
WORKING PAPER
N. 221
MAY 2024

Economic Shocks and Assimilation Policies: Phylloxera and Educational Expansion in French Algeria

Mara P. Squicciarini

Gianandrea Lanzara

Sara Lazzaroni

Paolo Masella

This Paper can be downloaded without charge from The Social Science
Research Network Electronic Paper Collection:
<https://ssrn.com/abstract=4816096>



**Università
Bocconi**

BAFFI
Centre on Economics,
Finance and Regulation

Economic Shocks and Assimilation Policies: Phylloxera and Educational Expansion in French Algeria

Gianandrea Lanzara

Bocconi University

Sara Lazzaroni

University of Bologna

Paolo Masella

University of Bologna

Mara P. Squicciarini

Bocconi University *

May 2, 2024

Abstract

This paper studies the impact of a shock, the phylloxera crisis in 19th-century France, on assimilation policies towards the native population of French Algeria. In particular, assembling a novel dataset on French MPs and their parliamentary speeches, we find that MPs coming from areas hit harder by the phylloxera, were more likely to: i) use keywords related to the impact of phylloxera and wine production; ii) express greater interest and support towards policies aimed at educating the native population. The latter pattern becomes visible approximately ten years after the phylloxera crisis, consistent with the view that organizing production requires time.

JEL-Classification: J15, Z10, I25.

Keywords: assimilation, minorities, education, 19th-century France.

*We thank Caterina Alfonzo and Matteo Pograxha for excellent assistance throughout the construction of the data. Financial support from the PRIN Grant # PRIN 2017ATLJHB is gratefully acknowledged.

1 Introduction

Throughout history, the occurrence of negative shocks (e.g., economic shocks, natural disasters, wars, and pandemics) has represented a major challenge to societies, often having an impact on policies and attitudes towards minorities and their assimilation. A large literature in economics and political sciences has shown that these shocks may lead to scapegoating behavior and discrimination against minorities (Becker and Pascali, 2019; Anderson et al., 2020; Doerr et al., 2022; Anderson et al., 2017; Fouka, 2019; Jedwab et al., 2019; Abraham et al., 2024). However, there is scant empirical evidence on: i) whether (and under which conditions) negative shocks may have positive externalities and foster assimilation policies; and ii) how this happens.

In this paper, we exploit an exemplary historical setting, i.e., the French phylloxera crisis in late 19th-century France and its impact on educational policies towards Algerians. Phylloxera is an insect from the aphid family that attacks grapevines, causing their decline and death. Its spread in France from 1863 represented an extremely negative shock to the French economy, which at the time was predominantly agricultural. In particular, wine production made up a significant part of agricultural output and employment. To compensate for the decline in domestic production and the resulting income losses, French families involved in the wine industry emigrated to Algeria, then a French colony, where some recent technological innovations had made wine production feasible in spite of the high temperatures. The economic development of Algeria required the training of local labor capable of integrating into the economic activities of French colonists. From 1890, educational policies aimed at the native population became an important discussion topic in the French Parliament, with the goal of promoting knowledge of French and the development of manual skills useful for agricultural work.

To carry out the empirical analysis, we assembled a novel dataset from historical sources, which includes information on all members of the French Parliament (MPs) and the texts of the Parliamentary debates during the period of the Third Republic (1870-1942). In particular, we first downloaded from official websites the biographical information of all MPs elected. For each MP, we observe the date and place of birth, as well as the electoral district, the election date, the term of office for each mandate (an MP may be in charge for several mandates), and a one-page biographical note. We then matched biographies to speakers in our dataset of Parliamentary speeches. This dataset provides details on the speeches of each MP for a total of 1,463,560 speeches. Finally, we manually digitized the data in Galet (1957) to recover measures of the phylloxera shock at the department level.

In order to measure MPs' attitudes towards the native population of Algeria, we use two

distinct yet complementary approaches to text analysis: dictionary-based methods and topic analysis. We find that MPs from departments affected by phylloxera use keywords related to the impact of phylloxera and wine production more frequently following the phylloxera shock (compared to other MPs). We also find that MPs from departments affected by phylloxera express greater interest and support towards policies aimed at educating and training the native population compared to unaffected department. This pattern becomes visible approximately ten years after the phylloxera crisis, consistent with the view that relocating families engaged in wine production to Algeria and organizing production in a new environment are processes that require time.

This paper contributes to two strand of the literature. First, it relates to those works analyzing the impact of a negative shock on minorities' assimilation. Gould and Klor (2016) study the assimilation behavior of Muslim immigrants in the U.S. after the 9/11 attacks caused a spike in anti-Muslim hate crimes. They find that in states with a higher incidence of hate crimes Muslims decreased intermarriages, increased fertility, and reduced female labor force participation and their proficiency in English. By contrast, Fouka (2019) investigates the effect of anti-Germanism in the United States during WorldWar I and show that the German minority increased its assimilation by changing naming patterns towards Americanization and by filing more petitions for naturalization. We provide novel empirical evidence that a shock may positively affect the assimilation policies towards minorities when the latter become complementary to the well-functioning of the economy. In the Algerian case, policies of educational expansion were put into place to favor natives' accumulation of human capital.

Then, it contributes to the literature on the effects of phylloxera. Banerjee et al. (2010) study its impact on adult height, health, and life expectancy of children born in phylloxera-affected departments during the crisis. Bignon et al. (2017) investigate the effect of the phylloxera shock on crime. This paper takes a novel angle and shows the impact of the agricultural (and economic) shock on the educational policies towards the Algerian population.

2 Historical background

2.1 The phylloxera crisis and the wine industry in Algeria

The phylloxera insect, originating from America, was first detected in French vineyards near the Rhine estuary in 1863, having been transported to Europe on commercial vessels. The disease, whose origin was initially unknown, attacked the roots of vines, impairing their ability to produce grapes and ultimately leading to the plants' death. Subsequently, the

pest spread throughout France, inflicting substantial damage on both vineyards and the broader wine production sector. It was first observed affecting the roots of vines in the Gard department, in the South of France, in 1868, yet it was not officially identified as the cause of the disease until 1875. During the 1870s, the insect reached the departments of Gard, Hérault, and Vaucluse. By the late 1880s, phylloxera had invaded 37 of the 40 departments where, according to Banerjee et al. (2010), wine constituted more than 15% of agricultural production.

In the 19th century, the wine industry represented roughly one-sixth of the agricultural production value in France and served as a vital source of livelihood for numerous families. The diffusion of phylloxera destroyed one-third of the vineyards in the country, resulting in a dramatic decline in wine production by approximately 70% from 1875 to 1889, as documented by Meloni and Swinnen (2014). Bignon et al. (2017) further estimate that wine production in wine-intensive departments decreased by almost 50% during the 1850–1905 time period. However, the French demand for wine was not falling, so it was crucial to find new suppliers to keep up with it (Isnard and Labadie, 1959).

Moving wine production to Algeria had two main advantages. First, it offered French consumers an alternative good wine against the recent practice of altering French wine by adding a cheaper one to the original to meet the demand, thus allowing to maintain quality standards. Second, it prevented imports from Italy and Spain, France’s direct competitors. As a result, settlers switched from wheat to wine: the area of European vineyards rose from 15,000 ha. in 1878 to 110,000 in 1890 (Ageron, 1991) and by 1914, vineyards accounted for 44% of the European real property, with wine representing one-third of the value of Algerian exports (Ruedy, 2005).

2.2 Education policies in Algeria and the integration of the indigenous people

The replacement of wheat with wine increased dramatically the demand for labor as “a hectare of vineyard require[d] eighty days’ work per year, eight times as much as wheat” (Isnard and Labadie, 1959, p.73). This led the settlers to increasingly rely upon the mass of rural natives, which had to quickly “learn the techniques of grape cultivation and wine-making” (Isnard and Labadie, 1959, p.73). Hence, the need of very specific education policies that could serve two main purposes: (i) acquire proficiency in the French language and writing to enhance interactions and connections with the settlers, and (ii) teaching to the natives skills they can use and pass on, especially those necessary for the agricultural work. This was largely discussed in parliamentary debates of the time.

As an example, let us quote a speech by Antoine Rozet. Antoin Rozet was elected in the department of *Haute-Marne*, a wine-growing department, in the North-East of the country, affected by phylloxera. On December 11th, 1896, he delivered the following speech at the *Chambre des Deputées*.

In his deposition before the Senate's committee of eighteen in 1891, the distinguished rector of the Academy of Algiers [...], the honorable Mr. Jeanmaire, made a deposition in which he said: "One essential thing is to find a way to make French education useful to the natives; until now, we have placed ourselves too much from the French point of view; we must succeed in interesting the native in school by teaching him manual trades that he could teach in turn, such as locksmithing, carpentry, masonry. We must even more tend towards agricultural education by organizing agricultural schools attached to French schools, where simple and practical instruction would be given. The initiative should be given to the governor-general. [...] Agricultural work in the garden or in the demonstration field of the school will not truly be the apprenticeship of the cultivator's trade, but an application of the elementary theoretical notions that should be given to students on the cultivation of cereals, vegetables, fruit trees, and vines. The teacher will strive to give students a taste for gardening, to spread among the natives the vegetables they do not know and whose acclimatization and cultivation will contribute to improving their diet. If he could propagate in Kabylie, for example, the planting of fruit trees from France, the use of potatoes and chestnuts, which thrive very well there, it would be a great service that he would have rendered to the indigenous population. The chestnut wood, which would be used for making casks increasingly necessary in Algeria due to the expansion of wine production, would become a wealth in itself". [...] Therefore, it is essential that the natives learn French to strengthen various relationships with the settlers. In my opinion, the limited education for the natives should consist mainly of teaching French and writing, with a bit of arithmetic, and also a lot of manual and agricultural work.

Similarly, "Colons, who had always been critical of the attempt to create a special school system for natives who did not want it, were quick, after 1870, to dismantle the Arab-French system. They claimed that Muslims were now free to attend the French schools [...] They insisted that whatever education was offered should be vocationally oriented, focus on the needs of the agricultural labor market, and avoid academic subjects that might tend to politicize the natives. While there were several versions of the native school curriculum

from the 1880s, all courses of study shared an emphasis on the practical rather than the theoretical, and all were taught almost exclusively in French.” (Ruedy, 2005, p. 104-5)

3 Data

3.0.1 Historical and institutional background

The history of France from 1830 to 1870 is traditionally divided into three sub-periods: the *Monarchie de Juillet* (July Monarchy, 1830-1848), the *Seconde République* (Second Republic, 1848-1852), and the *Second Empire* (Second Empire, 1852-1870).

Table 1: Legend

Historical Periods:

<i>Monarchie de Juillet</i>	MJ
<i>Seconde République</i>	SR
<i>Second Empire</i>	SE
<i>Troisième République</i>	TE

Newspapers:

<i>Le Moniteur Universel</i>	MU
<i>Journal officiel de la République française</i>	JORF

3.1 1830-1870

The *Monarchie de Juillet* began on July 30, 1830, with the accession of Louis Philippe of Orleans to the throne. The constitutional form was a constitutional monarchy. The hereditary principle was replaced by the principle of popular sovereignty. The parliament was bicameral, comprising the *Chambre des Députés* and the *Chambre des Pairs*. The *Chambre des Députés* was elected through limited suffrage—restricted to men over 25 who paid a minimum tax threshold. The king appointed the members of the *Chambre des Pairs*.

The *Seconde République* began on February 23, 1848, following the overthrow of Louis Philippe during the wave of European revolutions that swept through France as well. The constitutional form was a parliamentary republic. The parliament was composed of a single chamber, the *Assemblée Nationale*, elected by direct universal suffrage by citizens over the age of 21. Louis Bonaparte was elected President of the Republic.

In December 1851, Louis Bonaparte successfully carried out a coup d’état, leading to the establishment of the *Second Empire* in January 1852 and the promulgation of a new constitution. The constitutional form was still republican although with strong authoritarian

traits, with powers concentrated in the figure of Louis Bonaparte, who was proclaimed emperor in December 1852 under the name Napoleon III. The parliament was composed of two chambers: the *Corps législatif*, elected by universal male suffrage, and the *Sénat*, which included members from the clergy, the military, and other imperial appointees.

