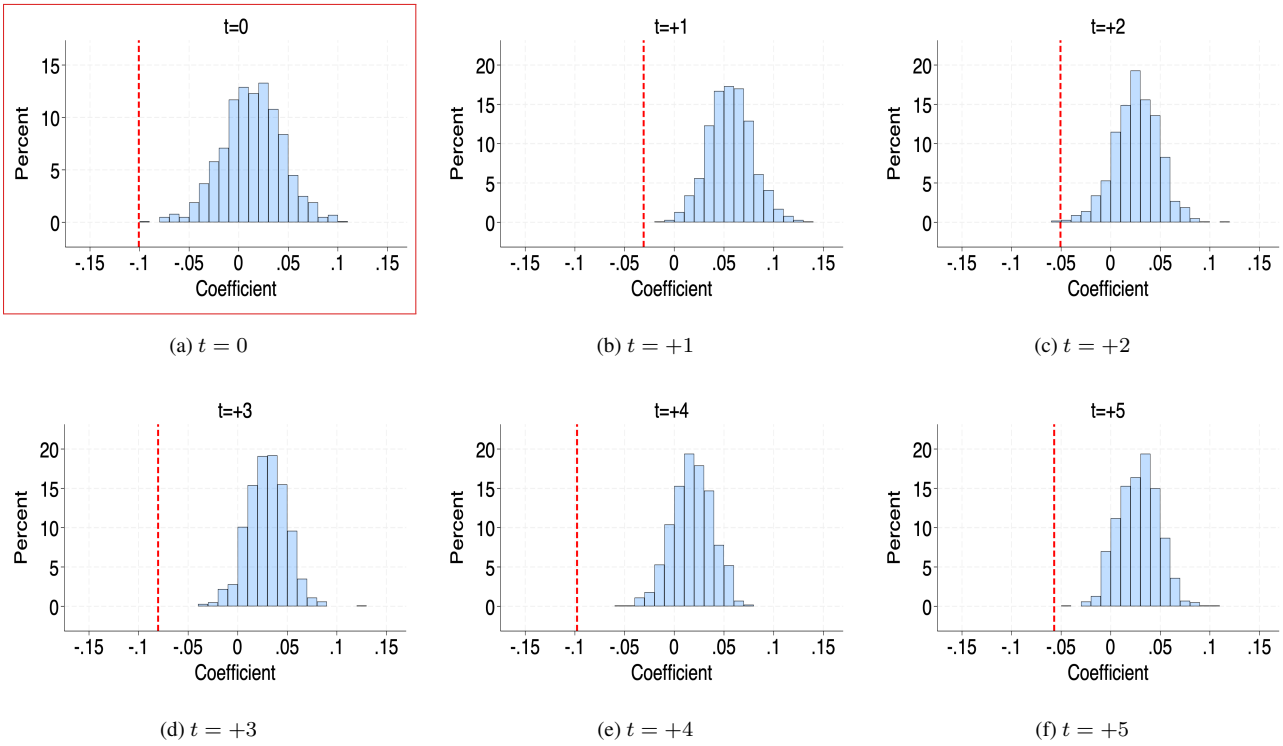


Figure 6: Event Study: JeMPs Reduce the Polarization of Subsequent Speakers

Notes: The figure shows that the partisan polarization of the parliamentary debate declines following a JeMP speech. The polarization score measures the relative distance, in a word-embedding space, between each speech and year–topic-specific dictionaries constructed from left-wing and right-wing parliamentarians' speeches, where topics are identified using a Latent Dirichlet Allocation algorithm. The x-axis indexes speech order relative to the first JeMP intervention ($t = 0$), and the y-axis plots the estimated β_k coefficients from equation (6), along with 90% confidence intervals, relative to the omitted category $k = -1$. The event-study specification includes speaker $\times$ discussion fixed effects, as well as controls for quartiles of speech length, an indicator for procedural speeches, and fixed effects for the speaker's speech count within a discussion. Standard errors are clustered at the discussion level.

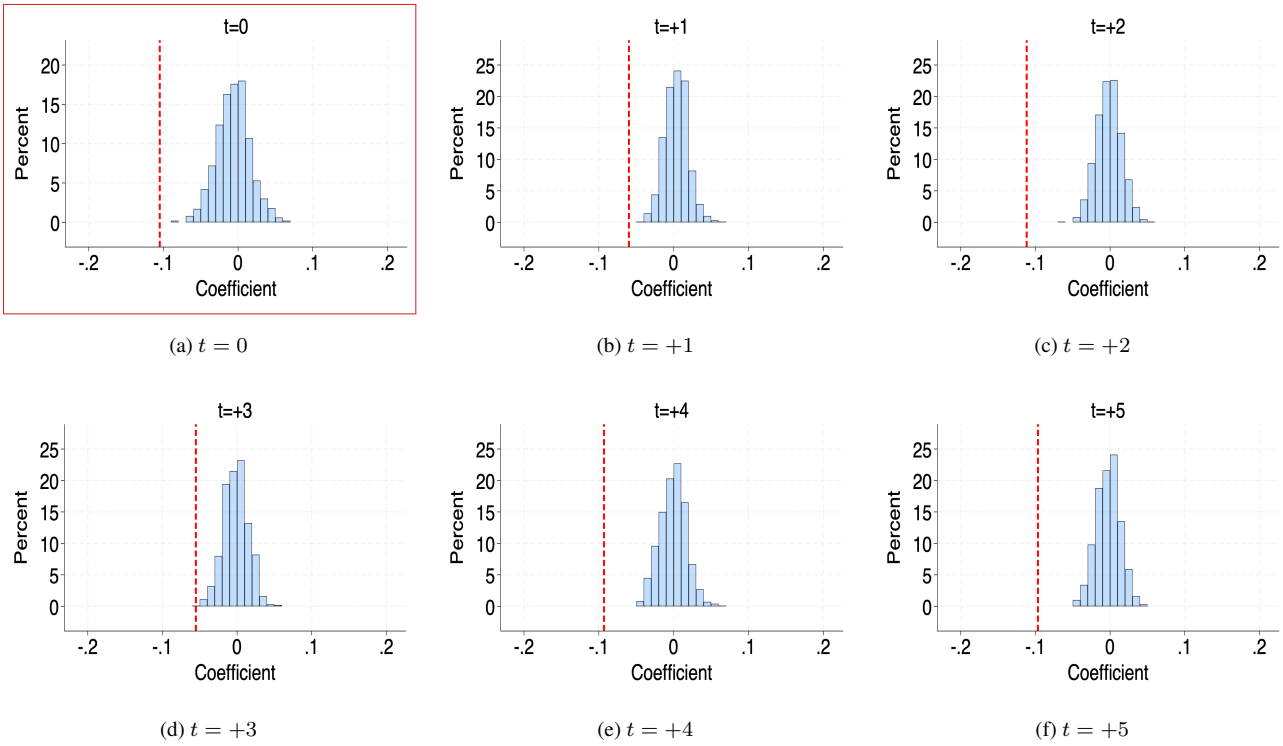


(a) $t = 0$ (b) $t = +1$ (c) $t = +2$

(d) $t = +3$ (e) $t = +4$ (f) $t = +5$

Figure 7: Placebo Event-Study for Emotionality: Random Samples of (non-JeMP) MPs

Notes: The figure repeats the event study from Figure 5 using 1,000 random samples of "Placebo JeMPs" as treatment. To construct those: (i) we exclude actual JeMPs, the President of the Chamber, and members of government; (ii) we randomly draw the 5% MPs of the remaining MPs. Panel (a) corresponds to the treatment period, and Panels (b)-(f) correspond to each of the 5 post-treatment periods. In each panel, the histogram represents the distribution of the estimated coefficients on the corresponding event-time dummy across the 1,000 simulations. The dashed red line represents the coefficient on the actual JeMPs from the event-study in Figure 5. See the note to Figure 5 for detail on the event study specification.

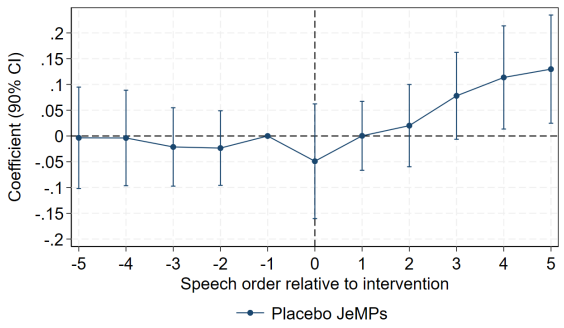


(a) $t = 0$ (b) $t = +1$ (c) $t = +2$

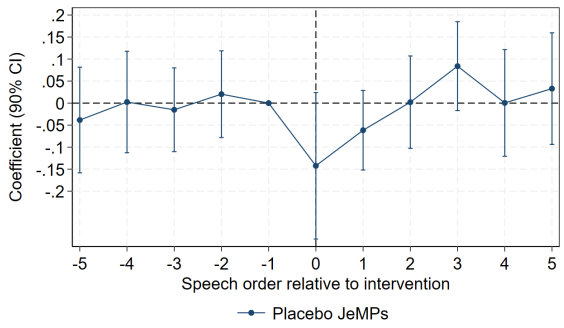
(d) $t = +3$ (e) $t = +4$ (f) $t = +5$

Figure 8: Placebo Event-Study for Polarization: Random Samples of (non-JeMP) MPs

Notes: The figure repeats the event study from Figure 6 using 1,000 random samples of "Placebo JeMPs" as treatment. To construct those: (i) we exclude actual JeMPs, the President of the Chamber, and members of government; (ii) we randomly draw the 5% MPs of the remaining MPs. Panel (a) corresponds to the treatment period, and Panels (b)-(f) correspond to each of the 5 post-treatment period. In each panel, the histogram represents the distribution of the estimated coefficients on the corresponding event-time dummy across the 1,000 simulations, and the dashed red line represents the coefficient on the actual JeMPs from the event-study in Figure 6. See the note to Figure 5 for detail on the event study specification.



