This is the author's accepted version of the article. The final version published by IEEE is O.E. Martínez-Durive, I. Ucar, Z. Smoreda, E. Moro and M. Fiore, "Mobile App Consumption and Political Orientation," IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2024, pp. 1-2, doi: 10.1109/INFOCOMWKSHPS61880.2024.10620873

Mobile App Consumption and Political Orientation

Orlando E. Martínez-Durive*†, Iñaki Ucar†, Zbigniew Smoreda‡, Esteban Moro†§¶, Marco Fiore*
*IMDEA Networks Institute, †Universidad Carlos III de Madrid, ‡Orange Innovation, §Northeastern University, ¶MIT
{orlando.martinez,marco.fiore}@imdea.org, inaki.ucar@uc3m.es, zbigniew.smoreda@orange.com, e.moroegido@northeastern.edu

***Abstract*—Elections are a cornerstone of democratic societies, and their outcome has important implications on the lives of citizens and on the interior and foreign politics of a country. Understanding biases in the political orientation of the electorate plays a key role in assessing the health of the voting process and the reasons underlying the preferences of voters. Traditionally, political orientation has been studied through the lenses of the socioeconomic status of voters, *e.g.*, their education level, type of occupation, wealth, or age. In this work, we take an original perspective and factor in mobile app usage as a different yet primary indicator of the vote decision. To this end, we explore the relationship between the 2019 European parliamentary election results in approximately 4,000 urban communes in France and the associated consumption of a wide range of mobile services. Our results show how app usage provides complementary information to the socioeconomic status and can feed a Dirichlet regression that is up to 21% more accurate in predicting the multiparty election outcome.**

I. Introduction

French elections for European parliamentary seats were held on May 26, 2019, and 34 political parties participated in the election. Only six of them reached the 5% threshold needed to secure seats in the parliament, and we focus our analysis on these six major parties. The list of parties and the percentage of votes they attracted are shown in Table I.

Recently, the influential newspaper Le Monde reported on the expensive Facebook ads campaign launched by the Identity and Democracy group to support Rassemblement National and how La République en Marche ads especially targeted young adults and those over 65 years old. Previous works also mentioned how Instagram has been employed for communicating political messages [1]. In this emerging context, we aim to study the relationship between app consumption and election outcomes and how such correlations relate to traditional socioeconomic explanatory factors.

II. Data

A. Election results

Election results are available on the French Ministry of Interior and include the region code, total number of valid, blank, and invalid votes, party name, candidate, and the number of valid votes obtained by each of the 34 parties in a given area. Our study focuses on the percentage of valid votes each party received per region. At first, the results are collected at individual polling stations and later aggregated at the level of commune, department, and region. To minimize the spatial mismatch between election results and socioeconomic indicators, we rely on the highest resolution of aggregation possible, *i.e.*, the most populated 4,000 communes in France.

Party	% of votes
Rassemblement National (RN)	23.34
Coal. Renaissance (LREM-MoDem)	22.42
Europe écologie-Les verts (EELV)	13.47
Coal. Les Républicains (LR-LC)	8.48
La France Insoumise (LFI)	6.31
Coal. Envie d'Europe écologique (SP-PP-RDG-ND)	6.19
Others	19.79

TABLE I
Results of the 2019 European elections in France.

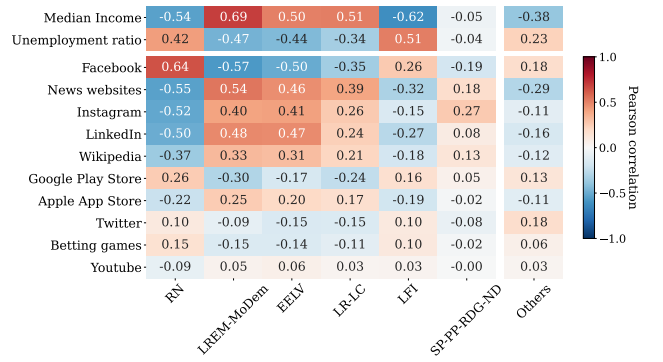

Fig. 1. Pearson's correlation of socioeconomic indicators and a subset of mobile services with votes to each major party.

B. Socioeconomic indicators

We include three classical socioeconomic indicators in our analysis, *i.e.*, median income, unemployment ratio, and population distribution by age, whose indicators are provided by the National Institute of Statistics and Economic Studies. Fig. 1 shows the correlation of median income and unemployment ratio with the percentage of votes each party received: preferences for moderate right and centrist parties correlate with higher income, while far-right and far-left parties got more votes in areas with lower income and higher unemployment.

C. Mobile apps consumption

Mobile service usage is derived from data collected by Orange in 2019, containing the byte counts of the demands generated by each app with 1-minute granularity at over 17,000 base stations in France. Orange is an incumbent mobile network operator in the country, with a customer base accounting for a 31.9% nationwide market share; this provides strong guarantees about the statistical validity of the study and about the representativeness of its results. As we are especially interested in app consumption in proximity of the election period, we focus on the week from May 19 to May 25, and select the 45 most popular mobile services for analysis.