3.1.1 Data sources

Speeches Throughout this period, the stenographic transcripts of the parliamentary debates are found in *Le Moniteur Universel* (MU), a newspaper founded during the French Revolution with the aim of publishing the debates of the revolutionary National Assembly. In 1834, the newspaper received the formal assignment of publishing the parliamentary transcripts. During this period, it served as the official gazette of France, providing coverage of the approved laws and other public acts. Starting from January 1st, 1869, Napoleon III replaced it with the *Journal officiel de l'Empire français* (later to become *Journal officiel de la République française*, but the MU kept publishing the parliamentary transcripts throughout 1869 and 1970. It then survived as an independent newspaper until 1901.

We downloaded all issues of the MU from 1830 to 1870, in JPEG format, from website Retronews (retronews.fr), the press website of the French National Library (see Figure ??).

Biographies The websites of both the *Assemblée Nationale* and the *Sénat* provide, on dedicated portals, information on the personal and political biographies of the parliamentarians throughout this period. See websites www2.assemblee-nationale.fr/sycomore/recherche for the *Assemblée Nationale* and www.senat.fr/senateurs-2nd-empire/senat1.html for the senators of the Second Empire.

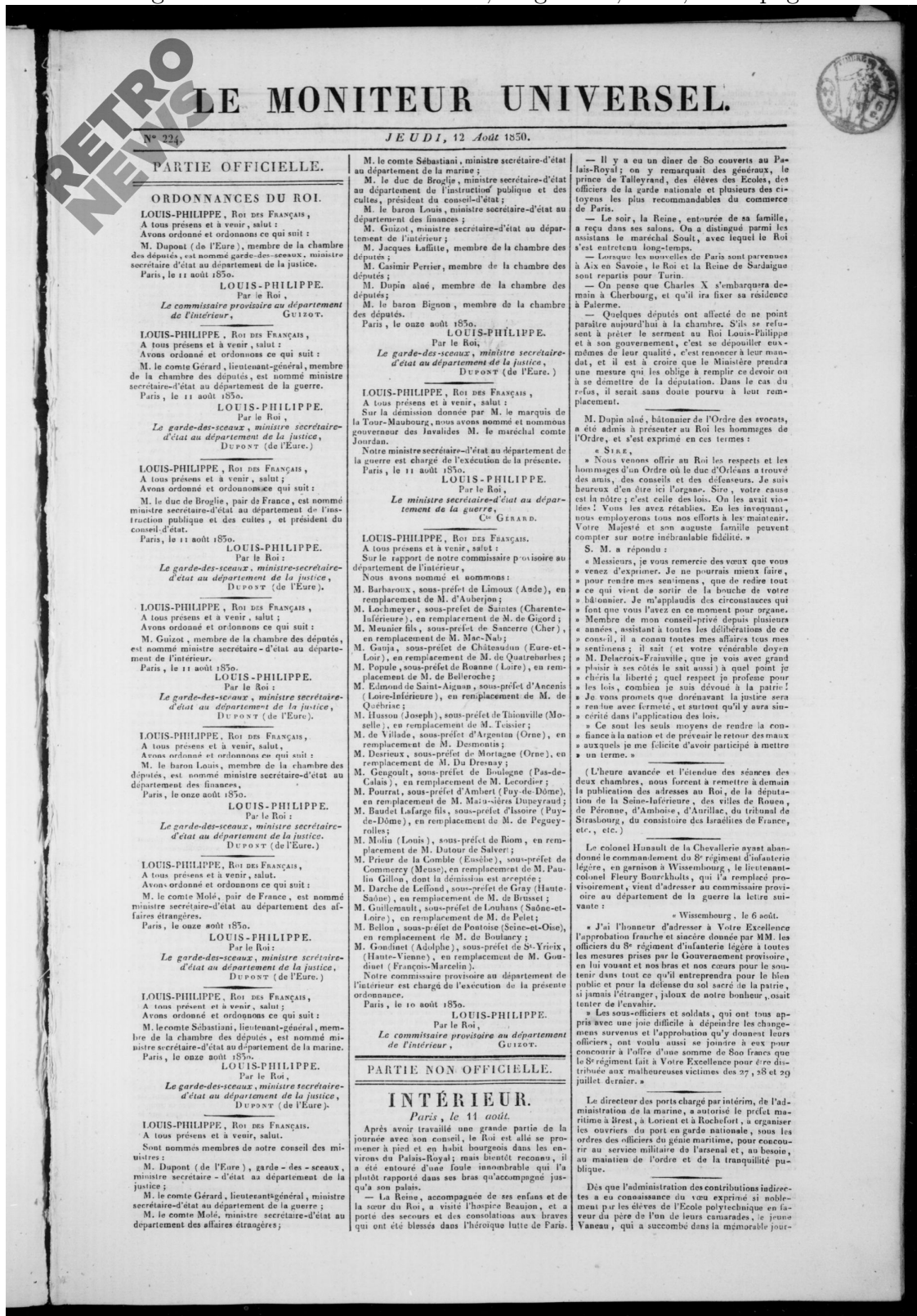
3.1.2 Speeches

3.1.2.1 OCR

Because the text files provided by Retronews were of insufficient quality to be used in the analysis, we opted to download all pages in JPEG format and leverage Google Vision to convert each newspaper issue into a single text file. The layout of the MU featured multiple columns, and Optical Character Recognition (OCR) algorithms often struggle to accurately read column breaks. To address this, we first used Detectron2, Facebook's library for image detection, to split each page into its respective columns. Then, we processed each column separately through the OCR algorithm.

We can outline the key steps of the procedure as follows:

Figure 1: *Le Moniteur universel*, August 12, 1930, front page



1. annotate a sample of pages from each period
2. train an image detection model on the annotated sample of pages
3. use the model to predict and crop all other pages from each issue of *Le Moniteur Universel* into individual columns
4. send each column to Google Vision for OCR
5. paste all columns for each issue into a single text file

We executed this procedure separately for the *Monarchie de Juillet*, the *Seconde Republique*, and the *Second Empire*, to account for variations in the page structure across periods. For step (1), we used Label Studio (labelstud.io) to manually delineate columns for thirty randomly selected issues – ten issues for each of the three historical sub-periods. Because the number of pages varies from issue to issue, we ended up with 76 pages for the *Monarchie de Juillet*, 100 pages for the *Seconde Republique*, and 46 pages for the *Second Empire*.

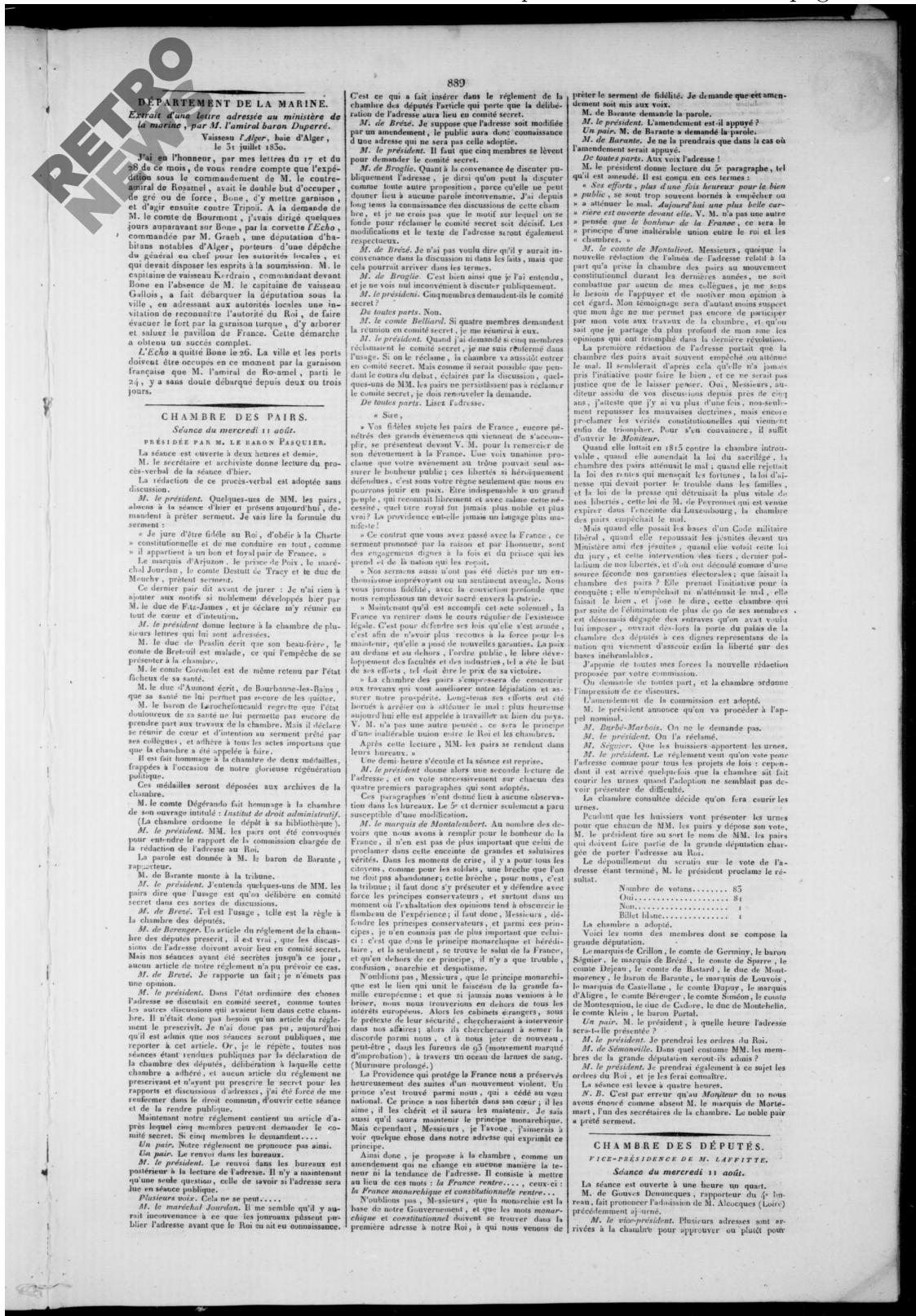
For the *Monarchie de Juillet*, we annotated a single “column” category. Because during this period the text tends to be reported without interruptions, columns always span the page from top to bottom. For the *Seconde Republique*, we also annotated a single “column” category. In this case, however, columns are allowed to have breaks in between, because tables occasionally appear during this period, often horizontally spanning multiple columns. At the same time, tables are too unfrequent for the algorithm to accurately detect them. Finally, for the *Second Empire* we annotated three different categories: “column”, “table”, and “dead zone”. In this period, since tables are more frequent, the algorithm does a good job at detecting them.

3.1.2.2 Extracting debates

The OCR procedure delivers raw text files for each issue of the MU. As the MU contained a variety of materials beyond those related to parliamentary activities, our primary task is to extract the section of the MU containing the parliamentary debates.

During the MJ, there were two branches of Parliament: *Chambre des Deputés* and *Chambre des Pairs*. During the SR, there was only one branch, the *Assemblée Nationale*. During the SE, there were again two branches: *Chambre des Deputés* and *Sénat*. For each of branch-period, we devise a procedure to locate and extract the debate.

Figure 2: *Le Moniteur universel*, August 12, 1930, page 3. The headers for the debate sections of the *Chambre des Pairs* and the *Chambre des députés* are visible on the page



3.1.2.2.1 Monarchie de Juillet

Beginning of the debate The parliamentary debates of each branch are located under the headers “CHAMBRE DES DÉPUTÉS” and “CHAMBRE DES PAIRS” (see Figure 2). Consequently, we build a regex pattern to identify this string in the text.

End of the debate All parliamentary sessions end with the phrase “the sitting is adjourned”, “la séance est levée” in French. Therefore, we use regex variations of this string to capture the end of the debate.