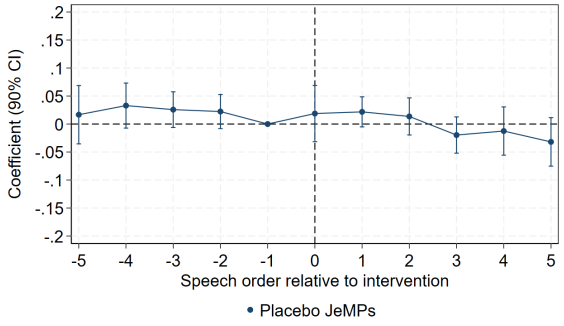
(a) Emotionality Score



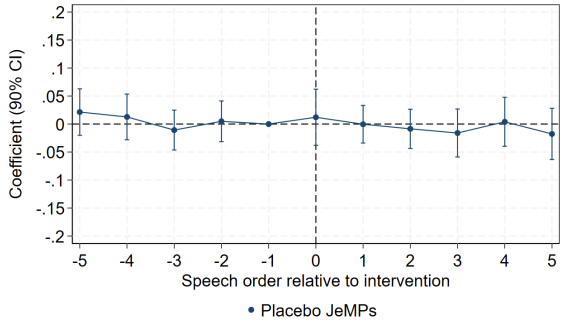
(b) Polarization Score

Figure 9: Placebo Event Study: Least Emotional or Polarized (non-JeMP) MPs

Notes: The figure repeats the event studies from Figures 5 and 6, using low-emotionality and low-polarization "Placebo JeMPs" as treatment. Those are constructed as follows: (i) we exclude actual JeMPs, the President of the Chamber, and members of government; (ii) for each year, we regress the emotionality score and the polarization score on speaker fixed effects, including the controls from col 4 in Table 1; (iii) within each year, we select the 5% MPs with the lowest estimated speaker fixed effects. The left panel in the figure uses "Placebo JeMPs" selected from the emotionality regressions, and the right panel, from the polarization regressions. See the note to Figure 5 for detail on the event study specification.



(a) Emotionality Score



(b) Polarization Score

Figure 10: Placebo Event Study: (non-JeMP) MPs from Least Emotional or Polarized Departments

Notes: The figure repeats the event studies from Figures 5 and 6, using speakers elected in the least emotional or polarized departments as "Placebo JeMPs." Those are constructed as follows: (i) we exclude actual JeMPs, the President of the Chamber, and members of government; (ii) for each year, we regress the emotionality score and the polarization score on department fixed effects, including the controls from col 5 in Table 1; (iii) within each year, we select the 10% MPs with the lowest estimated department fixed effects (iv) to approximately reflect the proportion of actual JeMPs, we randomly select 40% of these MPs to form the "Placebo JeMP" group, repeating this procedure in 1,000 simulations. See the note to Figure 5 for detail on the event study specification.

Online Appendix

De-Polarizing Politics: Faith Meets Science in the French Third Republic

Gianandrea Lanzara Mara P. Squicciarini Nico Voigtländer

A Data

A.1 Students Enrolled in Jesuit Schools in France, 1850-1901

A.1.1 Digitizing Student Registries

Data on students enrolled in Jesuit schools in France are available at the *Archives Jésuites* in Vanves near Paris.[1] The Jesuit archives are organized into four collections, each corresponding to one of the four Jesuit Provinces established in France: Paris, Lyon, Toulouse and Champagne. They contain the student registers for 30 of the 32 Jesuit schools that operated in France between 1850 and 1901.[2] The student registers are preserved in 142 separate documents dating from 1850 to 1901. For 37 of these documents, only the handwritten (hereafter HW) version has survived.[3]

Registers from each Jesuit school typically report the list of students enrolled by school year, with each document often spanning several school years. In most cases, these documents also report additional information on each student's grade, section, division (a group of students spanning multiple grades who shared the same boarding facilities), and boarding status. As an exception to the standard format, eight documents report only the name and year of entry for students.

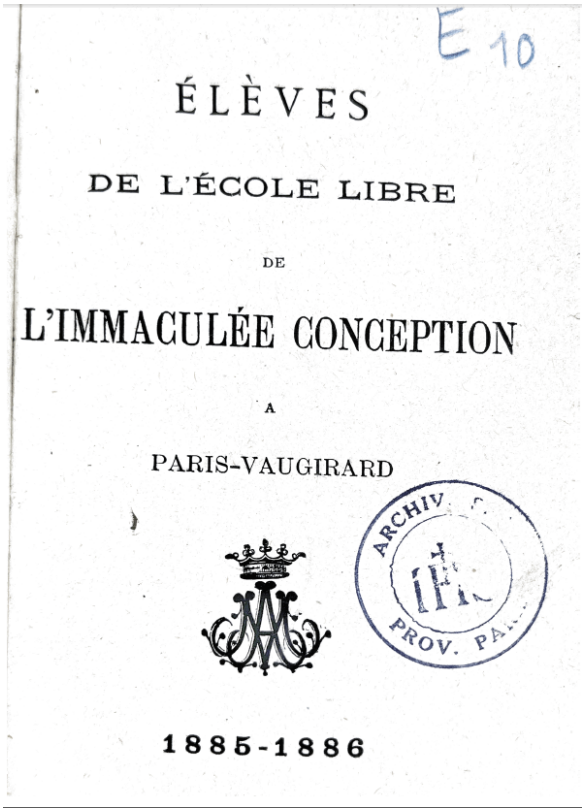
Figure A.1 illustrates the student registry for school year 1885-86 at the *Collège de l'Immaculée-Conception*, located in *Rue de Vaugirard* in Paris. The left panel shows the front page of the document, while the right panel displays the list of students in alphabetical order. The columns on this page, from left to right, indicate: (i) student number (a unique numeric identifier); (ii) student name; (iii) class (i.e., grade and section); and (iv) the division. The codes used for class, section, and division vary across schools and years, and each document typically includes the legends to explain them. Figure A.2 presents the relevant legend for this specific school and school year. To illustrate, Fernand

[1] See https://www.jesuites.com/contact/bureaux-archives-jesuites for an overview.

[2] The Provinces were, and still are, the main administrative divisions of the Society of Jesus. The archives lack documentation for the schools in Montauban and Mende, part of the Province of Toulouse, as well as for those in Algiers and Oran, located in colonial Algeria and associated with the Province of Lyon.

[3] According to private conversations with the archive staff, Jesuit school officials typically compiled the handwritten versions and sent them to a printing house.

Adhémar, Josse Allard, and Henri d'Anchald were all enrolled in the "Rhetoric" grade ("Rh" in the legend, short for *Rhétorique*), although each of them attended one of the three different sections that the school offered for this grade.[4] All three belonged to the first division. However, while Fernand and Josse were boarders, the asterisk indicates that Henri was a half-boarder. In our main analysis, we will only use the grade to infer the student's age. Below, we provide further details on the grade levels and curricular structure of Jesuit schools.



