



HAL
open science

Unraveling Citizen Division and Polarization during Elections

Carlos Camilo Navarrete Lizama

► **To cite this version:**

Carlos Camilo Navarrete Lizama. Unraveling Citizen Division and Polarization during Elections. Political science. Université Paul Sabatier - Toulouse III, 2023. English. NNT : 2023TOU30219 . tel-04471468

HAL Id: tel-04471468

<https://theses.hal.science/tel-04471468>

Submitted on 21 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par
Carlos Camilo NAVARRETE LIZAMA

Le 23 novembre 2023

Démêler la division et la polarisation des citoyens lors des élections

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Informatique et Télécommunications**

Unité de recherche :
TSE-R - Toulouse School of Economics - Recherche

Thèse dirigée par
Umberto GRANDI

Jury

Mme Stefania MILAN, Rapporteur
M. Evangelos POURNARAS, Rapporteur
M. Stéphane AIRIAU, Examineur
M. Jérôme MENGIN, Examineur
M. Umberto GRANDI, Directeur de thèse
M. Frédéric AMBLARD, Président

Resumé

Les accords et les désaccords font partie intégrante de nos interactions sociales. Bien qu'il existe un consensus général sur le fait que les désaccords sont essentiels à notre coexistence démocratique, les récents événements politiques nous font nous interroger sur l'hypothèse selon laquelle des désaccords systématiques au sein de la population conduisent à une société plus polarisée que par le passé. Par conséquent, comprendre le degré de polarisation de notre société est aujourd'hui crucial. Étant donné que la plupart des études dans cette discipline se rapportent au système américain, nous voyons une opportunité d'appliquer des techniques informatiques pour développer de nouvelles approches agnostiques et axées sur les données afin de caractériser et d'étudier la division et la polarisation au fil du temps.

Cette thèse évalue systématiquement les mesures d'accord, de division et de polarisation dans la société, en particulier lors des élections. Nous présentons trois chapitres empiriques démontrant le potentiel des données collectées via les plateformes en ligne et les résultats des élections comme formes fiables pour saisir la division et la polarisation des citoyens. Ce travail est avant tout interdisciplinaire, visant à contribuer à la communauté grandissante des sciences sociales computationnelles.

Notre contribution principale comprend deux propositions méthodologiques pour cartographier la division des sujets et les niveaux de polarisation dans la société. Premièrement, nous rapportons les conclusions de deux systèmes de démocratie numérique lancés par notre équipe en France et au Brésil avant leurs élections présidentielles respectives de 2022. Nous avons utilisé les données collectées pour introduire une mesure de division agnostique aux fonctions d'agrégation et capable de cartographier les sujets porteurs de division. Nous avons démontré que les valeurs de notre mesure de division ne sont pas corrélées à celles obtenues à partir de sa fonction d'agrégation respective, et ces résultats sont généralisés aux ensembles de données politiques et non politiques, ainsi qu'aux données synthétiques et réelles. Deuxièmement, motivés par les récents résultats électoraux inattendus à l'échelle mondiale, nous avons inventé le terme « polarisation électorale » (EP) pour désigner le niveau de division des citoyens le jour des élections. Nous avons divisé notre définition en compétitivité (Entre-EP) et divisions géographiques (Intra-EP). Nous proposons une méthode qui repose exclusivement sur les données électorales pour évaluer l'EP. En utilisant des données d'élections présidentielles synthétiques et réelles de France, du Chili et des États-Unis, nos résultats suggèrent que nous pouvons savoir si une élection est polarisée en sachant simplement si deux conditions sont remplies : des taux de participation plus élevés et des schémas de vote regroupés par régions.

Notre contribution finale analyse les données collectées par nos systèmes expérimentaux de démocratie numérique afin de déterminer dans quelle mesure les questions litigieuses et l'interface utilisateur conditionnent l'implication des citoyens dans un contexte anonyme et, par conséquent, les décisions collectives. Notre motivation est née de l'hypothèse selon laquelle les individus préfèrent ne pas parler de sujets litigieux lors de discussions en face à face afin d'éviter les confrontations. Nous avons constaté que le simple fait de lire une question litigieuse dans les premiers stades de la participation n'augmentait pas le taux d'abandon. Ensuite, en empruntant aux études comportementales des mesures de négligence, nous montrons que les caractéristiques démographiques des participants sont associées aux mesures traditionnelles de participation imprudente. Par exemple, les femmes (un groupe minoritaire qui représentait environ 30

En résumé, notre travail se concentre sur la compréhension de la division et de la polarisation des citoyens en utilisant la diffusion d'applications web et des techniques de science des données. Au total, nous proposons une approche interdisciplinaire qui nous permet de postuler que nos contributions méthodologiques sont suffisamment robustes pour être appliquées dans des travaux ultérieurs dans la discipline, contribuant ainsi à la littérature sur la participation numérique et la polarisation géographique.

Abstract

Agreements and disagreements are part of our social interactions. While there is a main consensus that disagreements are essential for our democratic coexistence, recent political events make us question whether systematic disagreements in the population lead to a more polarized society than in the past. Therefore, understanding the extent to which our society is polarized is crucial today. As most studies in this discipline speak to the U.S. system, we see an opportunity to apply computer science techniques to develop new, agnostic, and data-driven approaches to characterize and study divisiveness and polarization over time.

This dissertation systematically assesses metrics of agreement, divisiveness, and polarization in society, specifically during elections. We present three empirical chapters demonstrating the potential of data collected through online platforms and election outcomes as reliable forms to capture citizen division and polarization. This work is primarily interdisciplinary, aiming to contribute to the growing computational social science community.

Our main contribution includes two methodological proposals for mapping the divisiveness of issues and levels of polarization in society. First, we report the findings of two digital democracy systems released by our team in France and Brazil preceding their respective 2022 presidential elections. We used the data collected to introduce a metric of divisiveness that is agnostic to aggregation functions and can map divisive issues. We demonstrated that the values of our measure for divisiveness are uncorrelated to the one obtained from its respective aggregation function, and these results are generalized both for political and non-political datasets and for synthetic and real data. Second, motivated by recent unexpected election outcomes worldwide, we coined the term Election Polarization (EP) to refer to the division level of citizens on Election Day. We split our definition into competitiveness (Between-EP) and geographical divisions (Within-EP). We propose a method that relies exclusively on election data to assess the EP. Using synthetic and presidential election data from France, Chile, and the United States, our results suggest that we can know whether an election is polarized by simply knowing whether two conditions are satisfied: greater participation levels and voting patterns clustered by regions.

Our concluding contribution analyzes the data collected by our experimental digital democracy systems to determine to what extent contentious issues and the user interface condition citizen involvement in an anonymous setup and, consequently, collective decisions. Our motivation came from the hypothesis that individuals in face-to-face discussions prefer not to talk about contentious issues in order to avoid confrontations. We found that merely reading a divisive issue in the early stages of participation did not increase the dropout. Then, by borrowing measures of carelessness from behavioral studies, we show that the demographic characteristics of participants are associated with traditional measures of careless participation. For instance, female participants (a minoritarian group that represented around 30% of the participants in our platforms) abstained more than male participants by responding to approval ballots. This posits evidence that just removing careless participants without any context (for instance, using the naïve approach of removing abstentions) exacerbates disparities in the representation of men/women or young/old in digital democracy setups.

To sum up, our work focuses on understanding citizen division and polarization by using the release of web applications and data science techniques. Altogether, we provide an inter-

disciplinary approach that allows us to posit evidence that our methodological contributions are robust enough to be applied in further work in the discipline, contributing to the literature on digital participation and geographical polarization.

Acknowledgments

I remember a good friend once told me that the most challenging part of pursuing a Ph.D. was not the Ph.D. itself but the resilience in dealing with difficult situations. Three years after this conversation, I can guarantee that this comment is true. This journey was a bit more turbulent than I expected, but all my decisions made me feel proud of the person I have become.

I am deeply grateful to my supervisor, Umberto Grandi, who kindly accepted me into his group in 2023 after facing difficult times in my initial research group. His support and insights were invaluable in shaping me into a better researcher. Thank you to Evangelos Pournaras and Stefania Milan for graciously accepting to be reviewers of this manuscript and Frédéric Amblard, Jérôme Mengin, and Stéphane Airiau for being part of my jury at defense day. Finally, I am deeply indebted to Mariana Macedo for taking the time to read and comment on an early version of this dissertation and listen to my rehearsals.

I would like to express my gratitude to the good people I met in Toulouse: Johannes, Lea, Mariana, Melanie, Myroslava, Philipp, Rachael, Sophie, and Viktor. An apart sentence is for the most significant person I met in this city: Jingling. I do not have words to describe how thankful I am for all her help, contention, and friendship. She deserves the best things that can ever happen to a person.

My Ph.D. journey had two international adventures. First, thank you, Guillermo Cabrera, for accepting me as part of your research group during my four-month stay at Universidad of Concepcion. Second, I would like to thank Pierre-Alexander Balland, Ron Boschma, and Carolina Castaldi for hosting me at the Department of Economic Geography at Utrecht University. The support of Pierre-Alex in the last months of my Ph.D. gave me the motivation to finish my dissertation and continue pursuing my dream of becoming a doctor. Thank you to all the people I met at Utrecht University: Andrea, Charlotte, Eduardo, Masoumeh, Santiago, Sara, Sayantan, Suzanne, Tingting, Wanlin, Yuanyuan, and Yuting.

To my friends and family. Without their support, this dream would never have been possible. To Camila, she was likely the most important person in this step of my life. Her friendship, love, and patience backed me up in my darkest moments. Thanks for letting me share the road to your side.

As I conclude my own Ph.D. journey, I have a last message for grad students: You are intelligent and worthy of achieving your dream of becoming a doctor; never let anyone, including yourself, make you doubt that. My doctoral experience has taught me that you will always find good people who will be willing to help and listen to you along the road.

Table of Contents

1	Introduction	13
1.1	Background	13
1.1.1	Computational Social Science	13
1.1.2	Mapping Polarization in Society	14
1.1.3	Issues' Divisiveness and Political Disagreements	16
1.2	Research Questions	16
1.3	Methodology	20
1.4	Chapter Overview	21
1.4.1	Chapter 2: Understanding Issues' Agreements and Divisiveness in French and Brazilian Presidential Elections	21
1.4.2	Chapter 3: Mapping Citizen Division Through Elections	25
1.4.3	Chapter 4: Online Involvement in Two Real-World Digital Democracy Systems	26
1.4.4	Dissertation Resources	28
1.4.5	Summary	29
2	Understanding Issues Divisiveness through Online Participation: Evidence from the 2022 French and Brazilian Presidential Elections	31
2.1	Introduction	32
2.2	Literature Review	32
2.2.1	Digital Technologies in Political Episodes	32
2.2.2	Design of Online Participation	33
2.2.3	Issues Divisiveness and Polarization	34
2.3	Methods	36
2.3.1	The Platform	36
2.3.2	Bot Detection System	38

2.3.3	Data Curation	39
2.3.4	Ranking of Agreements	40
2.3.5	Political Responsiveness	40
2.3.6	Ranking of Divisiveness	41
2.3.7	Multidimensional Divisiveness	42
2.4	Results	42
2.4.1	Demographics of Online Participation	44
2.4.2	Ranking of Agreements	47
2.4.3	Ranking of Divisiveness	48
2.4.4	Analysis of Divisiveness	50
2.4.5	Empirical Boundaries of Divisiveness	55
2.4.6	Robustness of Key Results from Divisiveness	58
2.5	Discussion	59
3	Mapping Citizen Divisions Through Elections	61
3.1	Introduction	62
3.2	Literature Review	63
3.2.1	Unexpected election outcomes	63
3.2.2	The Footprint of Voting Patterns	64
3.2.3	Polarization in Social Science	64
3.3	Methods	65
3.3.1	Definition of Antagonism	65
3.3.2	Definition of Election Polarization (EP)	67
3.3.3	Numerical Examples	68
3.3.4	Election Data Schema	70
3.3.5	Election Data	71
3.3.6	Polarization Measurements	72

<i>TABLE OF CONTENTS</i>	11
3.4 Results	74
3.4.1 Numerical Properties of Election Polarization	74
3.4.2 EP in Real Data	76
3.4.3 Political Polarization, Political Interest, and Election Polarization	79
3.4.4 Robustness across Aggregation Levels	82
3.5 Discussion	84
4 Online Involvement in Two Real-World Digital Democracy Systems	87
4.1 Introduction	88
4.2 Literature Review and Hypothesis Formulation	88
4.2.1 Contentious issues and Anonymity	89
4.2.2 Collecting preferences in online participation	90
4.3 Methods	91
4.3.1 The Platforms	91
4.3.2 Mapping contentious issues in ballots	95
4.3.3 Measuring data quantity and quality	95
4.3.4 Econometric models	97
4.3.5 Interplay of data quantity and quality	99
4.4 Results	100
4.4.1 Divisive issues	100
4.4.2 Data quantity	103
4.4.3 Data quality	105
4.4.4 Interaction data quantity/quality	108
4.5 Discussion	109
5 Conclusions and Perspectives	111
Bibliography	115

A Appendix for Chapter 2	127
A.1 Robustness ranking of agreements	127
A.2 Robustness ranking of divisiveness	134
B Appendix for Chapter 3	137
B.1 Socio-demographic characteristics	137
B.1.1 United States	137
B.1.2 France	138
B.1.3 Chile	138
B.2 Mapping Political Interest	139
B.3 Data sampling	139
B.4 Additional Regression Models	140

Introduction

Agreements and disagreements are fundamental principles of our social interactions, essential to democracy (Prothro & Grigg, 1960). Nevertheless, there is a point in which the expressed disagreements provoke feelings of anger, mistrust, or simply discomfort with our peers, which can induce misunderstandings, feelings of threats, conflicts, and even civil wars in extreme cases (e.g., Reynal-Querol, 2002; Simons and Green, 2018). Related literature in social science has framed this line of argument from the lens of polarization. This term encompasses a large collection of theoretical and empirical work that has received wide attention, especially in election periods. From a computer science perspective, the potential of computational advancements to collect and analyze massive data sets can empirically enrich and strengthen these social science debates.