3.1.2.2.2 Seconde République

Beginning of the debate The parliamentary debates are located under the header “ASSEMBLÉE NATIONALE” (see Figure 3). Consequently, we build a regex pattern to identify this string in the text. Our search may return multiple matches, or a unique yet incorrect match. To make sure that we only retain the correct match, we implement an additional check. We observe that in the opening lines of the debate section: (i) an uppercase header announces the name of the president chairing the session; (ii) the date of the session is reported; (iii) a short summary section precedes the actual transcript. Therefore, we only retain the match when regex variations of “PRESIDENCE”, “*Séance du* [date]”, and “SOMMAIRE” are found in the subsequent five lines. In case these checks still leave us with multiple potential matches, we grab the first one.

End of the debate All parliamentary sessions end with the phrase “the sitting is adjourned”, “la séance est levée” in French. Therefore, we use regex variations of this string to capture the end of the debate.

3.1.2.2.3 Second Empire, after 1861

Beginning of the debate The parliamentary debates are located under the headers “CORPS LÉGISLATIF” and “SÉNAT” (see Figure 4 for the *Corps législatif*). Consequently, we build branch-specific regex patterns to identify these strings in the text. We first use a “strict” pattern that only allows for minimal variations. If no result is found, we then use a more flexible pattern. This strategy helps minimize errors, especially for the Senate pattern, which is very short. Our search may return multiple matches, or a unique yet incorrect match. To make sure that we only retain the correct match, we implement three additional checks. First, we observe that in the opening lines of the debate section: (i) an uppercase header announces the name of the president chairing the session; (ii) the date of the session

Figure 3: *Le Moniteur universel*, May 5, 1848, page 2. The headers for the debate section of the *Assemblée nationale* is visible on the page

946

Il est également indispensable que les bordures ou listels soient variés pour chaque métal.

Les couronnes ou encadrements du revers devront, en outre, affecter dans la forme ou les dispositions des feuilles et des branches, des différences pour les trois métaux.

Les couronnes ne perdront pas de vue que la gravure des coins pour les monnaies diffère essentiellement de celle des coins destinés à frapper des médailles, et que celles-ci, devant présenter un grand relief, ne peuvent recevoir une bonne exécution que par plusieurs coups de balancier, tandis que les pièces de monnaie ne reçoivent qu'un seul coup.

On doit leur faire observer encore que les pièces de monnaie, étant ordinairement composées en plaques, doivent se débiter des fontaines qui ne seraient d'aucun élan à l'empreinte, si elles présentaient des reliefs aussi saillants que ceux des médailles. Dans les monnaies, ces reliefs doivent être classés au point qu'une règle puisse se reposer sur les bordes.

Il leur est recommandé de disposer les matrices de manière à préserver les bords des pièces par un liseré et, en outre, afin de préserver l'empreinte de l'alésage qui arrive promptement, lorsque les bordes de la pièce s'écartent d'une saillie tranchante.

Les ayants aux prix font parvenir leurs poinçons, matrices, coins, etc., sous cachet, à la commission des monnaies, à Paris, dans le délai fixé par le programme de concours.

Il leur sera donné un récépissé portant la date du dépôt et le numéro d'enregistrement.

A l'expiration du terme assigné au concours, la commission des monnaies informera les concurrents du jour où ils devront s'assembler pour désigner les sept membres du jury dont le choix leur est délégué aux termes de l'art. 4 du décret n° 2, ainsi que du jour fixé pour le jugement des épreuves.

La commission des monnaies :
Le président,
F. LAURE.

SITUATION DE LA BANQUE DE FRANCE.

Le jeudi 3 mai 1848, au soir.

ACTIF.		PASSIF.	
Argent monnayé et lingots	56,049,879 57	Capital	67,900,000 »
Numéraire dans les caisses	54,288,262 50	Réserves	14,000,000 »
Effets en circulation	18,722,973 47	Id. des caisses	4,000,000 »
Portefeuille de Paris, dont 35,875,000 fr.		Billets au porteur en circulation	293,151,800 »
Et c. provenant des coupures	166,858,588 64	Id. des caisses	14,416,500 »
Portefeuille des caisses, effets sur place, etc.	77,153,026 18	Id. à ordre	1,499,123 30
Avances sur lingots et monnaies	6,295,714 »	Comptes courants de trésor, créditeurs	22,024,718 35
Avances sur effets publics français	10,669,962 80	Comptes courants de débiteurs	16,465,517 51
Rentes de la réserve, pour leurs intérêts et amortissement	14,418,500 »	Réceptifs payables à vue	1,371,590 »
Rentes de la caisse	10,044,000 »	Remboursement du dernier semestre	728,032 87
Rentes, fonds disponibles	11,660,197 89	Dividendes à payer	214,228 23
Hôtel et mobilier de la banque	4,000,000 »	Escomptes, intérêts divers et dépenses précomptées	3,386,369 19
Intérêt dans le comptoir d'Alger	1,000,000 »	Comptoir d'Alger, somme non encore employée en bons du trésor	1,121,477 62
Intérêt dans le comptoir national d'escompte	300,000 »	Divers	195,189 27
Avance à l'Etat au titre du trésor	23,579,489 39	Travaux des comptoirs à payer	486,367,909 69
Billets à encaisser provenant de la vente de rentes à l'Etat	81,524 92		
Depenses d'administration	511,911 51		
Divers	729,846 82		
Total	50,000,000 »		
Republique	486,367,909 69		

Certifié :
Le gouverneur de la Banque de France,
V. ANGOULT.

PARIS DIVERS. — PARIS. — Avant-hier deux ouvriers, un nommé Goussier et Picard, étaient en train de poser un chapeau au dessus de la porte du passage Pluvinet, dominant sur la rue de Lorient, pour la descente des eaux pluviales. Ils s'élevèrent l'un sur un chevron qui existait là depuis plus de trente ans, lorsque tout à coup les ferretiers qui maintenaient le chevron contre le mur se retirèrent, et il se fléchit sous le poids des deux ouvriers qui ont été entraînés.

Goussier est tombé sur le pavé, où il s'est fracturé la tête; il a été porté à l'hospice Saint-Louis dans un état désespéré. Quant à Picard, il a eu plus de chance au moment où il se relevait à terre, un de ses canons se cassa par ses vêtements et amorta la chute de manière à ce qu'il en ait été quitte pour une simple contusion.

DEPARTEMENTS. — On lit dans le *Mémorial de la République* : La commune de Narp a planté, le 21, son arbre de la Liberté. 31 Devises, dessinées, après avoir bûché le drapeau, à l'adresse au peuple un dieux chère. Un bouquet à l'adresse au peuple un dieux chère.

LE FAUCON. — On lit dans le *Mémorial de la République* : Le citoyen D. Olivier est parti hier, à trois heures, pour se rendre à Paris. Au moment de monter en voiture, il a remis autour de lui le poste d'honneur de la garde nationale à la préfecture, et dans une allocution vive et chaleureuse, il a remercié la garde nationale de Marseille des sentiments et de la confiance qu'elle a eue de donner des preuves de sa reorganisation.

Les gardes nationaux présents, rendant les interprètes des sentiments de la garde nationale tout entière, ont prié le citoyen D. Olivier de porter au Gouvernement l'assurance du concours de la garde nationale de Marseille, lui présentant toujours pour maintenir l'ordre et la tranquillité, si nécessaires à la consolidation de notre jeune République. On s'est séparé aux cris de *Vive la République!*

PUBLICATIONS. — Les plus gros livres ne sont pas les meilleurs; voici un opuscule de 70 pages qui en dit plus que bien des volumes : c'est la notice historique des *Assemblées législatives de France*, par Jules Lagrange, avocat à la Cour royale. Cette brochure pourra être de quelques secours aux personnes qui vont s'occuper de la nouvelle constitution. Elle est en vente chez Goussier, libraire, Palais-National, et chez Lagrange, rue de Bourgogne, n° 3, près la Chambre des députés. Prix : 50 centimes.

ASSEMBLÉE NATIONALE.

Séance du jeudi 3 mai.

PRÉSIDENCE DU CITOYEN ADRIEN DE FUYRAVAU, DOYEN D'ÂGE.

SOMMAIRE. — Formation des bureaux provisoires. — Proposition de Gouvernement provisoire. — Réunion dans les bureaux pour l'état des pouvoirs. — Reprise de la séance publique. — Vérification des pouvoirs. — Incident sur le serment à la République; les citoyens Hérisson (Bismont), Gilman, ministre de la justice, Frenon et Douau. — Suite de la vérification des pouvoirs. — Proposition de citoyens Berger, Gilman, Thomas, Germain, Serret, Dacot, Remondet, Hays, Depoix, Ollivier, Barbé, Vigorin et Ferrat. — Reprise de la vérification des pouvoirs. — Proposition de général Goussier; les citoyens Bédou-Lafitte et Flandin. — Proposition de la République devant le peuple. — Reprise de la séance. — Continuation de la vérification des pouvoirs.

A une heure et un quart une salle d'attente annonce l'arrivée des membres du Gouvernement provisoire dans le palais de l'Assemblée nationale.

Le citoyen Aubry de Fuyravau, doyen d'âge, monte au fauteuil de la présidence.

Les six plus jeunes membres de l'Assemblée, les citoyens Frenon, Lagrange, Avastin, Ferrat, et Ollivier, prennent place au bureau en qualité de secrétaires provisoires.

À une heure et demie les citoyens membres du Gouvernement provisoire et les citoyens ministres provisoires sont introduits dans la salle des délibérations, qui sont accueillis par les cris unanimes de *Vive la République!*

Les colonnes et lieutenants-colonnes des légions du département de la Seine, qui occupent les membres du Gouvernement provisoire, prennent place dans les couloirs à droite et à gauche de la tribune.

LE CITOYEN PRÉSIDENT. La parole est au citoyen Dupont de l'Éure, président du Gouvernement provisoire, ministre de l'Éducation nationale.

LE CITOYEN DÉPUTÉ DE LA SEINE. Citoyens représentants du peuple, le Gouvernement provisoire de la République vient solliciter devant vous, et rendre un honneur immense au pouvoir suprême dont vous êtes investis.

Elle est le peuple! soyez les bureaux dans la grande capitale, où votre présence fait naître un sentiment de bonheur et d'espérance qui ne sera pas trompé.

Disponibles de la souveraineté nationale, vous allez fonder nos institutions nouvelles sur les bases de la démocratie, et donner à la France la seule constitution qui puisse lui assurer son avenir, son indépendance, et son unité.

Mais, après avoir proclamé la grande loi qui va constituer définitivement le pays, comme nous, citoyens représentants, vous vous occupez de régler l'action possible et efficace du Gouvernement dans les rapports que la nécessité du travail établit entre tous les citoyens, et qui doivent servir pour tous les intérêts de la justice et de la fraternité. Applaudissons et nous en félicitons.

Enfin le moment est arrivé pour le Gouvernement provisoire de déposer entre vos mains le pouvoir ultime dont la résolution l'avait investi. Vous savez si, pour nous, cette détermination a été autre chose qu'une puissance morale, au milieu des circonstances difficiles que nous avons traversées.

Fidèles à notre origine et à nos convictions personnelles, nous nous sommes pas hâtés à proclamer la République nationale de Février. (Les cris de *Vive la République!* et *Vive le Gouvernement provisoire!*) et les applaudissements éclatent sur tous les bancs et dans les tribunes publiques.)

Aujourd'hui nous inaugurons les travaux de l'Assemblée nationale à ce cri qui doit toujours la valoir : *Vive la République!* (Les acclamations et les applaudissements recommencent avec une nouvelle énergie.)

LE CITOYEN GÉNÉRAL, membre du Gouvernement provisoire, ministre de la Justice. Citoyens représentants du peuple, au nom du Gouvernement provisoire, vos travaux sont ouverts, l'Assemblée nationale les commence aujourd'hui.

J'invie le citoyen président provisoire à vouloir bien entreprendre pour y procéder à la vérification des pouvoirs.

LE CITOYEN PRÉSIDENT. J'invie les représentants à se retirer dans leurs bureaux respectifs pour procéder à l'examen des pouvoirs.

Plusieurs membres. Comment chacun connaît-il son bureau?