ÉLÈVES

DE L'ÉCOLE LIBRE

DE

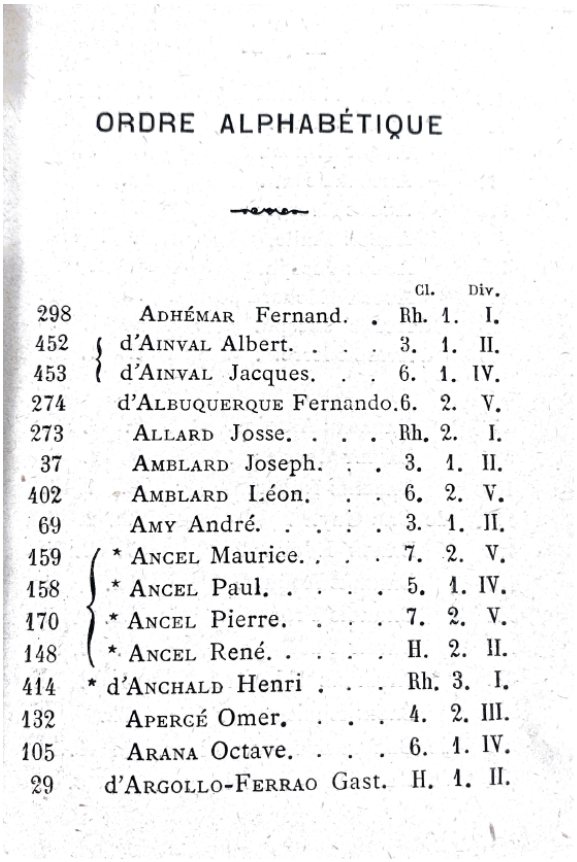
L'IMMACULÉE CONCEPTION

A

PARIS-VAUGIRARD

1885-1886

(a) Front Page



ORDRE ALPHABÉTIQUE

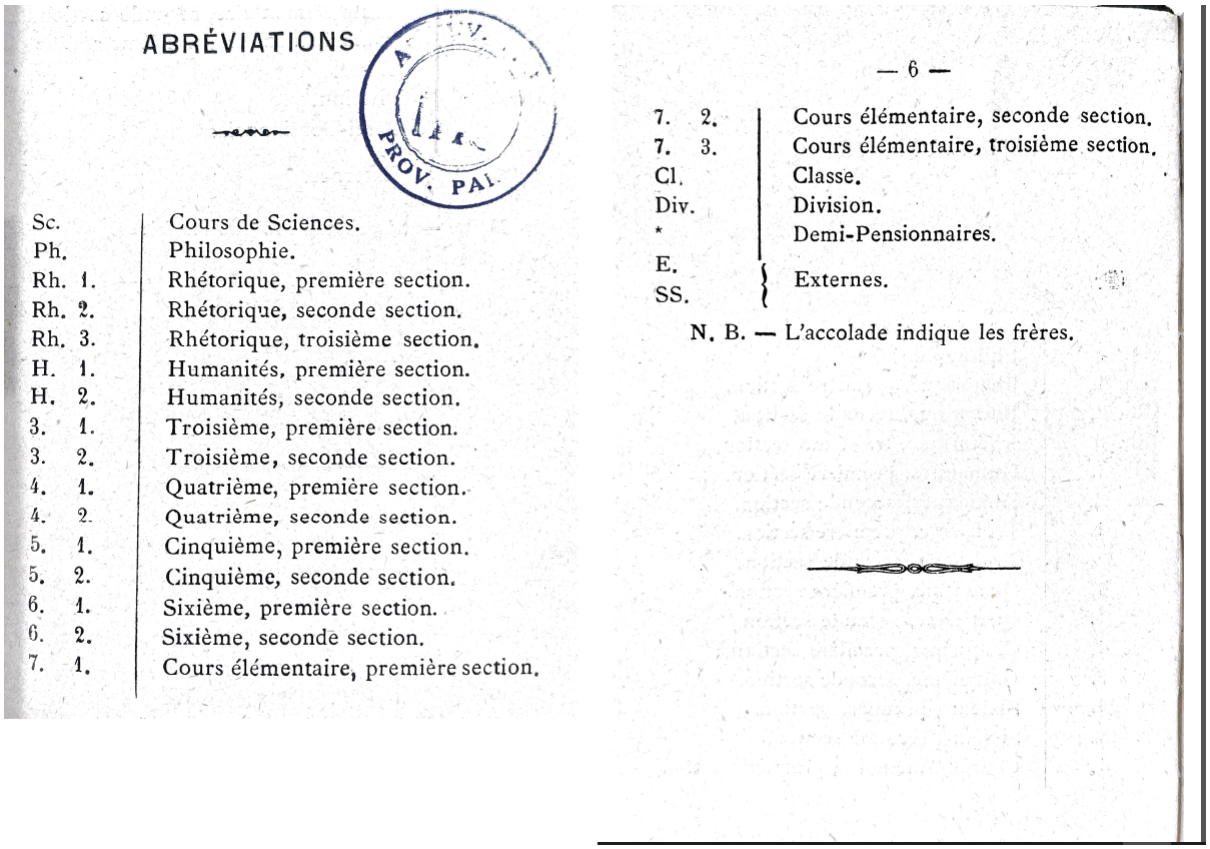
		Cl.	Div.
298	Adhémar Fernand. .	Rh.	1. I.
452	d'Ainval Albert. . . .	3.	1. II.
453	d'Ainval Jacques. . .	6.	1. IV.
274	d'Albuquerque Fernando.	6.	2. V.
273	Allard Josse. . . .	Rh.	2. I.
37	Amblard Joseph. . .	3.	1. II.
402	Amblard Léon. . .	6.	2. V.
69	Amy André.	3.	1. II.
159	* Ancel Maurice. . . .	7.	2. V.
158	* Ancel Paul.	5.	1. IV.
170	* Ancel Pierre. . . .	7.	2. V.
148	* Ancel René.	H.	2. II.
414	* d'Anchald Henri . . .	Rh.	3. I.
132	Apercé Omer. . . .	4.	2. III.
105	Arana Octave. . . .	6.	1. IV.
29	d'Argollo-Ferrao Gast.	H.	1. II.

(b) List of students in alphabetical order

Figure A.1: Student Registry – *Collège de l'Immaculée-Conception* (Paris), school year 1885-86

Note: The left panel shows the front page of the registry, while the right panel displays the list of students in alphabetical order. The columns from left to right indicate: student number, student name, class, and division (a group of students, spanning multiple grades, who shared the same boarding facilities). Asterisks indicate half-boarders, whereas curly brackets indicate brothers (see the legend in Figure A.2).

[4] At Jesuit schools, grades ran in reverse order, with *Rhétorique* corresponding to the final year of the standard curriculum. Some schools, including this one, offered additional post-secondary classes in Philosophy and Science.



ABRÉVIATIONS

Sc.	Cours de Sciences.
Ph.	Philosophie.
Rh. 1.	Rhétorique, première section.
Rh. 2.	Rhétorique, seconde section.
Rh. 3.	Rhétorique, troisième section.
H. 1.	Humanités, première section.
H. 2.	Humanités, seconde section.
3. 1.	Troisième, première section.
3. 2.	Troisième, seconde section.
4. 1.	Quatrième, première section.
4. 2.	Quatrième, seconde section.
5. 1.	Cinquième, première section.
5. 2.	Cinquième, seconde section.
6. 1.	Sixième, première section.
6. 2.	Sixième, seconde section.
7. 1.	Cours élémentaire, première section.

7. 2.	Cours élémentaire, seconde section.
7. 3.	Cours élémentaire, troisième section.
Cl.	Classe.
Div.	Division.
*	Demi-Pensionnaires.
E.	Externes.
SS.	Externes.

N. B. — L'accolade indique les frères.

Figure A.2: Legend for *Collège de l'Immaculée-Conception* (Paris), year 1885-86.

Note: From "Sc" to "7.3", the legend maps the class codes to the corresponding grade-section. For example, "Rh. 1." indicates the Rhetoric grade, first section (*Rhétorique, première section*). "Cl" and "Div" explain the names of the additional columns in the registry, respectively Class and Division (see Figure A.1). From "*" to "SS", the legend defines boarding status: half-boarders (*Demi-Pensionnaires*), and day students (*Externes*), while boarders are listed without a symbol. The final note specifies that curly brackets indicate brothers.

For printed school registries (as shown above), our digitization process consists of four main steps:

1. We manually scan all front pages, legends, and pages reporting the lists of students in alphabetical order;
2. We perform Optical Character Recognition (OCR) on each document using Google Document AI. A key advantage of Google Document AI is its ability to detect and preserve the tabular structure of the pages;
3. We parse and standardize the OCR output using custom R code, with two main objectives:
    - Associate each student list with the corresponding school year, typically extracted from the front page;
    - Infer and harmonize column types (e.g., name, student number, class, year of entry) de-

spite substantial variation in formatting. At this stage, grade, section, division, and boarding status, when available, are grouped into a single "class information" field.

This step includes numerous document-specific exceptions to accommodate irregular layouts and atypical reporting practices;

4. With the help of research assistants, we manually reviewed all final Excel files to identify and correct potential OCR errors.

For handwritten (HW) documents, we enter the data manually for the entire document, relying on a French research assistant with a background in history and prior experience reading historical French calligraphy.

At the end of this process, we obtain a dataset with 268,216 student records. Some of these records belong to the same student observed in multiple school years, while others refer to students who appear only once in the dataset along with their year of entry at their school.

A.1.2 Linking Student Records

We reconstruct the school histories of Jesuit-school students based on their names and the years during which they appear in the records. The main objective is to determine the earliest and latest years during which each student attended a Jesuit school. This information, in turn, helps us estimate the student's age, which is ultimately needed to match student records to parliamentarians (see Section A.1.3 below).

First, we clean the student names with the following steps: (i) we expand common abbreviations to full names (e.g., *Pre* $\rightarrow$ *Pierre*, *Lt* $\rightarrow$ *Laurent*); (ii) we remove all punctuation marks, which typically reflect OCR or transcription errors, with the exception of apostrophes and dashes (unless these appear at the end of the string); (iii) we remove all non-ASCII characters; (iv) we set the text to lower case.