This dissertation systematically explores agreements, divisiveness, and polarization in society, specifically surrounding electoral times. We present three empirical chapters demonstrating the potential of using data collected through online platforms and existent instruments (e.g., election data) as a reliable form to capture citizen divisions and polarization. This work is primarily interdisciplinary, looking to tie political science and economics research with computer science. Therefore, this dissertation aims to contribute to the growing community of computational social science (Lazer et al., 2009), specifically in the line of argument of computational (or data-driven) political science (Weber et al., 2013).

1.1 Background

The literature in social science, notably about divisive issues and political polarization, encompasses an extensive collection of work analyzing their causes, determinants, and consequences. This can be exemplified by work on how disparities in access to globalization benefits might explain unforeseen election outcomes (Rodrik, 2018; Rodriguez-Pose, 2018) or why political candidates emphasize their campaign in divisive issues (Ash et al., 2017). Likely, there is not another contemporary debate that receives more attention than political divisions among citizens, especially to explain the increasing support for extreme political parties in recent years. Nevertheless, this discussion is not bound to scholars from social science disciplines.

1.1.1 Computational Social Science

Fourteen years ago, Lazer et al., 2009 coined the term Computational Social Science (CSS) to refer to an emerging discipline at the intersection of Computer and Social Sciences. This term can be defined as the “development and application of computational methods to complex,

typically large-scale, human (sometimes simulated) behavioral data” (Lazer et al., 2020), and considers our capacity to collect and analyze massive data sets as an opportunity to revisit old social science questions, contrasting with traditional research that mainly relies on “one-time self-reported data on relationships.” (Lazer et al., 2009) To date, thousands of research articles associated with CSS have been presented in mainstream computer science conferences (e.g., Chomel et al., 2022; Hanna et al., 2013) and scientific journals (e.g., Edelman et al., 2020; Mann, 2016). Despite this rapid growth, some drawbacks need to be pondered. Scholars have emphasized that these efforts should not only test and revisit social science questions but also advance social science theory (Chetty et al., 2022; Conte et al., 2012). Furthermore, challenges on ethical guidance for researchers should be developed concerning the emerging risks of using sensitive personal data or unethical behaviors (Lazer et al., 2020).

Today, the maturing of computational advancements is quite different than it was in 2009; however, the inspiring and transformative idea behind CSS remains on. This dissertation contributes a line of argument in CSS, specifically in political science, concerning the identification of divisive issues and mapping levels of polarization. We aim to fill a gap in the literature concerning the adoption of agnostic and universal approaches to map these phenomena. To this end, we use massive data sets collected by experimental digital democracy systems or publicly available on the Web.

1.1.2 Mapping Polarization in Society

Defining polarization is not simple. In a broad sense, it can be addressed as systematic divisions of opposite or contrasting groups in society. Scholars have carefully emphasized that, depending on the circumstances and characteristics, this term must be divided into several dimensions, such as geographical, political, mass, elite, or affective, to name a few (Iyengar et al., 2012; Rohla et al., 2018; Tucker et al., 2018). Living in a more polarized society than our parents has mixed consequences in our political institutions. Nevertheless, the narrative that the population become more polarized is not new. In fact, journalists often tell this story in an attempt to explain unexpected election results, such as 2016 Trump’s victory in the U.S. or Brexit in the U.K.; nevertheless, these statements lack an empirical foundation in most cases. Prior work has established a clear distinction of polarization if we refer to partisans (elite polarization) or the population (mass polarization). To date, dozens of articles from the most respected scholars in political science in the U.S. have enlightened an active debate about whether the mass is more polarized (e.g., Abramowitz and Saunders, 2008; Fiorina and Abrams, 2008; Lelkes, 2016; Levendusky, 2009), without a clear consensus in this debate.

Nevertheless, elections represent a common denominator. Elections are the channel through which citizens express their voices concerning policies and political candidates (Makarenko, 2015). Nonetheless, they also reveal layers of segregation and polarization in society (e.g., Layton et al., 2021 in Brazil). In light of these distinctions, which facet of polarization is reflected by elections? Intuitively, when we think about political contexts and polarization, we tacitly assume that we refer to political polarization. According to Tucker et al., 2018, political polarization can be defined as the extent to which different political parties offer ideologically distant policy platforms. Since most work in this literature speaks to the U.S. political system, an intuitive way to interpret political polarization could be the ideological distance between the Republican and Democratic parties. The reason behind the asymmetric development of this line of argument relies upon a long-standing tradition of national election

instruments in the U.S., such as the American National Election Studies (ANES), the Cooperative Election Study (CCES), or the General Social Survey (GSS), that has allowed scholars to propose and follow-up forms of polarization in the citizenship along with their electoral dynamics in a period of over 70 years. Since it is expensive to maintain these instruments over time, only a few countries can afford this cost. Likewise, political polarization is rarely studied at the regional level (e.g., states or counties) because most of these surveys represent nothing other than the national scale. Scholars, in an attempt to study political polarization worldwide and on a more granular scale, have proxied it using the voting percentage of extreme political parties (Casal Bértoa & Rama, 2021; Winkler, 2019).

Besides being a common denominator, the use of election data summons a general agreement among scholars, considering it as an invaluable input both for theoretical and empirical discussions. To exemplify this agreement, we borrow the “Big Sort” hypothesis. The American journalist Bill Bishop coined this term to describe how within-country migration flows in the U.S. are partially influenced by the search to live in neighborhoods with like-minded people, sorting geographies into strongholds for Republicans (Democrats) in Rural (Urban) America (Bishop, 2008). To address this hypothesis, Bishop used the 1976 and 2004 presidential election results as input. Later, Abrams and Fiorina, 2012 argue that this statement, beyond anecdotal correlations, lacks substantial evidence, but they do not discard the use of election data; rather, they challenge the fact of cherry-picking just two presidential elections, not validating with additional election types (e.g., congressional results) and the absence of extensive temporal analysis. Since then, a lively debate emerged examining the “Big Sort”, combining election results and pre-post/election polls (Johnston et al., 2016; Sussell, 2013).

Having established that elections play an essential role in comprehending citizen divisions, they also inspire recent lines of arguments concerning their causes. Rodriguez-Pose, 2018 and Rodrik, 2018 shed light on the theoretical reasons behind unexpected election results, raising the hypothesis that places mainly affected by globalization and concentration of knowledge and opportunities “take revenge” of their marginalization versus big cities, voting for anti-establishment candidates or populisms. Thus, the surprising performance of anti-establishment candidates around the world, such as Donald Trump in the U.S., Jair Bolsonaro in Brazil, and Marine Le Pen in France, have raised concerns about whether these candidates exploit latent layers of segregation (Lamont et al., 2017; Layton et al., 2021).

Studying polarization extends beyond political science and elections. Related literature in economics has focused on establishing theoretical foundations for properly measuring polarization. This distinction is exemplified in studies motivated by the “disappearing middle class” phenomenon, which yields that societies with a decreasing middle-class group shift towards the appearance of two contrasting groups, dividing it into rich and poor (Esteban & Ray, 1994; Wang & Tsui, 2000; Wolfson, 1994). These works consider this phenomenon should be examined by polarization measurements since they are more suitable than the classical framework of inequality measures. The seminal contribution of Esteban and Ray provide an axiomatic perspective of how a polarization measure should behave. They introduce the identification-alienation framework, postulating that individuals identify with people from the same group (in their seminal paper, they address the same income level) and feel alienated from others who are “far away” (or with a different income level). Since then, we have found a flourish of approaches to characterizing social phenomena under the umbrella of polarization measurements, such as divisions in religious groups (Reynal-Querol, 2002) and spatial inequality in healthcare and education (Zhang & Kanbur, 2005), to name a few.

1.1.3 Issues' Divisiveness and Political Disagreements

The term polarization is not exclusively studied for masses or groups. A strand of literature has focused on “polarizing” or divisive issues and their relationship with political polarization. Can we argue that citizen disagreements over an issue are equivalent to polarization? The answer is clearly no since diversity in point of view is even healthy for our democratic co-existence (Huckfeldt et al., 2002). The importance of studying divisive or confronting issues is their potential to unravel polarization. In fact, a line of research has explored whether attitudes change toward “family issues” (Family Gender Roles, Women’s Roles, Abortion, Sexuality, Sex Education, and Divorce Law) can be framed to explain a rising polarization (DiMaggio et al., 1996; Hoffmann & Miller, 1998). Their theoretical ground is that society becomes polarized whether the distribution of opinions over an issue tends to be bimodal and the opinions range is highly dispersed. Nevertheless, divisive issues always exist. According to Gerber et al., 2013, people could perceive that discussing political issues is a double-edged sword since divulging one’s choices may expose them to social pressures, leading to feelings of threat and anxiety (Simons & Green, 2018). Similarly, Chen and Berger, 2013 found that higher levels of controversy in issues decrease their discussion.

Drawing upon electoral periods, there is evidence that political campaigns on social media micro-target voters with ads on divisive issues (Kim et al., 2018). Furthermore, political candidates spend more time on divisive issues when they are up for election and in response to higher news transparency (Ansolabehere, Puy, et al., 2015; Ash et al., 2017). Issues that divide the government from the opposition are significantly more likely to be prioritized on the policy agenda (Martin, 2004), and issue-specific cabinet conflicts are positively associated with the issue attention in coalition agreement (Klüver & Bäck, 2019). The causal explanation is that divisive issues are better at unraveling candidates’ political stances—a key aspect in electoral campaigns that agreements cannot provide. For instance, “increase the minimum wage” or “improve healthcare system” are often included in government programs of candidates running in presidential elections; nevertheless, these issues cannot substantially define their positions as “abortion” or “migration policies” do. Thus, partisans associate more salience to issues aligned with their parties (RePass, 1971), and legislators are likelier to follow the discussion of public issues on Twitter (Barberá et al., 2019).

Whether from political science, economics, or computer science, it seems reasonable to argue that electoral periods are relevant for the literature on polarization, either studying unexpected election outcomes, mapping political polarization, or exploring divisive/controversial issues. This dissertation is motivated by the use of massive data sets associated with election times as a robust approach to understanding polarization. Specifically, we will contribute to the literature on issues divisiveness and political polarization from a computational social science perspective.

1.2 Research Questions

The preceding discussion addressed the lack of quantitative methods to examine the shorthand hypothesis that the “population has become more polarized than before.” The main limitation concerns the collection of reliable data in order to tackle this hypothesis. Here we consider the premise that elections represent the voice of citizens as worthy of investigating deeply in order to establish a comparable measure of polarization around the world. We sub-

sequently demonstrated that elections themselves can help to answer the latter premise. We split our contribution into three main topics: issues divisiveness, election (and) polarization, and unraveling online participation. All these topics are tied by electoral periods.

Issues’ Divisiveness The idea that divergent opinions over societal issues can be interpreted as a form of polarization sparks our curiosity to explore this topic in depth. Even though previous work has already established that shifting in issue positions can be framed to explain mass polarization (DiMaggio et al., 1996; Hoffmann & Miller, 1998), a drawback in this line of argument refers to the fact that divisive issues must be identified beforehand. Even though it is evident, for example, that same-sex marriage was a controversial issue 30 years ago, social dynamics flow and evolve. How can we systematically identify those “new” divisive issues? While traditional approaches to aggregate individual preferences have focused on agreements and identifying controversial issues, which have been framed by finding divisions within groups, little work has explored how these issues can systematically be mapped without relying on self-reported data. The reasons why divisive issues cannot be captured by aggregation functions based on agreements are simple: First, a controversial issue will never be top-prioritized since a minority may always try to vote against it. This does not seem to be a problem when we need to categorize a relatively small number of issues, in which, even by hand, we can label them as divisive or not. However, considering a large set of issues ($N > 1000$), this task is almost impossible to achieve, and those issues in the middle of the ranking could be labeled as divisive or indifferent. Even though related literature has characterized the identification of these issues using statistical measures such as the variance or skewness of opinions, this approach does not seem to apply to all contexts. Second, as mentioned by Gerber et al., 2013, citizens prefer not to reveal their preferences on divisive issues to avoid exposure to social pressures, especially in electoral periods, becoming a challenging task asking explicitly for their *real* opinion over those issues. Our first research question calls to this issue, i.e.:

RQ.1: Can a data-driven metric help us to identify divisive issues within a population without relying on individual socio-demographic characteristics?

