Digital Consumption and Political Alignment: Evidence from Mobile Data Across Two European Elections in France

Orlando E. Martínez-Durive
IMDEA Networks Institute
Madrid, Spain
orlando.martinez@networks.imdea.org

Iñaki Úcar
Universidad Carlos III de Madrid
Madrid, Spain
inaki.ucar@uc3m.es

Zbigniew Smoreda
Orange Research
Paris, France
zbigniew.smoreda@orange.com

Esteban Moro
Northeastern University
Boston, MA, USA
e.moroegido@northeastern.edu

Marco Fiore
IMDEA Networks Institute
Madrid, Spain
marco.fiore@networks.imdea.org

Abstract

Understanding how digital platform consumption relates to political alignment is paramount in contemporary democracies, yet large-scale observational evidence across platforms and elections remains scarce. We analyze passive mobile network metadata from a major French Mobile Network Operator (MNO) covering ∼31% of the market across urban and suburban metropolitan France during the 2019 and 2024 European Parliament elections. Integrating demands for tens of mobile services, including social media, news, messaging, and streaming, with socioeconomic indicators, we model vote shares via Dirichlet regression. We find that digital consumption provides independent, complementary signals of political alignment that, for several parties, match or exceed traditional socioeconomic indicators; combining both increases explanatory ability by up to 28.23%. For instance, right-wing support is associated with higher Facebook and TikTok consumption, whereas online news, Twitter, and Instagram correlate with centrist and progressive vote. These associations are consistent across 2019 and 2024, giving a population-scale view of how mobile platform use relates to political alignment.

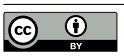
CCS Concepts

• **Networks** → **Network measurement**; *Mobile networks*; • **Applied computing** → *Sociology*.

Keywords

network measurement, mobile networks, digital consumption, political alignment, elections, Dirichlet regression

ACM Reference Format:
Orlando E. Martínez-Durive, Iñaki Úcar, Zbigniew Smoreda, Esteban Moro, and Marco Fiore. 2026. Digital Consumption and Political Alignment: Evidence from Mobile Data Across Two European Elections in France. In *Proceedings of the 2026 ACM Internet Measurement Conference (IMC '26), October 12–16, 2026, Karlsruhe, Germany*. ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3777912.3839827



This work is licensed under a Creative Commons Attribution 4.0 International License.
IMC '26, Karlsruhe, Germany
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2327-8/26/10
https://doi.org/10.1145/3777912.3839827

1 Introduction

The outcome of democratic elections has been traditionally studied in relation to socioeconomic indicators such as education, employment, income, and age, which are known to correlate strongly with political orientation and voter behavior [45, 70, 71, 85, 109, 119]. Media outlets frequently highlight these demographic divides through polling and survey data: for instance, The New York Times has reported on gender and age gaps in voter preferences [110, 111], links between income and candidate support [113], and the effects of unemployment on political leanings [112].

Yet, as citizens increasingly consume information through digital channels, political preferences are increasingly associated not only with socioeconomic structures but also with patterns of news exposure, social media engagement, and online communication–dimensions that have only recently become measurable at scale. Social media, in particular, has provided researchers with unprecedented access to behavioral data and studies have revealed strong links between political polarization, electoral outcomes, and online activity in popular platforms like Twitter [63, 107, 114, 115], Facebook [74], and Reddit [26, 28, 101], suggesting that digital signals can serve as behavioral proxies for vote preferences.

In fact, political actors themselves have recognized digital platforms as powerful tools for voter outreach and influence. Since the infamous Cambdrige Analytica scandal, where data from 80 million users fraudulently captured through Facebook was leveraged to influence the 2016 U.S. presidential elections [52], news outlets have reported numerous cases of significant expenditures in sponsorship campaigns [68], microtargeting strategies tailored to specific demographics [67], or circulation of suspicious content seeking to influence voters [92] in platforms that include X, TikTok, or Facebook. Prior research has in fact proven that Instagram and YouTube have become key spaces for political communication and agenda-setting [36, 65, 84, 106], as also demonstrated by the highly unanticipated, massive voters' endorsement of social media influencers with little prior political visibility [11, 95].

The trends above also raise growing concerns about fairness and manipulation in democratic processes, due to increasing risks that malicious actors exploit online ecosystems to amplify divisive narratives and misinformation, particularly during election periods [23, 25, 31, 73]. The severity of the matter is leading to dedicated regulatory frameworks like the Digital Services Act (DSA) in the European Union (EU) [66], as well as to targeted countermeasures

such as transparency and content moderation measures enacted by social media companies [82], public-private partnerships established to educate voters on misinformation [18], or transparency tools deployed to identify digital advertisement-based disinformation campaigns [13, 14, 22, 94].

All these considerations emphasize the importance and urgency of understanding the complex relationships between digital content consumption and political alignment. A large body of work is in fact already available on the topic, inquiring into political engagement and participation [16, 40, 44, 58, 61, 71], polarization [7, 9, 12, 31, 46, 57, 64, 69, 117] populism [16, 19, 59, 93, 105], micro-targeting and personalized political advertising [14, 55, 78, 103, 120], extremism and radicalization [15, 50, 51, 86], or misinformation[1, 22, 32, 122]. Political science research on the links between digital media and electoral behavior has taken two main forms. On the one hand, causal studies have exploited variations in the mobile communication infrastructure and service availability to unveil that broadband penetration in Italian and German municipalities increased populist party support [104]; 3G expansion widened partisan polarization in the U.S. [79]; 3G adoption reduced confidence in governments [53]; Twitter's staggered U.S. rollout shifted moderate voters' preferences toward liberal positions [42]; or Facebook activity by the AfD party caused increases in hate crime incidence in Germany [83]. On the other hand, observational analyses revealed asymmetric news exposure among over 200 million Facebook users, with conservative audiences disproportionately encountering misinformation [49]; widespread fake news display during the 2016 U.S. election in Facebook, although with limited individual persuasion effects [3]; or accurate predictions of political orientations of 1,000 German voters via machine learning of their web browsing patterns [62]. However, as detailed in Appendix F, no prior work explored the relationship of vote intents of large populations with a wide range of mobile application categories simultaneously.

In this paper, we contribute to closing the gap above by leveraging large-scale measurements of mobile service consumption from a major MNO in all urban and suburban areas of France during the 2019 and 2024 European Parliament elections. We use geographical demands for tens of applications such as Facebook, Instagram, or online news (illustrated in Figure 1b–d) and assess if and to what extent they are associated with political vote across these communes (depicted in Figure 1a and Figure 1e for two major competing parties). We also juxtapose and combine digital consumption with traditional socioeconomic indicators like income, employment, or urbanization levels (shown in Figure 1f–h) to compare the associative strength of each class of feature. Indeed, a qualitative look at Figure 1 already suggests that positive correlations may exist between, *e.g.*, a political preference for *Rassemblement National* and a combination of higher usage of Facebook and lower income, or that local favor for *La France Insoumise* may be associated to a higher consumption of news or significant unemployment.

Our approach is observational: we characterize population-scale associations between digital consumption and political alignment, without claiming causal mechanisms. By exploring in a systematic way these associations, our work yields the following contributions.

- We carry out a large-scale quantitative analysis of the relationships between a wide range of mobile services and political alignment across urban and suburban communes in metropolitan France, using passive traffic data from a single MNO covering ∼31% of the French mobile market. To the best of our knowledge, this is the first study to characterize multi-platform mobile application consumption and its associations with electoral outcomes at high spatial resolution and over millions of voters.
- We introduce a dedicated correlation analysis framework that aligns mobile service consumption with legacy socioeconomic indicators and uses Dirichlet regression to model electoral outcomes, enabling a multivariate analysis of vote shares across political parties.
- We show that mobile service consumption substantially enhances the explanatory power of conventional socioeconomic models, improving by 28.23% the captured variance across all parties in the 2024 elections. Our results provide evidence that digital usage carries behavioral and contextual signals about political alignment that traditional indicators do not fully capture, and suggest a growing role across the two studied elections, with the improvement in model accuracy increasing by 7.6% from 2019 to 2024.
- We uncover several interpretable relationships between specific digital behaviors and political orientation, such as the positive association of Facebook usage with *Rassemblement National* votes or the alignment of Instagram demands with centrist and progressive parties. We also reveal temporal shifts in the predictive role of platforms like Twitter[1], which transitioned from centrist engagement in 2019 to a higher consumption in areas where *La France Insoumise* had a higher votes shares in 2024.

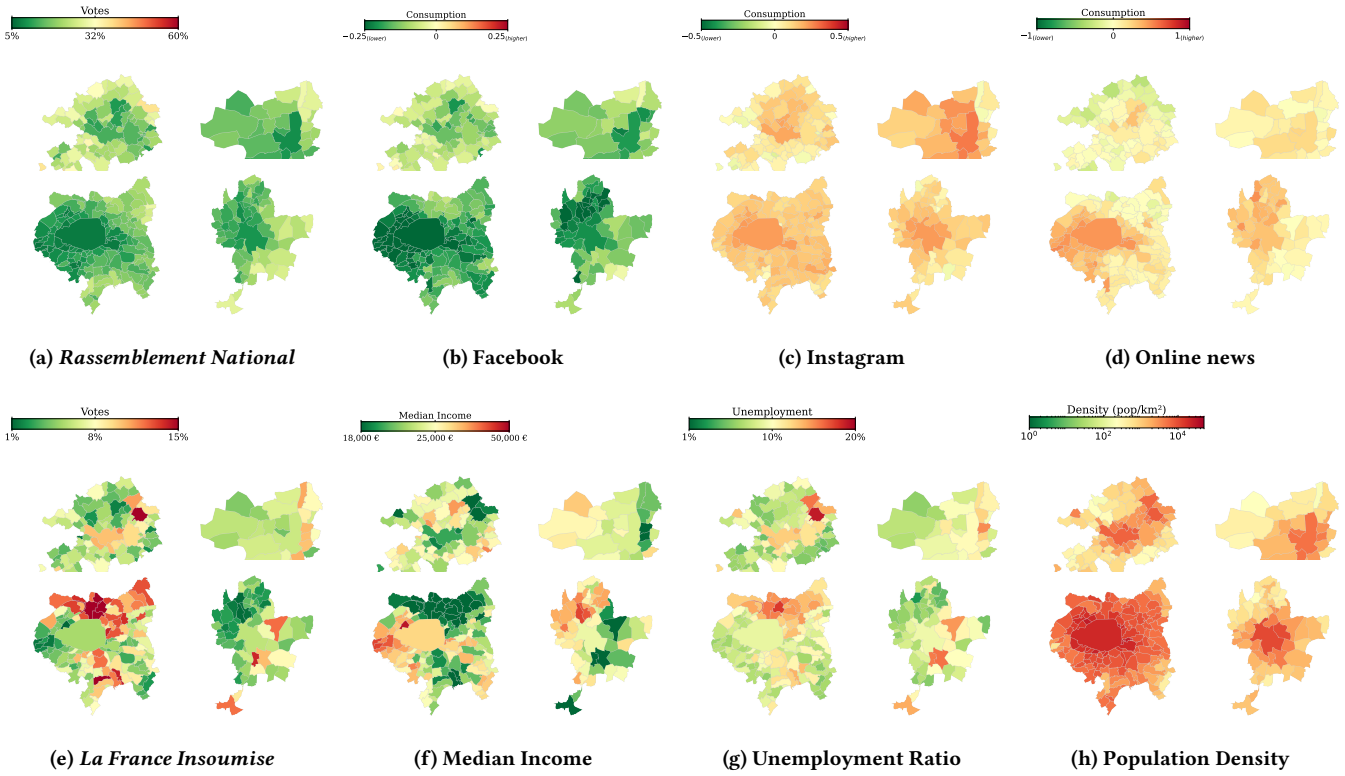
2 European Elections in France

France holds European Parliamentary elections every five years alongside all EU member states. Parties must obtain at least 5% of total votes to secure seats; we focus on the two most recent elections, held in 2019 and 2024.

Throughout, we place each qualifying party on a Left-Right (LR) axis (0 = far Left, 10 = far Right) drawn from the ParlGov dataset [34, 35] based on the Chapel Hill Expert Survey [96]; Figure 2 summarizes the results of both elections along this axis.

Six parties cleared the 5% threshold in 2019 and seven in 2024, hence we focus on those and aggregate all remaining parties into a single *others* category to account for 100% of votes. As also shown in Figure 2, between the two elections the French political landscape experienced a modest rise in polarization, with growing support for the ideological extremes (*La France Insoumise* and *Rassemblement National*), the emergence of *Coalition Réveiller* and *Coalition La France fière*, and the relative weakening of centrist political forces. Yet the positioning of all main groups remained fairly stable, which lets us juxtapose our analyses for the two years and carry out a meaningful longitudinal study. An extended electoral context is provided in Appendix I for readers unfamiliar with the French political landscape.

[1]Twitter underwent a change in ownership through a private acquisition on October 27, 2022, and was re-branded as X in 2023. For consistency across the 2019 and 2024 elections, we retain the name Twitter for the service.



(a) *Rassemblement National* (b) Facebook (c) Instagram (d) Online news

(e) *La France Insoumise* (f) Median Income (g) Unemployment Ratio (h) Population Density

Figure 1: Spatial distributions across a representative subset of urban and suburban French communes of the percentage of votes for *Rassemblement National* (a) and *La France Insoumise* (e) in the 2019 European Parliamentary election; symmetrical Revealed Comparative Advantage (sRCA) for (b) Facebook, (c) Instagram, (d) online news in the week preceding the election; and (f) median income, (g) unemployment ratio, and (h) population density in 2019. Within each panel, the four maps correspond to communes in the regions of Lille (top-left), Bordeaux (top-right), Paris (bottom-left), and Lyon (bottom-right). Details on the sRCA transformation are in Section 3.3.

3 Data

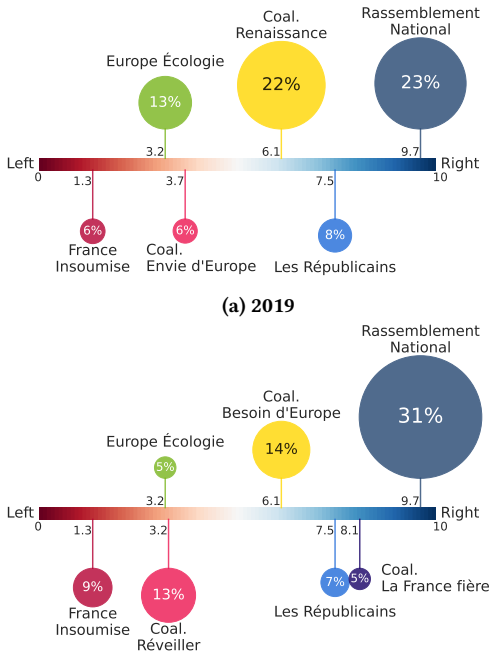
Our study hinges upon the integration of multiple data sources: socioeconomic indicators, election results, communes boundaries, and raw mobile network traffic, which are presented next. In order to make our results reproducible and to further research, we provide a curated version of the dataset used in this work together with the analysis and modeling code in a public repository[2], see Appendix C for a detailed description of the data and code availability. The raw mobile traffic data remain under Non-Disclosure Agreement (NDA).

These diverse data types yield high heterogeneity in terms of geographical resolution and scale representation, which requires alignments in the spatiotemporal dimensions as well as in the scale rendering, as discussed in Sections 3.2 and 3.3, respectively. Table 1 consolidates the key characteristics of the resulting dataset.