Second, we extract the first name and the last name from the name field. To do so, we must determine whether names in a given document-year follow a "firstname lastname" or "lastname firstname" format. We check the position of common French first names (e.g., *Jean*, *Louis*, *Henri*) within the name string. Based on the predominant pattern, we extract the first name either from the beginning or end of the cleaned name string. For HW documents, we directly classify first and last names in transcribing the handwritten records.

Third, we remove all nobiliary particles such as *de*, *du*, *la*, and *le* from surnames, except when they are part of a hyphenated compound name (e.g., *d'aubigny* $\rightarrow$ *aubigny*, but *mery-le-beuve* $\rightarrow$ *mery-le-beuve*).

Fourth, we construct the variables $y_r^{\min}$ and $y_r^{\max}$ for each *record* r as follows: (i) if the document reports the school year in standard format, e.g., 18XX–18(XX+1), we assign 18XX to $y_r^{\min}$ and 18(XX+1) to $y_r^{\max}$; (ii) if the document only reports the year of entry, we assign this value to both $y_r^{\min}$

and $y_r^{\max}$.[5]

Once we have the cleaned first name, last name, and the narrowest possible time window for each record, we implement an iterative procedure to link student records across school years. As multiple records are linked, we update $y_i^{\min}$ and $y_i^{\max}$ for each *student* i accordingly. Let $\mathcal{R}_i^k$ be the set of records associated with student i at the end of iteration k; then,

$$y_i^{\min,k+1} = \min_{r \in \mathcal{R}_i^k} y_r^{\min,k}, \quad \text{and} \quad y_i^{\max,k+1} = \max_{r \in \mathcal{R}_i^k} y_r^{\max,k}.$$

Once all records for student i are linked, $y_i^{\min}$ and $y_i^{\max}$ represent the earliest and latest years during which student i is observed in a Jesuit school.

The iterative procedure works as follows. As an initial step, we link records with an exact name match within the same school, provided that the student numbers as reported in the original documents coincide (as there could be cases of multiple names within the same schools).[6] Next, we iterate through the following steps:

- We consider all possible pairs of student records within the same school, under the restriction that their enrollment intervals in the current iteration – defined by their respective $y_i^{\min,k}$ and $y_i^{\max,k}$ – overlap within a four-year window;[7]
- If both the first and last names are identical, we link the records;
- We then consider four additional, and strictly looser, criteria for linking records:
  - One surname is a subset of the other, and the first names are identical;
  - One first name is a prefix of the other, and the last names are identical;
  - The first names have a Jaro distance below 0.1, and the last names are identical[8];
  - The last names have a Jaro distance below 0.1, and the first names are identical.

[5]Two documents, one for the school in Mongre and one for the school in Saint-Etienne, also report the year of exit. In these cases, we assign the year of entry to $y_r^{\min}$ and the year of exit to $y_r^{\max}$. These records are included in the iterative procedure, as additional records for the same students may exist in other documents from the same school. If the entry and exit years produced by the iterative procedure differ from those reported in the original document, we retain the former.

[6]Although these individual identifiers were intended to track students across school years, they exhibit too many inconsistencies to be reliably used on their own as linking variables. Moreover, they are not available for all entries. As a result, we use them only in this initial step to reduce the dimensionality of the name-matching problem.

[7]This restriction ensures that we do not link records with enrollment periods that are implausibly far apart, even if the student names match. For example, suppose we have a record for Jean Dubois in "1870-1871" and another one in the same school for "1880-1881," with no information on the student in between. These records are unlikely to belong to the same student, although the names match. Our four-year time-window restriction ensures that when comparing pairs of records At the same time, the four-year window is flexible enough to allow links between records that are reasonably close in time, allowing for temporary school absence, missing records, or OCR errors leading to gaps in student records.

[8]The Jaro distance measures the similarity between two sequences, taking into account the number of matching characters. The score is normalized to range from 0 (exact match) to 1 (completely dissimilar). A threshold of 0.1 is conservative: for example, "Smith" and "Smyth" have Jaro distance of 0.13.

- We update $y_i^{\min,k+1}$ and $y_i^{\max,k+1}$ for all linked records and repeat the procedure.

Because the enrollment window may expand after each round of matching, new records may be linked in subsequent iterations. The process continues until no further links are formed.

A.1.3 Determining the age of students

We use class information reported in the original documents, whenever available, to estimate the age of each student. As a first step, we harmonize class codes across schools and years using the legends provided in the school registries (see Figure A.2). At Jesuit schools, following the French school system of the time, grades ran in reverse order, with *Rhétorique* corresponding to the final year of the standard curriculum (coded as grade 1, which corresponds to grade 12 in modern terminology). Some schools, such as the one in *Rue de Vaugirard* in Paris (illustrated in Figure A.2), offered additional post-secondary classes in Philosophy and Science to prepare students for the *Grandes Écoles*. We code these as grades 0 and -1, respectively[9].

As a validation for both our student-record-linkage procedure and our coding of class levels, we compute grade-to-grade transition matrices for all students observed in consecutive years. Reassuringly, 96% of students either advanced or repeated the grade (with 82% advancing and 14% repeating). As an example, Figure A.3 below shows the transition matrix for students enrolled at the *Collège de l'Immaculée-Conception* on *Rue de Vaugirard* in Paris.

[9]Entry-level grades at Jesuit schools varied: while some schools began at grade 7 (i.e., grade 6 in today's upward-counting terminology for school years), others offered preparatory courses starting as early as grade 9 (corresponding to modern fourth grade).

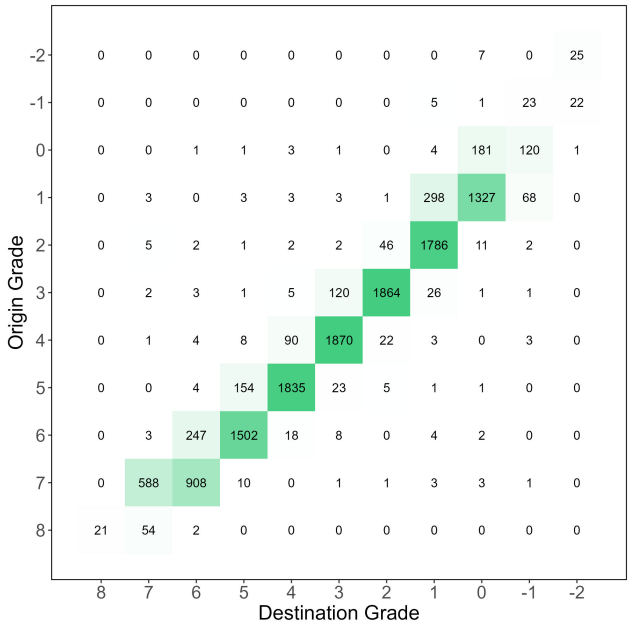


Figure A.3: Grade-to-grade transition matrix for the *Collège de l'Immaculée-Conception* (Paris)

Note: The figure displays the grade-to-grade transition matrix for students observed in at least two consecutive school years at the *Collège de l'Immaculée-Conception* on *Rue de Vaugirard* in Paris. Each row corresponds to the grade a student attended in year t, and each column to the grade attended in year $t+1$. The cells of the matrix report the number of observed transitions from a given grade to another grade in the subsequent year.

As mentioned above, in the French school system of the time grades ran in reverse order. Elementary school started at age 6 in grade 11. Accordingly, we then estimate the implied age for each record r using the formula: $Age_r = 6 + (11 - Grade_r)$.

A.2 Matching Students with Parliamentarians

We match students from Jesuit schools and parliamentarians based on their names and year of birth. For parliamentarians, we know the exact date of birth. For students, we approximate their year of birth using either information on their imputed age, where available (see Section A.1.3), or the number of years during which they are observed in a Jesuit school. Note that grade-based age information is available for approximately 90% of all records and students.

When age information is available for a specific record, we estimate the implied year of birth for record r as $y_r^{\text{min}} - \text{age}_r$. We then aggregate all records linked to the same student (see Section A.1.2 for details on the linkage procedure) and set $ybirth_i^{min}$ and $ybirth_i^{max}$ equal to, respectively, the minimum and the maximum implied years of birth across all records for the same students.

For the remaining 10% of students whose age information is missing from all linked records, we impute a plausible birth year range based on the years during which the student is observed in a Jesuit

school. Jesuit students typically attended school between the ages of 8 and 18, with a full trajectory lasting 11 years. Based on this, we compute a conservative interval for the year of birth:

$$ybirth_i^{\text{min}} = \min(y_i^{\text{min}} - 19 + n,\ y_i^{\text{max}} - 7 - n),$$
$$ybirth_i^{\text{max}} = \max(y_i^{\text{min}} - 19 + n,\ y_i^{\text{max}} - 7 - n),$$

where n is the number of years the student appears in the data, capped at 11. This formula incorporates the idea that the more years a student is observed, the more accurately we can infer their age. To see this, consider two edge cases:

- If a student is observed for the full eleven years ($n = 11$), then we can assume that the earliest observation corresponds to age 8 and the latest to age 18. This implies a year of birth between $y_i^{\text{min}} - 8$ and $y_i^{\text{max}} - 18$.
- If a student is observed for only one year ($n = 1$), this record may correspond either to the beginning or the end of their school career, leading to a wider range of possible birth years.