While traditional instruments largely depend on asking for sensitive data from individuals in order to characterize divisive issues between and within groups—e.g., political orientation, race, or religious belief—the view driven by *RQ.1* considers that it is possible to identify these controversial topics for the population without relying on these features. Specifically, we consider that the structure of pairwise comparison data, which compares the relative preferences of individuals, can help differentiate groups within a population. To tackle this question, we use data collected via two real-world digital democracy systems released by our research team in France and Brazil, preceding their respective 2022 presidential elections. We hypothesize that mapping these divisive issues should not be exclusively associated with a particular metric. Of course, we need to define a *metric*. Still, our view is that traditional aggregation functions can be related to the notion of the mean/variance relationship, where the latter corresponds to the second mathematical moment of the mean. In this vein, we interpret that aggregation functions can behave similarly, where the score represents the first moment and a metric of Divisiveness the second moment. *RQ.2* formalizes this assumption as follows:

RQ.2: Can a “Divisiveness” metric be applied by using any aggregation function?

The reasons underlying the latter question consider that aggregation functions—the way that preferences are aggregated to achieve a collective decision—can have a second interpretation of their outcomes. We argue that violations of axiomatic properties from social choice expose collective divisions (Arrow, 1950; De Condorcet, 1785), and these divisions can be interpreted similarly as the standard deviation is for the mean.

Election (and) Polarization Most polarization measurements rely on one-time surveys collected from a small population sample. As pointed out before, because these instruments are expensive to maintain over time, the greatest part of this line of argument speaks to the U.S. data. Despite finding efforts from scholars proxying measures of polarization using the voting percentage of extreme political parties, this approach is not blind to the election type-/political system or number of candidates, and cannot be generalized. A counter-example can be found in political systems with two main political parties, in which the voting percentage of protest parties is generally low. Having previously established that election results enlighten debates on polarization, we interpret that election data can also be studied as a form to map polarization. Specifically, our third research question can be formalized as the following:

RQ.3: Can we conceptualize a metric of polarization by exclusively using election data?

This problem is important for several reasons. First, massive and fine-grained administrative records collected on Election Day likely correspond to the most representative data set over citizen preferences. Second, since related literature speaks to the bi-party political system in the U.S., there is a lack of work exploring measures of “Election Polarization” around the world. Third, since it is unfeasible to afford the cost of electoral instruments for the majority of countries—especially low- and middle-income ones—the study of polarization in political contexts has been extensively confined to the case of developed countries.

In this sense, characterizing a measure of polarization derived from election outcomes seems reasonable to achieve, taking into account the vast evidence concerning the interplay of elections and polarization. Furthermore, the combination of two features: access to massive data sets collected from voting records and advancements in data science techniques, seems to be enough conditions to establish a measure of “Election” Polarization (the extent of the division level of citizens on Election Day). This dissertation does not argue that elections are a measure of ideological polarization; rather, we frame that voting patterns characterize a comparable measure of “how polarized the election was”. Nevertheless, this measure of Election Polarization needs to be further examined in order to validate its scope. This issue is worth investigating since no approach systematically examines the relationship of this form of polarization with theoretical expectations of polarized elections. As a corollary to our previous research question, we formulate the following, i.e.:

RQ.4: Can a metric of “Election Polarization” successfully label theoretical expectations of polarized elections?

On the one hand, we examine whether a measure mapped from election data is robust and is not caused by a spurious correlation. For instance, theoretical and empirical work has

suggested that political polarization is negatively associated with the perception of corruption and positively associated with income inequality. Whether Election Polarization is a reliable form of quasi-political polarization should behave similarly to those theoretical expectations. On the other hand, we test that our approach cannot be manipulated by the particularities of an election, such as the number of candidates, the level of abstentions, or the data aggregation using different scales.

Unraveling Digital Participation Thus far, our research questions have framed the necessity of having agnostic and universal methods to map divisiveness and polarization through the use of massive data sets. Nevertheless, we still lack an understanding of our digital democracy platforms in terms of how their content and interface condition online involvement. For instance, we briefly mentioned in the background that people prefer avoiding talking about controversial topics (Chen & Berger, 2013; Gerber et al., 2013; Simons & Green, 2018); nevertheless, no evidence posits this behavior in anonymized participation of digital democracy systems. This line of argument is important since it complements our efforts in establishing how divisive issues must be afforded to design online tools to increase citizen participation. Taking into account some recommendations from behavioral studies in psychology (Berry et al., 1992; Meade & Craig, 2012), we study the nature of data collected by our online experiments (See *RQ.1* and *RQ.2* for details.) Specifically, here we address two additional research questions, i.e.:

RQ.5: To what degree do divisive issues influence the participation cycle of a user in an online platform?

RQ.6: To what extent does the participation design and mechanism impact the quality, quantity, and percentage of responses obtained on online platforms?

Our platforms faced the challenge of asking participants about a large number of issues (120 issues in France and 67 in Brazil) in a short period and under uncontrolled conditions. Undoubtedly, it was impossible to achieve, at the same time, collect a complete set of preferences from individuals and provide a simple user interface. Likewise, the topics included a wide set of viewpoints that some participants could perceive as disgusting. *RQ.5* aims to examine the extent to which divisive issues affect the likelihood of exiting from the online platform, the response rate, and the percentage of abstentions in approval ballots, among other characteristics. In parallel, *RQ.6* aims to empirically analyze the balance between a fine-grained data set and more accessible participation frameworks. There is evidence that the attention of participants in online participation decreases over time (Berry et al., 1992) and that the input mechanism matters by reaching a collective decision (Benade et al., 2018). Nevertheless, there is no clear distinction in how this decrease in attention affects the results emitted from a digital democracy system. The main limitation of this type of study is the access to data; nevertheless, our platforms performed an A/B analysis of two elicitation methods: Fallback Voting and Pairwise Comparison. This is significant and worthy of investigation because it allows us to advance research on the design space of collaborative and consensus-oriented digital democracy systems (Helbing et al., 2023).

Final Remarks To summarize, this dissertation addresses six research questions using massive data collected from two real-world digital democracy systems in France and Brazil (devel-

oped by our research group) and presidential/parliamentary election outcomes at a granular scale (e.g., precincts and polling stations) downloaded and curated from Chile (2013-2021), France (2002-2022) and the United States (2000-2020). All the results presented in this dissertation are of an empirical kind, and their presentation aims to construct an interdisciplinary overview concerning metrics to capture divisiveness and polarization from computational social science.

1.3 Methodology

This dissertation empirically studies citizen division, polarization, and digital participation in electoral times. We ground our decision to investigate these topics around elections for several reasons. First, people exhibit greater levels of interest in political issues since they are actively reading/talking about contentious topics, leading us to argue that they are more prone to participate in online platforms that address them. Second, the shorthand hypothesis that the population become more polarized than in the past gets recurrent in debates that aim to explain election results. However, as previously mentioned, it lacks an empirical foundation. Third, elections represent a valuable and representative source of information about citizens' preferences. Because transparency is crucial for any free-and-fair election, public institutions usually deposit election data at granular scales.

While the lack identified in the literature concerns suitable alternatives for computing citizen divisions or polarization, we visualize that, foremost, it is necessary to have access to reliable and fine-grained data over political preferences from citizens. To study our first two research questions, we address the creation of experimental digital democracy systems. The releasement of a piece of software itself can be methodologically decomposed. This leads us to pursue an empirical study on the impact of content and interface to understand online participation better, responding to our RQ.5 and RQ.6.

Although the creation of web applications can give us valuable insights to understand agreements and divisiveness, these one-time collected snapshots via the Internet are only suitable to represent the individuals who voluntarily accepted to participate in them, not speaking to our RQ.3 and RQ.4. Of course, we could eventually argue that the releasement of a platform asking for the same issues in each region of the world can create an index of polarization. Nevertheless, maintaining digital democracy systems over time is almost impossible to accomplish. To be precise, whether citizens do not see tangible "results" derived from their participation, which is the incentive to enter an online platform systematically? Here we sought an alternative in using open data. To respond to RQ.3/RQ.4, we pursue a methodological approach that explodes the use of election data and its structure. To accurately represent a measure of polarization, we ground on common characteristics of elections, such as the voting percentage per precinct and number of votes for candidates, and tie these features with theoretical polarization measurements.

To sum up, Chapters 2 and 3 suggest methodologies for mapping divisiveness and polarization, whereas Chapter 4 assesses the impact of divisive issues and the role played by elicitation methods in online participation. This is summarized in Figure 1.1.

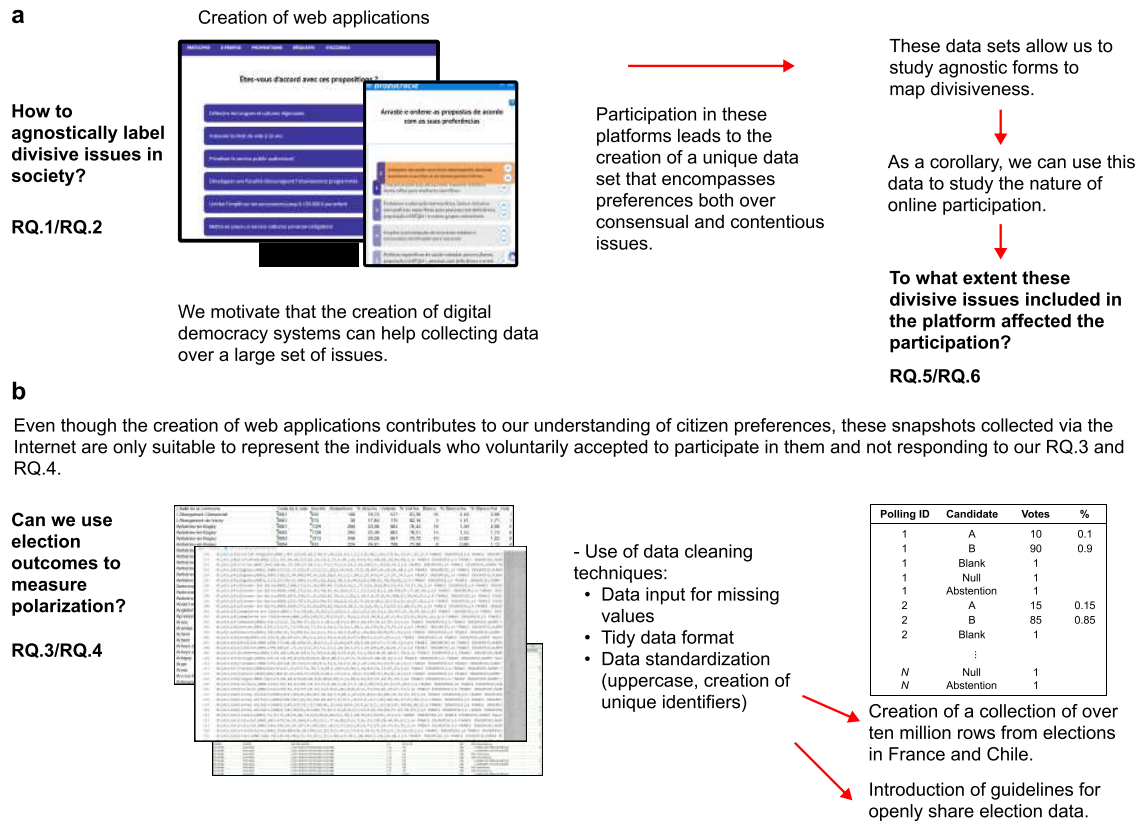


Figure 1.1: Interplay between our research questions and methodology. (a) Creation of web applications as a form to map citizen preferences. (b) Use of administrative records from elections to map polarization.

1.4 Chapter Overview

The remainder of this dissertation is summarized as follows: Chapter 2 studies the issues' agreements and divisiveness of France and Brazil in the context of presidential elections. Chapter 3 introduces Election Polarization as a framework to map citizen divisions on Election Day. Chapter 4 explores how the content and interface affect online involvement. Figure 1.2 illustrates the electoral period from the announcement of candidacies until the election day and how this dissertation contributes in different steps. Figure 1.3 illustrates our contribution in terms of examining election data.

1.4.1 Chapter 2: Understanding Issues' Agreements and Divisiveness in French and Brazilian Presidential Elections

Digital technologies have been used as a medium to enhance citizen participation. From participatory budgeting (Cabannes, 2004), online deliberation tools (Hsiao et al., 2018; Small et al., 2021) to voting advice applications (VAAs) (Garzia & Marschall, 2016), we find a surge of interest in understanding this form of participation and enlight how it promotes citizen involvement (For recent reviews, Congge et al., 2023; Helbing et al., 2023).