3.1 Sources

Mobile traffic data. Two measurement campaigns were conducted in 2019 and 2024 in collaboration with a major MNO that operates under two commercial brands: its main brand and a low-cost subsidiary, together serving approximately 22 million mobile lines (around 3.9 million of them under the low-cost brand in 2019) and representing a 31.9% share of the French mobile market in 2019, a share that remained at 31.5% in 2024 [89, 91]. Both brands share the same backbone infrastructure where the measurement probes are deployed, yielding a *single* dataset that captures the full combined subscriber base with *no sampling*, in contrast to platform-API-based approaches, where even large samples can introduce systematic bias [33]. The campaigns collected byte counts of mobile service demands at a 1-minute granularity across approximately 17,600 and 25,900 base stations (cell sites) nationwide, and covering 22 and 27 popular mobile services, in 2019 and 2024 respectively, identified by proprietary Deep Packet Inspection (DPI) classifiers; the additional services in 2024 (*e.g.*, TikTok, Disney+) reflect platforms that gained prominence between the two elections. For our analysis, we group these services into five categories: *Social Media*, *News*, *Streaming*, *Messaging*, and *Mobile Services* (*i.e.*, the union of all tracked services). Figure 4 reports every individual service under its category header for both elections. Complete technical information about the network architecture, probe deployment, and DPI classification is in Appendix B. Since our focus is on mobile consumption leading up to the elections, the campaigns measured data from the week

[2]https://github.com/nds-group/digital-consumption-political-alignment-france



(a) 2019

(b) 2024

Figure 2: European Parliamentary election results in France, for parties meeting the 5% vote threshold.

Table 1: Mobile traffic dataset at a glance. The MNO operates two commercial brands, that share a single backbone and are captured as one unsampled dataset. Mobile services are split across thematic categories. Base-station counts denote physical radio cell sites.

	2019	2024
Mobile lines	~21.8 M	~22.0 M
of which low-cost brand	3.9 M	4.8 M
Mobile market share	31.9%	31.5%
Base stations (nationwide)	17,567	25,894
in analyzed communes	12,871	15,832
Mobile services (total)	22	27
Social Media	6	7
News	6	6
Messaging	3	5
Streaming	7	9
Urban & suburban communes	4,102	
Coverage (registered voters)	≈66%	
Measurement period	19–25 May	2–9 Jun
Analysis hours	weekday 20:00–07:00	

preceding polling days in each year, *i.e.*, May 19–25 in 2019 and June 2–9 in 2024.

Socioeconomic indicators. To understand how socioeconomic factors compare to mobile service demand and how much they can explain election outcomes, we incorporate median income, unemployment rate, and population age structure (*i.e.*, the number of inhabitants within specific age ranges). These three variables consistently dominate electoral models in France and Western Europe [17, 47, 108] as discussed in detail in Section 6. The indicators were obtained from the National Institute of Statistics and Economic Studies [88], which collects and produces data about the French economy and society. Figures 1f–h illustrate the spatial distribution of median income, unemployment ratio, and population density for sample communes in 2019.

Election results. Electoral results were obtained from the French Ministry of the Interior's open data portal [29, 30]. The datasets include the regional code, number of valid, blank, and invalid votes, party names, candidates, and the number of valid votes obtained by each of the competing parties at each polling station.

3.2 Spatiotemporal Consolidation

The datasets used in our study exhibit differences in their temporal and spatial granularity. Resolving these discrepancies is critical to accurately measuring the relationships among the target variables of our study, as discussed next.

Areal consolidation. Election results are collected at the *bureau de vote* (*i.e.*, polling station) level and subsequently aggregated at commune, department, regional, and national levels. Similarly, census data and socioeconomic indicators are collected for individual households and aggregated into the different administrative units, the most accurate being the Islets Regrouped for Statistical Information (IRIS). For privacy reasons, income and unemployment data are only published for administrative areas with more than 2,000 residents [88]. Only a subset of IRIS zones meet this threshold, whereas most communes, which represent the second most granular geographical zoning, do. Therefore, we adopt communes as the spatial unit for our analysis, which lets us (*i*) access complete data for the vast majority of the French territory unlike what happens with IRIS zones, (*ii*) immediately relate electoral outcomes and socioeconomic indicators that natively employ communes as aggregation levels, and (*iii*) achieve a suitable trade-off between the preference for higher geographical resolutions and the need for the spatial interpolation of data to be dependable.

The last point above relates to the fact that mobile service traffic is geo-referenced at the level of radio access antennas, whose location is known but whose precise coverage area is not disclosed by the MNO for confidentiality reasons. In order to perform a spatial mapping of the mobile service demands recorded at each antenna and align the digital consumption to the other data sources, we rely on the widely adopted practice of relaying on a Voronoi tessellation of space [77, 81, 97, 116]. This approach is illustrated in Figure 3 and consists in partitioning the target space (whole France[3] in our case) into Voronoi cells (middle layer in the figure) based on the location of the radio access antennas (lower layer). Voronoi cells are then spatially intersected with the commune boundaries (top layer) to determine the fraction of each cell overlapping with every commune. Since individual subscriber positions are unavailable, we follow the standard practice in MNO spatial analysis of assuming uniform traffic distribution within each Voronoi cell [10, 39, 48, 76, 100];

[3]European elections are hold in all French territories, which include several overseas departments like French Guiana or Guadeloupe. Yet our analysis focuses exclusively on so-called *metropolitan* France, as the mobile network data does not cover French territories globally.

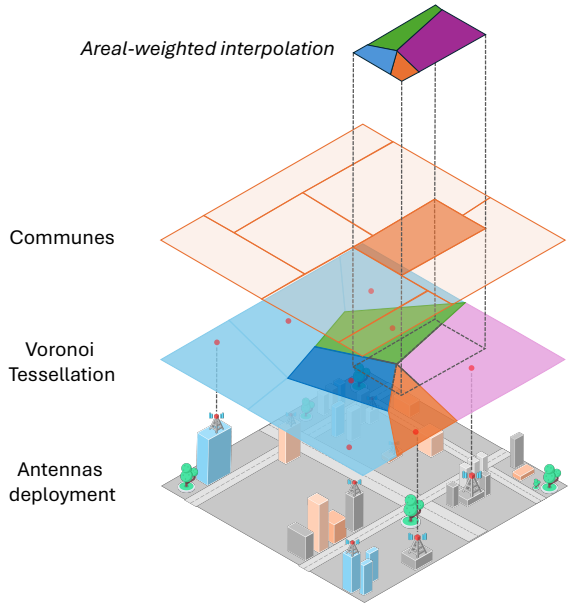


Figure 3: Illustration of areal-weighted interpolation between Voronoi tesselation and administrative units.

the overlap fractions serve as weights to proportionally allocate antenna traffic to communes. Cell-level misattribution from this assumption is expected to average out upon aggregation across the multiple cells that overlap each commune.

This areal-weighted interpolation process is reliable when the radio access infrastructure is dense but introduces artifacts in underpopulated regions where antennas are sparsely deployed [80]. To avoid the problem, we restrict our analysis to *suburban* and *urban* communes [87] where the cellular network is composed of a large number of closely located sites each supporting multiple antennas. The selected 4,102 communes (identical across both election years) account for approximately 66% of the eligible voting population and generate over 70% of the mobile demand, which lets us look at more stable mobile service usage patterns and reinforces the statistical dependability of our analysis.

Temporal consolidation. A temporal mismatch also exists between the mobile traffic data and the socioeconomic indicators. While mobile service consumption is highly dynamic, census data is relatively static, updated only during periodic survey campaigns. For this study, we use the 2024 edition of the census dataset [88], whose indicators date from 2019 and 2021, the latter being the closest available to the 2024 election at the time of writing.

Mobile service demands measured within a given area over the whole observation period of one week before each election may not exclusively reflect the digital behavior of the local inhabitants, as they include contributions from commuters, temporary visitors, or passers-by. To improve the alignment between the local voting population and the monitored digital consumption, we restrict our investigation to traffic usage during evening and night hours, *i.e.*, when people are most likely to be at home. Specifically, we consider mobile network traffic generated between `20:00` and `07:00` on weekdays only, since previous work [20] shows that mobile application usage peaks around `20:00` and remains relatively stable throughout the evening. We show in Appendix G that this choice does not affect the analysis in significant ways in our case and similar results are obtained under different daytime or full-day windowing.

3.3 Scale Consolidation

The variables underpinning our analysis are heterogeneous and span very different value ranges. For instance, median income is measured in Euros and can take any value from 11,000 to 52,000, unemployment is expressed as a percentage over the local population, and mobile traffic demands are byte counts per minute. This calls for re-scaling the variables so that their correlations are best exposed in the regression.

The issue is especially severe for traffic demands, where the volumes induced by different mobile services vary significantly. For example, a video streaming application inherently generates much higher traffic than a messaging one, even if they feature similar customer populations and per-user screen time. As a result, in our mobile traffic dataset, the top 10 applications in 2024 are responsible for more than 87% of the total mobile demand, and scrolling through TikTok, the single service generating the most data, causes over $10^4\times$ higher traffic than browsing online news, which may however be a more important service to explain political alignment. To ensure comparability across mobile services, we employ a modified version of the Revealed Comparative Advantage (RCA) index [8], originally developed in economics to measure a country's relative advantage or disadvantage in the production of specific goods or services based on trade flows. The RCA, and its application to network traffic characterization, has also been established in the literature [48, 80, 116]. More precisely, we reinterpret this metric to capture a *service advantage* across communes, *i.e.*, to expose communes where the usage of a service is substantially different than the national average.

Formally, for each commune and mobile service, we first aggregate the total byte counts (accounting for both uplink and downlink traffic) across all evening and night hours of the reference week before each election. Next, we normalize the obtained values by the number of local inhabitants to obtain the per-capita consumption T_{ij} of a given service j in commune i. The associated RCA index is then computed as:

$$\mathrm{RCA}_{ij} = \frac{T_{ij}/\sum_k T_{ik}}{\sum_h T_{hj}/\sum_h \sum_k T_{hk}}, \tag{1}$$

where $\sum_k T_{ik}$ is the total traffic per inhabitant in commune i, aggregated over all services, $\sum_h T_{hj}$ is the mean traffic per inhabitant for service j observed across all communes, and $\sum_h \sum_k T_{hk}$ is the total traffic generated by the average inhabitant of France. Therefore, the enumerator captures the mean prevalence of a specific service j within the total mobile demand in a commune i, whereas the denominator represents the same metric but at a national scale, *i.e.*, when aggregating digital usage data from all communes.

A service j is said to have a *comparative advantage* in commune i if $\mathrm{RCA}_{ij} > 1$, implying a higher-than-ordinary usage of service j in that commune; conversely if $\mathrm{RCA}_{ij} < 1$ suggests that service j presents a *comparative disadvantage*, *i.e.*, a reduced adoption in commune i with respect to the national average. However, the RCA may

still yield extreme values due to large variations in service usage across locations, as well as non-comparable trends for advantage (characterized by an unbounded linear increase as one moves away from the origin) and disadvantage (featuring non-linear harmonic decline bounded by -1) cases. To mitigate these effects, we use a normalized, symmetric version of RCA, named sRCA:

$$\mathrm{sRCA}_{ij} = \frac{\mathrm{RCA}_{ij} - 1}{\mathrm{RCA}_{ij} + 1}, \tag{2}$$

which rescales the RCA values to the range $(-1, 1)$, ensuring symmetry around zero and reducing the influence of outliers. The interpretation remains the same with positive values indicating *comparative advantage* while negative values indicate a *comparative disadvantage*. Overall, these transformations enable a fair comparison of mobile service demand across applications, reducing distortions from intrinsic differences in per-application content type and data volumes. Appendix E provides a visual walkthrough of the raw-to-sRCA transformation for representative services.

Finally, all mobile service and socioeconomic variables are standardized via a division by their sample standard deviation. This transformation brings all variables onto a comparable numerical scale and prevents those with larger magnitudes from dominating the regression in Section 4.2.

4 Explaining Election Results

We introduce a two-step methodology to assess the extent to which mobile service consumption is associated with election outcomes, relative to traditional socioeconomic indicators. First, in Section 4.1 we analyze the pairwise correlation between vote shares and both individual socioeconomic indicators and mobile service usages. Second, in Section 4.2 we employ a regression model adapted to our target problem and to guarantee interpretability.

4.1 Correlation Analysis

We start by carrying out an exploratory analysis of the correlations between vote and individual variables in both the traditional socioeconomic status space as well as in the digital consumption domain. Figure 4 summarizes the Pearson's coefficient ρ of pairwise linear correlations of each target variable, *i.e.*, the share of votes obtained by each political party across communes in France (as rows), to every input variable, *i.e.*, individual socioeconomic indicators or mobile application demands (as columns).

Overall, the initial correlation analysis shows promise, with significant concurrence in the values of specific digital usage input and target vote variables that lead to ρ as high as 0.63, for the case of Facebook and *Rassemblement National*. More generally, applications such as the same Facebook, Instagram, LinkedIn, Snapchat, or online news (which gathers traffic generated by access to the websites of newspapers) show explanatory capabilities that are on par with those of customary socioeconomic predictors like income or the level of unemployment. Importantly, correlations appears largely stable across the 2019 and 2024 elections, corroborating the solidity of the observed relationships.

Delving into the sign and values of ρ reveals how a higher consumption of some mobile services is associated to a stronger *political polarization*, *i.e.*, a tendency to vote for far Left or far Right parties: this is the case of Facebook or Snapchat. The opposite also happens, with other applications signaling a local preference for more moderate parties: examples include Instagram or online news. Yet, we observe the same effects in socioeconomic indicators, with, *e.g.*, unemployment favoring polarization and income more centrist parties. Therefore, a legitimate doubt is whether the effects we observe for the mobile service consumption variables are rightful bonds to political alignment or just proxies for the socioeconomic ones, *e.g.*, because more affluent individuals also use LinkedIn or read online news more often.

The key question is whether digital usages *add* predictive power beyond conventional socioeconomic indicators. Simple pairwise regression cannot answer that question, which instead requires multivariate analyses that compare the goodness of fit of models using diverse combination of input variables. However, straightforward multivariate models that use subsets of all input socioeconomic and mobile demand variables to predict the vote share of a single party are not suitable for our analysis either. The reason is that election outcomes, represented as party vote shares, are compositional in nature, since the proportions across parties are non-negative and sum to a fixed 100% total. In a pure two-party system, this constraint can be ignored: modeling one party's vote share automatically determines the other [56]. However, in multiparty elections, neglecting this compositional structure can lead to inconsistent or logically impossible predictions where the total vote share exceeds 100%. Properly regressing political alignment therefore requires a dedicated methodology, which we develop next.

4.2 Regressing Vote in Multiparty Elections

Our strategy to explaining political alignment in multiparty elections hinges upon Dirichlet Regression [56, 75] as a suitable modeling framework for compositional data. This approach estimates the parameters of a continuous multivariate Dirichlet distribution, ensuring that the predicted vote shares across all parties sum to 100%. Formally, the Dirichlet distribution is defined as:

$$\mathcal{D}(\mathcal{V}, \alpha) = \frac{1}{\mathcal{B}(\alpha)} \prod_{l=1}^{N} v_l^{(\alpha_l - 1)}, \tag{3}$$

where $\mathcal{B}(\alpha)$ is the multivariate Beta function, and α is a *concentration vector* of length N that determines the location of the center of mass of the distribution across its N dimensions. In our case, each dimension corresponds to a political party, and v_l represents the vote share of party l. Our goal is to estimate the components α_p of the vector α that best fit the observed vote share distribution in the n-dimensional space.