For each parliamentarian, we then look for a match in the dataset of linked students. Our matching procedure consists of four steps.

- first, as a minimal consistency requirement, we restrict attention to students who appear in a record when the parliamentarian was older than 4 and younger than 22.
- second, we only retain students whose birth window overlaps with the MP's birth year, within a margin of 4 years on either side.
- third, we identify students whose last name is similar to the MP's last name using the Jaro string distance metric, with a cutoff of 0.1. Because, after linking the student records, we may observe different variants for a student's last name, we consider each of the variants.
- fourth, for students whose last names matched the MP's last name in step 3, we apply a similar fuzzy matching procedure to the first names.[10]

Finally, we manually review all fuzzy matches and retain only those where discrepancies clearly stem from minor typos or spelling variations that do not affect French pronunciation (e.g., *Louis Du Bochet* and *Louis Dubochet*). This procedure identifies 117 parliamentarians – 108 exact matches and 9 fuzzy matches – who served in the *Chambre des députés* between January 1881 and April 1914.

[10]For parliamentarians with multiple given names, we consider only the first given name.

A.3 Parliamentary Speeches, 1881–1914

A.3.1 Digitizing Parliamentary Debates

The transcripts of parliamentary debates of the French Third Republic were published in the official gazette (*Journal officiel de la République française*) for both the Chamber of Deputies (*Chambre des députés*) and the Senate (*Sénat*). These publications are available online in PDF format and as an OCR-generated text layer on Gallica, the digital repository of the French National Library[11] While Gallica provides a text version, we found that OCR errors were too frequent to allow the reliable parsing of individual speeches; we therefore opted for re-digitizing the documents from scratch to obtain a higher-quality text extraction.

Figure A.4 shows the front page of the parliamentary transcript for the sitting of April 5th, 1905 at the Chambre of Deputies. While this page illustrates the typical three-column structure of the *Journal officiel*, some pages embed numerical tables that disrupt the three-column layout. Because OCR software may fail to reliably detect column breaks, we use Detectron2 image detection algorithms to segment each page into text columns/blocks and tables, and then apply Google Cloud Vision for optical character recognition on each detected text region.[12]

More specifically, our digitization process consists of six main steps:

1. We download all PDFs for both branches of Parliament from 1881 to 1942. For dates with multiple PDFs, we manually verify these cases and discard duplicates when they contain transcripts of the same parliamentary sitting.

2. We convert all PDF pages into high-resolution JPEG images.

3. We train an image-detection model to identify pages containing tables, using a manually annotated sample of 60 pages. When we identify a table, we strip it from the page before further processing by setting its pixels to white.

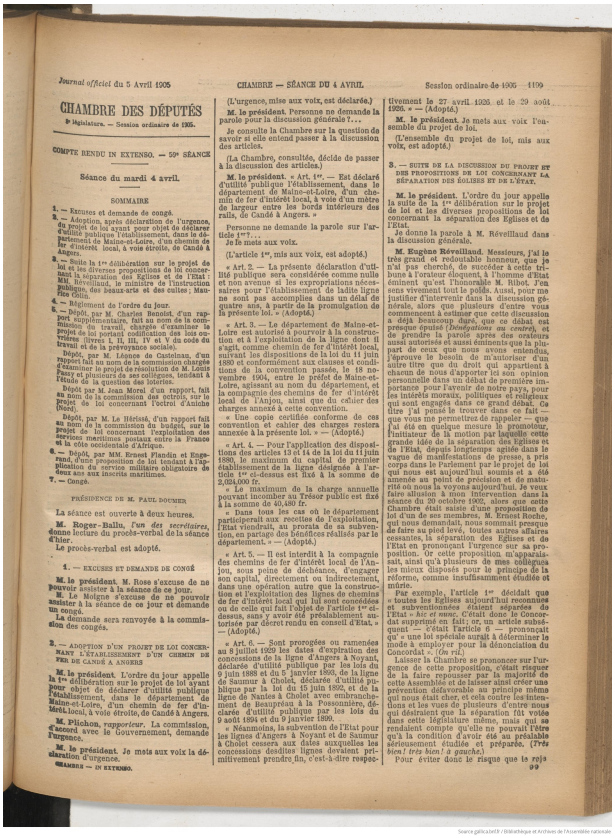
4. We train two layout models to detect individual text columns/blocks: one for standard three-column pages and one for pages containing tables. Each model is trained on a sample of 50 pages with the corresponding structure.

5. We use Google Cloud Vision to perform optical character recognition on each detected text column/block.

[11] See https://gallica.bnf.fr.

[12] Detectron2 is an image detection library developed by Facebook AI, see https://github.com/facebookresearch/detectron2. For information on Google Cloud Vision, see https://docs.cloud.google.com/vision/docs/ocr.

6. Finally, we reassemble columns/blocks (using their coordinates) to reconstruct each page, and then concatenate pages to reconstruct each document.



Journal officiel du 5 Avril 1905 — CHAMBRE — SÉANCE DU 4 AVRIL — Session ordinaire de 1905 — 1199

CHAMBRE DES DÉPUTÉS

9e législature. — Session ordinaire de 1905.

COMPTE RENDU IN EXTENSO. — 59e SÉANCE

Séance du mardi 4 avril.

SOMMAIRE

1. — Excuses et demande de congé.
2. — Adoption, après déclaration de l'urgence, du projet de loi ayant pour objet de déclarer d'utilité publique l'établissement, dans le département de Maine-et-Loire, d'un chemin de fer d'intérêt local, à voie étroite, de Candé à Angers.
3. — Suite de la 1re délibération sur le projet de loi et les diverses propositions de loi concernant la séparation des Eglises et de l'Etat : M. Réveillaud, le ministre de l'instruction publique, des beaux-arts et des cultes ; Marcel Sibille.
4. — Règlement de l'ordre du jour.
5. — Dépôt, par M. Charles Benoist, d'un rapport supplémentaire, fait au nom de la commission du travail, chargée d'examiner la proposition de loi portant codification des lois ouvrières (livres I, II, III, IV et V du code du travail et de la prévoyance sociale).

Dépôt, par M. Léonce de Castelnau, d'un rapport fait au nom de la commission chargée d'examiner le projet de résolution de M. Louis Passy et plusieurs de ses collègues, tendant à l'étude de la question des loteries.

Dépôt par M. Jean Morel d'un rapport, fait au nom de la commission des octrois, sur le projet de loi concernant l'octroi d'Asnières (Seine).

Dépôt, par M. Le Hérissé, d'un rapport fait au nom de la commission du budget, sur le projet de loi concernant l'exploitation des services maritimes postaux entre la France et la côte occidentale d'Afrique.

6. — Dépôt, par MM. Ernest Flandin et Engerand, d'une proposition de loi tendant à l'application du service militaire obligatoire de deux ans aux inscrits maritimes.
7. — Congé.

PRÉSIDENCE DE M. PAUL DOUMER

La séance est ouverte à deux heures.

M. Roger-Ballu, l'un des secrétaires, donne lecture du procès-verbal de la séance d'hier.

Le procès-verbal est adopté.

1. — EXCUSES ET DEMANDE DE CONGÉ

M. le président. M. Rose s'excuse de ne pouvoir assister à la séance de ce jour.

M. Le Moigne s'excuse de ne pouvoir assister à la séance de ce jour et demande un congé.

La demande sera renvoyée à la commission des congés.

2. — ADOPTION D'UN PROJET DE LOI CONCERNANT L'ÉTABLISSEMENT D'UN CHEMIN DE FER DE CANDÉ A ANGERS

M. le président. L'ordre du jour appelle la 1re délibération sur le projet de loi ayant pour objet de déclarer d'utilité publique l'établissement, dans le département de Maine-et-Loire, d'un chemin de fer d'intérêt local, à voie étroite, de Candé à Angers.

M. Pichon, rapporteur. La commission, d'accord avec le Gouvernement, demande l'urgence.

M. le président. Je mets aux voix la déclaration d'urgence.

CHAMBRE — IN EXTENSO.

(L'urgence, mise aux voix, est déclarée.)

M. le président. Personne ne demande la parole pour la discussion générale ?...

Je consulte la Chambre sur la question de savoir si elle entend passer à la discussion des articles.

(La Chambre, consultée, décide de passer à la discussion des articles.)

M. le président. « Art. 1er. — Est déclaré d'utilité publique l'établissement, dans le département de Maine-et-Loire, d'un chemin de fer d'intérêt local, à voie d'un mètre de largeur entre les bords intérieurs des rails, de Candé à Angers. »

Personne ne demande la parole sur l'article 1er ?...