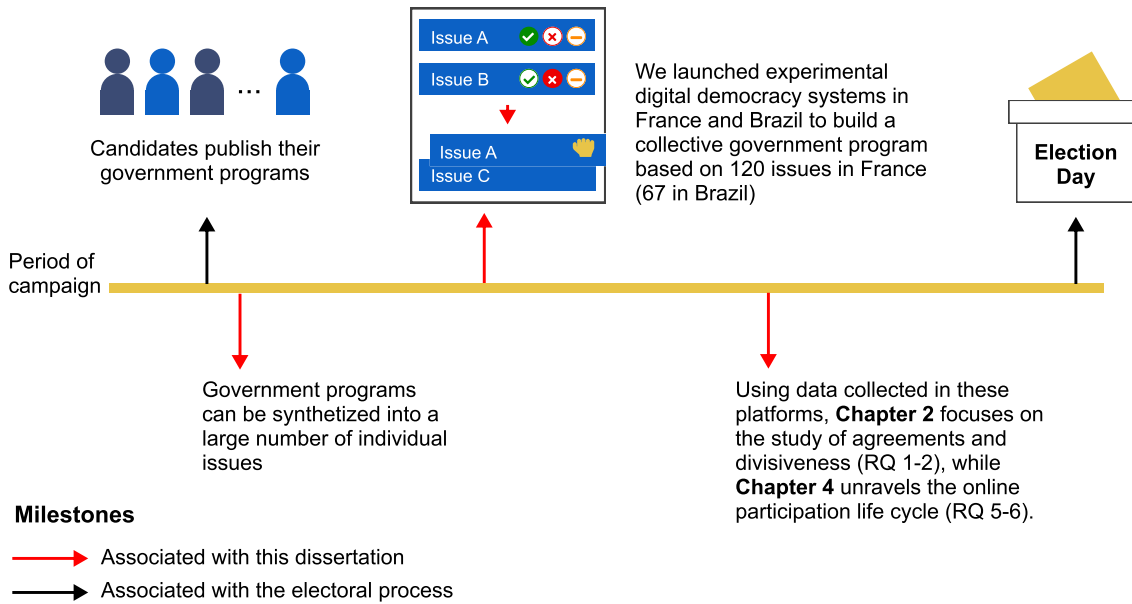


Figure 1.2: The election journey. Milestones from the official candidacies and the Election Day. Red arrows represent the parts studied by this dissertation.

This Chapter uses data collected from two real-world collaborative government program builders released in the context of the 2022 Brazilian and French presidential elections. This data allowed us to examine political agreements and divisiveness between participants. Both platforms—*Mon Programme* (<https://monprogramme2022.org>) and *Brazucracia* (<https://brazucracia.org>)—were developed by our research group in France in collaboration with a local team in Brazil. Notably, the data collected helped to initially address our *RQ.1*.

The Platforms and Demography of Digital Participation

We ask participants about their preferences over a large set of issues included in the official government programs of presidential candidates (120 issues in France, 67 issues in Brazil.) To represent a broad pool of points of view, our team manually curated these issues. The platform design is initially inspired by pairwise comparison technique (Thurstone, 1927) and due to the cognitive burden involved in asking a large set of issues in this setup, we developed an online version of an elicitation method called Fallback Voting (Brams & Sanver, 2009). This mechanism splits the participation into two steps: approval and ranking screens. The platforms were built in accordance with protocols for personal data and following ethical guidelines. As a side note, Chapter 4 thoroughly explores the data collected by our platforms regarding interaction and careless participation.

We start exploring the demographics of online participation. We found that participation in both platforms was skewed toward young, well-educated men from urban areas. Furthermore, we observed that participants were mainly located near the institutions involved in the projects (Toulouse in France and Pernambuco in Brazil.) Despite this evident skewness of participation, previous work has established that digital participation still contributes valuable information about participants’ preferences for research purposes (Ramsey et al., 2016). Moreover, we do not claim representativeness of the population; instead, we focus on under-

standing the preferences collected from these users who decided to participate in our platforms voluntarily.

Agreements, Disagreements, and Divisiveness

First, we calculate a collective ranking of preferences (we named it a ranking of agreements henceforth) using a simple technique, such as the Win Percentage (fraction of times that an issue was prioritized over all their comparisons). We find that the main priorities were “Plan to use 100% renewable energies by 2050”, “Increase personnel in public hospitals”; “Increase the minimum wage” in France, and “Valorize the minimum salary”, “Expanding the guarantee of citizenship for the most vulnerable and brings a universal minimum income”, and “Invest on the management of the healthcare system” in Brazil. From using a straightforward aggregation function to more sophisticated approaches, such as AHP or TrueSkill, we find that the ranking of agreements was virtually the same. Despite the main agreements reached a high consensus among participants, they did not represent the most discussed issues by political candidates on social media and their campaigns. It seems reasonable to think, in light of theoretical evidence, that divisive issues—that by definition would not be in the top-ranked for any aggregation function—should provide more information over citizen divisions.

Second, we move our focus to analyze divisiveness. To this end, we split the population into six groups: Political Orientation (Left/Right), Sex (Male/Female), Location (Capital/Region), Age (Younger/Older than the median), Education (Undergraduate or more/Less than undergraduate), and Zone (Rural/Urban), relying us on self-reported data. We intend to compare differences in the ranking of agreements according to groups. Our findings suggest that most divisive issues per group were more correlated to the ones discussed on social media, in line with literature in issues salience (Ash et al., 2017). Nevertheless, this analysis is biased towards the self-reported socio-demographic characteristics asked by our experiment. For instance, our data do not cover additional factors that could explain divisiveness, such as religious affiliation or race in Brazil (Layton et al., 2021). In this sense, we introduce an agnostic approach to identify divisive issues that we call *Divisiveness*. The main advantage of our method w.r.t the state-of-the-art is that it can be applied for any aggregation function.

Intuitively, Divisiveness is inspired by a design feature of pairwise comparison: in each comparison, a user must select either issues A or B . Our assumption is that the population that chooses issue A is statistically different from the population that chooses issue B . By repeating this procedure for all issue pairs, we formalize the Divisiveness of issue i as its average difference in score across all issue pairs, i.e.:

$$D_i = \frac{1}{N-1} \sum_{i \neq j}^N \sqrt{(S_i(P_i > P_j) - S_i(P_j > P_i))^2} \quad (1.1)$$

Where S_i is the score of i of a given aggregation function S , $P_i > P_j$ represents the preferences of the population that selected issue i over j , and N is the number of issues.

Overall, we find that Divisiveness scores are uncorrelated to scores obtained by its corresponding aggregation function, providing new information that is not captured purely by these functions.

Multidimensional Divisiveness What is striking in our analysis is that divisive issues labeled by our approach can be related to specific groups. For example, “Revocation of the spending ceiling” in Brazil is more divisive for Location and Age dimensions. However, the divisiveness of “Expand the privatization of state-owned companies and national infrastructure concessions” is mainly driven by the political dimension. Conducting a regression analysis, our findings suggest that divisiveness can be interpreted as a multidimensional metric, finding that, as expected, political orientation is the dimension with the most explanatory power. Due to the fact that our data set does not contain additional relevant characteristics, such as religious groups or race, we highlight that our approach can still identify controversial issues from those groups.

Then, we borrow some axiomatic and matrixial properties of social choice to explore our ranking agreements and divisiveness. For this, we study Pairwise Efficiency: the fraction of times that an aggregated-higher ranked issue beats an aggregated-lower ranked issue in their individual comparison and the Independence of Irrelevant Alternatives (IIA): how much the ranking changes after removing an “irrelevant” alternative. We observe that most violations are located in the middle of the ranking. By relaxing our criteria, we tend to observe a steady increase in satisfaction with these properties. Nevertheless, top scores issues from both rankings are consistent and satisfy the axioms in almost all cases. This result is interesting since it gives us the glance that despite the pessimistic views that it is impossible to fairly aggregate preferences (Arrow, 1950), especially under a large set of issues (in our case, even 120), our findings suggest that most violations detected on the ranking of agreements are in fact, part of the most divisive issues. Next, inspired by techniques of matrixial decomposition from linear algebra, we use Singular Value Decomposition (SVD) to compute the eigenvectors of our pairwise comparison matrix. Our view is that the unitary vectors of this decomposition can be correlated to measures of agreements and divisiveness, similar to the idea that PCA can mainly explain factors in dimensionality reduction. We find that the first and second eigenvectors are positively associated with measures of agreement, and higher-order eigenvectors (third in France and fifth in Brazil) are associated with our measure of divisiveness.

To conclude, we expand the divisiveness analysis towards using benchmark data sets from computational social choice and synthetic data. Our assumption is whether Divisiveness is a new dimension for any aggregation function and does not depend on a particular context (in this case, election periods); its properties should behave similarly to the results presented using the French and Brazilian data. We show that the key results presented throughout the chapter were reproduced in synthetic data and benchmark data sets from Computational Social Choice. In other words, the divisiveness scores are uncorrelated to the output of any aggregation function and seem to provide new information over divisions that are not captured primarily by aggregation functions. In light of the evidence, we interpret Divisiveness as the second central moment of any aggregation function, similar to the relationship between mean and standard deviation.

1.4.2 Chapter 3: Mapping Citizen Division Through Elections

Chapter 3 introduces the concept of Election Polarization (EP) to measure the division level among citizens on Election Day. Given that the main political consequence of an election is dividing the electorate into winners and losers (Wojtasik et al., 2013) and motivated by methodological approaches from economics (Esteban & Ray, 1994; Wolfson, 1994), we consider that these conditions can be put together into the EP. Specifically, we borrow the axiomatic definition of polarization proposed by Esteban and Ray that stated that “the total polarization in the society is postulated to be the sum of all the effective antagonism,” to argue that each political candidate (or party) represents an effective antagonist, and the sum of all their antagonisms represents the election polarization. Then, inspired by previous work in statistics and polarization, we consider that this polarization can be decomposed into the performance of candidates within a geography (Within-EP), the interaction (or competitiveness) between candidates (Between-EP), and the candidates’ relative relevance. We also present an agnostic approach that relies exclusively on election data. We subsequently show the significance of this distinction by addressing the phenomena captured by elections.

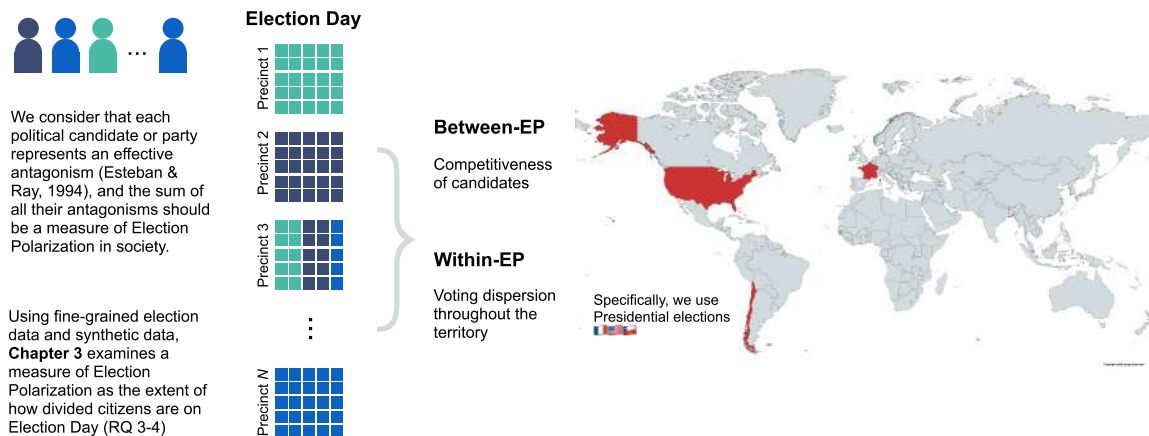


Figure 1.3: Conceptualizing the Election Polarization.

Additionally, we suggest a basic guideline for the structure of election data. This guideline follows recommendations from tidy data and continues the work from Baltz et al., 2022, who systematically curated election results in the United States. As part of our contribution, we deposited on GitHub over ten million data points collected and curated from presidential elections in France and Chile (<https://github.com/cnavarreteliz/election-polarization>).

Numerical Properties of EP & Empirical Validation

First, we explore the numerical properties of our approach. We simulated synthetic elections held for two and three candidates. We show that our method successfully labeled theoretical expectations of “polarized” elections: those with both high voting dispersion and competitiveness. In addition, we compare our results with traditional polarization measurements, such as Esteban and Ray, Wang and Tsui, and Reynal-Querol, finding that even when we observe a strong positive association between those techniques, they tend to fail as the voters are more clustered concentrated or under highly dispersion scenarios.

Second, we use presidential election data from Chile (2013-2021), France (2002-2022), and the United States (2000-2020), to empirically explore our method. Our assumption is that whether the EP is a proxy for political polarization, it should behave similarly to its theoretical associations. We find that our method successfully labeled polarized elections and regions, both at the national and regional levels. A few examples are the 2016 presidential election in the U.S. and a growing trend in France from 2017 to 2022. Next, we find that state-level political polarization data from the United States correlates to Within-EP and political interest–proxied by voting turnout–positively correlates to Between-EP. Altogether, we suggest that an election can be considered “polarized” by simply knowing whether two conditions are satisfied: candidates’ strongholds are geographically clustered, and the election reaches high levels of participation.

Robustness of Election Polarization

Next, we perform five robustness analyses to validate our method by examining edge cases that could potentially bias the EP. First, two elections held on the same day should return the same result. For this, we compare the state-level EP calculated using the U.S.’s 2016-2020 presidential and senate elections, finding a strong positive correlation in all comparisons ($\rho > 0.8$.) Second, aggregating election data at granularity should not alter the results. Here, we calculate the region-level EP for France and the U.S. by aggregating data on precinct and commune scales, showing the results are virtually the same. Third, abstentions and spoilt votes should not modify the EP. We consider these types of votes as additional candidates, finding that high levels of abstention bias the Between-EP. Fourth, since our approach cannot compare elections with a different number of candidates, we study the convergence of EP using a fraction of all candidates running in the election. We find that the EP computed using at least 80% of preferences is strongly associated ($\rho > 0.8$) with the one calculated using data from all preferences. Fifth, EP computed from runoffs in France and Chile are mostly uncorrelated to the first rounds despite the short period between rounds.