Figure 5 shows a simplified illustration of the empirical distribution for the 2019 election results using a ternary plot. Two sides are associated with the two main parties, *Rassemblement National* and *Coalition Renaissance*, while all other parties are grouped under a single *others* category. For this example, the α vector that best fits the distribution is $\alpha = (2.24, 1.89, 2.88)$, meaning that the center of mass of the distribution is concentrated primarily on the third dimension (others), followed by the first dimension (*Rassemblement National*), with lower probability on the second dimension (*Coalition Renaissance*). The objective of the Dirichlet regression is then to model the α vector based on the socioeconomic and digital usage indicators for each commune.

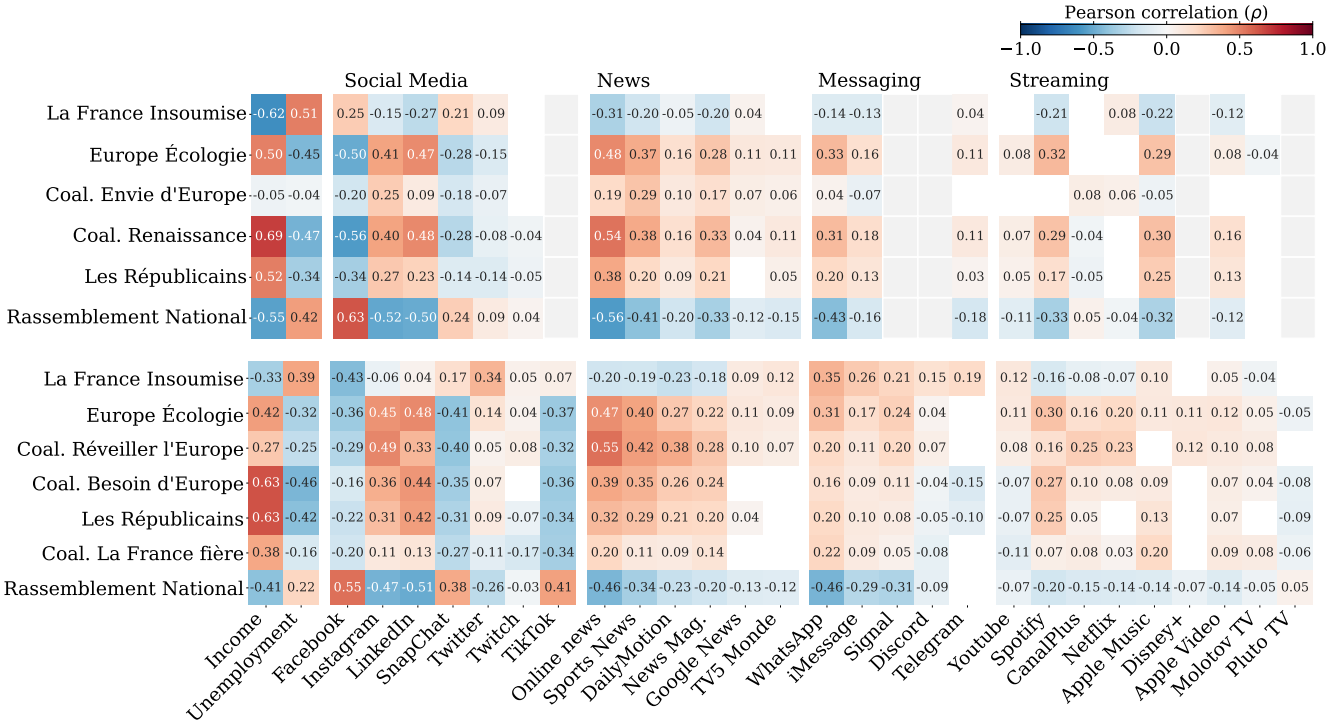


Figure 4: Correlation between variables and political parties' vote shares in 2019 (top) and 2024 (bottom). Empty cells denote non-significant correlations (p-value > 0.05), grey cells indicate missing services in 2019.

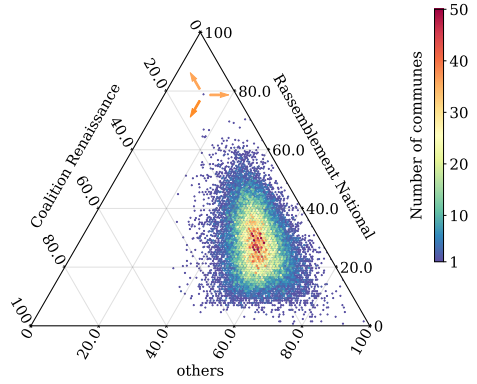


Figure 5: Dirichlet distribution as a ternary plot for a simplified case where votes in the 2019 elections are shared among three groups, *i.e., Rassemblement National, Coalition Renaissance*, and any other party.

Coefficients. We fit the model in (3) via maximum likelihood estimation [75]. Here, each α_p component is expressed as a function of a set of predictor variables, which can take many forms; for the sake of *interpretability* of the fitted model, we restrict our analysis to linear relationships, formally:

$$\alpha_p = \sum_q \beta_q x_q + \gamma_p, \tag{4}$$

Table 2: Sets of variables used for Dirichlet regression.

Sample variables	Population	Unemployment	Income	Socioeconomic	Streaming	News	Messaging	Social Media	Mobile Services	All
Fraction of population in five age intervals, breaking at 15, 30, 45, 60	✓			✓						✓
Fraction of unemployed inhabitants		✓		✓						✓
Median household income			✓	✓						✓
YouTube, Spotify, Netflix, Disney+					✓				✓	✓
Online news, DailyMotion						✓			✓	✓
WhatsApp, Signal, Telegram							✓		✓	✓
Facebook, Instagram, TikTok								✓	✓	✓

where β_q are the regression coefficients, x_q represent the predictor variables, such as mobile service consumption and socioeconomic indicators, and γ_p is the intercept. This formulation enables a straightforward interpretation of coefficients, analogous to that of a standard linear regression.

Predictors. To assess the explanatory power of socioeconomic indicators and mobile service demand, we evaluate 10 different sets of predictor variables. These sets range from individual variables to all socioeconomic and mobile service variables combined (Table 2). This lets us quantify each group's contribution to model performance and how much mobile service usage complements traditional socioeconomic factors in explaining voting behavior.

Goodness of fit. Standard R^2 measures are not directly applicable to maximum-likelihood models such as Dirichlet regression. We therefore compute a predictive R^2, defined as the coefficient of determination ($1 - SS_{res}/SS_{tot}$) between observed and predicted

vote shares, which quantifies the proportion of variance in voting outcomes explained by the model. Since this metric increases monotonically as more predictors are added, we apply the Adjusted R^2 (Adj. R^2 hereinafter) correction [60] to enable a fair comparison across models with different numbers of variables. Specifically, this adjustment penalizes the inclusion of additional predictors and prevents overestimating the model's explanatory power. We note that the OLS Adj. R^2 formula is not formally derived for maximum-likelihood models; we apply it here as a practical complexity penalty, and complement it with the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and log-likelihood in Appendix H. Formally, our Dirichlet goodness-of-fit metric is:

$$\text{Adj. } R^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - p - 1}, \tag{5}$$

where R^2 is the predictive R^2, n is the number of observations, and p is the number of explanatory variables. Higher Adj. R^2 indicates better model fit, with values near 1 indicating the model captures most of the variation in voting patterns.

Multicollinearity. Considering the large set of variables employed, we control for multicollinearity, quantified using the Variance Inflation Factor (VIF) [60] defined as:

$$\text{VIF}(x_m) = \frac{1}{1 - R^2(x_m)}, \tag{6}$$

where $R^2(x_m)$ is obtained by regressing variable x_m on all other predictors. High VIF values indicate strong collinearity; values above 10 are typically considered problematic [121]. The mean VIF for 2019 is 1.68, with a maximum of 3.53 for Facebook, while in 2024, the mean is 2.20, and the maximum is 5.50 for TV5MONDE. All values remain below the reference threshold, suggesting that although some mobile service variables are moderately correlated, they can each contribute distinct information. Based on this outcome, we retain all variables in the subsequent regression analysis.

5 Results

In this section, we evaluate how well the different sets of predictors explain political alignment, with a focus on the contribution of mobile service consumption indicators. We present goodness-of-fit results in Section 5.1, visuals about the spatial quality of the model predictions in Section 5.2, and a discussion of the regression coefficients in Section 5.3.

5.1 Goodness of Fit

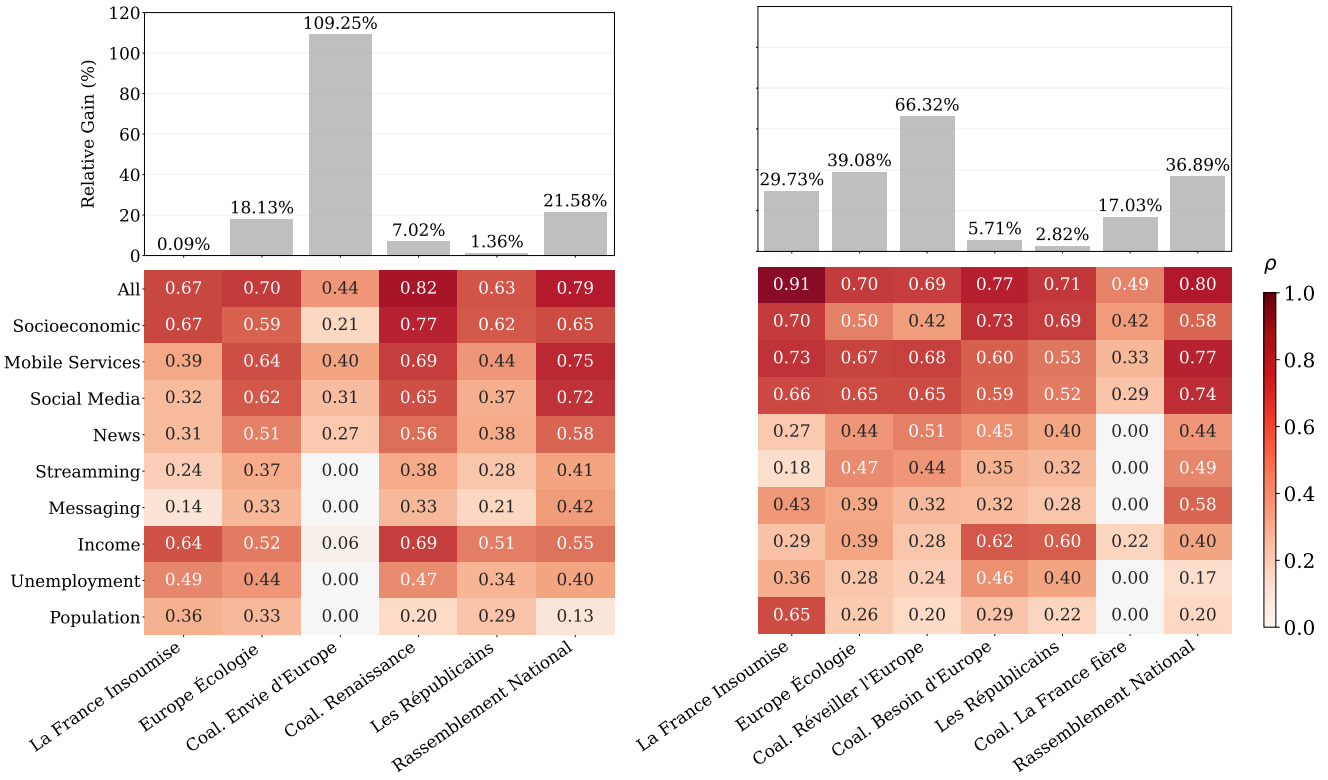
Our Dirichlet regressor maximizes likelihood expressed in terms of Adj. R^2 so as to abide by best practices in multivariate statistical analysis. To render the results of the regression more understandable and easily comparable with those attained by the simple pairwise univariate linear models employed in Section 4.1, we transform the obtained Adj. R^2 values into Pearson's coefficient ρ. Figure 6 reports the correlation obtained for each regression model based on the sets of predictor variables summarized in Table 2 and political groups across the two elections.

The bottom part of the figure shows the complete matrices of correlation values and lets us observe how the *All* model, including the full set of socioeconomic and mobile usage predictors, consistently achieves the highest figures across all parties, attaining $\rho = 0.82$ for *Coalition Renaissance* in 2019, $\rho = 0.91$ for *La France Insoumise* in 2024, and $\rho = 0.80$ for *Rassemblement National* in 2024. Also, the ρ values for the *All* model are substantially stronger than those reported for univariate approaches in Figure 4, which highlights the explanatory advantage of the multivariate Dirichlet regression we adopt. These gains persist after penalizing for additional predictors in (5), and the AIC, BIC, and log-likelihood in Appendix H corroborate the same ranking.

Indeed, these results allow us to quantify the added value of mobile service demands in capturing multiparty political alignment. To single out that specific contribution, the top panels of Figure 6 show the relative increase in correlation when mobile service variables are added to the *Socioeconomic* baseline, for all parties and in the two target electoral events. Essentially, the plots show the percent increment in ρ obtained by the model informed by predictors in the *All* set over that achieved by the *Socioeconomic* indicator set. In 2019, as shown in Figure 6a, the gains for *La France Insoumise* and *Les Républicains* were marginal; however, the improvements for the rest of the parties were more pronounced with increases between 7.02% to 109.13%. Notably, despite the relatively low correlation of *Coalition Envie d'Europe écologique*, the majority of its explained variance came from the *Mobile Services* variables, suggesting that digital behavior played a more important role than *Socioeconomic* variables in explaining its vote distribution. The 2024 results present an even greater impact of mobile application consumption on the explanatory power of the model: except for the *Coalition Envie d'Europe écologique* 2019 outlier, incorporating digital usages significantly (between 2% and 66%) increases the correlation for all parties. When averaging the gains over all parties, digital usage is associated with a 9.6% and 28.23% increase in mean correlation with vote shares in 2019 and 2024, respectively. For 2019, this figure excludes the *Coalition Envie d'Europe écologique* list, whose very low socioeconomic baseline ($\rho \approx 0.21$) turns a modest absolute improvement into an outlying per-party relative gain; including it raises the 2019 average to 26.24%, overstating the result.

Our results also enable a longitudinal analysis about how the added explanatory capability of mobile application consumption has evolved in the five years that separate the two target elections. Figure 7 shows the average correlation across all parties for the *Socioeconomic*, *Mobile Services*, and *All* models (on the left axis), plus the relative gain in explainability between elections (on the right axis). The striking result is that the *Mobile Services* model yields a significant improvement of more than 10% from 2019 to 2024, far outpacing the *Socioeconomic* indicators, whose explanatory power remains essentially flat; this growth in the digital signal is responsible for the higher *All* model accuracy. Likelihood-based metrics corroborate this trend and reveal an additional qualitative shift: AIC and BIC results in Appendix H show that *Social Media* alone outperforms the *Socioeconomic* baseline on all three metrics in 2024, a reversal of the 2019 ranking.

Takeaways. *Mobile service usage is associated with facets of political alignment not captured by conventional indicators, improving the accuracy of voting models in multiparty elections. Importantly, this effect seems amplified in recent years, with digital consumption capturing a growing share of voter behavior, while legacy socioeconomic predictors remain essentially stable.*



(a) 2019 (b) 2024

Figure 6: Goodness-of-fit across predictor sets and parties. Bottom: Correlation (ρ) values for each model and party. Top: Relative gain in ρ when adding mobile service variables to the *Socioeconomic* baseline.