D. Spatiotemporal and scale consolidation

Traffic data and election results are related using the well-known Voronoi tessellation strategy to spatially map traffic

information associated to individual network base stations. As the resulting base station coverage areas have a different geometry than that of the communes used for voting statistics and socioeconomic indicators, we employ an area-weighted interpolation [2] to associate app usage to communes.

A mismatch between traffic and socioeconomic indicators also exists in the temporal dimension because of the inherently mobile nature of app consumption as opposed to the static character of voting and socioeconomic indicators. We resolve the discrepancy by only considering the mobile traffic generated by users when they tend to be at their residence locations, *i.e.*, from 20:00 to 7:00 on working days only.

Finally, traffic demands can vary considerably among apps due to the intrinsic nature of the services: for instance, YouTube generates byte counts that are orders of magnitude higher than those produced by WhatsApp, even if the two are used for the same amount of time. To avoid that the scale difference in the traffic of apps biases the results, we first, compute the average hourly consumption of each app in every commune and obtain a per-capita value via a normalization by the number of inhabitants in the commune. Then, we derive the Reveal Comparative Advantage (RCA) index [2] of the per-capita app usage, which measures fluctuations in the utilization of a specific service in a commune with respect to what is typically observed in all other communes.

As can be see in Fig. 1, the RCA of the consumption of representative mobile apps strongly correlates with the percentage of votes obtained by multiple parties.

III. Dirichlet regression

Election results are an instance of compositional data, where the sum of the shares of votes that each party receives is a defined value of 100%. The Dirichlet regression is well suited in compositional scenarios as it allows estimating a continuous multivariate Dirichlet distribution where the sum of its random variate is always one. The distribution is parametrized via the concentration parameters of an α vector of length k that defines where the center of mass is located over k categories; in our case, those are the major parties plus the 'Others' category that aggregates the remaining votes. The regression derives the α vector via Maximum Likelihood Estimation.

IV. Results

The regressed α vector can be expressed as a function of the predictor variables, *i.e.*, the socioeconomic indicators and the RCA of the consumption of each app. The relationship between the concentration parameters and the predictor variables can take any form; for the sake of explainability, we focus on linear models that ease interpretation as in linear regression.

To understand the *explainability capabilities* of the predictor variables, we test them in 6 different combinations, from individual types of variables to a complete model with all the available variables. Fig. 2 shows the adjusted R^2 obtained by the different models and sets forth the following considerations: (i) a model based solely on app usage outperforms all models informed by individual or combined socioeconomic indicators

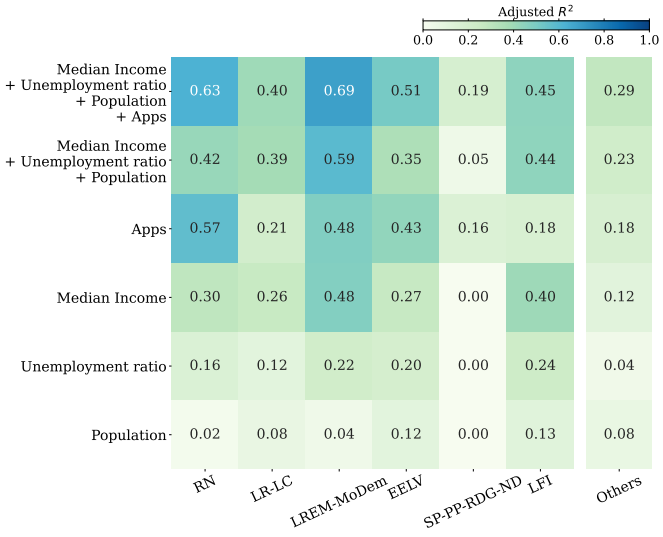

Fig. 2. Adjusted R^2 attained by Dirichlet regressions informed by diverse predictor variables (rows) for each party (columns).

when it comes to explaining the percentage of votes that RN, EELV, and SP-PP-RDG-ND received; (ii) only in the LREM-MoDem case, the combined socioeconomic indicators explain the election outcome better than app consumption; (iii) mobile app utilization always offers complementary information that is not present in the socioeconomic variables, such that when all variables are combined in a complete model, the adjusted R^2 invariably increases. The added explainability power of mobile services seems especially relevant to capture better the vote percentages attracted by all of the three major parties and by the SP-PP-RDG-ND coalition.

V. Conclusions

Our preliminary analysis shows that app usage information can complement socioeconomic indicators to predict election results better. In other words, there are behaviors in political orientations that are explained by mobile service consumption but not by socioeconomic status. We plan to build on these early results to conduct an in-depth investigation of the impact that the usage of specific apps can have on the political polarization of voters and verify the generality and robustness of the method across different elections.

Acknowledgment

This work was supported by the Comunidad de Madrid under Talent Attraction grant no. 2019-T1/TIC-16037 and 2023-5A/TIC-28944 (NetSense), and the research project CoCo5G (Traffic Collection, Contextual Analysis, Data-driven Optimization for 5G), grant no. ANR-22-CE25-0016, funded by the French National Research Agency (ANR).

References

[1] A. O. Larsson, "The rise of instagram as a tool for political communication: A longitudinal study of european political parties and their followers," *New Media & Society*, 2023.

[2] I. Ucar *et al.*, "News or social media? socio-economic divide of mobile service consumption," *Journal of The Royal Society Interface*, 2021.