Our results yield a series of considerations that must be followed in order to use the EP. First, the geographical hierarchy of an election represents the largest aggregation level to calculate the EP. The reason is that local candidates’ quality could bias the party’s support in a location. For example, the House of Representatives election outcomes in the United States do not work for computing state-level EP since citizens in two congressional districts from the same state cast different ballots. Second, our method is not valid for non-free-and-fair elections—a referendum with 99% approval in a dictatorship does not mean that society is not polarized at all. Finally, run-offs would be unable to map EP since it is documented that candidates moderate their political positions in second rounds, and we find that in some cases, the EP from the first and second rounds are uncorrelated.

1.4.3 Chapter 4: Online Involvement in Two Real-World Digital Democracy Systems

Having introduced methodological contributions to mapping divisiveness and polarization in our previous chapters, we still miss an extensive understanding of an interesting feature collected by our platforms: A/B tests for interface in France (screens with different numbers of issues) and Brazil (Fallback voting vs. Pairwise comparison). Specifically, Chapter 4 examines

data collected by our two platforms from the lens of behavioral studies in psychology. We borrow some definitions from this discipline to study how the content and interface condition early dropouts, data collection, and careless participants in online involvement. Notably, this chapter speaks to our *RQ.5* and *RQ.6*.

In France, we implemented an adapted version of Fallback Voting that split the number of issues displayed per screen (4, 5, or 6). In Brazil, we compared Fallback Voting (uniquely of 5 points per screen) with the classical pairwise comparison framework. We randomly assigned participants to different scenarios, and they were unaware of the existence of others. Furthermore, the order of questions was randomized.

Contentious Issues in Online Participation

The fact that our platforms randomized the order of issues displayed to participants suits a good natural experiment to examine how the content affected online involvement. Participants could decide when to leave the experiment because staying on the platform was voluntary. In this sense, we first assessed whether the mere reading of a divisive issue in the first approval screen increased the odds of early dropouts. We reported that reading contentious issues in anonymous environments did not increase the odds of early dropouts; nevertheless, some endogenous characteristics of the screen, such as the length of issues in Brazil, increased the retention.

Exploring Data Quantity and Quality

Next, we compare the number of issue pairs collected via fallback voting to the classical pairwise comparison framework. We show that fallback voting could dramatically increase data collection, which is determined by the nature of participation in approval screens. Whether the platform planners expect greater approval ratios, more traditional vias of involvement, such as only rank screens or pairwise comparison, are better alternatives. Conversely, approval screens represent a good framework for data collection in cases with a balanced rate between approvals and disapproval. The reason is that it is more difficult to distinguish between relative preferences under greater approvals.

Then, we focus on understanding our data in terms of data quality. We proxied this concept by using six measures from behavioral studies and social choice to assess carelessness: percentage of abstentions in approval screens, longstring index (the longest chain of consecutive responses over the same choice in approval screens), interaction index (a dummy value to represent whether participants modified the randomly assigned order of issues in rank panels), response speed, consistency, and transitivity. While increased abstention is the rule in both platforms over time, these differences are relatively marginal. On average, participants increased around 3% their abstentions every 10 minutes. The most significant change is concerning the interaction index, showing that participants decreased by around 10% at the end of the experiment. No concluding evidence is seen in response speed or longstring index.

Interestingly, we found that the self-reported demographic characteristics of participants were associated with some measures of carelessness. For instance, female participants were more likely to abstain over an issue in approval screens than male participants. In the same

vein, older participants modified a lower percentage of the randomly assigned order of issues in rank screens than younger participants. Even though we recognize that these findings are evidence of carelessness by these features, we claim that this is not enough evidence to suggest carelessness from these groups. On the one hand, the platform itself was not intuitive enough, as we believed, framing that the drag-and-drop feature could explain the lack of interaction of older participants rather than merely carelessness. On the other hand, simply abstaining in a greater percentage does not mean that female participants were careless.

Finally, we raise a series of recommendations for platform planners grounded in our results. First, we demonstrated that the measures from behavioral studies contribute to understanding participation in digital democracy systems. Second, we suggest that straightforward decisions, including engaging instructions and attention checks, increase careful participation. Third, we point out that the interpretations of results and use of behavioral studies measures must be seen beyond a statistical perspective. For instance, the *Mahalanobis distance* is well-studied in this literature to measure the distance to the average response and thus flag outliers. Nevertheless, in our view, this index is unsuitable for an environment focused on democracy.

1.4.4 Dissertation Resources

The source code, data curated, and algorithms implemented are deposited and maintained in the following repositories and interactive resources:

Chapter 2

- Data set: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi%3A10.7910%2FDVN%2F8E0EA4>
- Released website in France: <https://monprogramme2022.org> (Accessed on September 17th, 2023)
- Released website in Brazil: <https://brazucracia.org> (Accessed on September 17th, 2023)

Chapter 3

- Interactive platform: <https://electionpolarization.com> (Accessed on September 17th, 2023)
- Data set: <https://github.com/cnavarreteliz/election-polarization>

Chapter 4

- Opencracia (Javascript library): <https://github.com/CenterForCollectiveLearning/opencracia>

Miscellaneous

- Comchoice (Python library for the aggregation of preferences): <https://github.com/CenterForCollectiveLearning/comchoice>

1.4.5 Summary

The study of citizen division and polarization has received considerable interest in political science, economics, and computer science, especially after unforeseen political events, such as the 2016 U.S. presidential election and the surprisingly good performance of the far right in Europe. This dissertation offers an interdisciplinary view to estimate issues' divisiveness and election polarization by relying on data collected by experimental digital democracy experiments and administrative records from elections.

Understanding Issues Divisiveness through Online Participation: Evidence from the 2022 French and Brazilian Presidential Elections

Abstract

Digital technologies have contributed to enhancing citizen participation around the world. While a vast literature examines collective decisions in terms of agreements, little work has assessed citizen disagreements from online involvement. Here we present data collected from two real-world digital democracy systems released during the 2022 presidential elections in France and Brazil. These systems aimed to create a collaborative government program between participants. We analyze agreements and disagreements within the data and propose a metric of Divisiveness, agnostic to any aggregation function and socio-demographic characteristics of participants, to identify divisive issues. Our analyses raise that Divisiveness can provide additional information about citizen preferences not captured by traditional aggregation functions, and their interpretations can be applied beyond online participation data.

This chapter is based on the following articles:

Navarrete, C; Macedo, M; Colley, R; Zhang, J; Ferrada, N; Mello, M; Lira, R; Bastos-Filho, C; Grandi, U; Lang, J; Hidalgo, C. (2023). Understanding political divisiveness using online participation data from the 2022 French and Brazilian presidential elections. *Nature Human Behaviour*.

Colley, R; Grandi, U; Hidalgo, C; Macedo, M; Navarrete, C. (2023). Measuring and Controlling Divisiveness in Rank Aggregation. *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023)*.

2.1 Introduction

The spread of digital democracy systems worldwide has dramatically contributed to increasing citizen involvement. Nevertheless, moving from sporadic to more active and continuous forms of participation represents a challenge for citizens and our public institutions. While a vast literature examines citizen participation in terms of agreements, little work has carefully examined divisions that emerged from online interactions. Here, we use the 2022 presidential elections in France and Brazil as a chance to study citizen agreements and divisiveness. To this end, we released two real-world collaborative government program builders. These tools, Mon Programme (<https://monprogramme2022.org>) and Brazucracia (<https://brazucracia.org>), encompassed a broad collection of 120 issues in France and 67 issues in Brazil. These platforms allowed us to study citizen priorities, though they also provided evidence of political divisiveness in electoral periods.

This chapter also coined the term *Divisiveness* to refer us to an agnostic method to map divisive issues. This approach can be applied to any particular aggregation function and does not require socio-demographic characteristics of individuals, contrasting with existing work that pre-defines controversial issues and highly depends on self-reported information of individuals. We subsequently show that our approach provides new information on issues mostly uncorrelated with its corresponding aggregation function, suggesting that it can be understood as the second mathematical moment of them, analogously to the relationship between mean and variance. We explore the numerical properties of Divisiveness using both data collected from our platforms and benchmark data sets from Computational Social Choice. Our results show that Divisiveness could systematically label controversial issues, and the results are consistent under several scenarios, such as using non-political data sets or synthetic preferences data.

Overall, our study contributes to Computational Social Science (Lazer et al., 2009) by providing data collected in two electoral periods and an agnostic approach to characterize controversial issues. This Chapter is structured as follows. Section 2 provides an overview of online participation, divisive issues, and the study of polarization measurements. Section 3 introduces our methods to compute divisiveness. Section 4 reports our main findings, and Section 5 discusses the role of online participation in understanding citizen preferences.

2.2 Literature Review

This section provides an overview of the use of digital technologies in citizen participation and the relevance of examining divisive issues in electoral periods. Moreover, we discuss how these topics can be utilized to map polarization in society.

2.2.1 Digital Technologies in Political Episodes

The spread of Internet access and recent advancements in digital technologies have radically shaped our social interactions (Hoehe & Thibaut, 2022). Today, nobody doubts that our democracy has been touched by this medium, leading to the emergence of digital democracy. According to Hacker and van Dijk, 2000, digital democracy can be defined as the use of

information and communication technology for purposes of enhancing political democracy or the participation of citizens in democratic communications. This discipline has seen flourishing interest from academia and digital activists (See, e.g., Helbing and Pournaras, 2015; Small et al., 2021) that have perceived digital technologies as a form of being involved with decision-making processes in public institutions transitioning from sporadic participation events to more systematic forms of communicating with the government.

Undoubtedly, digital technologies have impacted our forms of interaction in society. Nonetheless, digital democracy is not constrained to online platforms released in episodes of political upheaval. Two popular applications include Participatory Budgeting (PB) and Voting Advice Applications (VAAs). On the one hand, Participatory Budgeting can be defined as a process in which participants collaboratively decide how to spend a certain budget in a territory. The origins of PB can be traced back to the 80s in Brazil (Cabannes, 2004) and today cover more than three thousand cities worldwide (Su, 2017). The success of PB is grounded in its simplicity: a public institution (e.g., a city council) allocates an annual budget to spend on a set of projects. These potential projects can either be submitted by citizens or have already been pre-selected by the institution. Next, citizens are invited to participate in an online platform to up/downvote those projects for a certain period. Its success can be associated with the feature of “seeing” implemented in the winning projects. On the other hand, Voting Advice Applications (VAAs) corresponds to a family of recommender system that guides citizens regarding their similarities or “matches” with political candidates and parties in electoral periods (Garzia & Marschall, 2012). There is likely no other family of digital democracy systems with a global adoption other than VAAs today. Despite its massive adoption around the world (See, e.g., Vote Compass: Lees-Marshment et al., 2015), data collected via these platforms are rarely used in research since most of these efforts are released by newspapers, representing in some cases a black box both for participants and scholars interested in this line of argument.

It should be highlighted that an under-mentioned characteristic of these online participation systems is their interdiscipline. This can be exemplified in a Participatory Budgeting platform. On the one hand, team members with public policy expertise can be implicated in pre-selecting issues of concern. On the other hand, research communities related to the study of preferences (e.g., social choice) can be involved in conceptualizing efficient methods to aggregate these preferences (Benade et al., 2021; Peters et al., 2021). Despite this “Golden Age” for digital participation, the role played by digital technologies in democracy is seen with concern in some scholarship circles (Fukuyama, 2021; Persily, 2017), that perceive the risk of, for example, how social media can lead the formation of echo chambers (Cinelli et al., 2021), conditioning the proliferation of polarization. For example, Kim et al., 2018 found that political campaigns on social media micro-targeted voters with ads on divisive issues.

2.2.2 Design of Online Participation

Thus far, we have introduced some examples of how digital technologies have contributed to citizen involvement in such diverse political episodes as social movements and electoral periods. Now, we briefly examine some mechanisms used to collect and analyze citizen preferences in these episodes. Specifically, we position ourselves in the study of the aggregation of preferences. This line of argument can be traced back to the Enlightenment period (Borda, 1784; Condorcet, 1785) and is motivated by the interest of intellectuals in understanding paradoxes and methods for aggregating individual preferences. Centuries later, the seminal

work from Kenneth Arrow inspired a new strand of literature in economics named “social choice theory”, whose breakthrough contributions have emphasized the relevance (or rather impossibility) of finding a fair method to reach a collective decision (Arrow, 1950; Gibbard, 1973; Satterthwaite, 1975). Despite this pessimistic point of view, a line of argument in computer science recognizes that beyond theoretical impossibilities, there remains potential for exploring how manipulable the aggregation methods are using computational techniques (Brandt et al., 2012). This research community, named Computational Social Choice (COM-SOC), has enlightened recent discussions of how elicitation methods studied in this literature can contribute to digital democracy systems (Brill, 2021), finding either experimental online participation systems (Boudou et al., 2022), the use of crowd-sourcing for collective decisions (Lee et al., 2014), more efficient voting rules for Participatory Budgeting (Benade et al., 2021; Peters et al., 2021), to name a few.