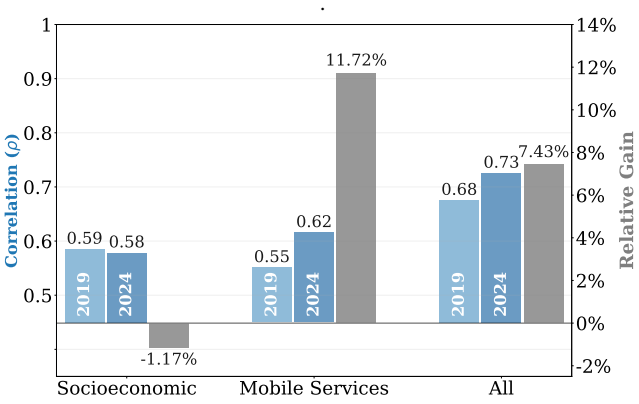


Figure 7: Change in explainability across elections. Average correlation (right axis) and relative increase in explainability (left axis) from 2019 to 2024.

5.2 Spatial Analysis

Beyond the quantitative view provided by the Goodness of fit, Figure 8 provides a complementary, qualitative visualization of the capability of the different indicators to capture political alignment. The top row shows results for *Coalition Renaissance* in 2019, while the bottom row shows results for *La France Insoumise* in 2024 across different models. While the *All* model captures much of the spatial structure in the observed vote shares, the maps also make clear that the contributions of each variable set vary substantially across parties, years, and geographic areas.

The case of *Coalition Renaissance* in 2019 illustrates that the *Mobile Services* model alone captures part of the spatial pattern of observed votes, mainly driven by *Social Media* variables, yet the *Socioeconomic* model more accurately identifies specific communes where *Coalition Renaissance* received stronger support. The combination of both in the *All* model synthesizes these complementary signals. In contrast, for *La France Insoumise* in 2024, the *Socioeconomic* model tends to underestimate vote shares, whereas *Mobile Services* appears more informative, successfully capturing communes with high levels of support for *La France Insoumise*.

Takeaways. *The capability of diverse indicator sets to explain political vote is not homogeneous over space. Instead, specific variables are more or less relevant across geographical areas, paving the way to finer-grained explorations of the added value of mobile application consumption in urban and suburban statistical zones.*

5.3 Coefficient Analysis

As outlined in Section 4.2, expressing α as a linear combination of predictor variables provides intrinsic interpretability, similar to standard linear regression. This design choice lets us derive Figure 9, which presents the regression coefficients β_q of the *All*

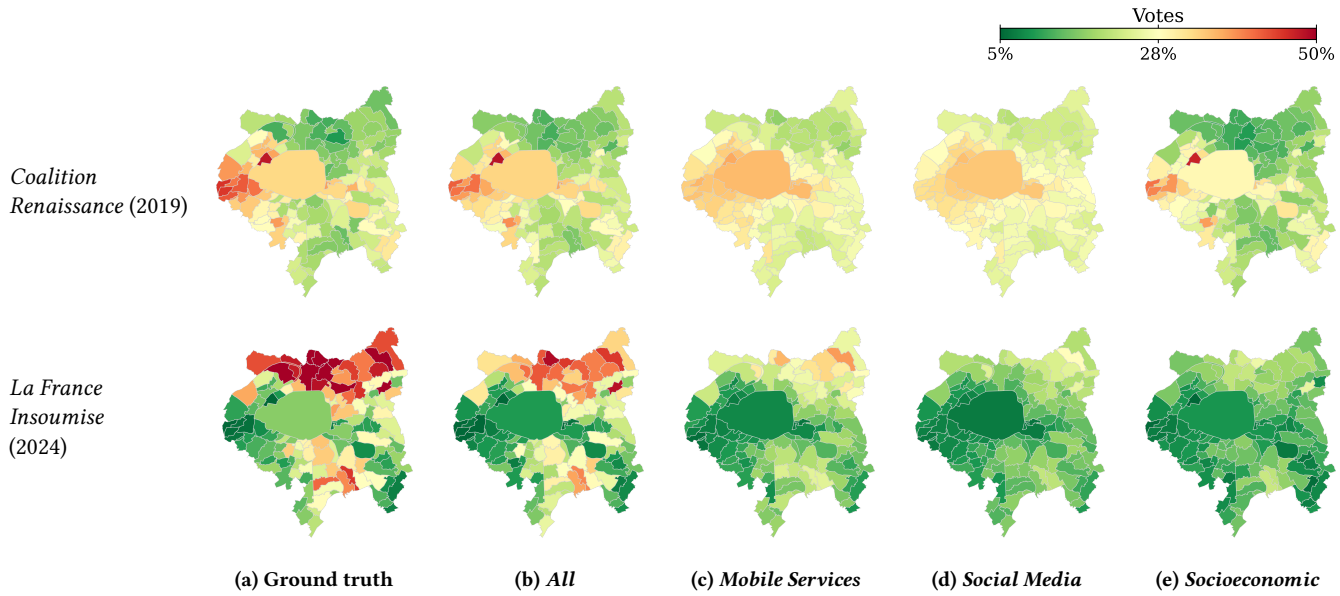


Figure 8: Predicted electoral support in the European Parliamentary elections within the Grand Paris region for two representative parties in 2019 and 2024, respectively. Panels present (a) the actual vote shares, (b)–(e) predictions from models using four different variable sets from Table 2.

model that integrates both socioeconomic indicators and mobile service consumption. By examining these coefficients, we can assess how each variable relates to party vote shares and identify which services are associated with higher or lower electoral support.

Since the Dirichlet regression maximizes the log-likelihood, the resulting formulation can be interpreted analogously to a linear model in log-space, where the vertical line at $x = 0$ marks the threshold for coefficient significance. Positive coefficients indicate a direct relationship with vote share (higher consumption, more votes), while negative ones indicate the opposite.

To account for statistical uncertainty, we report 95% confidence intervals (horizontal error bars) and exclude variables whose intervals overlap zero, as they do not significantly impact any party's vote share. Because socioeconomic variables are explicitly included, the coefficients associated with mobile service demand reflect the unique contribution of each application, disentangled from its potential effects as proxies to the underlying socioeconomic conditions. From the results in Figure 9, we draw the following points.

Median Income. Household median income plays a significant role in explaining political alignment and is a key variable for all parties except *Coalition Réveiller* in 2024. While relatively consistent across both elections, it relates strongly to *La France Insoumise* and *Les Républicains*, as well as *Coalition Renaissance* in 2019 and *Coalition Besoin d'Europe* in 2024.

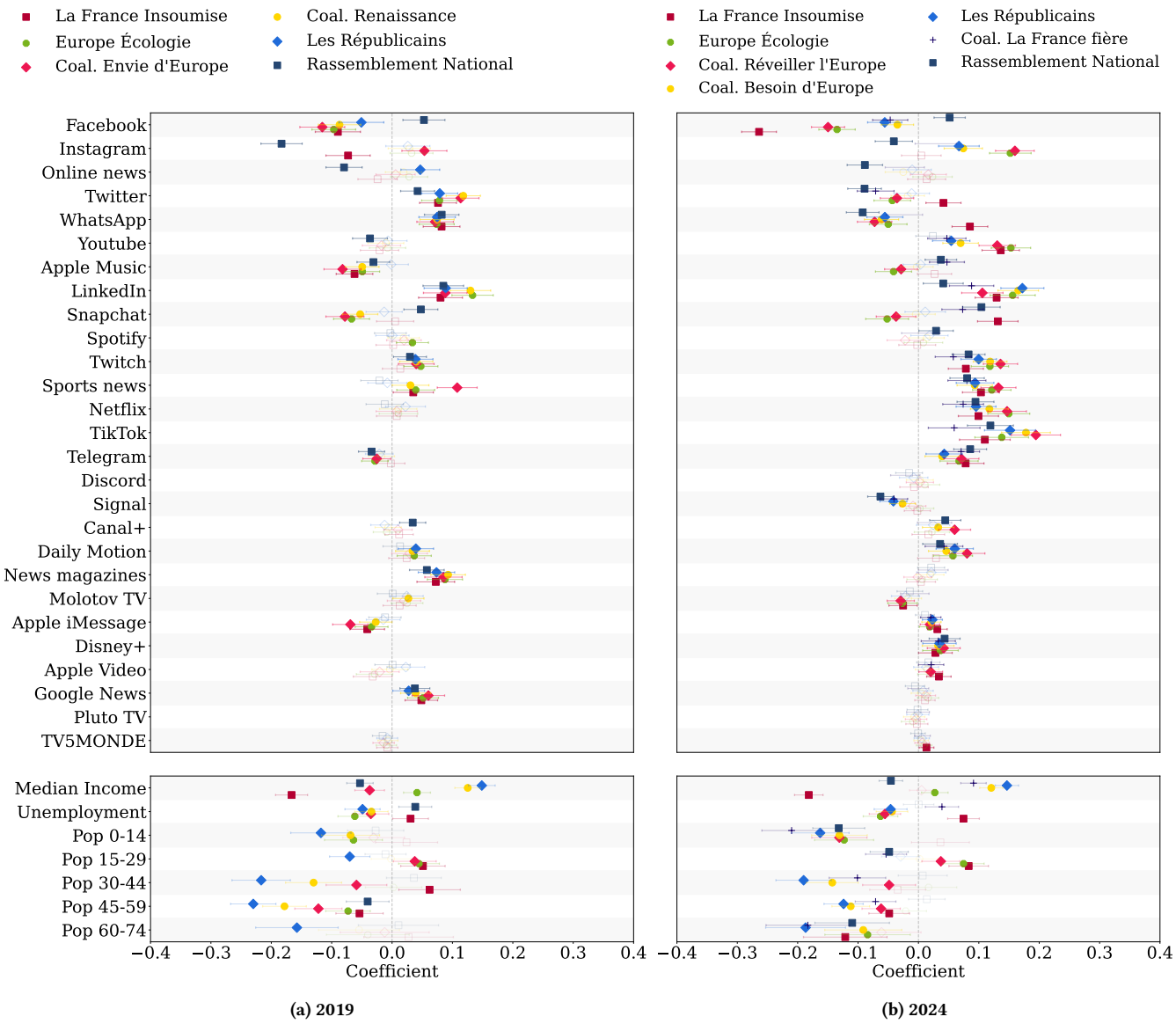
Facebook and Instagram. The social media platforms that play a major role in political alignment are Facebook and Instagram, with relatively positive behavior. On one hand, higher Facebook usage consistently correlates with more votes for *Rassemblement National* across both elections, while it is negatively associated with the vote shares of all other parties. On the other hand, areas with higher Instagram usage tend to favor moderate and centrist parties, such as *Coalition Envie d'Europe écologique* in 2019 and *Coalition Réveiller*, *Europe Écologie-Les Verts*, *Coalition Besoin d'Europe*, and *Les Républicains* in 2024. In a complementary manner, lower Instagram demand is associated to higher political alignment for extremist Left and Right parties like *Rassemblement National* and *La France Insoumise*.

Several of these temporal shifts admit plausible interpretations; we offer them as hypotheses for future empirical validation rather than established findings.

Twitter. Twitter exhibits a notable shift across elections. In 2019, it is a significant predictor for all parties, being particularly popular among centrist parties like *Coalition Renaissance* and *Europe Écologie-Les Verts*. However, in 2024, Twitter loses its alignment with centrist parties and instead shows a clear polarized pattern: higher use of the application becomes associated with *La France Insoumise*, while areas with lower Twitter consumption match those where *Coalition La France fière* and *Rassemblement National* received more votes. Previous work [54] has reported on the successful Twitter campaign launched by Manon Aubry after the 2019 elections, based on the concept of *raising awareness* and promoting among users of the platform her image as a disruptive outsider in the French political scenario, which may explain the observed shift.

TikTok. While we lack a 2019 reference point for TikTok, the 2024 results show it is most strongly associated with *Coalition Réveiller* (moderate-left) and least with *Rassemblement National* and the other extreme parties. Once socioeconomic factors are controlled for, the application, widespread among younger voters, is not specific to these parties, despite its links to polarizing political alignment [2, 24].

Netflix and Spotify. Some apps are not significant predictors in 2019 but become relevant for certain parties in 2024. Notably, Netflix is insignificant to all parties in 2019 but becomes significant



(a) 2019 **(b) 2024**

Figure 9: Dirichlet regression coefficients for 2019 and 2024. Values above (below) zero indicate positive (negative) associations with political alignment. Non-significant coefficients are greyed out.

for all of them in 2024, while Spotify is significant only for *Europe Écologie-Les Verts* in 2019 and *Rassemblement National* in 2024. One possible explanation is the growing emphasis on podcasts on Spotify, supported by major deals with popular political commentators [98] and a collaboration agreement with Radio France, the leading podcast producer in the country [41]. These initiatives have promoted podcast consumption in France, potentially contributing to the increasing presence of political content on the platform [21]. As for Netflix, in 2024, areas with higher demand for the platform were those where *Europe Écologie-Les Verts* and *Coalition Réveiller* received more votes. This suggests a possible shift in the platform's content that attracted voters from these parties, though further analyses would be needed to confirm this hypothesis.

Takeaways. *By controlling for socioeconomic factors and using confidence intervals for significance, our framework analyzes platform-party associations across major and niche services. Beyond identifying stable relationships (e.g., Facebook with Rassemblement National, Instagram with moderates), our longitudinal comparison reveals temporal dynamics: Twitter shifted from centrist to polarized usage between elections, while Netflix emerged as a meaningful predictor in 2024, matching evolving content strategies and shifting user behavior.*

6 Limitations of the Study

Our findings reveal associations between digital usage and electoral outcomes; the following limitations bear on their interpretation and generalization.

Causality. We identify associations between mobile service consumption and voting patterns, but do not claim causal relationships. Our observational data cannot distinguish whether usage of a given platform directly influences political views, or political views drive platform selection, or third factors explain both. Our two-election panel offers temporal replication, not causal identification, since two periods provide no exogenous variation in digital consumption. Studies that have achieved causal identification in related settings exploited precisely such variation, via 3G expansion [79] or platform outages [83], none of which are directly observable in our data. We provide population-scale associations and temporal dynamics that serve as empirical foundations for future causal work [72].

Ecological inference. Our analysis operates at the commune level; all associations are therefore population-level relationships. Platform usage and voting preferences cannot be assumed to co-occur within individuals, limiting conclusions about individual behavior. Analyses relating socioeconomic or behavioral indicators to vote are conventionally conducted at the aggregate level, since ballots are secret; our commune-level design follows this standard. The commune is a spatially fine unit, with urban communes typically containing thousands of residents, helping commune-level patterns largely reflect local tendencies rather than compositional artifacts, though the ecological inference gap remains.

Limited socioeconomic variables. We use age distribution, unemployment, and income: the three variables that consistently emerge as dominant commune-level predictors of electoral behavior in France and Western Europe [17, 47, 102, 108, 118]. Gender, while a significant predictor in individual-level surveys, is uninformative at the commune level, where the sex ratio is approximately 0.50 everywhere with near-zero spatial variance. This baseline is not a shortcut; our goal is to assess whether mobile service patterns add independent explanatory power beyond this established set. Whether the digital signal persists when additional controls, including education, occupation, and migration status, are incorporated remains a natural direction for future work.

Generalizability beyond France. Our findings are specific to France, including its media ecosystem, political landscape, and platform usage. We do not claim these associations extend to other countries; for instance, TikTok's correlation with certain parties may be French-specific. However, the broader finding that mobile service consumption reveals systematic political information beyond socioeconomic factors likely extends to other developed democracies with comparable digital adoption and multiparty systems, as in Western and Northern Europe.