Je le mets aux voix.

(L'article 1er, mis aux voix, est adopté.)

« Art. 2. — La présente déclaration d'utilité publique sera considérée comme nulle et non avenue si les expropriations nécessaires pour l'établissement de ladite ligne ne sont pas accomplies dans un délai de quatre ans, à partir de la promulgation de la présente loi. » (Adopté.)

« Art. 3. — Le département de Maine-et-Loire est autorisé à pourvoir à la construction et à l'exploitation de la ligne dont il s'agit, comme chemin de fer d'intérêt local, suivant les dispositions de la loi du 11 juin 1880 et conformément aux clauses et conditions de la convention passée, le 18 novembre 1904, entre le préfet de Maine-et-Loire, agissant au nom du département, et la compagnie des chemins de fer d'intérêt local de l'Anjou, ainsi que du cahier des charges annexé à cette convention.

« Une copie certifiée conforme de ces convention et cahier des charges restera annexée à la présente loi. » — (Adopté.)

« Art. 4. — Pour l'application des dispositions des articles 13 et 14 de la loi du 11 juin 1880, le maximum du capital de premier établissement de la ligne désignée à l'article 1er ci-dessus est fixé à la somme de 2,054,000 fr.

« Le maximum de la charge annuelle pouvant incomber au Trésor public est fixé à la somme de 40,480 fr.

« Dans tous les cas où le département participerait aux recettes de l'exploitation, l'Etat viendrait, au prorata de sa subvention, en partage des bénéfices réalisés par le département. » — (Adopté.)

« Art. 5. — Il est interdit à la compagnie des chemins de fer d'intérêt local de l'Anjou, sous peine de déchéance, d'engager son capital, directement ou indirectement, dans une opération autre que la construction et l'exploitation des lignes de chemins de fer d'intérêt local qui lui sont concédées ou de celle qui fait l'objet de l'article 1er ci-dessus, sans y avoir été préalablement autorisée par décret rendu en conseil d'Etat. » — (Adopté.)

« Art. 6. — Sont prorogées ou ramenées au 9 juillet 1929 les dates d'expiration des concessions de la ligne d'Angers à Noyant, déclarée d'utilité publique par les lois du 9 juin 1888 et du 5 janvier 1893, de la ligne de Saumur à Cholet, déclarée d'utilité publique par la loi du 15 juin 1895, et de la ligne de Nantes à Cholet avec embranchement de Beaupréau à la Possonnière, déclarée d'utilité publique par les lois du 9 août 1894 et du 9 janvier 1899.

« Néanmoins, la subvention de l'Etat pour les lignes d'Angers à Noyant et de Saumur à Cholet cessera aux dates auxquelles les concessions desdites lignes devaient primitivement prendre fin, c'est-à-dire respectivement le 27 avril 1926 et le 29 août 1926. » — (Adopté.)

M. le président. Je mets aux voix l'ensemble du projet de loi.

(L'ensemble du projet de loi, mis aux voix, est adopté.)

3. — SUITE DE LA DISCUSSION DU PROJET ET DES PROPOSITIONS DE LOI CONCERNANT LA SÉPARATION DES ÉGLISES ET DE L'ÉTAT.

M. le président. L'ordre du jour appelle la suite de la 1re délibération sur le projet de loi et les diverses propositions de loi concernant la séparation des Eglises et de l'Etat.

Je donne la parole à M. Réveillaud dans la discussion générale.

M. Eugène Réveillaud. Messieurs, j'ai le très grand et redoutable honneur, que je n'ai pas cherché, de succéder à cette tribune à l'orateur éloquent, à l'homme d'Etat éminent qu'est l'honorable M. Ribot. J'en sens vivement tout le poids. Aussi, pour me justifier d'intervenir dans la discussion générale, alors que plusieurs d'entre vous commencent à estimer que cette discussion a déjà beaucoup duré, que ce débat est presque épuisé (*Dénégations au centre*), et de prendre la parole après des orateurs aussi autorisés et aussi éminents que la plupart de ceux que nous avons entendus, j'éprouve le besoin de m'autoriser d'un autre titre que du droit qui appartient à chacun de nous d'apporter ici son opinion personnelle dans un débat de première importance pour l'avenir de notre pays, pour les intérêts moraux, politiques et religieux qui sont engagés dans ce grand débat. Ce titre j'ai pensé le trouver dans ce fait — que vous me permettrez de rappeler — que j'ai été en quelque mesure le promoteur, l'initiateur de la motion par laquelle cette grande idée de la séparation des Eglises et de l'Etat, depuis longtemps agitée dans le vague de manifestations de presse, a pris corps dans le Parlement par le projet de loi qui nous est aujourd'hui soumis et a été amenée au point de précision et de maturité où nous la voyons aujourd'hui. Je veux faire allusion à mon intervention dans la séance du 20 octobre 1902, alors que cette Chambre était saisie d'une proposition de loi d'un de ses membres, M. Ernest Roche, qui nous demandait, nous sommait presque de faire au plus tôt, toutes autres affaires cessantes, la séparation des Eglises et de l'Etat en prononçant l'urgence sur sa proposition. Or cette proposition m'apparaissait, ainsi qu'à plusieurs de mes collègues les mieux disposés pour le principe de la réforme, comme insuffisamment étudiée et mûrie.

Par exemple, l'article 1er décidait que « toutes les Eglises aujourd'hui reconnues et subventionnées étaient séparées de l'Etat » *hic et nunc*. C'était donc le Concordat supprimé en fait; or, un article subséquent — c'était l'article 6 — prononçait qu' « une loi spéciale aurait à déterminer le mode à employer pour la dénonciation du Concordat ». (*On rit.*)

Laisser la Chambre se prononcer sur l'urgence de cette proposition, c'était risquer de la faire repousser par la majorité de cette Assemblée et de laisser ainsi créer une prévention défavorable au principe même qui nous était cher, et cela contre les intentions et les vues de plusieurs d'entre nous qui désiraient que la séparation fût votée dans cette législature même, mais qui se rendaient compte qu'elle ne pouvait l'être qu'à la condition d'avoir été au préalable sérieusement étudiée et préparée. (*Très bien ! très bien ! à gauche.*)

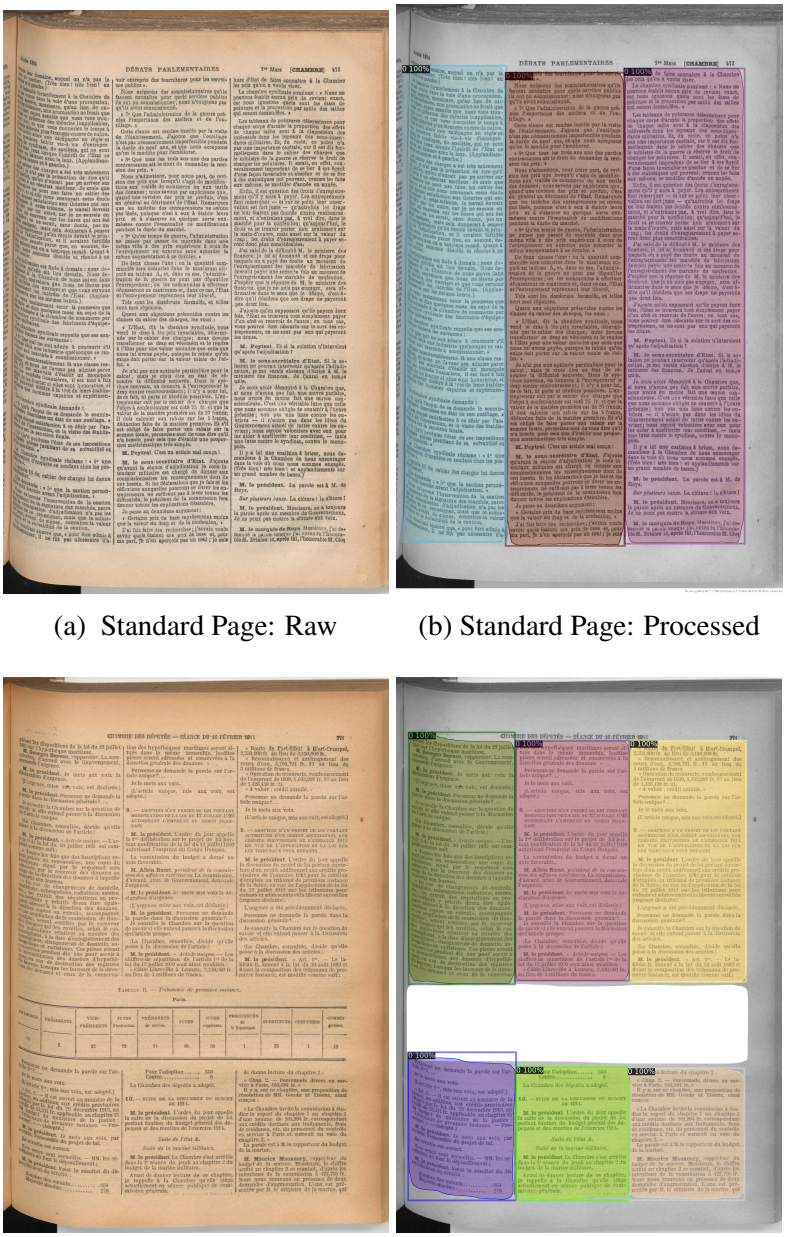
Pour éviter donc le risque que le rejet

99

Source gallica.bnf.fr / Bibliothèque et Archives de l'Assemblée nationale

Figure A.4: Parliamentary Transcript - Chamber of Deputies, April 5th, 1905

Note: The figure shows the front page of the parliamentary transcript for the sitting of April 5th, 1905 at the Chambre of Deputies.