We focus on a well-known technique to gather individual preferences called pairwise comparison. This method can be traced back to Thurstone, 1927. He suggested that relative comparisons between two alternatives can effectively create an aggregate ranking of preferences. Today, research based on this technique ranges several communities, such as the aggregation of preferences (See e.g., Copeland, 1951; Saaty, 1988), urban planning (e.g., Sallés et al., 2013) digital democracy (e.g., Salganik and Levy, 2015), social science (e.g., Awad et al., 2018), and even recommender systems (Kalloori et al., 2018). The reasons why the pairwise comparison is widely accepted can be addressed twofold: First, since it reduces the cognitive burden in online platforms, and second, data structured in relative comparisons is rich for data analysis. These reasons can be exemplified by the work of Awad et al., 2018. They built an online platform (For details, see <https://www.moralmachine.net/>) for gathering human perspectives on decisions made by self-driving cars. Asking for decisions that a machine should make in relative terms (e.g., “What should the self-driving cars do?”), they yielded a study to enlighten the discussion over the dilemmas involved in this new technological change. Thus, Lee et al., 2014 suggest that only using a small fraction of pairwise comparison preferences is possible to approximate the ranking using complete data.

2.2.3 Issues Divisiveness and Polarization

While efforts in digital democracy have emphasized the relevance of building systems based on consensus and agreements (Helbing et al., 2023), little attention has received the salience of characterizing divisive issues. This is relevant since disagreements are part of our social interactions and are considered healthy for our democratic coexistence (Huckfeldt et al., 2002). Nevertheless, there is a point at which these controversies can trigger feelings of anger that, consequently, polarize our society.

Nevertheless, defining polarization is not straightforward. While an intuitive definition can be the existence of two or more contrasting groups in society, scholars have emphasized that this concept must be distinguished into several branches, such as affective (Iyengar et al., 2012), political (Tucker et al., 2018), issues (DiMaggio et al., 1996), mass/elite (Abramowitz & Saunders, 2008; Fiorina & Abrams, 2008; Levendusky, 2009), to name a few. The reasons for these distinctions rely on examining their effects according to the contexts. Here, we focus on establishing the polarization (or divisions) concerning issues. The interest in issues is grounded by its interpretation as a form of polarization. This literature has, for example, linked opinion variations over divisive issues with the study of polarization. A line of argument has explored whether attitudes change toward “family issues” (Family Gender Roles, Women’s

Roles, Abortion, Sexuality, Sex Education, and Divorce Law) can be framed to explain a rising polarization (DiMaggio et al., 1996; Hoffmann & Miller, 1998). Their causal explanation is tied to the fact that society becomes polarized whether the distribution of opinions over an issue tends to be bimodal and the opinions range is highly dispersed. Even though divisive issues always exist, they can be hidden by our social interactions. According to Gerber et al., 2013, people could perceive that discussing political issues is a double-edged sword since divulging one's choices may expose them to social pressures, leading to feelings of threat and anxiety (Simons & Green, 2018). Similarly, Chen and Berger, 2013 found that higher levels of controversy in issues decrease their discussion.

There are good reasons to believe that divergent opinions are a salient condition for candidates to differentiate their political platforms in electoral campaigns. According to Ash et al., 2017, political candidates spend more time on divisive issues when they are up for election and in response to higher news transparency. Similarly, Ansolabehere, Puy, et al., 2015 suggest that the relative relevance of issues for the electorate can be differentiated by their salience and divisiveness. Furthermore, issues that divide the government from the opposition are significantly more likely to be prioritized on the policy agenda (Martin, 2004), and issue-specific cabinet conflicts are positively associated with the issue attention in coalition agreement (Klüver & Bäck, 2019). A potential explanation refers to the fact that divisive issues are better at unraveling candidates' political stances—a key aspect in electoral campaigns that agreements cannot provide. For instance, “increase the minimum wage” or “improve health-care system” are often included in government programs of candidates running in presidential elections; nevertheless, these issues cannot substantially define their positions as “abortion” or “migration policies” do. In this sense, Kim et al., 2018 established that political campaigns on social media micro-target voters with ads on divisive issues, and Barberá et al., 2019 show that legislators are likelier to follow the discussion of public issues on Twitter. In fact, partisans associate more salience to issues aligned with their parties (RePass, 1971) and have a higher level of knowledge of facts that confirm their beliefs (Jerit & Barabas, 2012). The previous discussion opens the question of how divisive issues can be tied to the study of online participation. The salience of investigating this point touches on an essential angle of social interactions. We consider that digital participation is just a mirror of our beliefs and preferences. Hence, data collected via digital democracy systems should contain information on both agreements and controversies. Despite the relevance of this argument, little work has explored the divisiveness of preferences in online tools. The work of Cantador et al., 2020 using data collected from the Participatory Budgeting project Decide Madrid, shows that the most divisive issues do not always receive the same level of attention as issues that target large majorities and their nature can be related to ideological and socio-economic factors.

Having established that mapping the divisiveness of preferences can be, in fact, a form of polarization, we proceed to describe its measurements. For this, we turn to a line of argument from economics concerning social phenomena that can be addressed under this umbrella. In the decade of the 90s, scholars got interested in the hypothesis that societies had become divided into rich and poor groups due to a shrinking of the middle-class group (e.g., Esteban and Ray, 1994; Wolfson, 1994). Nevertheless, they considered that classical measurements of inequality, such as the Gini coefficient or Theil index, were not suitable to describe this phenomenon, drawing new methods based on the bimodal distribution of groups in society. Here, we highlight the seminal contribution Esteban and Ray, 1994, which propose an axiomatic perspective for defining a polarization measure, introducing two statements that we consider relevant for this chapter. First, they presented the identification-alienation framework, which suggests that individuals feel identified with those from the same group

but are alienated from those in other groups. To exemplify this framework, they referred to the poor/rich distribution of income, stating that individuals with greater income feel similar to the ones with greater income, though they feel alienated from the ones with lower income. Second, they suggested that the total polarization in society is equivalent to the sum of all effective antagonisms. That is, considering that we have a certain number of groups, the sum of all their antagonisms is the measure of polarization. The axiomatic perspective of Esteban and Ray (ER) provided a fresh outlook to the study of polarization. Since then, several works in this literature have addressed other conditions of polarization, with notable examples from Reynal-Querol, 2002 in the study of religious polarization and its relationship with civil wars and Zhang and Kanbur, 2005 leveraging the work from ER in order to characterize a Within/Between polarization in the study of income inequality. Furthermore, related literature in computer science has studied the preferences in terms of polarization (Can et al., 2015), diversity (Hashemi & Endriss, 2014), and cohesiveness (Alcalde-Unzu & Vorsatz, 2013).

In light of the evidence, despite the distance between political science and economics scholarships to address the study of polarization, these pieces of literature are tied by a general agreement: we are in the presence of polarization whether the distribution of a condition (issues' opinions, income groups, etc.), presents a bimodal distribution. Taken together, this chapter speaks to digital democracy and issues divisiveness communities. First, by sharing data collected via real-world consensus-based online platforms released by our research team in France and Brazil. Second, providing evidence on the interaction of agreements/divisiveness of participants in an electoral period. Third, adopting an agnostic approach to measure Divisiveness. Fourth, we further validate divisiveness as a second momentum of any aggregation functions, exploring their properties and extending their use with benchmark data sets from social choice.

2.3 Methods

The central hypothesis addressed in this Chapter is that digital technologies contribute to understanding citizens' agreements and disagreements. Notably, methodological approaches can benefit from data collected through these tools.

2.3.1 The Platform

Thus far, this chapter has argued that digital technologies enhance citizen participation. Drawing upon the literature, we assume that a consensus-based online platform can help us understand agreements and divisiveness in political contexts. For this, we use the 2022 presidential elections in France and Brazil as an opportunity to examine this argument. We released two online platforms in France (<https://monprogramme2022.org>) and Brazil (<https://brazucracia.org>) in 2022 to allow participants to create a collaborative government program. These platforms were launched two weeks before the first round of their respective elections and continued collecting preferences until the runoff.

Platform Design The platform design took into account the coverage of a large and broad number of policy issues (e.g., more than 100 in France). Due to this, we faced the challenge of

developing an online platform that, at the same time, collects a fine-grained set of preferences from users and does not introduce friction in participation. Therefore, we focus on elicitation methods that allow us to collect as much data as possible in a short time period. Here we implemented an online adaptation of Fallback Voting (FV) (Brams & Sanver, 2009). In FV method, users first approve or disapprove of a subset of issues, then only rank issues they approved previously. To the best of our knowledge, this is the first time FV has been implemented in a real-world digital participation setting. This design was chosen to overcome some of the limitations of pairwise comparisons. An election of 120 (France) and 67 (Brazil) issues involve 7,140 and 2,211 possible pairs to get complete information on participants (i.e., $N(N-1)/2$ issues pairs for a given N number of issues). The use of FV involves fewer screens than the traditional framework of pairwise comparison and theoretically maximizes our data collection in a time slot (This point is stressed in Chapter 4).

Universes Nevertheless, asking participants to rank a large number of issues on one mobile screen is cognitively a time-consuming task and involves non-user-friendly interfaces. To tackle this challenge, we implemented an A/B test to examine the nature of online participation and its consequences regarding data collection. In France, we divided the issues displaying sub-ballots of k issues per screen, where $k \in \{4, 5, 6\}$. Conversely, we split participants into a classical pairwise comparison framework and sub-ballots of 5 issues in Brazil (See more details in Chapter 4). We call these conditions as universes, represented by k . A JavaScript routine randomly assigned participants to a universe, and they were unaware of other universes' existence. In terms of our research questions, the additional analysis associated with the universes is studied in Chapter 4. Figure 2.1 illustrates the mobile screen for (a) Fallback Voting (four issues per panel) and (b) Pairwise Comparison. We address this discussion in detail in Chapter 4.

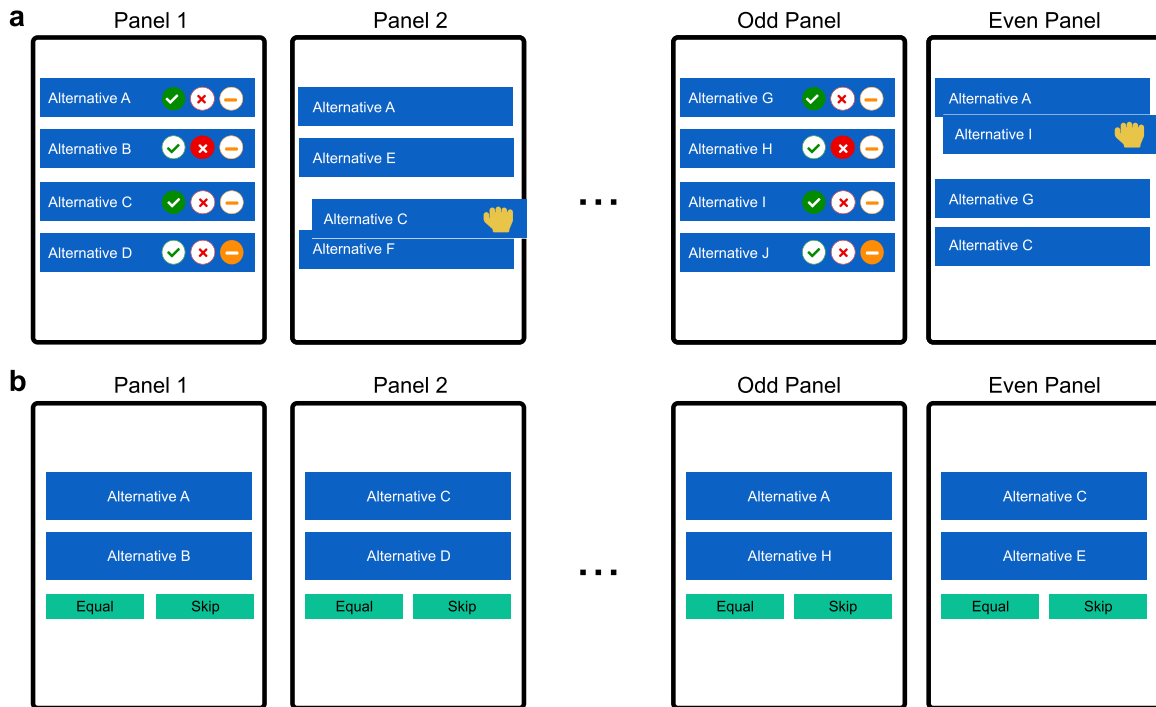


Figure 2.1: Illustrative examples for (a) Fallback Voting and (b) Pairwise Comparison. The maximum number of panels is given by $\lfloor N/k \rfloor$, where N is the number of issues and k the number of issues per screen.

Issues Selection To reduce potential bias toward issues from a specific group, we aim to cover a large spectrum of points of view. For this, we collected and curated 120 policy issues from all candidates in France and 67 policy issues from the six top candidates (according to pre-election polls: Luiz Inácio Lula da Silva, Jair Bolsonaro, Ciro Gomes, Luiz Felipe d’Avila, Soraya Thronicke, Simone Tebet) in Brazil. These issues were extracted from the official government program of candidates published on their websites. This curation was led by a lawyer and validated by senior team members.