Urban and suburban scope. Our analysis is restricted to urban and suburban communes (approximately 66% of eligible voters and over 70% of mobile demand), where antenna density supports accurate traffic interpolation. Mobile traffic signatures vary systematically with urban fabric and land use [43], and mobile service consumption is markedly uneven across France [80]; in rural communes, low antenna density and greater reliance on services in nearby towns plausibly dilute locally distinctive signals. We therefore do not assume our associations extend to rural communes, where the digital signal may be weaker or differently shaped, and we leave the full national territory to future work.

Sample representativeness. Our study analyzes data from a major MNO representing 31.9% of the French mobile market in 2019 and 31.5% in 2024. All findings in this paper are conditioned on this single operator; multi-operator validation would require data access that is practically infeasible given privacy, regulatory, and proprietary constraints. The MNO's customer base spans all French departments and includes diverse socioeconomic segments through both its primary brand and its low-cost subsidiary, as detailed in Appendix D. Given this market share and near-uniform smartphone ownership across demographic groups, any operator-specific bias would need to be very substantial to affect our conclusions.

Volume versus engagement. The sRCA normalization neutralizes the coarsest volume confound across services: it measures whether a service is over-represented in a commune relative to the national average. Therefore we compare geographic footprints rather than raw byte totals, ensuring that a data-heavy service such as video streaming does not dominate a data-light one such as a news outlet. However, our signal cannot separate active engagement, such as scrolling, from background synchronization, nor distinguish content types within an application. Disentangling these would require device-side instrumentation or application- and platform-level telemetry, none of which a passive MNO vantage based on DPI provides; we leave this to future work.

Limitations in service classification. Mobile services are classified into broad categories *e.g.*, *online news*, limiting our ability to distinguish outlets with different political leanings. Finer-grained classification could offer more insight but requires balancing analytical benefits against privacy risks.

7 Conclusions

We present a large-scale measurement study analyzing mobile service consumption and electoral outcomes across urban and suburban communes in France, using passive traffic data from a major MNO covering ∼31% of the French mobile market. Our Dirichlet regression framework provides evidence that mobile service usage patterns are associated with distinct signals of political alignment and reveals specific platform-party associations. Longitudinal analysis across 2019 and 2024 reveals strengthening relationships suggesting a growing role for passive network measurements as large-scale behavioral proxies for understanding social and political dynamics. Our findings are not aimed at campaign microtargeting, which would be both ethically inappropriate and infeasible from passive, aggregate measurements; they instead serve three audiences. Regulators and oversight bodies gain an independent, population-scale proxy for digital exposure, enabling monitoring of political information flows across platforms without platform cooperation, complementing Digital Services Act transparency obligations [33]. Political researchers gain a large-scale passive alternative to surveys for tracking digital-political alignment between elections. The networking and measurement community gains a new instrument for studying the digital information environment longitudinally, including cross-platform dynamics and coordinated campaigns.

Acknowledgments

This work was partially supported by the CoCo5G project (grant no. ANR-22-CE25-0016), funded by the French National Research Agency (ANR). We thank our shepherd and the anonymous reviewers for their careful and constructive feedback, which substantially improved this work.

References

[1] Esma Aïmeur, Sabrine Amri, and Gilles Brassard. 2023. Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining* 13, 1 (2023), 30.

[2] Yann Algan, Thomas Renault, and Hugo Subtil. 2025. La Fièvre parlementaire: ce monde où l'on catche! *Colère, polarisation et politique TikTok à l'Assemblée nationale. Note de recherche, Observatoire du Bien-être, CEPREMAP* (2025).

[3] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 2 (2017), 211–236.

[4] Hunt Allcott, Matthew Gentzkow, Winter Mason, Alexandra Wilkins, Pablo Barberá, Taylor Brown, et al. 2024. The effects of Facebook and Instagram on the 2020 election: A deactivation experiment. *Proceedings of the National Academy of Sciences* 121, 21 (2024), e2321584121.

[5] Arcep, Arcom, CGE, and ANCT. 2025. *Baromètre du numérique – Édition 2025.* Technical Report. Centre de Recherche pour l'Étude et l'Observation des Conditions de Vie (CREDOC). https://www.credoc.fr/publications/barometre-du-numerique-edition-2025 Survey conducted July–August 2024; $n = 4{,}066$, representative of the French population aged 12 and over.

[6] William Audureau. 2024. Snap elections in France: How it works, precedents, and consequences. https://www.lemonde.fr/en/les-decodeurs/article/2024/06/10/snap-elections-in-france-how-it-works-precedents-and-consequences_6674391_8.html Accessed: 2025-11-04.

[7] Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115, 37 (2018), 9216–9221.

[8] Bela Balassa and Marcus Noland. 1989. Revealed Comparative Advantage in Japan and the United States. *Journal of International Economic Integration* 4, 2 (1989), 8–22.

[9] Antoine Banks, Ernesto Calvo, David Karol, and Shibley Telhami. 2021. # polarizedfeeds: Three experiments on polarization, framing, and social media. *The International Journal of Press/Politics* 26, 3 (2021), 609–634.

[10] Filipe Batista e Silva, Sérgio Freire, Marcello Schiavina, Konštantín Rosina, Mario Alberto Marín-Herrera, Lukasz Ziemba, Massimo Craglia, Eric Koomen, and Carlo Lavalle. 2020. Uncovering temporal changes in Europe's population density patterns using a data fusion approach. *Nature Communications* 11, 1 (2020).

[11] BBC. 2024. YouTube prankster voted in as Cyprus MEP. https://www.bbc.com/news/articles/c4nnrwr72dqo. Accessed: 2025-11-17.

[12] Michael A Beam, Myiah J Hutchens, and Jay D Hmielowski. 2020. Facebook news and (de) polarization: Reinforcing spirals in the 2016 US election. In *Digital media, political polarization and challenges to democracy*. Routledge, 26–44.

[13] Márton Bene and Simon Kruschinski. 2021. Political advertising on Facebook. In *Campaigning on Facebook in the 2019 European Parliament Election: Informing, Interacting with, and Mobilising Voters*. Springer, 283–299.

[14] Davide Beraldo, Stefania Milan, Jeroen de Vos, Claudio Agosti, Bruno Nadalic Sotic, Rens Vliegenthart, Sanne Kruikemeier, Lukas P Otto, Susan AM Vermeer, Xiatong Chu, et al. 2021. Political advertising exposed: Tracking Facebook ads in the 2021 Dutch elections. *Internet Policy Review* 11, 1 (2021).

[15] Sreyasee Das Bhattacharjee, Bala Venkatram Balantrapu, William Tolone, and Ashit Talukder. 2017. Identifying extremism in social media with multi-view context-aware subset optimization. In *2017 IEEE international conference on big data (big data)*. IEEE, 3638–3647.

[16] Giuliano Bobba, Cristina Cremonesi, Moreno Mancosu, and Antonella Seddone. 2018. Populism and the gender gap: Comparing digital engagement with populist and non-populist Facebook pages in France, Italy, and Spain. *The International Journal of Press/Politics* 23, 4 (2018), 458–475.

[17] Diane Bolet. 2020. Local labour market competition and radical right voting: Evidence from France. *European Journal of Political Research* 59, 4 (2020), 817–841. https://doi.org/10.1111/1475-6765.12378

[18] Elisa Borry-Estrade. 2022. Elections françaises 2022 : une série d'initiatives dédiées sur Facebook, Instagram et WhatsApp pour aider les citoyens à lire et à décrypter l'information en ligne | À propos de Meta — about.fb.com. https://about.fb.com/fr/news/2022/02/elections-francaises-2022-une-serie-dinitiatives-dediees-sur-facebook-instagram-et-whatsapp-pour-aider-les-citoyens-a-lire-et-a-decrypter-linformation-en-ligne/. Accessed 08-04-2024.

[19] Shelley Boulianne, Karolina Koc-Michalska, and Bruce Bimber. 2020. Right-wing populism, social media and echo chambers in Western democracies. *New media & society* 22, 4 (2020), 683–699.

[20] Connor Brooke. [n. d.]. App Usage Peaks At 8 P.M. https://web.archive.org/web/20241202230228/https://www.business2community.com/mobile-apps/app-usage-peaks-8-p-m-01153062. Accessed: 2025-10-05.

[21] Christopher Buschow and Christian-Mathias Wellbrock. 2022. "Spotify for News"? User Perception of Subscription-Based Content Platforms for News Media. *Journalism and media* 4, 1 (2022), 1–15.

[22] Lorena Cano-Orón, Dafne Calvo, Guillermo López García, and Tomás Baviera. 2021. Disinformation in Facebook ads in the 2019 Spanish general election campaigns. *Media and Communication* 9 (2021).

[23] Qiang Cao, Xiaowei Yang, Jieqi Yu, and Christopher Palow. 2014. Uncovering large groups of active malicious accounts in online social networks. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*. 477–488.

[24] Manuel J Cartes-Barroso, Noelia García-Estévez, and Sandra Méndez-Muros. 2025. Attracting the Vote on TikTok: Far-Right Parties' Emotional Communication Strategies in the 2024 European Elections. *Journalism and Media* 6, 1 (2025), 33.

[25] Matteo Castiglioni, Diodato Ferraioli, and Nicola Gatti. 2020. Election Control in Social Networks via Edge Addition or Removal. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 02 (Apr. 2020), 1878–1885. https://doi.org/10.1609/aaai.v34i02.5556

[26] Daniele Cirulli, Antonio Desiderio, Giulio Cimini, and Fabio Saracco. 2025. Polarization and echo chambers in Reddit's political discourse. *arXiv preprint arXiv:2510.27467* (2025).

[27] CREDOC. 2019. *Baromètre du numérique – Rapport 2019.* Technical Report. Centre de Recherche pour l'Étude et l'Observation des Conditions de Vie (CREDOC). https://www.credoc.fr/publications/barometre-du-numerique-2019 Commissioned by Arcep, CGIET, ANCT. Survey conducted June 2019; $n = 2{,}259$, representative of the French population aged 12 and over.

[28] Gianmarco De Francisci Morales, Corrado Monti, and Michele Starnini. 2021. No echo in the chambers of political interactions on Reddit. *Scientific reports* 11, 1 (2021), 2818.

[29] Ministère de l'Intérieur. 2019. Résultats des élections européennes 2019. https://www.data.gouv.fr/fr/datasets/resultats-des-elections-europeennes-2019 Accessed on November 30, 2023.

[30] Ministère de l'Intérieur. 2024. Résultats des élections européennes du 9 juin 2024. https://www.data.gouv.fr/datasets/resultats-des-elections-europeennes-du-9-juin-2024/ Accessed on October 20, 2024.

[31] Wouter De Nooy and Jan Kleinnijenhuis. 2013. Polarization in the media during an election campaign: A dynamic network model predicting support and attack among political actors. *Political communication* 30, 1 (2013), 117–138.

[32] Ashok Deb, Luca Luceri, Adam Badaway, and Emilio Ferrara. 2019. Perils and challenges of social media and election manipulation analysis: The 2018 us midterms. In *Companion proceedings of the 2019 world wide web conference*. 237–247.

[33] Gabriele Di Bona, Emma Fraxanet, Björn Komander, Andrea Lo Sasso, Virginia Morini, Antoine Vendeville, Max Falkenberg, and Alessandro Galeazzi. 2024. Sampled datasets risk substantial bias in the identification of political polarization on social media. *arXiv preprint arXiv:2406.19867* (2024).

[34] Holger Döring. 2025. Advancing data infrastructures on democratic representation: an open-source version of ParlGov. https://doi.org/10.26092/elib/4362

[35] Holger Döring and Philip Manow. 2020. ParlGov 2020 Release. https://doi.org/10.7910/DVN/Q6CVHX

[36] Robert Epstein and Alex Flores. 2024. The Video Manipulation Effect (VME): A quantification of the possible impact that the ordering of YouTube videos might have on opinions and voting preferences. *PLoS One* 19, 11 (2024), e0303036.

[37] Ericsson. 2019. *Ericsson Mobility Report.* Technical Report. Ericsson. https://www.ericsson.com/en/reports-and-papers/mobility-report/reports

[38] Ericsson. 2024. *Ericsson Mobility Report.* Technical Report. Ericsson. https://www.ericsson.com/en/reports-and-papers/mobility-report/reports

[39] Zhihan Fang, Guang Wang, Shuai Wang, Chaoji Zuo, Fan Zhang, and Desheng Zhang. 2020. CellRep: Usage Representativeness Modeling and Correction Based on Multiple City-Scale Cellular Networks. In *The Web Conference (WWW)*. https://doi.org/10.1145/3366423.3380141

[40] Patrick Ferrucci, Toby Hopp, and Chris J Vargo. 2020. Civic engagement, social capital, and ideological extremity: Exploring online political engagement and political expression on Facebook. *New Media & Society* 22, 6 (2020), 1095–1115.

[41] Radio France. 2020. Radio France et Spotify signent un nouvel accord pour la distribution des podcasts Radio France. https://www.radiofrance.com/presse/radio-france-et-spotify-signent-un-nouvel-accord-pour-la-distribution-des-podcasts-radio. Accessed: 2025-11-09.

[42] Thomas Fujiwara, Karsten Müller, and Carlo Schwarz. 2024. The effect of social media on elections: Evidence from the United States. *Journal of the European Economic Association* 22, 3 (2024), 1495–1539.

[43] Angelo Furno, Marco Fiore, Razvan Stanica, Cezary Ziemlicki, and Zbigniew Smoreda. 2017. A Tale of Ten Cities: Characterizing Signatures of Mobile Traffic in Urban Areas. *IEEE Transactions on Mobile Computing* 16, 10 (2017), 2682–2696. https://doi.org/10.1109/TMC.2016.2637901

[44] Jason Gainous, Jason P Abbott, and Kevin M Wagner. 2021. Active vs. passive social media engagement with critical information: Protest behavior in two Asian countries. *The International Journal of Press/Politics* 26, 2 (2021), 464–483.

[45] James K Galbraith and J Travis Hale. 2008. State income inequality and presidential election turnout and outcomes. *Social Science Quarterly* 89, 4 (2008), 887–901.

[46] Kiran Garimella, Tim Smith, Rebecca Weiss, and Robert West. 2021. Political polarization in online news consumption. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15. 152–162.

[47] Amory Gethin, Clara Martínez-Toledano, and Thomas Piketty. 2022. Brahmin left versus merchant right: Changing political cleavages in 21 western democracies, 1948–2020. *The Quarterly Journal of Economics* 137, 1 (2022), 1–48. https://doi.org/10.1093/qje/qjab036

[48] Rahul Goel, Angelo Furno, and Rajesh Sharma. 2025. Predicting socio-economic well-being using mobile apps data: a case study of France. *Journal of Computational Social Science* 8, 4 (2025), 97.

[49] Sandra González-Bailón, David Lazer, Pablo Barberá, Meiqing Zhang, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, Deen Freelon, Matthew Gentzkow, Andrew M. Guess, Shanto Iyengar, Young Mie Kim, Neil Malhotra, Devra Moehler, Brendan Nyhan, Jennifer Pan, Carlos Velasco Rivera, Jaime Settle, Emily Thorson, Rebekah Tromble, Arjun Wilkins, Magdalena Wojcieszak, Chad Kiewiet de Jonge, Annie Franco, Winter Mason, Natalie Jomini Stroud, and Joshua A. Tucker. 2023. Asymmetric ideological segregation in exposure to political news on Facebook. *Science* 381, 6656 (2023), 392–398. https://doi.org/10.1126/science.ade7138 arXiv:https://www.science.org/doi/pdf/10.1126/science.ade7138

[50] Jarod Govers, Philip Feldman, Aaron Dant, and Panos Patros. 2023. Down the rabbit hole: Detecting online extremism, radicalisation, and politicised hate speech. *Comput. Surveys* 55, 14s (2023), 1–35.