LE CITOYEN PRÉSIDENT. La liste des représentants est affichée dans la salle des conférences et à la porte de chaque bureau; chacun pourra, par conséquent, connaître le bureau auquel il appartient.

LE CITOYEN GÉNÉRAL. La liste est dressée département par département et par ordre alphabétique.

LE CITOYEN PARLEMENTAIRE. A quelle heure la séance sera-t-elle reprise?

LE CITOYEN PRÉSIDENT. La séance sera reprise à trois heures.

(Les représentants se retirent dans leurs bureaux au cri de *Vive la République!*)

La séance est suspendue à deux heures moins vingt minutes. (La séance est reprise à trois heures et demie.)

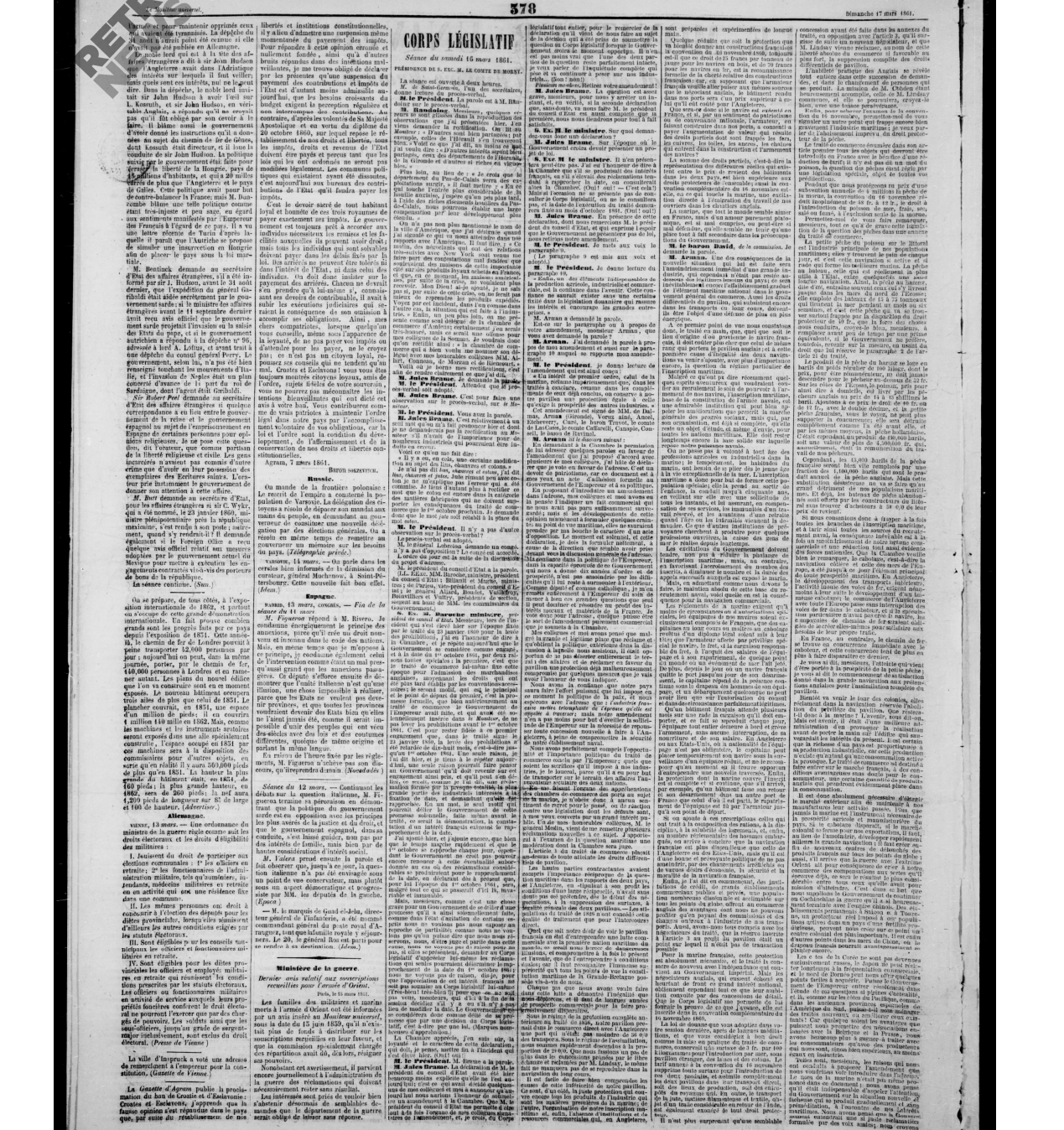
LE CITOYEN PRÉSIDENT. L'ordre du jour appelle la vérification des pouvoirs. Le rapporteur du 1^{er} bureau a la parole.

LE CITOYEN MARCHELAIN. J'ai l'honneur de soumettre à l'Assemblée le résultat des vérifications auxquelles le 1^{er} bureau vient de se livrer. Son examen a porté sur six députés-membres.

Ainsi Ont été élus les citoyens :
Fruichaut,
Sarran jeune,
Bayard,
Barbes,
André, Paris,
Marc-Saillier,
Joly fils.

Les opérations ont été trouvées complètes, régulières, pour tous. Seulement, un de ces représentants, le citoyen Bayard, 1^{er} bureau propose l'admission des citoyens :
Fruichaut,

Figure 4: *Le Moniteur universel*, March 17, 1861, page 2. The header for the debate section of the *Corps législatif* is visible on the page.



is reported; (iii) a short summary section precedes the actual transcript. Therefore, we only retain the match when regex variations of “PRESIDENCE”, “*Séance du [date]*” or “*Suite de la séance de [date]*”, and “SOMMAIRE” are found in the subsequent five lines. Second, because in these years the Senate pattern also appears in other section of the MU, we explicitly exclude these cases. Third, we exclude matches preceded by the names of foreign States or Capitals, because in these years the MU often reports summaries of the parliamentary debates of other countries. In case these checks still leave us with multiple potential matches, we additionally search the string “*compte rendu*” in the five lines following each candidate match. If not found, we grab the first match.

End of the debate All parliamentary sessions end with the phrase “the sitting is adjourned”, “la séance est levée” in French. Therefore, we use regex variations of this string to capture the end of the debate. As a sanity check before exporting the text file with the debate, we make sure that the debate contains the pattern “*M.*”, which stands for *Messieur* and always introduces a speaker.

3.1.2.3 Extracting speeches

Our next task is to split each debate file into individual speeches.

Beginning of the speeches As a preliminary step, we cut the short initial section summarizing the content of the parliamentary session. We use regex variations of the string “SOMMAIRE” to detect the beginning of the summary section. In case multiple matches are found, we grab the first one. We then look for regex variations of the phrase “la séance est ouverte ‘a...heures” or “le procès verbal est adopté”. Again, in case multiple matches are found we grab the first one. This is the start of the debate section. In case no match is found, we assume that debate starts from line one. In some cases, we fail to detect the start of the summary section. In this scenario, we start from the line that follows the string “PRESIDENCE”, in case it appears in the initial ten lines. Otherwise, we assume that the debate starts from line one.

Splitting pattern We then build a regex pattern to detect the end of a speech and the beginning of a new one. During the MJ and the SE, in its simplest form, a succession of speakers is identified by the presence of the following structure in the text :

...fin.

M. Untel. J’ai l’honneur...

The same is true for the SR, except that the abbreviation for *Messieur* was replaced with “*Le citoyen*” during this period. We will need to build a rather complex regex pattern in order to account for the many variations to this basic structure as well as the frequent OCR errors. We now describe the building blocks of our splitting pattern.

- **before_speaker** is a new line marker, potentially followed by one or two spaces.
- **before_M** matches “*S. Exc.*” the abbreviation for “*Son excellence*”, or His excellency, that some speakers bear before their name.
- **M** matches the pattern *M.*, which stands for *Messieur* and precedes all speaker names. We allow both for *M.* and for *M*, to account for OCR mistakes.
- **Le** matches variations of definite article “Le”. This pattern is used in conjunction with “*citoyen*” during the SR, and with “*président*”
- **after_M** matches zero to five “words” — more precisely, bundles of alphanumeric characters separated by spaces. This rather broad pattern encompasses such cases as *M. le duc de...*, and other situations where the speaker’s name is preceded by aristocratic or military titles.
- **titles** matches variations of *président*, *ministre*, and *rapporteur*. In each case, we allow these titles to be followed by zero to three words. In combination with `after_M`, this pattern allows us to match speakers such as, e.g., *M. le président* and *M. le ministre de la guerre*.
- **lastname** matches any word starting with an uppercase letter. **lastname2** is a laxer pattern that matches any sequence of nonspace characters.
- **other** matches additional text that occasionally appears after the last name, before the concluding period (see the example above). This text further qualifies the speaker or his posture and always follows a comma. For instance, we often encounter cases like: “M. Untel, rapporteur”. With some variations across years, we may also find such instances as, e.g., “M. Untel, de sa place” (i.e., from his seat), or “M. Untel, continuant” (i.e., continuing). Clearly, we must be very specific to minimize the probability that we match names occurring within other speeches.
- **dot** matches a dot. **dot2** matches a dot, a colon, or a semicolon. **dot3** matches a dot or a comma.

- **after_speaker** ensures that the preceding pattern is not followed by numbers, reducing the risk of matching cases where election outcomes (and the number of votes) are reported.

3.1.3 Biographies

3.1.4 Matching speeches and biographies

Our final task is to link parliamentarians who speak during a parliamentary session, i.e., “speakers”, to their respective entries in the dataset with the biographies, i.e., “MPs”.

For each parliamentary debate, we extract from the dataset of parliamentary speeches a list of speakers who took the floor. Then, we retrieve from the biographies the parliamentarians who were in office in that branch of parliament and on that date, using the start and end dates of their terms. The matching process compares the surnames (and, where available, the first names) of these two sets of observations.

We can outline the key steps of the procedure as follows:

1. we compute the matrix of Levenshtein distances from speakers’ lastnames to the MPs’ lastnames. Then, we associate each speaker to the MP with the lowest Levenshtein distance. In case of ties, we keep both matches.
2. for each matched speaker-MP pair from the first step, we calculate the matrix of Levenshtein distances for the other components of the name, i.e., first names. if only the speaker’s last name is available, this step has no effect. Otherwise, we match each speaker’s first name to the closest corresponding first name in the MP’s list. After matching, we remove that first name from the MP’s list to prevent double-counting. A first name is considered successfully matched if the Levenshtein distance is less than two. Finally, we retain the pair with largest number of successful first names matches. In case of ties, we randomly select one of the two matches.

3.2 1871-1942

3.2.1 Historical and institutional background

The Third French Republic lasted from the conclusion of the Franco-Prussian War in 1871, which ended with France’s defeat and the capture of Napoleon III, to the German invasion of 1940. The constitutional form was a parliamentary republic. From February 19, 1871, to December 31, 1875, the legislature was composed of a single chamber, the *Assemblée Nationale*, which was elected by universal male suffrage. The Assembly’s responsibilities

included signing a peace treaty with Prussia and drafting the forthcoming constitution. The constitutional laws, which took effect from January 1876, established a bicameral parliament: the *Chambre des Députés*, elected by universal male suffrage, and the *Sénat*, elected by a departmental college consisting of deputies and other local representatives.

3.2.2 Data sources

Speeches During the French Third Republic, the *Journal Officiel de la République Française*¹ (JORF) was responsible for publishing the stenographic reports of the parliamentary debates.

All the issues of the JORF from 1870 to 1942 can be freely downloaded from the Gallica website, the digital portal of the National Library of France (<https://gallica.bnf.fr>). The documents are available both in PDF and in text format (see Figure 5).