User Identifier Participation in our platforms was anonymous. We created a random UUID to identify participants, storing this identifier across browser sessions (We use the javascript property `localStorage()`). Similarly, the IP address was hashed immediately and automatically using a one-way hash. We set up Google reCAPTCHA V3 (Google, 2023) to mitigate suspicious participation.

Platform Languages *Mon Programme* was available in French, English, and Spanish; *Brazucracia* was available in Portuguese, English, and Spanish. The issues’ labels displayed in Tables 2.1 and 2.2 represent shorter versions of the ones presented on the website (Full labels can be found in the official websites).

Self-Reported Socio-demographic Data We asked participants to voluntarily self-report six socio-demographic characteristics—sex, age range, political orientation, zone, education, and location—through a popup questionnaire. The popup was displayed automatically after a pre-defined number of panels (France: after 5, 4, and 3 panels for universes 4, 5, and 6, respectively; Brazil: after two panels in all scenarios). Yet, participants could also access the questionnaire manually. In terms of data curation, we removed duplicated records associated with the same UUID, maintaining the last record (based on timestamp). Since filling out this popup was voluntary, this data set does not include records of all the participants in the platform. Specifically, we collected self-reported information from 72% of participants in France and 71.3% of participants in Brazil.

2.3.2 Bot Detection System

To mitigate concerns over the influence of suspicious participation in our data set, we implemented reCAPTCHA V3—an API provided by Google that helps detect non-human participation by verifying requests that come from a participant. Scores range from 0 to 1, where 0 means a highly suspicious account. Even though we recognize that reCAPTCHA is not infallible in detecting suspicious actions, this score contributes to flagging potential attacks.

We define six criteria (independent of each other) to target participants that exhibited abnormal behavior regarding the volume or frequency:

1. **Unknown Universe.** Participants who voted in a universe not included in the experiment. As we mentioned before, the values accepted for universes are 2, 4, 5, and 6. Nevertheless, we identified participants associated to universes not defined in the experiment.

2. **Participant not Registered in the Consent Form.** We coded our front-end in order to prevent participants accessed to the website without accepting the consent form.
3. **Static Rank Screen.** Participants that updated the random proposal of issues in the ranking panel less than 10% of the time. Here we consider participants that responded at least three times.
4. **Suspicious Accounts Detected by Google reCAPTCHA.** Participants with an average reCAPTCHA score of less than 0.7.
5. **Suspicious IP Addresses detected by Google reCAPTCHA.** Participants belonging to IP addresses with an average reCAPTCHA score of less than 0.7.
6. **Over-Participation.** Participants registered preferences in the approval table for more than the number of issues included in the experiment. By design, this is an impossible scenario.

If a participant is part of at least one of these criteria, we flag her as a suspicious account. Following these criteria, we removed 57 suspicious participants in France and 37 in Brazil.

2.3.3 Data Curation

As mentioned above, our elicitation method splits the participation into approval and ranking panels. Here, we focus our main analysis on using data collected via ranking ballots. This decision is motivated by several reasons. First, this sample implicitly contains information about approvals since we only asked participants to rank issues approved previously. Second, given that our method for measuring divisiveness splits participants into sub-populations (we will give more details about this approach in subsequent sections), we cannot compute this measure just using pairs from approval ballots. This is due to highly approved or rejected issues being unable to split the population into groups in contrast to ranking ballots, in which we collect relative preferences. Third, we consider that using a mix of approval and rank ballots might bias the results and interpretation of divisiveness in favor of the approval data since it represents around 90% of issue pairs. Specifically, we removed data points using two criteria to reduce potential distortions:

- Participants labeled as suspicious by a simple Bot Detection System.
- Duplicated preferences of a participant on the same issues pair. We only kept the last data point based on the **datetime**. This condition mitigates the “power” of bots to distort the results. This step is not associated with the robustness analysis of our data, such as consistency and transitivity, that are presented in Chapter 4.

Additionally, we test our main findings using all data and not removing suspicious participation in the Appendix. Nevertheless, this chapter does not contest the quality of responses from users, assuming that all preferences were according to their *real* preferences (A more detailed analysis can be found in Chapter 4).

2.3.4 Ranking of Agreements

Our first goal is to understand the nature of agreements in online participation. For this, we compute a ranking of agreements using a straightforward method as is the Win Percentage (W_i)—or winning percentage of an issue i . This is a well-known technique to quantify performance in sports competitions. Let W_i be the fraction of times an issue i was selected over all its appearances. In a pairwise comparison data set, let x_{ij} be the total number of wins of issue i over j in the entire data set. The formula to calculate the win percentage is the following:

$$W_i = \frac{\sum_{j=1}^N x_{ij}}{\sum_{j=1}^N (x_{ij} + x_{ji})} \quad (2.1)$$

Beyond its simplicity, Win Percentage is useful for our purposes since it can compute a ranking of agreements for incomplete data, as is our case. This corresponds to Borda’s score when all preferences are complete.

As a side note, the issues IDs used in this Chapter correspond to the issues’ ranking computed using Win Percentage.

Bootstrapping In order to calculate confidence intervals for our results, we bootstrapped the data set 30 times, sampling half of the data in each iteration. For instance, suppose we bootstrapped our data set three times to calculate the Win Percentage. For an issue i , its score in each bootstrap is given by $\{0.5, 0.6, 0.7\}$. In this case, the score of i is $W = 0.6$, and the 95% confidence interval is $[0.487, 0.713]$.

2.3.5 Political Responsiveness

Then, we analyze the responsiveness of political issues concerning participants’ political orientation. Our assumption is that participants should show a greater attachment toward issues associated with like-minded presidential candidates. For this, we use the political orientation as a proxy. Specifically, the criteria to categorize the issues are the following:

France

We consider that an issue is associated with the political left when 50% or more of the candidates labeled as left-wing (Mélenchon, Jadot, Roussel, Hidalgo, Poutou, Arthaud) and less than 50% of the candidates labeled as right-wing (Zemmour, Dupont-Aignan, Pécresse, Le Pen, Lassalle) included it in their government program—vice versa for the right wing. In the case of Emmanuel Macron (a well-known “centrist”), we tested three scenarios: (i) excluding Macron from both groups, (ii) grouping Macron with right-wing candidates, and (iii) grouping Macron with left-wing candidates.

Right-Wing Candidates Eric Zemmour, Nicolas Dupont-Aignan, Valérie Pécresse, Jean Lassalle, Marine Le Pen.

Left-Wing Candidates Anne Hidalgo, Yannick Jadot, Fabien Roussel, Jean-Luc Mélenchon, Phillipe Poutou, Nathalie Arthaud.

Brazil

Unlike France, we asked participants in our Brazilian platform to self-report their intended candidate preference. Here, we focus on Luiz Inácio Lula da Silva (henceforth Lula) and Jair Bolsonaro (henceforth Bolsonaro) since they represented more than 91% valid votes in the first round. In contrast, the remaining candidates accounted for less than 5% of the electorate. We consider that an issue is from Bolsonaro whether he is included in his government program and is not included by Lula, and vice versa.

2.3.6 Ranking of Divisiveness

There are good reasons to believe that the ranking of agreements does not tell us the complete story behind citizen priorities. Having established that divisive issues provide valuable information and are associated with political engagement, we consider examining them as key. In this sense, we start decomposing a simple measure of divisiveness using self-reported socio-demographic data, i.e.:

$$d_i = S_i(A) - S_i(\tilde{A}) \quad (2.2)$$

Where $S_i(A)$ is the score of issue i for population A and $S_i(\tilde{A})$ the score of issue i for population \tilde{A} . In addition, by comparing the scores from two sub-populations (e.g., Male/Female), we can rely on measures of correlation or coefficient of determination (R^2) to estimate the level of agreement in the population. According to Palacios-González and Garcia-Fernández, 2012, the determinant coefficient can be used to measure polarization ($1 - R^2$). Nevertheless, this approach to measuring divisiveness has some limitations. First, relying on self-reported information involves participants providing real information. Second, some characteristics are not controlled by our experiment and could be relevant, such as religious affiliation or race in Brazil (Layton et al., 2021). Third, this method cannot quantify the relative divisiveness of an issue according to each other in a multidimensional spectrum.

Next, we introduce an agnostic approach to estimate divisiveness. Let i, j be two issues. The idea is that each issue's pair $\{P_i, P_j\}$ split the population into two groups. Our assumption is that the population that chooses issue i feels alienated from the ones that select issue j (Duclos et al., 2004; Esteban & Ray, 1994), though they feel identified with the ones that chose the same issue. Then, divisiveness can be formalized as the average difference in score of an issue i across all issue pairs, i.e.:

$$D_i = \frac{1}{N-1} \sum_{i \neq j}^N \sqrt{(S_i(P_i > P_j) - S_i(P_j > P_i))^2} \quad (2.3)$$

Where $P_i > P_j$ represents the population that preferred issue i over issue j , S_i represents the score of a given aggregation function S , and N is the number of issues. This can be interpreted as the second momentum of any aggregation function S . In the following sections, we ground the reasons behind this interpretation.

It should be noted that our approach could be seen from the lens of Esteban and Ray, 1994. For us, the divisiveness of an issue can be interpreted as an effective antagonism. Theoretically, the sum of the antagonisms of all issues that concerns a society should map its total polarization. Nevertheless, we do not control by the population size, as stated by one of the axioms of Esteban and Ray, in order to zoom in on the divisive issues from minorities. Even though this statement goes beyond the scope of this dissertation, in related work we further explored Divisiveness by taking into consideration the size of populations (Colley et al., 2023).

2.3.7 Multidimensional Divisiveness

Next, we hypothesize that an aggregated measure of Divisiveness derived from a population must encompass the aggregated effects of ideological and socio-demographic characteristics. To test this assumption, we set up a regression model presented in eq. 2.4:

$$y_i = \beta_0 + \beta_1 W_i + \beta_2 X_i^T + \epsilon_i \quad (2.4)$$

Where y_i is the dependent variable (in our case, D_i defined in eq. 2.3), W_i is the Win Percentage of issue i (defined in eq. 2.1), X_i^T is a vector term that encompasses the absolute value of disagreement (defined in eq. 2.2) of issue i with respect to six dimensions: Political Orientation (Left vs. Right), Sex (Male vs. Female), Location (Capital vs. Region), Age (Younger vs. Older than the median), Educational Attainment (Undergraduate or more/Less than undergraduate), and Zone (Urban vs. Rural) as control variables and ϵ_i is the error term. The variables have been normalized by subtracting the average and dividing by the standard deviation before running the regressions.

2.4 Results

This section outlines the main findings of this chapter. We emphasized the study of Divisiveness and its corresponding associations.