[51] Ted Grover and Gloria Mark. 2019. Detecting potential warning behaviors of ideological radicalization in an alt-right subreddit. In *Proceedings of the international AAAI conference on web and social media*, Vol. 13. 193–204.

[52] The Guardian. 2018. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election. [Accessed 04-06-2026].

[53] Sergei Guriev, Nikita Melnikov, and Ekaterina Zhuravskaya. 2021. 3G internet and confidence in government. *The Quarterly Journal of Economics* 136, 4 (2021), 2533–2613.

[54] Taru Haapala. 2023. Disrupting European politics? Rhetorical analysis of MEP Manon Aubry's politicising strategies in Twitter. *Journal of Contemporary European Studies* (2023), 1–16.

[55] Katherine Haenschen and Jay Jennings. 2019. Mobilizing millennial voters with targeted internet advertisements: A field experiment. *Political Communication* 36, 3 (2019), 357–375.

[56] Chris Hanretty. 2021. Forecasting multiparty by-elections using Dirichlet regression. *International Journal of Forecasting* 37, 4 (2021), 1666–1676.

[57] James Hawdon, Shyam Ranganathan, Scotland Leman, Shane Bookhultz, and Tanushree Mitra. 2020. Social media use, political polarization, and social capital: is social media tearing the US apart?. In *International Conference on Human-Computer Interaction*. Springer, 243–260.

[58] Raffael Heiss, Johannes Knoll, and Jörg Matthes. 2020. Pathways to political (dis-) engagement: Motivations behind social media use and the role of incidental and intentional exposure modes in adolescents' political engagement. *Communications* 45, s1 (2020), 671–693.

[59] Raffael Heiss and Jörg Matthes. 2020. Stuck in a nativist spiral: Content, selection, and effects of right-wing populists' communication on Facebook. *Political Communication* 37, 3 (2020), 303–328.

[60] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, et al. 2013. *An introduction to statistical learning*. Vol. 112. Springer.

[61] Yonghwan Kim, Shih-Hsien Hsu, and Homero Gil de Zúñiga. 2013. Influence of social media use on discussion network heterogeneity and civic engagement: The moderating role of personality traits. *Journal of communication* 63, 3 (2013), 498–516.

[62] Nikolai Kirkizh, Roberto Ulloa, Sebastian Stier, and Jürgen Pfeffer. 2024. Predicting political attitudes from web tracking data: a machine learning approach. *Journal of Information Technology & Politics* 21, 3 (2024).

[63] Sanne Kruikemeier. 2014. How political candidates use Twitter and the impact on votes. *Computers in human behavior* 34 (2014), 131–139.

[64] Emily Kubin and Christian Von Sikorski. 2021. The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association* 45, 3 (2021), 188–206.

[65] Anders Olof Larsson. 2023. The rise of Instagram as a tool for political communication: A longitudinal study of European political parties and their followers. *New Media & Society* 25, 10 (2023), 2744–2762. https://doi.org/10.1177/14614448211034158 arXiv:https://doi.org/10.1177/14614448211034158

[66] I. Lella, C. Ciobanu, M. Theocharidou, E. Magonara, A. Malatras, R. Svetozarov Naydenov, and E. Tsekmezoglou. 2023. ENISA threat landscape 2023 – July 2022 to June 2023. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023

[67] Damien Leloup. 2022. Sur Facebook, La République En Marche et ses publicités un peu trop bien ciblées. https://www.lemonde.fr/pixels/article/2022/04/02/sur-facebook-la-republique-en-marche-et-ses-publicites-un-peu-trop-bien-ciblees_6120226_4408996.html

[68] Corentin Lesueur. 2023. La Très Chère publicité Facebook des Eurodéputés RN, Entre dépenses record et Prestataires Amis. https://www.lemonde.fr/politique/article/2023/06/02/la-tres-chere-publicite-facebook-des-eurodeputes-rn-entre-depenses-record-et-prestataires-amis_6175830_823448.html

[69] Ro'ee Levy. 2021. Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review* 111, 3 (2021), 831–870.

[70] Michael S Lewis-Beck and Mary Stegmaier. 2014. US presidential election forecasting: Introduction. *PS: Political Science & Politics* 47, 2 (2014), 284–288.

[71] Michael S Lewis-Beck and Charles Tien. 2018. Candidates and campaigns: How they alter election forecasts. *Electoral Studies* 54 (2018), 303–308.

[72] Philipp Lorenz-Spreen, Lisa Oswald, Stephan Lewandowsky, and Ralph Hertwig. 2023. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature human behaviour* 7, 1 (2023), 74–101.

[73] Luca Luceri, Eric Boniardi, and Emilio Ferrara. 2024. Leveraging large language models to detect influence campaigns on social media. In *Companion Proceedings of the ACM Web Conference 2024*. 1459–1467.

[74] Matthew C MacWilliams. 2015. Forecasting congressional elections using Facebook data. *PS: Political Science & Politics* 48, 4 (2015), 579–583.

[75] Marco Johannes Maier. 2014. *DirichletReg: Dirichlet Regression for Compositional Data in R*. Research Report Series / Department of Statistics and Mathematics 125. WU Vienna University of Economics and Business, Vienna. http://epub.wu.ac.at/4077/

[76] Orlando E Martínez-Durive, Stefanos Bakirtzis, Cezary Ziemlicki, Jie Zhang, Ian James Wassell, and Marco Fiore. 2024. DeepMEND: Reliable and Scalable Network Metadata Geolocation from Base Station Positions. In *2024 21st Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 1–9.

[77] Orlando E Martínez-Durive, Theo Couturieux, Cezary Ziemlicki, and Marco Fiore. 2022. VoronoiBoost: Data-driven probabilistic spatial mapping of mobile network metadata. In *2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 100–108.

[78] Juan Carlos Medina Serrano, Orestis Papakyriakopoulos, and Simon Hegelich. 2020. Exploring political ad libraries for online advertising transparency: lessons from Germany and the 2019 European elections. In *International conference on social media and society*. 111–121.

[79] Nikita Melnikov. 2021. Mobile Internet and Political Polarization. (2021). Working paper. SSRN: https://ssrn.com/abstract=3937760.

[80] Sachit Mishra, Zbigniew Smoreda, and Marco Fiore. 2022. Second-level Digital Divide: A Longitudinal Study of Mobile Traffic Consumption Imbalance in France. In *Proceedings of the ACM Web Conference 2022* (Virtual Event, Lyon, France) *(WWW '22)*. Association for Computing Machinery, New York, NY, USA, 2532–2540.

[81] Sachit Mishra, André F Zanella, Orlando E Martínez-Durive, Diego Madariaga, Cezary Ziemlicki, and Marco Fiore. 2024. Characterizing 5g adoption and its impact on network traffic and mobile service consumption. In *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*. IEEE, 1531–1540.

[82] Le Monde. 2017. «Fausses informations»: Facebook annonce avoir supprimé 30000 comptes en France — lemonde.fr. https://www.lemonde.fr/pixels/article/2017/04/13/facebook-lance-une-campagne-publicitaire-contre-les-fausses-informations_5110899_4408996.html. [Accessed 08-04-2024].

[83] Karsten Müller and Carlo Schwarz. 2021. Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association* 19, 4 (2021), 2131–2167.

[84] Marie Neihouser and Julien Figeac. 2025. Bringing the campaign closer to the voters: Facebook in partisan-managed campaigning in France. *Communications* 50, 1 (2025), 85–104.

[85] Helmut Norpoth. 2014. The electoral cycle. *PS: Political Science & Politics* 47, 2 (2014), 332–335.

[86] Mariam Nouh, Jason RC Nurse, and Michael Goldsmith. 2019. Understanding the radical mind: Identifying signals to detect extremist content on twitter. In *2019 IEEE international conference on intelligence and security informatics (ISI)*. IEEE, 98–103.

[87] The National Institute of Statistics and Economic Studies. 2021. A new definition of rural to better reflect the realities of territories and their transformations. https://www.insee.fr/fr/statistiques/5039991?sommaire=5040030. Accessed: 2024-10-09.

[88] The National Institute of Statistics and Economic Studies. 2023. Base du dossier complet. https://www.insee.fr/fr/statistiques/5359146. Accessed: 2023-01-09.

[89] Orange. 2019. Annual financial report 2019. Universal registration document. https://www.orange.com/sites/orangecom/files/2020-09/ORANGE%5FURD%5F2019%5FVA.pdf Accessed: 2023-11-30.
[90] Orange. 2024. *Second Quarter 2024 Results: Databook (Key Performance Indicators).* Financial databook (spreadsheet). Orange S.A. https://www.orange.com/en/finance/financial-and-extra-financial-information France segment, "France – KPIs" sheet, mobile accesses by commercial offer: Sosh = 4,879 thousand at 31 March 2024 (Q1 2024). Accessed June 2026.
[91] Orange. 2025. Document d'enregistrement universel 2024 — Rapport Financier Annuel. https://gallery.orange.com/element?id=404737 Filed with the AMF on 27 March 2025. Mobile market share data in §1.4.1 (France)..
[92] Manuel G. Pascual. 2024. European elections: Which parties are trying to influence the vote through Facebook or Instagram? *El Pais* (April 2024). https://english.elpais.com/technology/2024-04-10/european-elections-which-parties-are-trying-to-influence-the-vote-through-facebook-or-instagram.html
[93] Concha Pérez-Curiel, Rubén Rivas-de Roca, and Mar García-Gordillo. 2021. Impact of Trump's digital rhetoric on the US elections: A view from worldwide far-right populism. *Social sciences* 10, 5 (2021), 152.
[94] Francesco Pierri. 2023. Political advertisement on Facebook and Instagram in the run up to 2022 Italian general election. In *Proceedings of the 15th ACM Web Science Conference 2023*. 13–22.
[95] POLITICO. 2024. The far-right 'hooligan' disrupting Spanish politics. https://www.politico.eu/article/spain-politics-luis-perez-salf-eu-vote-donald-trump-national-rally-germany-france/. Accessed: 2025-11-17.
[96] European Union Open Data Portal. 2022. Using Open Data to Understand Politics. https://data.europa.eu/en/publications/datastories/using-open-data-understand-politics Accessed: 2025-01-21.
[97] José N Portela and Marcelo S Alencar. 2008. Cellular coverage map as a Voronoi diagram. *Journal of Communication and Information Systems* 23, 1 (2008).
[98] Associated Press. 2024. Controversial podcast host Joe Rogan signs a new deal with Spotify for up to a reported 250 million. https://apnews.com/article/joe-rogan-spotify-deal-76fa0e2c9d4b137f510428528ea6226b. Accessed: 2025-11-09.
[99] The French public data platform. 2023. French municipal administrative division from OpenStreetMap. https://www.data.gouv.fr/datasets/decoupage-administratif-communal-francais-issu-d-openstreetmap. Accessed: 2024-10-09.
[100] Fabio Ricciato and Angelo Coluccia. 2021. On the estimation of spatial density from mobile network operator data. *IEEE Trans. on Mobile Computing* (2021).
[101] Alexander James Richardson. 2020. *Estimating the Impact of Political Propaganda on Reddit Users' Political Opinions.* Georgetown University.
[102] Álvaro Sánchez-García, Toni Rodon, and Maria Delgado-García. 2025. Where has everyone gone? Depopulation and voting behaviour in Spain. *European Journal of Political Research* 64, 1 (2025), 296–319. https://doi.org/10.1111/1475-6765.12702
[103] Piotr Sapiezynski, Levi Kaplan, Alan Mislove, and Aleksandra Korolova. 2024. On the Use of Proxies in Political Ad Targeting. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–31.
[104] Max Schaub and Davide Morisi. 2020. Voter mobilisation in the echo chamber: Broadband internet and the rise of populism in Europe. *European Journal of Political Research* 59, 4 (2020), 752–773.
[105] Sandy Schumann, Diana Boer, Katja Hanke, and James Liu. 2021. Social media use and support for populist radical right parties: assessing exposure and selection effects in a two-wave panel study. *Information, communication & society* 24, 7 (2021), 921–940.
[106] Alexander Shevtsov, Maria Oikonomidou, Despoina Antonakaki, Polyvios Pratikakis, and Sotiris Ioannidis. 2023. What Tweets and YouTube comments have in common? Sentiment and graph analysis on data related to US elections 2020. *PLOS One* 18, 1 (2023), e0270542.
[107] Alexandra A Siegel, Evgenii Nikitin, Pablo Barberá, Joanna Sterling, Bethany Pullen, Richard Bonneau, Jonathan Nagler, Joshua A Tucker, et al. 2021. Trumping hate on Twitter? Online hate speech in the 2016 US election campaign and its aftermath. *Quarterly Journal of Political Science* 16, 1 (2021), 71–104.
[108] Timo Sipma and Marcel Lubbers. 2020. Contextual-level unemployment and support for radical-right parties: A meta-analysis. *Acta Politica* 55 (2020), 1–27. https://doi.org/10.1057/s41269-018-0120-2
[109] Richard Stafford, Jenny Alexander, Stephen Axon, Zach Boakes, Elena Cantarello, Abigail R Croker, Marin Cvitanović, Victoria Dominguez Almela, Tilak Ginige, Natalie Harris, et al. 2025. Evaluating post 2024 election scenarios for the UK based on political party manifestos. *PLoS One* 20, 5 (2025), e0324942.
[110] New York Times. 2024. The gender election. https://web.archive.org/web/20251109011629/https://www.nytimes.com/2024/10/23/podcasts/the-daily/gender-election.html Accessed: 2025-11-04.
[111] New York Times. 2024. Why Gender May Be the Defining Issue of the Election. https://web.archive.org/web/20250213021232/https://www.nytimes.com/2024/10/23/us/politics/harris-trump-election-gender.html Accessed: 2025-11-04.
[112] New York Times. 2025. Election Uncertainty Looms Over Hiring. https://www.nytimes.com/2024/11/01/business/election-uncertainty-jobs-economy.html Accessed: 2025-11-04.
[113] New York Times. 2025. See Which Neighborhoods and Groups Drove Mamdani's Success in the N.Y.C. Primary. https://www.nytimes.com/interactive/2025/06/25/nyregion/nyc-mayor-election-results-map-mamdani-cuomo.html Accessed: 2025-11-04.
[114] Adam Tsakalidis, Symeon Papadopoulos, Alexandra I Cristea, and Yiannis Kompatsiaris. 2015. Predicting elections for multiple countries using Twitter and polls. *IEEE Intelligent Systems* 30, 2 (2015), 10–17.
[115] Andranik Tumasjan, Timm O Sprenger, Philipp G Sandner, and Isabell M Welpe. 2011. Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social science computer review* 29, 4 (2011), 402–418.
[116] Iñaki Ucar et al. 2021. News or social media? Socio-economic divide of mobile service consumption. *Journal of The Royal Society Interface* 18, 185 (2021), 20210350. https://doi.org/10.1098/rsif.2021.0350 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2021.0350
[117] Jay J Van Bavel, Steve Rathje, Elizabeth Harris, Claire Robertson, and Anni Sternisko. 2021. How social media shapes polarization. *Trends in Cognitive Sciences* 25, 11 (2021), 913–916.
[118] Pavlos Vasilopoulos, Haley McAvay, and Sylvain Brouard. 2021. Residential context and voting for the far right: The impact of immigration and unemployment on the 2017 French presidential election. *Political Behavior* 44, 4 (2021), 1703–1730. https://doi.org/10.1007/s11109-021-09676-z
[119] David A Walker. 2006. Predicting presidential election results. *Applied Economics* 38, 5 (2006), 483–490.
[120] Zheng Wei, Xian Xu, and Pan Hui. 2024. Digital democracy at crossroads: A meta-analysis of web and ai influence on global elections. In *Companion Proceedings of the ACM Web Conference 2024*. 1126–1129.
[121] Sanford Weisberg. 2005. *Applied linear regression*. Vol. 528. John Wiley & Sons.
[122] Liang Wu, Fred Morstatter, Kathleen M. Carley, and Huan Liu. 2019. Misinformation in Social Media: Definition, Manipulation, and Detection. *SIGKDD Explor. Newsl.* 21, 2 (Nov. 2019), 80–90. https://doi.org/10.1145/3373464.3373475

A Ethics

Our work hinges upon mobile network traffic generated by users of a nationwide cellular infrastructure and collected by the MNO for network management and research purposes. The measurements were processed within a secure platform at the MNO premises and by personnel of the operator so as to generate anonymous aggregates of service-specific traffic volumes at the antenna level. Such aggregates merge information from tens to hundreds of users, ensuring that no single data subject can be re-identified. Also, the aggregates cancel out all personal identifiers (*e.g.*, the International Mobile Subscriber Identifier (IMSI)) or sensitive data (*e.g.*, visited cell locations, consumed mobile services) about individual users. The data collection and processing were approved by the Data Protection Officer (DPO) of the MNO in the context of a collaborative research project. The researchers involved in the work presented in this paper only had access to the aggregates described above, which do not configure as personal data in the EU GDPR acceptation. As a result, our work does not raise any ethical or privacy issues.