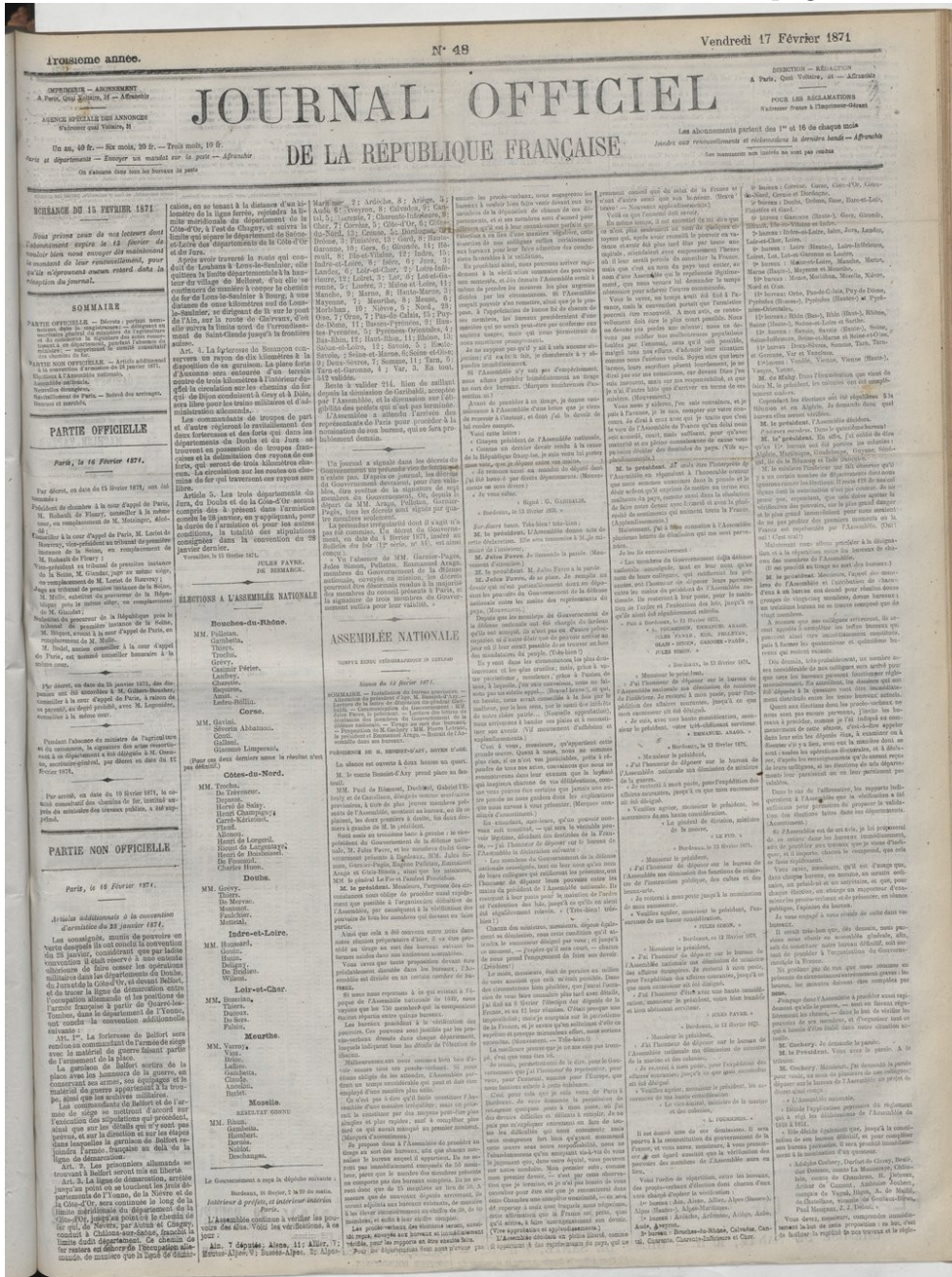
In the years from 1870 to 1880, it is necessary to download the entire JORF issue, while after 1880, the website provides a series of documents that exclusively contain the section of parliamentary debates. We downloaded all the documents in text format using the R software with packages `rvest` and `RSelenium`.

In some instances, multiple documents are associated with the same date. Between 1870 and 1880, this actually happens for most dates, and typically, these are multiple identical versions, of variable quality, with only one having an available text file. In the years following 1880, this happens rarely, but when it does, these are documents containing different material and need to be manually verified.

Biographies The websites of both the *Assemblée Nationale* and the *Sénat* provide, on dedicated portals, information on the personal and political biographies of the parliamentarians of the Third Republic. See websites www2.assemblee-nationale.fr/sycomore/recherche for the *Assemblée Nationale* and www.senat.fr/senateurs-3eme-republique/senat1.html for the senators of the Third Republic.

¹Napoleon III created the *Journal Officiel de la République Française* in 1868. A decree of November 5, 1870 granted this journal the monopoly to publish laws and regulations (Source: Wikipedia).

Figure 5: Journal officielle de la République Française, February 17, 1871, page 1. The header for the debate section of the Assemblée nationale is visible on the page



3.2.3 Speeches

3.2.3.1 Speeches: 1870-1880

3.2.3.1.1 Extracting debates

In this period, our starting point is the raw text files of the JORF. As the JORF contained a variety of materials beyond those related to parliamentary activities², our primary task is to extract the section of the JORF containing the parliamentary debates.

During the constitutional phase between 1870 and 1875, there was only one branch of Parliament, the *Assemblée Nationale*. Following the approval of the Third Republic's constitution in 1876, both the *Chambre des Députés* and the *Sénat* were established. For each of these three branches, we devise a procedure to locate and extract the debate. We can outline the key steps of the procedure as follows:

- come up with text patterns that unequivocally identify the beginning and the end of the debate
- devise variations of these patterns in the regex language to account for spelling mistakes in the raw text
- develop a series of checks to ensure that the matched text patterns do correspond to the desired outcome

Beginning of the debate From 1870 to 1875, the debat section appears under the header “ASSEMBLÉE NATIONALE”, while from 1876 to 1880 there are two sections that appear, respectively, under the headers “CHAMBRE DES DÉPUTÉS” and “SÉNAT”. Thus, we build year-specific regex patterns to locate these strings in the text. Because these headers also appear in the short summary section at the very beginning of the JORF, we make sure we cut off this section before we search the regex patterns.

Our search may return multiple matches, or a unique yet incorrect match. To make sure that we only retain the correct match, we implement two main additional checks. First, we check if a match is followed by additional patterns that strongly signal the presence of a parliamentary debate. Second, we exclude matches preceded by the names of foreign States or Capitals, because the JORF often reports summaries of the parliamentary debates of

²The first part, the official section, announced to the public new laws and regulations. The second part, in which the parliamentary debates are found, also featured sections dedicated to foreign policy news (*Nouvelles et correspondances étrangères*), cultural and social news (*Informations et Faits*), as well as materials related to the activities of other bodies such as the *Académie des Sciences* and the *Banque de France*.

other countries. In case these checks still leave us with multiple potential matches, we grab the first one.

End of the debate To capture the end of the debate section, we use three sets of patterns. First, all parliamentary sessions end with the phrase “the sitting is adjourned”, “*la séance est levée*” in French. Second, all parliamentary records are signed by the head of the stenographic service, “*le chef du service sténographique*” in French. Thus, we include variations of this phrase, as well as the year-specific names of the directors. Third, we include patterns to capture the headers of the sections that typically follow the parliamentary debates, in specific years, in the *Journal officiel*. As a sanity check before exporting the text file with the debate, we make sure that the debate contains the pattern “*M.*”, which stands for *Messieur* and always introduces a speaker.

3.2.3.1.2 Extracting speeches

Our next task is to split each debate file into individual speeches.

Beginning of the speeches As a preliminary step, we cut the short initial section summarizing the content of the parliamentary session. We assume the summary starts at line one since we used the beginning of the summary to extract the debate file. Following the summary, an uppercase header announces the name of president (or, at times, the vice-president) who is going to chair the session. Then, the opening phrase “*la séance est ouverte à...heures*” follows, meaning that the session started at a certain time. Thus, in order to match the end of the summary and the beginning of the debate, we use regex variations of PRESIDENCE and the opening phrase. Additionally, for certain years, it is useful to include the actual names of the President and the Vice-President.

Splitting pattern We then build a regex pattern to detect the end of a speech and the beginning of a new one. In its simplest form, a succession of speakers is identified by the presence of the following structure in the text :

...fin.

M. Untel. J’ai l’honneur...

However, we will need to build a rather complex regex pattern in order to account for the many variations to this basic structure as well as the frequent OCR errors. For instance, since the OCR struggles to correctly interpret line breaks, we do not use this marker. We now describe all the building blocks of our workhorse pattern.

- **before_speaker** matches all combinations of characters that may precede a speaker. These characters are the period (as in the example above), question mark and exclamation marks. We also account for some frequent OCR mistakes (e.g., 1 for !), as well as uppercase letters (excluding M), given that uppercase headers often introduce new sections of the debate. Note that `before_speaker` does not include the comma (although the OCR may occasionally interpret periods as commas); the reason is that the name of a speaker preceded by a comma is a string that may easily occur *within* the speech delivered by a parliamentarian.
- **M** matches the pattern *M.*, which stands for *Messieur* and precedes all speaker names. We allow both for *M.* and for *M*, to account for OCR mistakes.
- **after_M** matches zero to five “words” — more precisely, bundles of alphanumeric characters separated by spaces. This rather broad pattern encompasses such cases as *M. le duc de...*, and other situations where the speaker’s name is preceded by aristocratic or military titles.
- **titles** matches variations of *président*, *ministre*, and *rapporteur*. In each case, we allow these titles to be followed by zero to three words. In combination with `after_M`, this pattern allows us to match speakers such as, e.g., *M. le président* and *M. le ministre de la guerre*.
- **lastname** matches any word starting with an uppercase letter. We employ it as an alternative to the titles. Indeed, when a speaker is referred to by his title, the last name is not reported. In all other cases, while allowing for some flexibility with `after_M`, at least one word resembling a surname (i.e., starting with an uppercase letter) must appear not too far from `M`.
- **other** matches additional text that occasionally appears after the last name, before the concluding period (see the example above). This text further qualifies the speaker or his posture and always follows a comma. For instance, we often encounter cases like: “M. Untel, rapporteur”. With some variations across years, we may also find such instances as, e.g., “M. Untel, de sa place” (i.e., from his seat), or “M. Untel, continuant” (i.e., continuing). Clearly, we must be very specific to minimize the probability that we match names occurring within other speeches.
- **after_speaker** ensures that the preceding pattern is not followed by numbers, reducing the risk of matching cases where election outcomes (and the number of votes) are reported.

3.2.3.2 Speeches: 1881-1942

3.2.3.2.1 Extracting speeches

In this period, Gallica provides the parliamentary records for each branch of Parliament as individual documents (see Figure 6 for the *Chambre des Députés*). As a result, we are spared the trouble of extracting the debates from the JORF.