(a) Standard Page: Raw (b) Standard Page: Processed

(c) Page with Table: Raw (d) Page with Table: processed

Figure A.5: Parsing the Layout of Parliamentary Transcripts

Note: The figure illustrates our procedure for parsing the layout of the parliamentary transcripts. We first train an image-detection model to identify pages containing tables. For standard pages with a three-column structure (a), we then use an image-detection model to identify column breaks (b). For pages containing a table (c), we strip the table by setting its pixels to white and use an image-detection model to detect the remaining text regions (d)

A.3.2 Parsing Parliamentary Speeches

Our next task is to split the digitized transcripts into parliamentary sittings, agenda items, and individual speeches. First, we separate distinct sittings within each transcript document (when Parliament met multiple times on the same date, the corresponding record typically reports all sittings). Second, we partition each sitting into consecutive sections of the debate, which we refer to as discussions, each centered on a specific issue or bill that Parliament was called upon to examine. Finally, we split each discussion into individual speeches.

As illustrated in Figure A.4, each sitting begins with a short agenda summary under the header "SOMMAIRE." The summary is followed by a header naming the presiding officer (e.g., "PRESIDENCE DE M. PAUL DOUMER"), after which the chair announces the start of proceedings with the formula *la séeance est ouverte*. Each agenda item is then introduced by an uppercase header (for instance, the second item in Figure A.4 concerns the layout of a railway line between Cand'e and Angers). Finally, each intervention starts on a new line, typically after a period or an uppercase header, and the speaker is introduced by "M." followed by a surname (e.g., M. Plichon) or full name (e.g., M. Eugène Réveillaud), ending with a period. We leverage these conventions to construct regular-expression patterns that identify sitting boundaries, section headers, and to extract individual speeches along with the reported name of the speaker.

B Text Analysis

B.1 Text Preprocessing

We preprocess the text to reduce dimensionality (Gentzkow et al., 2019). First, we assign part-of-speech (POS) tags [13], retaining only adjectives, adverbs, nouns, and verbs from each speech. Second, we normalize the tokens by removing accents, converting all text to lowercase, stripping punctuation and digits, and keeping only words composed entirely of alphabetic characters. Third, we apply the implementation of the Porter2 stemming algorithm provided in the R package `SnowballC` Porter (2009) to reduce each word to its morphological root.

B.2 Removing Stopwords and Procedural Words

We further preprocess the text data to remove uninformative words. First, we remove short words consisting of fewer than three characters. Second, we filter out rare words that occur in five ore fewer documents or ten or fewer times in the entire corpus. Third, we remove stopwords using a custom list of French stopwords.[14] Next, we remove *procedural* language that is specific to parliamentary

[13] We use the French language model (`fr_core_news_lg`) from the Python library `SpaCy`

[14] We use the list of French stopwords available in the R package `stopwords`, and supplement it with the days of the week, months, and names of French departments, plus 12 frequent low-information terms: (e.g., *faire*, *parl*, *temps*, *nombre*, etc.)

discourse. Following Gentzkow et al. (2019), we digitized the Rules of Procedure (*Règlements*) of the Chamber of Deputies and the Senate during the Third Republic, as compiled in *Les Règlements des Assemblées Législatives de France depuis 1789* by Bonnard (1926). We use these texts to identify a set of procedural stopwords – terms frequently used in the context of legislative procedure. Specifically, we implement the following steps: (i) We pre-process the text of the Rules of Procedure using the same pipeline described in Section B.1 (except for the POS tagging). (ii) We train an auxiliary Word2Vec model on the full corpus of parliamentary speeches and use it to generate vector embeddings for each word in the Rules of Procedure. (iii) Using these embeddings, we construct a frequency-weighted centroid vector that captures the semantic content of procedural language. (iv) We compute the cosine distance between each (vectorized) word in the corpus of speeches and the procedural centroid. (v) We rank all words by this distance and remove those falling in the bottom 1% (i.e., most similar to the procedural language centroid). Figure A.6 shows a word cloud of the 100 words in the corpus with the smallest cosine distance to the centroid of procedural terms.

After removing short words, rare words, stopwords, and procedural words, our sample comprises 1,468,790 speeches with nonempty text.



Figure A.6: Procedural Lexicon Removed from the Corpus

Note: The figure shows a word cloud of the 100 words in the corpus with the smallest cosine distance to the centroid of procedural terms, computed using an auxiliary word-embedding model. The procedural centroid is defined as a frequency-weighted average of the (vectorized) words contained in the Rules of Procedure of the French Parliament during the Third Republic. After removing these procedural terms from the corpus, we train the main word-embedding model.

B.3 Word Embedding

We use this sample of preprocessed speeches to train a Word2vec model using the Python `gensim` package (Mikolov et al., 2013). We train the model with 300 dimensions and using a context window of 8 words. As a result, we obtain vector representations of all words in the corpus.

We then compute a document embedding for each speech as the weighted average of its word vectors. We apply inverse-frequency weighting to down-weigh frequent words in the corpus. Specif-

ically, for a given speech $\mathcal{D}_i$, the document vector $\vec{d_i}$ is computed as:

$$\vec{d_i} = \frac{1}{|\mathcal{D}_i|} \sum_{w \in \mathcal{D}_i} \frac{\alpha}{\alpha + p(w)} \vec{w} \tag{A.1}$$

where $\vec{w}$ is the vector representation of word w, $p(w)$ is the relative frequency of the word in the corpus, and $\alpha = 0.001$ is a smoothing parameter [15].

B.4 Topic Analysis

We apply Latent Dirichlet Allocation (LDA) to uncover the latent topics in our corpus of speeches (Blei et al., 2003). The LDA is an unsupervised topic modeling technique that identifies recurring semantic patterns across documents. Each document is modeled as a mixture of topics and each topic as a distribution over words.[16] Figure A.7 shows the word-topic matrix from our preferred specification with 20 topics. Each row corresponds to a topic, represented as a probability distribution over the entire vocabulary. The ten columns display the ten words with the highest probability weights in each topic, with darker shades indicating higher weights. Most topics have a clear interpretation and correspond to major policy domains. To mention just three examples, Topic 4 is associated with schooling, Topic 8 with labor, and Topic 11 with military affairs.

[15] We set the value of α following Gennaro and Ash (2022)

[16] We use the implementation provided in R package `topicmodels` (Grün and Hornik, 2011) with Gibbs sampling. Following Griffiths and Steyvers (2004), we set the Dirichlet prior $\alpha = 2.5$, and $\delta = 0.1$. Following Hansen, McMahon, and Prat (2018), we run the model for 4,000 iterations with a burn-in of 4,000 and a thinning interval of 50, and we average document-topic and word-topic distributions over iterations.