B Mobile Network Data

This section describes the technical collection process for the mobile service measurements. Figure 10 provides a simplified view of the target mobile network architecture, illustrating the evolution and integration of technologies from 2G to 5G Non-Standalone (NSA) as well as the probe deployment strategy adopted by the MNO. The architecture comprises the Radio Access Network (RAN), providing wireless connectivity to the end users, and the Core Network (CN), managing mobility, authentication, and data routing. Traffic destined for external networks traverses gateway nodes that interface with public data networks, primarily the Internet. Passive

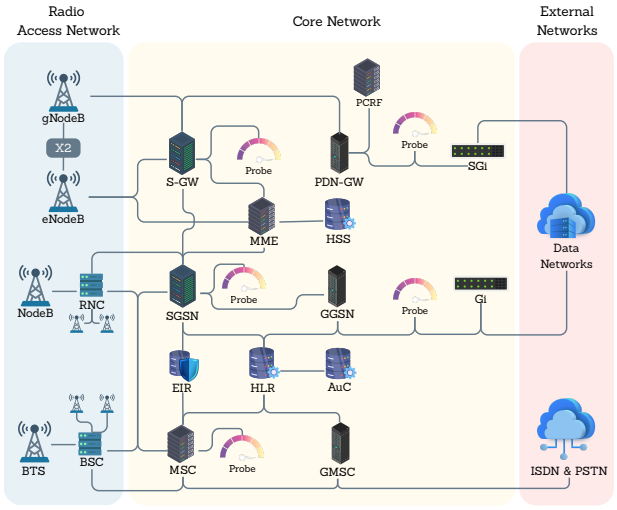


Figure 10: Diagram of mobile network architecture highlighting the main nodes and interconnections across 2G, 3G, 4G, and 5G NSA generations.

measurement probes are deployed by the MNO at two critical locations: at gateway interfaces (Gi Interface (Gi)/SGi Interface (SGi)) to capture traffic flows, and at signaling interfaces (S1) to capture control-plane events, enabling comprehensive monitoring of network performance and user behavior.

In Long Term Evolution (LTE), mobile networks adopt an all-IP architecture, integrating all communication into packet-switched domains. This also centralizes metadata collection at key network interfaces. Therefore, the Serving Gateway (S-GW) acts as a mobility anchor, routing data between base stations (Evolved Node B (eNodeB)) and external networks, maintaining metadata on handovers, session management, and data usage. The Packet Data Network Gateway (PDN-GW) connects mobile users to external IP networks and collects detailed metadata on traffic flows. Probes deployed at the SGi interface between the PDN-GW and external networks capture all data exchange, categorizing traffic by application type and providing comprehensive insights into service demand and network load.

Control-plane metadata is collected by probes at the S1 interface connecting base stations to the Mobility Management Entity (MME). These probes capture signaling events related to user authentication, mobility events, and paging requests. The collected Network Signaling Data (NSD) tracks attach and detach procedures, tracking area updates, and handovers. This signaling data enables precise attribution of traffic flows to serving base stations and provides insights into user mobility patterns, essential for both network optimization and research applications.

As far as the 5G portion of the network is concerned, that is only relevant for the 2024 elections since no 5G service was active in 2019. Also, the MNO only ran 5G in NSA configuration in 2024, as Standalone (SA) commercial services were launched only in early 2025. Therefore, 5G traffic flows in 2024 through existing LTE core network infrastructure, which lets the same passive probe deployment capture 5G Next Generation Node B (gNodeB) demands.

As expected, the network deployment also evolved between the two campaigns, growing from roughly 17,600 base-station sites in 2019 to about 25,900 in 2024, alongside the introduction of 5G NSA, with an increase of around 3,000 new sites in the target communes. A denser deployment affects only the absolute traffic volume captured at each site, with no impact on our analysis. Our per-site traffic volumes are redistributed at the commune level (see Figure 3), which is the spatial granularity of our analysis; the addition of new sites therefore translates to more sites contributing to the same commune relative to 2019, changing the absolute traffic volume attributed to each commune. However, the sRCA normalization is per election and at the commune level, and higher or lower absolute traffic volumes at a commune are transformed into relative, per-election values that do not propagate across campaigns.

To accurately measure individual mobile services, proprietary DPI classifiers run on the gateway probes, associating traffic flows with specific applications (*e.g.*, YouTube, WhatsApp, Facebook) based on packet headers, payload signatures, and behavioral patterns. These three families correspond, respectively, to header-based features (such as the Server Name Indication exposed in the TLS handshake), signature matching, and flow-level heuristics, which is what allows classification to operate on largely encrypted traffic. These classifiers enable granular analysis of network traffic, allowing detailed understanding of per-service demand. The resulting metadata consists of aggregated traffic volumes recorded at each antenna over time, accompanied by geolocation information. Beyond this description, the classifiers are the operator's production DPI systems, and their internal methodology and per-service ground-truth accuracy are not disclosable under the commercial-sensitivity and NDA terms that govern the data. Our per-app measurements therefore rely on the classifiers attributing traffic correctly, and we cannot independently validate that accuracy. We note, however, that these are the same classifiers the operator runs in production at national scale, a performance bar we consider well beyond what our aggregate, commune-level analysis requires.

To situate our measurements within the broader evolution of mobile consumption, we recall that global mobile data traffic grew from 38 to 157 EB per month between 2019 and 2024, with video rising from 63% to 74% of the total and average data per smartphone in Western Europe more than doubling to 23 GB per month [37, 38]. Our data reflect this expansion: the set of tracked services grew from 22 in 2019 to 27 in 2024, as platforms such as TikTok and Disney+ gained prominence (Table 1). Because the operator's classifiers and the absolute traffic volumes are not directly comparable across the two campaigns, we do not contrast raw demands across years; instead, we characterize each service through its commune-level sRCA (Section 3.3). Appendix E illustrates the resulting spatial consumption patterns across both elections for a representative service (Figure 13), and Section 5.3 quantifies how the per-service associations with vote shift between 2019 and 2024.

Table 3: Smartphone ownership in France by demographic group. Source: CREDOC, Baromètre du numérique [5, 27].

Dimension	2019	2024	Trend
Overall	77%	91%	+14 pp
Low income	72%	90%	+18 pp
High income	85%	95%	+10 pp
Income gap	13 pp	5 pp	halved
Age 18–24	98%	98%	stable
Age 70+	44%	70%	+26 pp
Rural communes	71%	84%	+13 pp
Paris agglomeration	83%	95%	+12 pp

C Data and Code Availability

To allow reproducibility, the aggregated data and code used in this study are available in a public GitHub repository[4], with the exception of the raw mobile traffic data, which are subject to a non-disclosure agreement (NDA) with the MNO and cannot be redistributed. Specifically, the repository contains:

- The electoral results obtained from the French Ministry of the Interior's open data portal [29, 30].
- The socioeconomic factors used: median income, unemployment rate, and population age structure [88]. The French public data platform does not link the specific snapshots, as these are replaced by newer versions; we therefore include our own copies, downloaded during the study period.
- The geospatial data, including the boundaries of each commune, their classification as urban, suburban, or rural [87], and the city boundaries used in our spatial analysis. This information is obtained from the French public data platform [99], and the specific snapshot used in this study is included in the repository.
- The sRCA values for each mobile service and commune, which are the main features used in the analysis. These values are derived from the raw mobile traffic data and the aggregated traffic volumes at the commune level.

D Sample Representativeness

The MNO operates under two brands: its main brand and a low-cost subsidiary, together serving approximately 22 million mobile lines (of which 3.9 million in 2019 and 4.8 million in 2024 under the low-cost brand), representing a 31.9% share of the French mobile market, a share that remained at 31.5% in 2024 [89–91]. Smartphone ownership in France is broadly uniform across income groups (72–85%) and age cohorts (80–98% for working-age adults; [5, 27]). With near-uniform penetration across the relevant demographics, the MNO's ∼31% mobile market share is unlikely to be severely skewed toward any particular socioeconomic group. The main known residual bias is toward older age groups (smartphone ownership was 44% for those aged 70+ in 2019, rising to 70% by 2024; Table 3), a pattern expected to affect all MNOs equally.

Figure 11 illustrates what the urban and suburban commune selection looks like at intra-city resolution, using as examples LinkedIn and Facebook sRCA values in 2019. For each service, the center map shows the full set of selected communes across metropolitan France, while the surrounding panels zoom into four metropolitan areas, revealing commune-level heterogeneity within each city. The two services display distinct but coherent spatial signatures, professional districts for LinkedIn and broad residential coverage for Facebook, confirming both the spatial resolution of the dataset and that service-level demand varies meaningfully within the selected communes.

E Representativeness of sRCA

Figure 12 illustrates the effect of transforming from Raw traffic distribution for Facebook and WhatsApp traffic to RCA and to sRCA. The raw traffic map (left) conflates two confounds: *i*) the intrinsic byte footprint of the service and *ii*) the population density of the commune, making high-traffic communes indistinguishable from communes with genuine relative preference for the service. Raw values are shown in anonymized units (multiplied by a constant 10^k per the NDA). The RCA map (center) normalizes away both confounds simultaneously: the numerator removes the service's national byte footprint and the denominator removes the commune's total traffic, revealing genuine spatial heterogeneity in service preference. The sRCA map (right) applies the symmetric rescaling to $(-1, 1)$, removing the asymmetry between advantage and disadvantage in raw RCA values and making the index comparable across services with very different traffic volumes.

Figure 13 shows commune-level sRCA patterns for a representative service (Instagram) across the six largest French metropolitan areas, comparing 2019 (top row) and 2024 (bottom row). Vertical comparison tracks the stability of consumption patterns across elections in a city; horizontal comparison reveals how the spatial footprint varies across cities.

F Political Science Literature

Table 4 places our work in the context of prior research on digital media and elections, comparing each study along its objective, dataset and identification strategy, coverage, and key finding. These studies span diverse designs, from survey panels and web-tracking to field and natural experiments, and most examine a single platform, predominantly in the United States. Where they rely on surveys, platform APIs, or single-platform traces, our study instead draws on passive, unsampled MNO traffic spanning tens of services and two consecutive elections at commune-level granularity, offering a complementary, population-scale vantage.

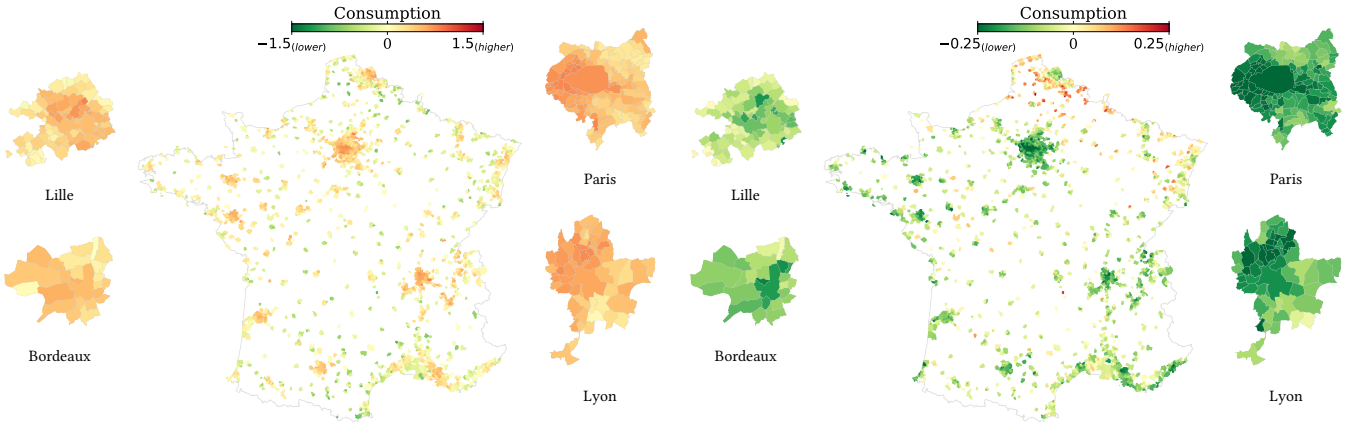
G Time-of-Day Sensitivity

Our analysis uses evening and night traffic (weekday `20:00–07:00`), when subscribers are most likely at home, so that demand corresponds to the resident population that votes in each commune (Section 3.2). To verify that our results do not depend on this choice, we recompute the sRCA of every service and refit the Dirichlet regression over four alternative weekday windows: the full day, the complementary daytime window (`07:00–20:00`), and its morning (`07:00–13:00`) and afternoon (`13:00–20:00`) halves. Table 5 reports, for each window, the relative gain in mean predictive correlation ρ of the *All* model over the *Socioeconomic* baseline, the