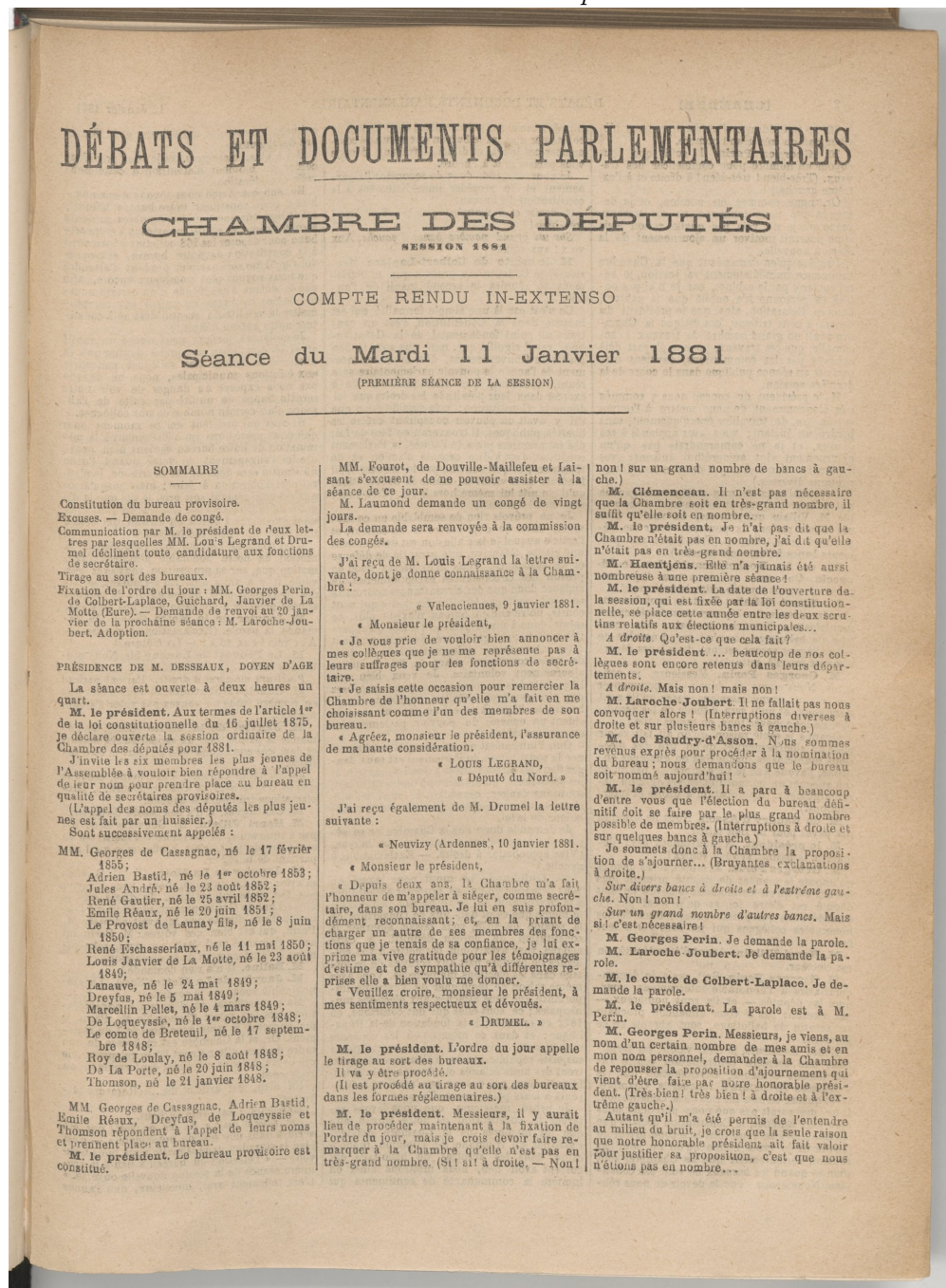
Beginning of the session During this period, especially from 1900 onward, the Parliament occasionally convened multiple times in a day (rarely more than twice). In such instances, a document will present the records of all the sessions, with each session introduced by a short summary. As a first step, we break the text into individual sessions using regex variations around the word “*SOMMAIRE*”. After excluding the initial 25 lines of introductory Gallica notes, which appear in the downloaded text file, we search for multiple sessions from the 125th line onward.

Beginning of the summary To cut the initial summary session, we use variations around string SOMMAIRE. In cases where this pattern is not detected (due, for instance, to OCR errors in the text), we assume the summary starts at line 4—this being the average starting line across the sessions where the pattern is found.

Beginning of the speeches Following the summary, an uppercase header announces the president (or, at times, the vice-president) who is going to chair the session. Then, the opening phrase “*la séance est ouverte à...heures*” follows, meaning that the session started at a certain time. Thus, in order to match the end of the summary and the beginning of the debate, we use regex variations of PRESIDENCE and the opening phrase. Additionally, for certain years, it is useful to include the actual names of the President and the Vice-President.

End of the speeches We still need to detect the end of the debate in order to separate the block of parliamentary speeches from the additional parliamentary material often included in the document. This material includes technical reports and, more prominently, the outcomes of roll call votes occurred during the session. To capture the end of the debate section, we use two sets of patterns. First, all parliamentary sessions end with the phrase “the sitting is adjourned”, “*la séance est levée*” in French. Second, all parliamentary records are signed by the head of the stenographic service, “*le chef du service sténographique*” in French. Thus, we include variations of this phrase, as well as the branch- and year-specific names of the directors. Third, after 1909, the written questions of the MPs systematically follow the

Figure 6: *Journal officiel de la République Française*, January 11, 1881, page 1. Parliamentary documents and debates for the *Chambre des Députés*.



DÉBATS ET DOCUMENTS PARLEMENTAIRES

CHAMBRE DES DÉPUTÉS

SESSION 1881

COMPTE RENDU IN-EXTENSO

Séance du Mardi 11 Janvier 1881

(PREMIÈRE SÉANCE DE LA SESSION)

SOMMAIRE

Constitution du bureau provisoire.

Excuses. — Demande de congé.

Communication par M. le président de r'œux lettres par lesquelles MM. Lou's Legrand et Drumel déclinent toute candidature aux fonctions de secrétaires.

Tirage au sort des bureaux.

Fixation de l'ordre du jour : MM. Georges Perin, de Colbert-Laplace, Guichard, Janvier de La Motte (Eure). — Demande de renvoi au 20 janvier de la prochaine séance : M. Laroche-Joubert. Adoption.

PRÉSIDENCE DE M. DESBREAUX, DOYEN D'ÂGE

La séance est ouverte à deux heures un quart.

M. le président. Aux termes de l'article 1^{er} de la loi constitutionnelle du 16 juillet 1875, je déclare ouverte la session ordinaire de la Chambre des députés pour 1881.

J'invite les six membres les plus jeunes de l'Assemblée à vouloir bien répondre à l'appel de leur nom pour prendre place au bureau en qualité de secrétaires provisoires.

(L'appel des noms des députés les plus jeunes est fait par un huissier.)

Sont successivement appelés :

- MM. Georges de Cassagnac, né le 17 février 1855;
- Adrien Bastid, né le 4^o octobre 1853;
- Jules André, né le 23 août 1852;
- René Gantier, né le 25 avril 1852;
- Emile Réaux, né le 20 juin 1851;
- Le Provost de Launay fils, né le 8 juin 1850;
- René Eschassériaux, né le 11 mai 1850;
- Louis Janvier de La Motte, né le 23 août 1849;
- Lanaave, né le 24 mai 1849;
- Dreyfus, né le 5 mai 1849;
- Marcellin Peltet, né le 4 mars 1849;
- De Loqueyssié, né le 1^{er} octobre 1848;
- Le comte de Breteuil, né le 17 septembre 1848;
- Roy de Loulay, né le 8 août 1848;
- De La Porte, né le 20 juin 1848;
- Thomson, né le 21 janvier 1848.

MM. Georges de Cassagnac, Adrien Bastid, Emile Réaux, Dreyfus, de Loqueyssié et Thomson répondent à l'appel de leurs noms et prennent place au bureau.

M. le président. Le bureau provisoire est constitué.

MM. Fourot, de Douville-Maillefeu et Laisant s'excusent de ne pouvoir assister à la séance de ce jour.

M. Laumond demande un congé de vingt jours.

La demande sera renvoyée à la commission des congés.

J'ai reçu de M. Louis Legrand la lettre suivante, dont je donne connaissance à la Chambre :

« Valenciennes, 9 janvier 1881.

« Monsieur le président,

« Je vous prie de vouloir bien annoncer à mes collègues que je ne me représente pas à leurs suffrages pour les fonctions de secrétaire.

« Je saisis cette occasion pour remercier la Chambre de l'honneur qu'elle m'a fait en me choisissant comme l'un des membres de son bureau.

« Agréez, monsieur le président, l'assurance de ma haute considération.

« LOUIS LEGRAND,

« Député du Nord. »

J'ai reçu également de M. Drumel la lettre suivante :

« Neuvizy (Ardennes), 10 janvier 1881.

« Monsieur le président,

« Depuis deux ans, la Chambre m'a fait l'honneur de m'appeler à siéger, comme secrétaire, dans son bureau. Je lui en suis profondément reconnaissant ; et, en la priant de charger un autre de ses membres des fonctions que je tenais de sa confiance, je lui exprime ma vive gratitude pour les témoignages d'estime et de sympathie qu'à différentes reprises elle a bien voulu me donner.

« Veuillez croire, monsieur le président, à mes sentiments respectueux et dévoués.

« DRUMEL. »

M. le président. L'ordre du jour appelle le tirage au sort des bureaux.

Il va y être procédé.

(Il est procédé au tirage au sort des bureaux dans les formes réglementaires.)

M. le président. Messieurs, il y aurait lieu de procéder maintenant à la fixation de l'ordre du jour, mais je crois devoir faire remarquer à la Chambre qu'elle n'est pas en très-grand nombre. (S'it sit à droite. — Non!

non! sur un grand nombre de bancs à gauche.)

M. Clémenceau. Il n'est pas nécessaire que la Chambre soit en très-grand nombre, il suffit qu'elle soit en nombre.

M. le président. Je n'ai pas dit que la Chambre n'était pas en nombre, j'ai dit qu'elle n'était pas en très-grand nombre.

M. Haentjens. Elle n'a jamais été aussi nombreuse à une première séance!

M. le président. La date de l'ouverture de la session, qui est fixée par la loi constitutionnelle, se place cette année entre les deux scrutins relatifs aux élections municipales...

A droite. Qu'est-ce que cela fait?

M. le président. ... beaucoup de nos collègues sont encore retenus dans leurs départements.

A droite. Mais non! mais non!

M. Laroche-Joubert. Il ne fallait pas nous convoquer alors! (Interruptions diverses à droite et sur plusieurs bancs à gauche.)

M. de Baudry-d'Asson. Nous sommes revenus exprès pour procéder à la nomination du bureau; nous demandons que le bureau soit nommé aujourd'hui!

M. le président. Il a paru à beaucoup d'entre vous que l'élection du bureau définitif doit se faire par le plus grand nombre possible de membres. (Interruptions à droite et sur quelques bancs à gauche.)

Je soumetts donc à la Chambre la proposition de s'ajourner... (Bruitantes exclamations à droite.)

Sur divers bancs à droite et à l'extrême gauche. Non! non!

Sur un grand nombre d'autres bancs. Mais s'il c'est nécessaire!

M. Georges Perin. Je demande la parole.

M. Laroche-Joubert. Je demande la parole.

M. le comte de Colbert-Laplace. Je demande la parole.

M. le président. La parole est à M. Perin.

M. Georges Perin. Messieurs, je viens, au nom d'un certain nombre de mes amis et en mon nom personnel, demander à la Chambre de repousser la proposition d'ajournement qui vient d'être faite par notre honorable président. (Très-bien! très bien! à droite et à l'extrême gauche.)

Autant qu'il m'a été permis de l'entendre au milieu du bruit, je crois que la seule raison que notre honorable président ait fait valoir pour justifier sa proposition, c'est que nous n'étions pas en nombre...

debate, thus we can include variations of the section’s header (*QUESTIONS ÉCRITES*). Finally, we include in the code a large number of exceptions to handle specific documents where we could not automatically detect the ending pattern.

Splitting pattern We then build a regex pattern to detect the end of a speech and the beginning of a new one. This pattern is identical to the one used for the period 1870-1880, and we refer the reader to Section 3.2.3.1.2 for details.

Exception for Senat in years 1922 to 1932 The Senate records from March 17th, 1922 to December 30th, 1932 have a different structure and we handle them in a different script. Specifically, the *sommaire* now follows, rather than precedes, the debate. Therefore, we look for the beginning of the debate right after the introductory Gallica notes. All the regex patterns are the same as for the general case.

3.2.4 Biographies

Chambre des députées The website of the *Assemblée Nationale* features a dedicated page for each member of the Chamber of Deputies. In addition to the name and surname, these pages provide their birth and death details, including dates and locations. Further, the pages list each term served in either the National Assembly or the Chamber of Deputies, specifying the start and end dates, the department of election, and the political affiliation of the deputy. Finally, these pages feature one or two biographical snippets sourced from two different references: “Dictionnaire des parlementaires français de 1789 à 1889” by Adolphe Robert and Gaston Cougny, and “Dictionnaire des parlementaires français de 1889 à 1940” by Jean Jolly.