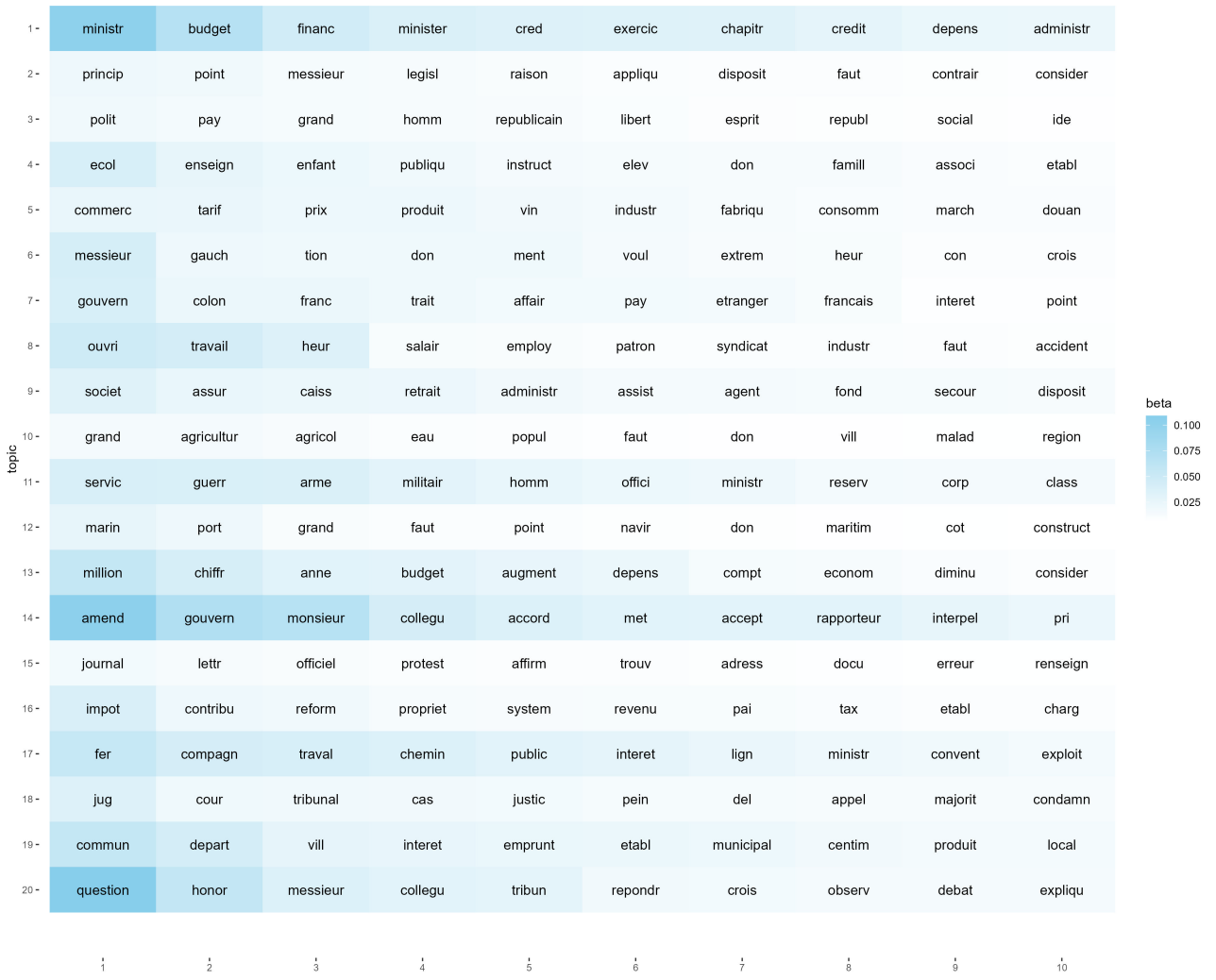


Figure A.7: Word-Topic Distributions

Note: The figure displays the 20 estimated topics from our Latent Dirichlet Allocation analysis. Topics are modeles as probability distributions over words in the vocabulary. Each row corresponds to a topic, and the ten columns display the ten words with the highest probability weights in each topic, with darker shades indicating higher weights.

B.5 Emotionality Score

We follow Gennaro and Ash (2022) to build a measure of emotionality in political speeches. First, we use the French version of the Linguistic Inquiry and Word Count (LIWC) to construct dictionaries of emotional and rational language Tausczik and Pennebaker (2010). LIWC dictionaries were developed to identify linguistic markers of emotions, cognition, and social processes. They are based on expert annotation and psycholinguistic validation of thousands of English words grouped by semantic

and psychological categories. We use the French version of these dictionaries available in the 2007 release. Following Gennaro and Ash (2022), we focus on the "Affective Processing" and "Cognitive Processing" categories to capture emotionality and reasoning, respectively. The Affective category includes 1,445 words related to positive and negative emotions and moods, such as anxiety, anger, and sadness. The Cognitive category includes 799 words related to causation, comparison, uncertainty, inclusion, and exclusion. We refine the dictionaries through the following steps: (i) wildcard entries are converted to regular expressions and matched to French lemmas in the WordNet lexical database (e.g., wildcard "pens*" is matched to "penser", "pensif", "penseur", etc.) ; (ii) words that appear in both affective and cognitive categories are removed; (iii) for each category, we compute the centroid of the word vectors and exclude the 25% of terms that are most semantically distant from the centroid; (iv) we normalize the remaining words by lower-casing, removing diacritics, and applying stemming (and remove again words shared between the two lists following this step); (v) finally, we eliminate all words found in our list of French stopwords.

Second, we compute the centroid vector of of the emotional and cognitive dictionaries in our word embedding, as the inverse-frequency weighted average of the individual word vectors. Let $\mathbf{E}$ and $\mathbf{C}$ denote the centroids of the emotional and cognitive dictionaries within the word embedding. Using these centroids, we compute the emotionality score of document $\mathcal{D}_i$ as:

$$escore_i \;=\; \frac{\mathrm{sim}(\vec{d_i}, \mathbf{E}) + 1}{\mathrm{sim}(\vec{d_i}, \mathbf{C}) + 1} \tag{A.2}$$

where $\vec{d_i}$ denotes the vector representation of document $\mathcal{D}_i$ as defined in (A.1) and sim denotes the cosine similarity. As an illustration, Figure A.8 shows the word clouds of the 100 words with the smallest cosine distance to, respectively, the emotional and rational centroids, excluding the words that directly appear in the LIWC dictionaries.



(a) Emotionality



(b) Rationality

Figure A.8: Emotional and Rational Lexicon

Note: The figure shows the word clouds of the 100 words in the corpus with the smallest cosine distance to the centroids of, respectively, the emotional (A.8a) and cognitive (A.8b) LIWC dictionaries, computed with our main word-embedding model. We exclude from the wordclouds the terms that appear in the corresponding dictionary.

B.6 Polarization Score

We adapt Gennaro and Ash (2022)'s methodology to construct a measure of political polarization in language. Broadly speaking, we use the Word2Vec model to compute a "right-wing" alignment score that quantifies the similarity of each speech to right-aligned versus left-aligned political discourse. We then define our *polarization score* as the absolute deviation from the median of this alignment score. This measure captures how much a speech departs from the median rhetorical position, whether toward the left or the right of the ideological spectrum.

Because no categorized dictionaries of politically charged language exist for the French Third Republic, we construct them from our own data, using the speeches of left- and right-aligned parliamentarians. Moreover, since polarized language likely varies over time and across topics, we allow these dictionaries to be year- and topic-specific. We assign speeches to a single topic based on the topic with the highest weight (see Section B.4 for details on the topic analysis).

First, for each year-topic group g, we select speeches by MPs with political affiliation "left" or "right", excluding the president of the chamber and members of government. Second, we compute the centroid vector of the left-wing and right-wing dictionaries for each group g in our word embedding, as the inverse-frequency-weighted average of the individual word vectors. Let $\mathbf{L}_g$ and $\mathbf{R}_g$ denote the inverse-frequency-weighted averages of word vectors appearing in left- and right-wing speeches within group g. Using these centroids, we compute the right-wing alignment score of document $\mathcal{D}_i$ in group g as:

$$rscore_i = \frac{\text{sim}(\vec{d_i}, \mathbf{R}_g) + 1}{\text{sim}(\vec{d_i}, \mathbf{L}_g) + 1}, \tag{A.3}$$

where sim denotes cosine similarity. The polarization score is then defined as the absolute deviation of this alignment from the group median:

$$pscore_i \;=\; \Big|\, rscore_i - \mathrm{median}(rscore) \,\Big|. \tag{A.4}$$



Left-wing Lexicon Right-wing Lexicon

Year 1909, Topic 4: Schooling



Left-wing Lexicon Right-wing Lexicon

Year 1881, Topic 8: Labor



Left-wing Lexicon Right-wing Lexicon

Year 1894, Topic 11: Military Affairs

Figure A.9: Word Clouds for Left- and Right-Wing Lexicons across Selected Years and Topics

Note: Note: The figure shows word clouds of the 100 words in the corpus with the smallest cosine distance to the centroids of left- and right-wing language in some selected years and topics, computed using our main word embedding model. The top panel shows the word clouds for Topic "Schooling" in year 1909; the middle panel shows the word clouds for Topic "Labor" in year 1881; finally, the bottom panel shows the word clouds for Topic "Military Affairs" in 1894.

References

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research 3*, 993–1022.

Bonnard, R. (1926). *Les règlements des assemblées législatives de la France depuis 1789: notices historiques et textes*. Paris: Société anonyme du Recueil Sirey.

Gennaro, G. and E. Ash (2022). Emotion and Reason in Political Language. *The Economic Journal 132*(643), 1037–1059.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as Data. *Journal of Economic Literature 57*(3), 535–574.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica 87*(4), 1307–1340.

Griffiths, T. L. and M. Steyvers (2004). Finding Scientific Topics. *Proceedings of the National Academy of Sciences 101*(suppl_1), 5228–5235.

Grün, B. and K. Hornik (2011). topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software 40*(13), 1–30.

Hansen, S., M. McMahon, and A. Prat (2018). Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics 133*(2), 801–870.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean (2013). Distributed Representations of Words and Phrases and Their Compositionality. *arXiv preprint arXiv:1310.4546*.

Porter, M. (2009). The English (Porter2) Stemming Algorithm. http://snowball.tartarus.org/algorithms/english/stemmer.html.

Tausczik, Y. R. and J. W. Pennebaker (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology 29*(1), 24–54.