[4] https://github.com/nds-group/digital-consumption-political-alignment-france

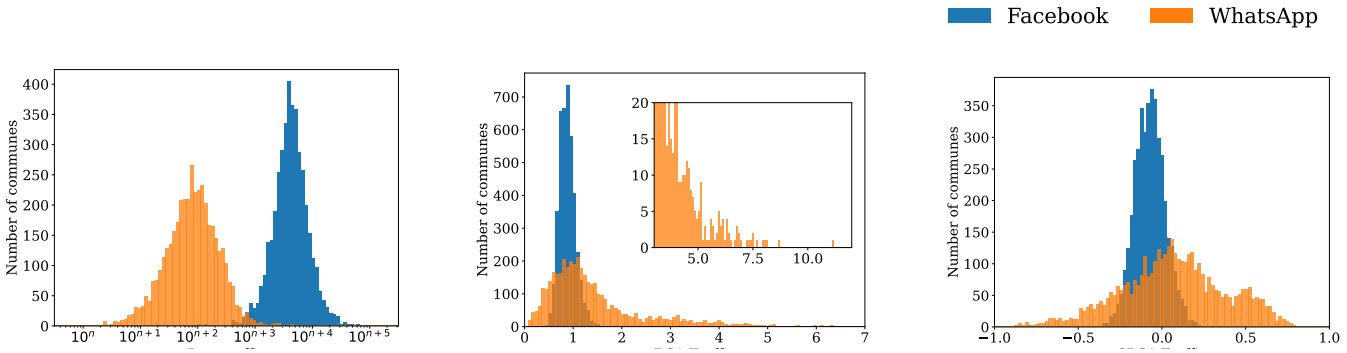
Study	Objective	Dataset (identification)	Coverage	Key finding
Allcott & Gentzkow (2017) [3]	Measure prevalence and persuasive effect of fake news	YouGov survey panel (~1,200) + web-traffic data + Facebook post database (observational)	US, 2016 Presidential	Fake-news exposure was widespread, but its persuasive effect on votes was limited
Schaub & Morisi (2020) [104]	Test whether broadband internet causally increases populist support	Municipal broadband coverage + survey data (quasi-exp., IV)	Italy & Germany, national elections (~2013–2018)	Broadband access increased populist party support
Guriev, Melnikov & Zhuravskaya (2021) [53]	Estimate effect of 3G expansion on confidence in government	Gallup World Poll + GSMA 3G coverage maps (quasi-exp., IV)	116 countries, 2008–2017	3G expansion reduced confidence in government
Müller & Schwarz (2021) [83]	Link right-wing Facebook activity to hate-crime incidence	AfD Facebook posts + German hate-crime registry + platform-outage instrument (quasi-exp., IV)	Germany, 2010–2017	Higher local Facebook activity increased hate-crime incidence
Melnikov (2021) [79]	Estimate causal effect of 3G mobile internet on political polarization	Gallup Daily Poll (1.76M individuals, 31,499 ZIP codes) + 3G rollout (quasi-exp., DiD)	US, 2008–2017	3G mobile internet increased partisan polarization
Levy (2021) [69]	Estimate the effect of social-media news exposure on news consumption and polarization	Facebook field experiment, randomized news-outlet subscriptions (field experiment)	US, ~17,000 users, 2018	Counter-attitudinal news exposure reduced out-party animosity; the feed algorithm limited such exposure
González-Bailón et al. (2023) [49]	Quantify ideological segregation in Facebook news exposure	Facebook internal data, ~208M US users (Meta Content Library; platform-internal)	US, 2020 election period	Conservative audiences were disproportionately exposed to misinformation
Allcott et al. (2024) [4]	Measure causal effect of Facebook/Instagram deactivation on political behavior	Randomized deactivation experiment (19,857 Facebook + 15,585 Instagram users; RCT)	US, 2020 Presidential	Deactivation reduced online participation but did not significantly change vote choice
Fujiwara, Müller & Schwarz (2024) [42]	Estimate causal effect of Twitter adoption on vote preferences	Twitter staggered county rollout + ANES survey (quasi-exp., IV)	US, elections 2006–2018	Twitter adoption shifted moderate voters toward liberal positions
Kirkizh et al. (2024) [62]	Predict individual political attitudes from web browsing	Web-tracking data (19M visits, 1,003 users) + survey (observational + ML)	Germany, no specific election	Browsing patterns predict political attitudes via machine learning
This paper (2026)	**Characterize associations between multi-platform mobile consumption and multiparty political alignment; quantify the gain over socioeconomic baselines**	**Passive MNO traffic, two brands (~22M mobile lines, ~31% market share, no sampling; observational)**	**France, urban & suburban communes (≈two-thirds of voters), 2019 & 2024 EP elections**	**Multi-platform consumption adds explanatory power beyond socioeconomic indicators, with stable and shifting platform–party associations across two elections**

Table 4: Comparison of prior work on digital media and elections along objective, dataset and identification strategy, coverage, and key finding. Our study is the only one that combines passive, platform-agnostic MNO traffic, multi-platform coverage across all application categories, full subscriber-base sampling, and two consecutive elections at commune-level granularity, without requiring platform cooperation.



(a) LinkedIn (b) Facebook

Figure 11: sRCA of (a) LinkedIn and (b) Facebook across urban and suburban communes (center map of each panel) and four metropolitan areas (2019). Commune boundaries are visible at city scale, confirming the spatial resolution of the dataset; the two services exhibit distinct spatial signatures within the same selected communes.



(a) Raw traffic **(b) RCA** **(c) sRCA**

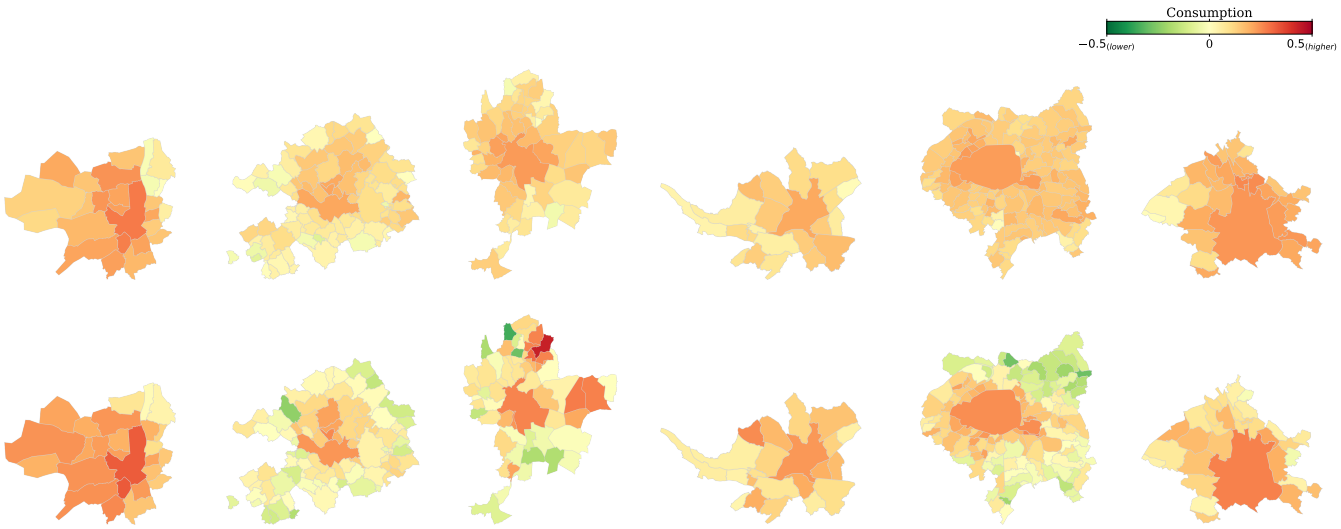
Figure 12: Effect of the Raw to RCA and sRCA transformation for Facebook and WhatsApp traffic across French communes. Raw traffic (left) reflects population density more than service preference; values are in anonymized units ($\times 10^n$) per the NDA. RCA (center) normalizes by commune size and national service share, removing both confounds. sRCA (right) symmetrizes the index to $(-1, 1)$, enabling direct cross-service comparison.

metric used in the main results; since the baseline is independent of the time window, this gain isolates the mobile-service contribution.

Table 5: Relative gain in mean predictive correlation ρ of the *All* model over the *Socioeconomic* baseline (%), by weekday window. For 2019, the value excludes the low-baseline *Envie d'Europe* list; the value including it is in parentheses.

Window	2019	2024
Full day (00–24)	11.2% (33.5%)	31.3%
Daytime (07–20)	11.2% (34.5%)	31.1%
Morning (07–13)	10.4% (32.9%)	29.6%
Afternoon (13–20)	10.8% (33.1%)	30.1%
Residential (20–07), adopted	9.6% (26.2%)	28.2%

Mobile services improve the predictive correlation under *every* window, in both elections, so the result is not an artifact of the hour selection: across the five windows the gain spans at most 1.6 percentage points in 2019 and 3.1 in 2024. The residential window we adopt yields the *smallest* gain of the five, so our reported results are a conservative lower bound rather than a window chosen to maximize fit. We retain the residential window on validity grounds, since evening and night traffic corresponds to residents at home, matching where they vote, whereas daytime windows mix in workplace and commuting geography. For 2019, these figures exclude the *Envie d'Europe* list, whose very low socioeconomic baseline ($\rho \approx 0.21$) turns a modest absolute improvement into an outlying relative gain; including it raises the 2019 values (in parentheses in Table 5) but leaves the ordering across windows unchanged. The 2024 election has no comparable low-baseline list.



(a) Bordeaux (b) Lille (c) Lyon (d) Nantes (e) Paris (f) Toulouse

Figure 13: Instagram sRCA across six French metropolitan areas in 2019 (top) and 2024 (bottom).

H Model Fit Statistics: AIC, BIC, and Log-Likelihood

The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are likelihood-based model-selection criteria that balance goodness of fit against model complexity: both reward a higher log-likelihood while penalizing the number of parameters p, with BIC imposing a stricter penalty that scales with the sample size. Table 6 reports the AIC, BIC, and log-likelihood for all predictor sets (Table 2) across both elections; lower (more negative) AIC and BIC indicate better fit, while a higher log-likelihood indicates a closer fit to the observed vote shares. Three findings complement the Adj. R^2 results in Section 5.1.

First, the AIC gain from adding mobile services to the *Socioeconomic* baseline nearly doubles from 2019 (ΔAIC = −6,548) to 2024 (ΔAIC = −12,537), consistent with the growing digital contribution in Figure 7. Second, *Social Media* alone crosses a qualitative threshold between elections: in 2019 it is outperformed by *Socioeconomic* on all three metrics; by 2024 it exceeds *Socioeconomic* on log-likelihood (63,885 vs. 62,871), AIC (−127,643 vs. −125,614), and BIC (−127,238 vs. −125,210), using the same number of parameters ($p = 64$). This is not a marginal improvement: even BIC, which penalizes model complexity, favors *Social Media* over *Socioeconomic* by 2024. Third, in 2019 the *Mobile Services* model's BIC (−105,872) falls marginally below *Socioeconomic* (−106,461), reflecting that BIC's penalty offsets the likelihood gain from 161 vs. 56 parameters; this reverses in 2024 (−128,741 vs. −125,210), confirming that the full set of digital signals earns its complexity by the second election.

I Elections Background

This appendix provides extended electoral context for readers unfamiliar with French politics.

The European Parliament

The European Parliament is one of the two legislative bodies of the EU, alongside the Council of the EU. It is responsible for: (*i*) adopting EU legislation; (*ii*) deciding on the annual EU budget together with the Council; (*iii*) overseeing EU institutions, including the European Commission; (*iv*) cooperating with national parliaments on European affairs; and (*v*) exercising full legislative power in over 40 policy areas, including agriculture, energy security, immigration, justice, and EU funds. Elections take place every five years, with approximately 400 million eligible voters, making it one of the largest democratic processes worldwide. As of 2024, the Parliament consists of 720 Members of the European Parliament (MEPs), with seats allocated proportionally to each country's population, then distributed among national parties by vote share.

France has been the second-largest country by population in the EU, giving it the second-highest seat allocation, only behind Germany. French MEPs therefore exert meaningful influence over EU decisions on topics such as energy policy and immigration. Nationally, the results of European Parliamentary elections serve as a barometer of public opinion between national electoral cycles. This was particularly evident in 2024, when *Rassemblement National*'s decisive victory prompted President Macron to dissolve the National Assembly and call snap legislative elections [6] originally scheduled for 2027.

2019 Election

The French European Parliamentary elections of 2019 were held on May 26. Thirty-four political groups participated, but only six cleared the 5% threshold required to secure seats, together accounting for more than 80% of total votes and distributing among them the 79 seats for French MEPs.

Table 6: Log-likelihood, AIC, and BIC for each predictor set (Table 2), for the 2019 and 2024 European Parliament elections. Arrows next to each metric indicate the direction of better fit: higher log-likelihood (↑), lower (more negative) AIC and BIC (↓). The bottom row reports the gain from adding mobile services to the *Socioeconomic* baseline (*i.e.*, *All* − *Socioeconomic*); all three values indicate improvement.

	2019				2024			
Predictor set	p	**LogLik (↑)**	**AIC (↓)**	**BIC (↓)**	p	**LogLik (↑)**	**AIC (↓)**	**BIC (↓)**
Intercept only	7	46,681	−93,349	−93,305	8	55,715	−111,415	−111,364
Population	42	47,710	−95,336	−95,071	48	58,431	−116,766	−116,462
Unemployment	14	49,018	−98,008	−97,920	16	57,652	−115,273	−115,172
Income	14	51,310	−102,591	−102,503	16	58,811	−117,589	−117,488
Socioeconomic	56	53,463	−106,814	−106,461	64	62,871	−125,614	−125,210
Streaming	56	48,471	−96,830	−96,476	80	58,557	−116,955	−116,449
News	49	50,526	−100,953	−100,644	56	58,749	−117,385	−117,031
Messaging	28	48,159	−96,262	−96,085	48	59,198	−118,300	−117,996
Social Media	49	52,167	−104,236	−103,926	64	63,885	−127,643	−127,238
Mobile Services	161	53,606	−106,889	−105,872	224	65,302	−130,156	−128,741
All	210	56,891	−113,362	−112,035	280	69,356	−138,151	−136,382
Δ (*All* − *Socioeconomic*)		+3,428	−6,548	−5,575		+6,484	−12,537	−11,172

- ***Rassemblement National***, 23.34%. During this election, *Rassemblement National* positioned itself as *"Take Power, the list supported by Marine Le Pen,"* emphasizing a narrative of change and continuity within the party's nationalist agenda.
- ***Coalition Renaissance***, 22.42%: a coalition of La République En Marche!, MoDem, Agir, and Mouvement Radical, Social et Libéral.
- ***Europe Écologie-Les Verts***, 13.47%.
- ***Les Républicains***, 8.48%.
- ***La France Insoumise***, 6.31%.
- ***Coalition Envie d'Europe écologique***, 6.19%: a coalition of the Parti Socialiste, Radicaux de Gauche, Place Publique, and Nouvelle Donne.

This election carried particular political weight: following *Rassemblement National* 's decisive victory, President Macron called snap legislative elections [6], originally scheduled for 2027.

Party positions on the Left-Right (LR) spectrum (0 = far Left, 10 = far Right) are drawn from the ParlGov dataset [34, 35], itself based on the Chapel Hill Expert Survey [96].

2024 Election

The French European Parliamentary elections of 2024 were held on June 9. Thirty-eight political parties participated; seven cleared the 5% threshold, together accounting for more than 87% of total votes. France gained two seats relative to 2019, electing 81 MEPs in total.

- ***Rassemblement National***, 31.37%.
- ***Coalition Besoin d'Europe***, 14.60%: an evolution of the centrist alliance behind Emmanuel Macron's pro-European vision, transitioning from *Coalition Renaissance* in 2019. Despite the rebranding and the inclusion of Horizons, the coalition's core political stance remained largely unchanged, with support from La République En Marche!, MoDem, Horizon, Union des Démocrates et Indépendants, and Parti Radical.
- ***Coalition Réveiller***, 13.83%: a coalition of the Parti Socialiste and Place Publique.
- ***La France Insoumise***, 9.89%.
- ***Les Républicains***, 7.25%.
- ***Europe Écologie-Les Verts***, 5.50%.
- ***Coalition La France fière***, 5.47%: a coalition of Reconquête and the Centre National des Indépendants et Paysans.

The same LR scoring methodology as in 2019 is applied to these seven parties. Between the two elections, the French political landscape experienced a modest rise in polarization, reflected in growing support for the ideological extremes (*La France Insoumise* and *Rassemblement National*), the emergence of *Coalition Réveiller* and *Coalition La France fière*, and the relative weakening of centrist forces. Yet the positioning of all main groups remained fairly stable, enabling a meaningful longitudinal comparison across both years.