The name of the parliamentarian is presented in the following format: the first names are separated by commas; after the last first name, the surname follows without a comma, e.g., “Georges, Eugène, Benjamin Clemenceau”. Thus, we construct a regex pattern to separate the first names and the surname. The pattern matches the first occurrence of the string “space plus an upper case letter” which does not follow a comma, along with all subsequent characters.

Sénat The website of the *Sénat* features a dedicated page for each senator. In addition to the name and surname, these pages provide their birth and death details, including dates and locations, their profession, and their department of election. Further, the pages list the dates of election, and the final date at which the senator was in office. Finally, these pages

feature one or two biographical snippets sourced from the same references as for the *Chambre des députées*.

The name of the parliamentarian is reported in the following format: the surname in uppercase letters followed by the first names, e.g., “CLEMENCEAU Georges”. Thus, we construct a regex pattern to separate the first name and the surname. The pattern matches all characters preceding a sequence of uppercase letters followed by a string indicating the first name, i.e., an uppercase letter followed by a sequence of lower case letter.

Also, we use the information on election dates to deduce the start and end dates of each term served by the senator. To assign a political affiliation to each senator, we build a regex pattern for each political group existed during the TR, and we search these patterns into the biographies.

3.2.5 Matching speeches and biographies

Our final task is to link parliamentarians who speak during a parliamentary session, i.e., “speakers”, to their respective entries in the dataset with the biographies, i.e., “MPs”.

For each parliamentary debate, we extract from the dataset of parliamentary speeches a list of speakers who took the floor. Then, we retrieve from the biographies the parliamentarians who were in office in that branch of parliament and on that date, using the start and end dates of their terms. The matching process compares the surnames (and, where available, the first names) of these two sets of observations.

We can outline the key steps of the procedure as follows:

1. we compute the matrix of Levenshtein distances from speakers’ lastnames to the MPs’ lastnames. Then, we associate each speaker to the MP with the lowest Levenshtein distance. In case of ties, we keep both matches.
2. for each matched speaker-MP pair from the first step, we calculate the matrix of Levenshtein distances for the other components of the name, i.e., first names. if only the speaker’s last name is available, this step has no effect. Otherwise, we match each speaker’s first name to the closest corresponding first name in the MP’s list. After matching, we remove that first name from the MP’s list to prevent double-counting. A first name is considered successfully matched if the Levenshtein distance is less than two. Finally, we retain the pair with largest number of successful first names matches. In case of ties, we randomly select one of the two matches.

4 Empirical Patterns

In this section, we present some empirical patterns from our textual analysis of parliamentary speeches, aimed at assessing the impact of phylloxera on integration policies directed towards the indigenous populations of Algeria. Our methodology employs two distinct yet complementary approaches to text analysis: dictionary-based methods and topic analysis. For an overview of text analysis techniques, see Gentzkow et al. (2019).

The dictionary-based approach searches the text for specific keywords that may indicate the subject matter or attitudes expressed in the text. While straightforward to apply and interpret, this approach has two main limitations: it is dependent on the researcher’s choice of keywords, and by neglecting the context in which these keywords appear, it may lead to classification errors.

Topic analysis makes use of unsupervised machine learning algorithms to uncover latent topics within the text. Specifically, we employ the Latent Dirichlet Allocation (LDA) algorithm, as developed by Blei et al. (2003). This method is based on a generative probabilistic model where each topic is represented as a distribution over all words in the text corpus, and each document (speech) is represented as a distribution over these topics. The number of topics and the parameters of the underlying distributions are determined by the researcher. All word-topic and document-topic probability weights are estimated through maximum likelihood.

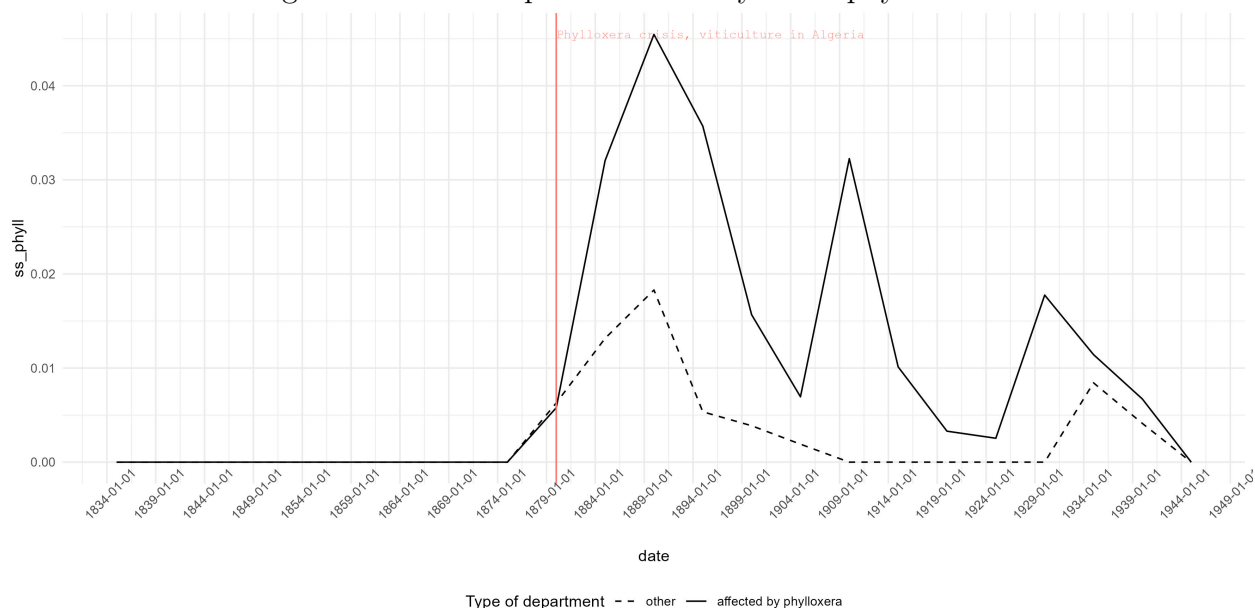
In the following, we restrict our analysis to the subset of parliamentary speeches including the keyword “Algeria”. This restriction allows us to capture more nuances in the debates pertaining to Algeria.

4.1 Keywords

As a first step, we search for occurrences of the keyword “phylloxera” within the parliamentary speeches. Figure 1 illustrates the share of parliamentary speeches, within five-year intervals, that include this keyword relative to the total number of parliamentary speeches in the same period. The solid line in the figure denotes parliamentarians elected in departments affected by phylloxera, whereas the dashed line corresponds to parliamentarians elected in other departments. Consistent with the historical evidence, we detect the initial occurrences of the keyword “phylloxera” in the early seventies, when the phylloxera insect was recognized as the cause of the disease devastating French vineyards. The share of speeches including this keyword reaches its peak between 1885 and 1890, at the height of the phylloxera crisis. While parliamentarians from both affected and unaffected departments mention this keyword,

its usage is more prevalent among parliamentarians from phylloxera-affected departments. In particular, the share of speeches that included the keyword “phylloxera” reaches 4.5 percent among this group of parliamentarians, whereas it remains below 2 percent among their counterparts from other departments. After reaching its peak, the usage of the keyword “phylloxera” gradually declined among all parliamentarians. However, references to phylloxera continued among parliamentarians from the affected departments until the end of the study period, only dropping to zero between 1940 and 1945.

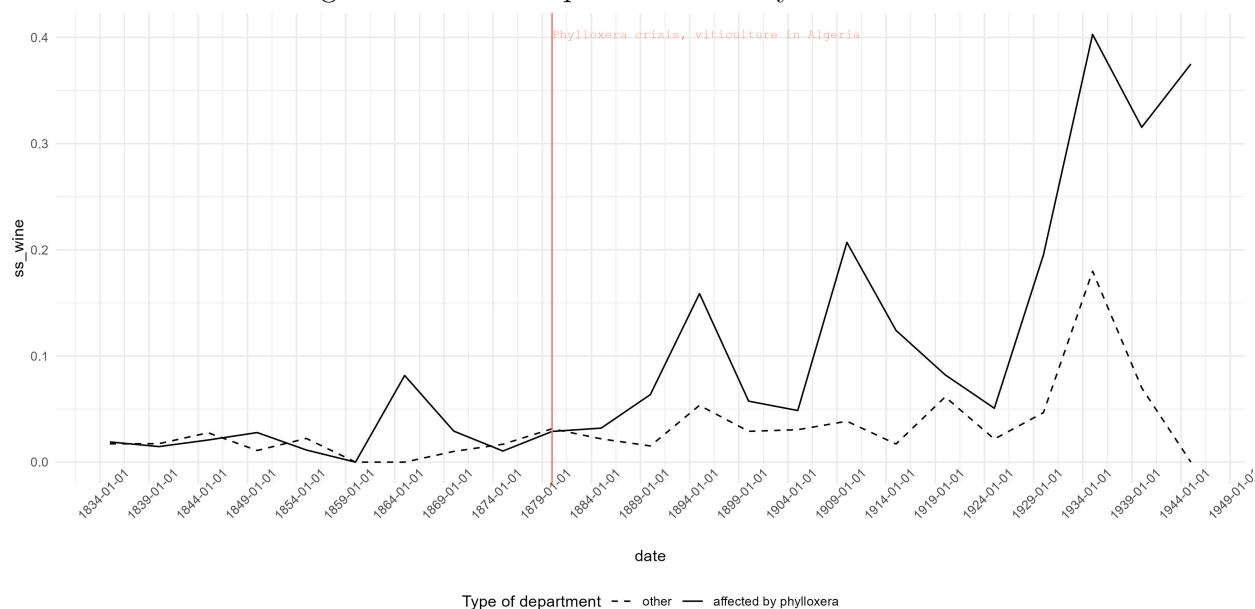
Figure 7: Share of speeches with keyword “phylloxera”



To investigate the connection between the diffusion of the phylloxera and its impact on wine production in the parliamentary speeches, we next search wine-related keywords. These include *vin* (wine), *vigne* (vineyard), and *viticole* (viticultural, or wine-growing). Similar to Figure 7, Figure 8 illustrates the share of speeches including these keywords over five-year intervals. In the decades leading up to the phylloxera crisis, approximately 2% of parliamentary speeches mentions wine, showing neither significant temporal variations nor notable differences between departments impacted by phylloxera and those unaffected. However, beginning in 1885-1890, there is a marked and steady increase in references to wine and viticulture in the speeches of parliamentarians from departments affected by phylloxera, in contrast to a stable pattern observed in other departments. Notably, although phylloxera appears as a temporary shock in the parliamentary discourse (see Figure 7), its influence on discussions related to wine is persistent and increasingly significant over time. Among parliamentarians from affected departments, the share of speeches containing wine-related

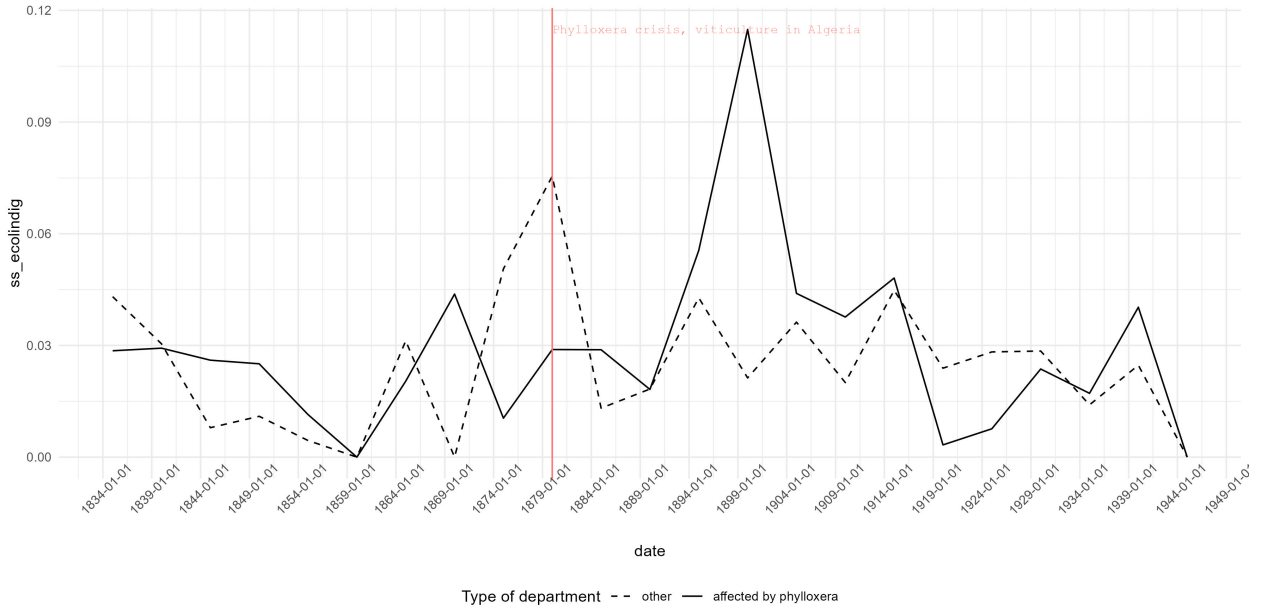
keywords peaked at 40% during the years 1935-1940. This probably reflects the persistent effects of the relocation of wine production to Algeria on competition and on the economic conditions of French wine-growers.

Figure 8: Share of speeches with keyword “wine”



Finally, we explore the relationship between the impact of phylloxera and the support for educational policies aimed at the indigenous populations of Algeria. In Figure 9, we present the share of speeches that include both the keywords *école* (school) and *indigène* (indigenous) across five-year intervals. Beginning in the 1890-1895 period, the use of these keywords is significantly more prevalent in departments affected by phylloxera compared to other departments, with a peak of approximately 12% of speeches during the 1895-1900 interval. This wedge persists up to the 1910-1915 period and then disappears. Interestingly, the divergence in support for educational policies targeting the indigenous population between affected and unaffected departments becomes apparent approximately ten years after a similar divergence in the usage of keywords associated with phylloxera and wine. This pattern is not surprising. The process of relocating families engaged in wine production to Algeria, acquiring and cultivating land, organizing production in a new environment, pinpointing new challenges and potential solutions are all steps that require time. Hence, we may expect that the discussion of policies aimed at training the local Algerian workforce display a lag relative to the initial discovery of phylloxera.

Figure 9: Share of speeches with keywords “school” and “indigeneous people”



4.2 Topic Analysis

We now present the findings of our topic analysis. The results reported below derive from an analysis involving 40 topics, although our findings remain robust when employing both fewer and greater numbers of topics. We chose this specific number because our own experimentation indicated that it offers an optimal balance between computational demands and the granularity of topic disaggregation. The parameters for the Dirichlet prior distributions have been set according to standard values recommended in Griffiths and Steyvers (2004).

Our results are consistent with the insights obtained through the dictionary-based methods. Specifically, we identified two topics that are easily interpretable as counterparts to the keywords for wine production and indigeneous schools used in section 4.1.

The topic “wine production” (identified as topic number 17 in the topic analysis) is depicted in Figure 10, and it follows exactly the same trend as Figure 8 concerning the share of speeches that include the keyword “wine”.

Similarly, the topic “indigenous schools” (topic 8 in the topic analysis) is shown in Figure 11 and mirrors the trend in Figure 9 for the share of speeches that include the keywords “ecole” and “indigeneous”.

Figure 10: Topic “Wine production”

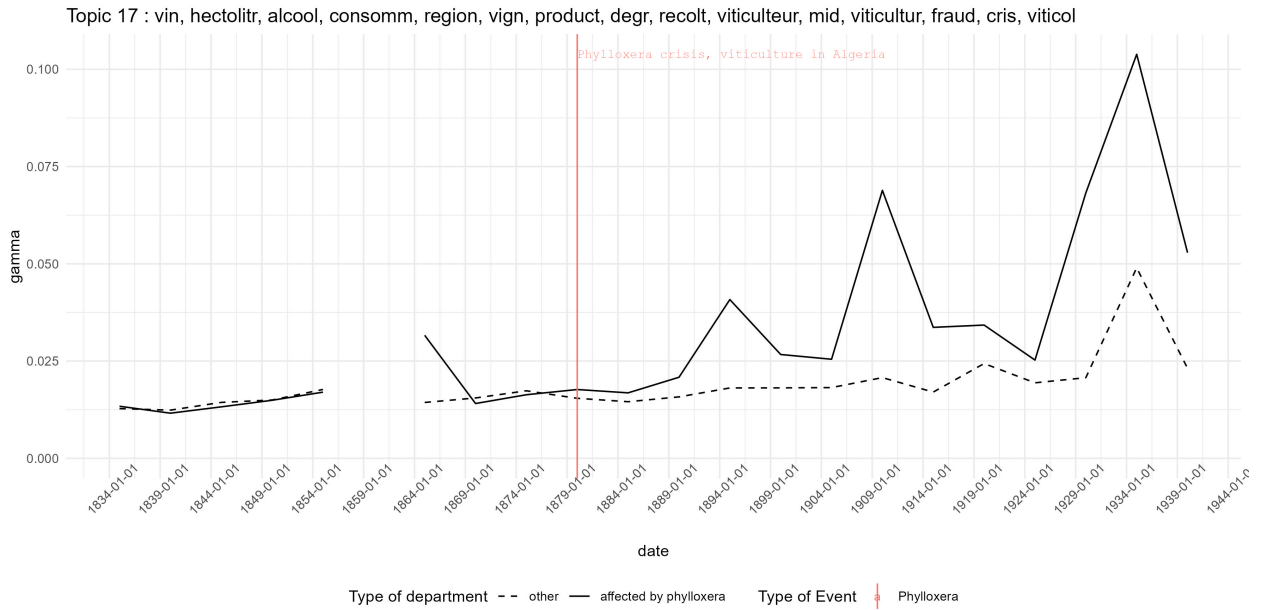
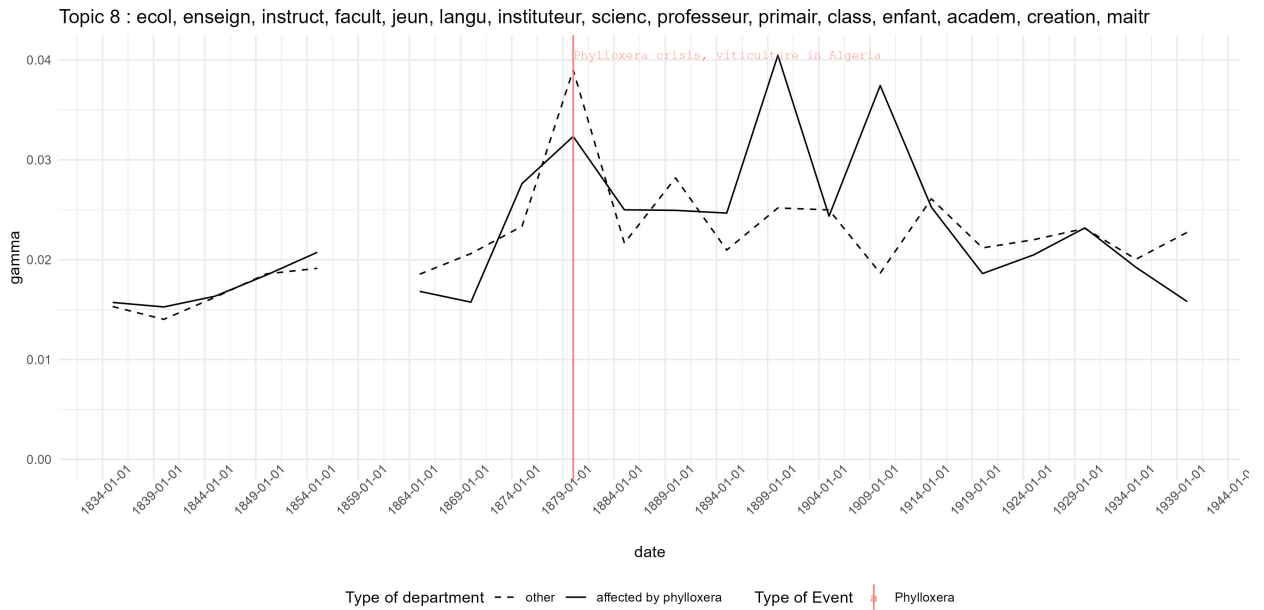


Figure 11: Topic “Indigeneous schools”



5 Conclusions

The occurrence of negative shocks may have an impact on the attitudes and policies towards minorities and on their assimilation. While a literature in economics and political sciences

has shown the negative impact of these shocks on minorities assimilation, there is scant empirical evidence on whether (and under which conditions) shocks may instead trigger assimilation policies favoring integration. This paper studies the impact of the phylloxera crisis in 19th-century France on policies of educational expansion towards the Algerians. In particular, assembling a novel dataset on French MPs and their parliamentary speeches, we find that MPs coming from areas hit harder by the phylloxera, were more likely to express interest and support towards policies aimed at educating the native population. This becomes visible approximately ten years after the phylloxera crisis, consistent with the view that relocating families engaged in wine production to Algeria and organizing production in a new environment are processes that require time.

Thus, these empirical patterns, together with the historical records, suggest that a shock may foster assimilation policies towards minorities when the latter become complementary to the well-functioning of the economy. In the Algerian case, policies of educational expansion were put into place to favor natives' accumulation of human capital.

References

- Abraham, S. S., G. Lanzara, S. Lazzaroni, P. Masella, and M. P. Squicciarini (2024). Spatial and historical drivers of fake news diffusion: Evidence from anti-muslim discrimination in india. *Journal of Urban Economics* 141, 103613.
- Ageron, C.-R. (1991). *Modern Algeria: A History from 1830 to the Present*. Hurst & Company.
- Anderson, D. M., B. Crost, and D. I. Rees (2020). Do economic downturns fuel racial animus? *Journal of Economic Behavior & Organization* 175, 9–18.
- Anderson, R. W., N. D. Johnson, and M. Koyama (2017). Jewish persecutions and weather shocks: 1100–1800. *The Economic Journal* 127(602), 924–958.
- Banerjee, A., E. Duflo, G. Postel-Vinay, and T. Watts (2010). Long-run health impacts of income shocks: Wine and phylloxera in nineteenth-century france. *The Review of Economics and Statistics* 92(4), 714–728.
- Becker, S. O. and L. Pascali (2019). Religion, division of labor, and conflict: Anti-semitism in germany over 600 years. *American Economic Review* 109(5), 1764–1804.
- Bignon, V., E. Caroli, and R. Galbiati (2017). Stealing to survive? crime and income shocks in nineteenth century france. *The Economic Journal* 127(599), 19–49.
- Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3(Jan), 993–1022.
- Doerr, S., S. Gissler, J.-L. Peydró, and H.-J. Voth (2022). Financial crises and political radicalization: How failing banks paved hitler’s path to power. *The Journal of Finance* 77(6), 3339–3372.
- Fouka, V. (2019). How do immigrants respond to discrimination? The case of Germans in the US during World War I. *American Political Science Review* 113(2), 405–422.
- Galet, P. (1957). *Cépages et vignobles de France*. Paul Déhan.
- Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as Data. *Journal of Economic Literature* 57(3), 535–574.

- Gould, E. D. and E. F. Klor (2016). The long-run effect of 9/11: Terrorism, backlash, and the assimilation of Muslim immigrants in the west. *The Economic Journal* 126(597), 2064–2114.
- Griffiths, T. L. and M. Steyvers (2004, April). Finding scientific topics. *Proceedings of the National Academy of Sciences* 101(suppl 1), 5228–5235.
- Isnard, H. and J. H. Labadie (1959). Vineyards and social structure in algeria. *Dio- genes* 7(27), 63–81.
- Jedwab, R., N. D. Johnson, and M. Koyama (2019). Negative shocks and mass persecutions: evidence from the black death. *Journal of Economic Growth* 24(4), 345–395.
- Meloni, G. and J. Swinnen (2014). The rise and fall of the world’s largest wine exporter—and its institutional legacy. *Journal of Wine Economics* 9(1), 3–33.
- Ruedy, J. (2005). *Modern Algeria: the origins and development of a nation*.