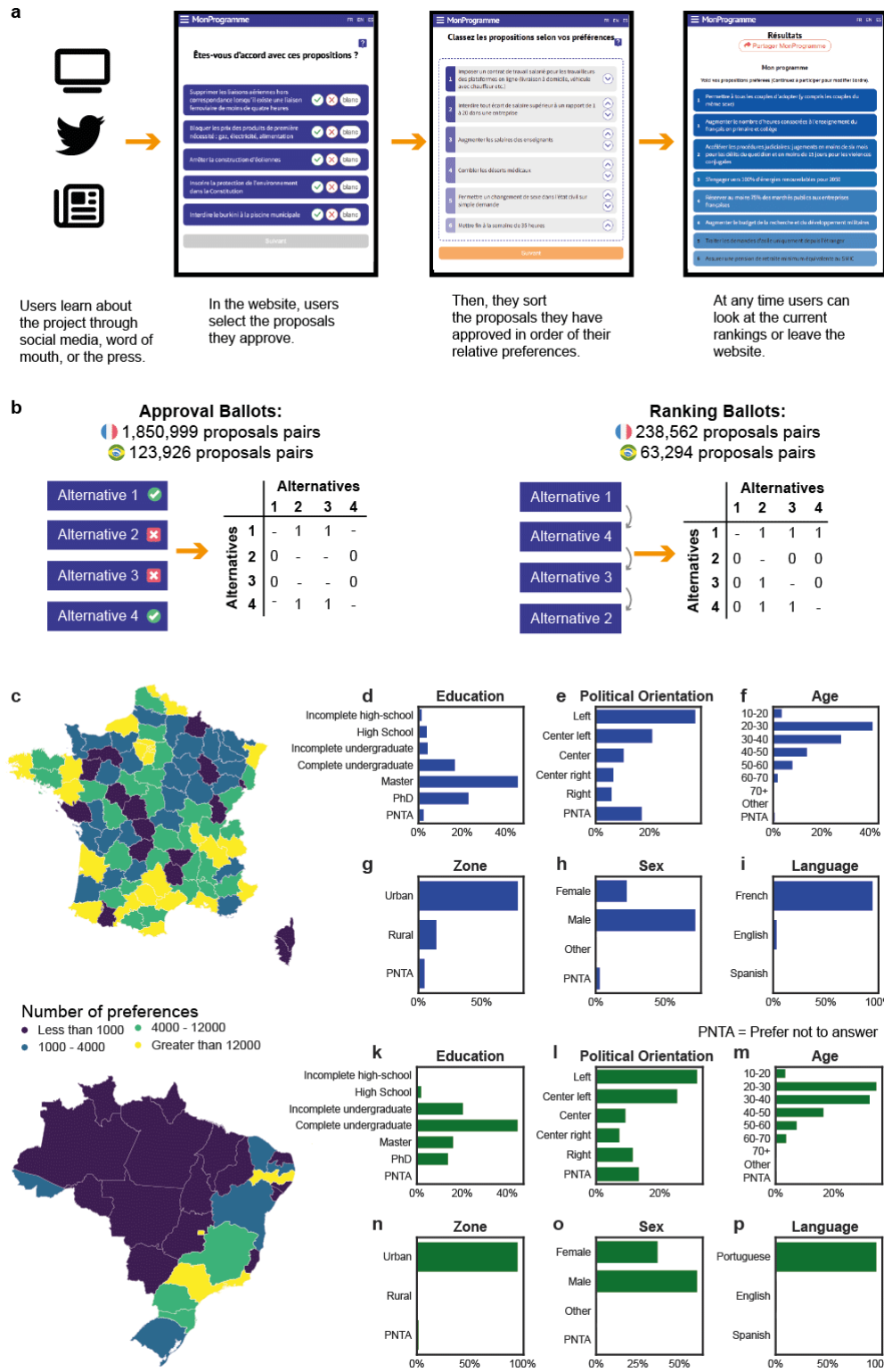


Figure 2.2: Participation platform and pairwise comparison data. a. Basic design of an approval + rank participation platform. Icons were obtained from Font Awesome Free Icons (v.5.0.0) under CC BY 4.0 License. b. Procedure to convert approval and rank data into pairwise comparison matrices. Further details of the method in the SM. (c-i) Demographics of digital participation in France ($N_{\text{Participants}}=1175$, $N_{\text{Pairs}}=1,705,104$) by (c) Geography: France Metropolitan by Departments (d) Education, (e) Political Orientation, (f) Age (g) Zone, (h) Sex, and (i) Language. (k-p) Demographics of digital participation in Brazil ($N_{\text{Participants}}=740$, $N_{\text{Pairs}}=157,280$) by (k) Geography: Brazilian States (l) Political Orientation, (m) Age, (n) Zone, (o) Sex, and (p) Language. X-axis values represent the percentage of self-reported participants. The number of participants corresponds to the ones that responded to the self-report questionnaire. In case a user has more than one response, we keep the most recent record. Additionally, the number of pairs excludes preference from users targeted as suspicious by our bot detection system.

To start, Figure 2.2 a illustrates the participation mechanism of our platforms. Participants learned about the projects both in social media and in local newspapers. As a side note, we promoted our Brazilian platform via Facebook ads to reach a wider audience. After giving their preferences over ten issues inside the platform, participants could access an individual ranking of preferences and see the collective ranking of agreements and disagreements displayed on the results page. Next, Figure 2.2 b summarizes our protocol to convert approval and ranking ballots from Fallback Voting into pairwise comparison data. As seen, the approval ballots contributed to around 90% of the issue pairs. We interpret that FV is efficient in terms of quantity since we collect around 2 million preferences for around a thousand participants. Nonetheless, an extensive analysis in terms of data quality can be found in Chapter 4.

2.4.1 Demographics of Online Participation

Now, we look into the demography of participation. Here, we limit our analysis to self-reported data voluntarily provided by participants, which corresponded to 72% of participants in France and 71.3% in Brazil. Figure 2.2 c and j show the Geography of Online Participation in France and Brazil, divided by departments and regions, respectively. Roughly, we observe that participation is concentrated in regions associated with the main affiliation of our research team (Toulouse in France and Pernambuco in Brazil) and most populated areas such as Paris in France or Rio de Janeiro and São Paulo in Brazil. Next, we decompose the participation into six dimensions: Educational level, political orientation, age range, zone, sex, and language displayed on the platform (Figure 2.2 d-i for France and Figure 2.2 k-p for Brazil). Here, we learn that most participants were young men from urban areas with high education, mainly from the political left. This finding is not surprising since previous efforts in online participation have reported self-selection bias towards similar socio-demographic groups (Best & Krueger, 2005; Mellon et al., 2017). Our study does not claim to be a representative sample of the population; instead, we focus on understanding the preferences of those who voluntarily self-selected to participate.

Id	Name	Win percentage
1	Plan to use 100% renewable energies by 2050	75.1%
2	Increase personnel in public hospitals	72.2%
3	Increase the minimum wage	71.5%
4	Further develop the French nuclear park	70.8%
5	Include ecology in the Constitution	68.9%
6	Reduce working hours to 32 hours per week	67.6%
7	Create a citizen income	66.8%
8	"Cap prices of essential products: gas, electricity, food"	66.5%
9	Prohibit any salary difference of more than 1 to 20 in a company	66.2%
10	Devote 3% of GDP to research and development	65.4%
11	Create a Constituent Assembly to pass to the Sixth Republic	65.4%
12	Lower retirement age to 60	65.0%
13	Increase the industrialization of the country	63.9%
14	Ensure a minimum pension is equivalent to the minimum wage	62.7%
15	Restore the solidarity tax on wealth (ISF)	62.5%
16	Increase number of doctors in rural underserved areas	62.3%
17	Increase teacher salaries	62.2%
18	Ban dangerous pesticides (eg neonicotinoides)	59.6%
19	Nationalize or renationalize some large companies (EG Telecom / Orange)	59.1%
20	Increase social housing	58.5%
21	Establish a Citizens' Initiative Referendum	58.3%
22	End the 35-hour working week	57.7%
23	Increase the retirement age	57.6%

24	Abolish fee-for-service pricing in hospitals	56.6%
25	Abolition of the law of the soil	56.4%
26	Add to the Constitution the right to voluntary termination of pregnancy	55.5%
27	Acquisition of French nationality only by descent or by merit	55.4%
28	Make those convicted of corruption unable to run for office	54.9%
29	Index pensions to inflation	54.8%
30	Develop a taxation to discourage programmed obsolescence	54.4%
31	Elect the National Assembly by proportional voting	53.7%
32	Prohibit single-use plastics	53.7%
33	Lower VAT on fuels	53.5%
34	Oppose intensive breeding and slaughter	53.3%
35	Recognition of the blank vote	53.2%
36	Restoration of border control by France leaving the Schengen agreements	53.1%
37	Guarantee the rights of asylum for refugees	52.9%
38	Engrave in the Constitution the superiority of French law over international law	52.5%
39	Only process asylum applications from abroad	51.9%
40	Reserve social security assistance only for people of French nationality	50.8%
41	Invest to maintain and rebuild local railway lines	50.3%
42	Increase the amount of APL (housing assistance)	50.1%
43	Increase paternity leave to match maternity leave	49.9%
44	"Introduce a universal rental guarantee to facilitate access to housing, financed by landlords and the State"	49.8%
45	Create police units and magistrates dedicated to violence against women	49.7%
46	Reduce legal proceedings time for everyday offenses and domestic violence	49.6%
47	Relocate the production of medicines to France	49.6%
48	Eliminate air routes when there is a rail connections taking less than 4 hours	49.6%
49	Continue the pension reform of the outgoing government	49.5%
50	Create spaces in EHPAD (accommodation institutions for dependent senior) and recruit staff	48.8%
51	Ban facial recognition in public spaces	48.6%
52	Introduce a tax on airline tickets when there is a comparable rail route	48.2%
53	Impose employment contracts for the workers of online platforms	48.0%
54	Introduce immigration quotas by profession and country	47.2%
55	Lift patents for Covid-19 vaccines	46.9%
56	"Allow all couples to adopt (married or not, including same-sex couples)"	46.6%
57	Reimburse medically assisted procreation for all women	46.5%
58	Require activities that promote employability to obtain the RSA (income support)	46.4%
59	Regulate undocumented migrants	46.3%
60	At least 75% of public contracts to be reserved for French companies	46.3%
61	Nationalization of the highways	46.0%
62	Suspend social benefits of parents of juvenile offenders	45.9%
63	Make taxes individual by removing marital quotient	45.9%
64	Increase the defense budget by at least 2% of GDP	45.9%
65	Eliminate university admission requirements	45.8%
66	Increase the military's research and development budget	45.5%
67	Establish a mandatory national civic service	45.5%
68	Pay adult disability benefits independently of their spouse's revenue	45.2%
69	Increase availability of space at day cares	45.0%
70	Create a large number of police and militarized police positions	44.6%
71	Leave NATO's Defence Planning Committee	44.5%
72	Disenfranchise those who attack those responsible for public authority	44.5%
73	Increase number of hours spent in French language education	44.4%
74	Limit family reunification	44.1%
75	Decriminalization of assisted death	43.6%
76	Teach a second language from primary school	43.5%
77	Deport foreign offenders at the end of their sentence	43.3%
78	Redirect a portion of a city's public funds to support rural areas	43.2%
79	Propose a simpler and more advantageous combination of employment and retirement	42.9%
80	Departmental elected officials will be the same as regional elected officials	42.9%
81	Pay a significant child allowance from the first child	42.7%
82	Generalize the teaching of computer code and digital uses from the 5th	42.7%
83	Submit foreign investments to the approval of a High Council for Economic and Digital Sovereignty	42.3%
84	Pay family allowance independently of family resources	42.2%
85	Lower charges for the self-employed	42.0%
86	"Limit inheritance tax up to €150,000 per child"	41.6%
87	Total corporate tax exemption for entrepreneurs under 30 years old	41.2%
88	Refrain from any military intervention without the mandate of the United Nations	40.3%

89	Do not send French soldiers to Ukraine	39.8%
90	Create a sex-offender registry	39.7%
91	Send weapons to Ukraine	39.7%
92	Offer student loans	39.7%
93	Expel foreigners whose behavior is part of radical Islamism and registered in the anti-terrorism files	39.1%
94	Reintroduce a seven-year presidential term	38.6%
95	Allow universities to have complete autonomy	38.1%
96	Raise the salaries of gendarmes and police officers	37.9%
97	Give 16 year olds the right to vote	37.4%
98	Restore the life sentences	36.8%
99	Close any mosque where sermons are hostile to France and the values of the Republic	36.6%
100	Allow non-Europeans to vote in local elections	36.5%
101	Lower the age of criminal responsibility from 18 to 16 years old	36.4%
102	Oppose Ukraine's entry into NATO	35.0%
103	Privatize the audiovisual public service	34.2%
104	Replace Parcoursup with a more transparent system	33.7%
105	Stop the construction of wind turbines	33.7%
106	Reimburse the transition of trans people	33.6%
107	Allow a change of sex in the civil status by a simple request	33.5%
108	Legalise cannabis	33.2%
109	Use facial recognition at the entrance to public transport to better identify wanted people	33.1%
110	Reduce the tax on real estate wealth by exempting it from 50% of the main residence	32.4%
111	Forbid hunting	31.1%
112	Add 2 more hours of sport per week in school	29.8%
113	Prohibit the wearing of the veil by school chaperones	29.7%
114	Remove of the TV licence fee	28.1%
115	Defend regional languages and cultures	27.6%
116	Guarantee the free movement of cars in cities	27.1%
117	Restore ENA (the National School of Administration)	23.9%
118	Establish full autonomy for Corsica	21.8%
119	Protect hunting and fishing rights	21.7%
120	Prohibit the burkini at municipal swimming pools	17.5%

Table 2.1: List of issues included in the French platform, ranked by win percentage.

Id	Name	Win percentage
1	Valorize the minimum salary to recuperate the purchasing power	73.7%
2	Create a program that expands the guarantee of citizenship for the most vulnerable and brings a universal minimum income	72.7%
3	Invest on the management of the SUS	72.0%
4	Strengthening career plans and valuing teachers	68.6%
5	Tax reform with change in burden reducing taxation on consumption and increasing income progressively so that the richest pay more	67.6%
6	Actions aimed at training and qualification of teachers	64.7%
7	Commitment to the goals stipulated by the National Education Plan	63.0%
8	Invest in the national system to promote technological development through funds and public agencies such as CNPq and CAPES	62.9%
9	No income tax for workers making up to 5 minimum wages.	62.9%
10	Strengthening the national vaccination program	62.6%
11	Equal pay policy between men and women performing the same function	60.2%
12	Invest in specific programs and actions aimed at the educational recovery of those affected by the pandemic	60.0%
13	Strengthen the popular pharmacy program	59.6%
14	Strengthen the energy supply with the expansion of clean and renewable sources	58.9%
15	Actions that guarantee internet access in public schools	58.4%
16	Increase transparency through compliance with the Access to Information Law	57.2%
17	Continue the policy of social and racial quotas for admission to higher education	57.0%
18	Investing in vocational education in line with labor market expectations	56.6%
19	Write a new labour legislation to include modern labour regulations and social protection	56.5%
20	Propose rules for the transparency of final beneficiaries of public resources	55.2%
21	"Strengthen democratic, secular and inclusive education with specific policies for people with disabilities, the LGBTQIA+ population and among other vulnerable groups"	54.0%
22	Increase the participation of women in politics and public management	53.5%
23	Reduce agricultural production costs and marketing price	53.2%

24	Recover lands deteriorated by predatory activities and reforestation of devastated areas	52.9%
25	Encouraging sustainable agricultural practices	51.9%
26	The complete opening of banking and fiscal secrecy of first and second level positions in the Executive Power.	51.7%
27	Curb drug mining and money laundering in the Amazon by increasing the number of ecological bases	50.3%
28	Policies and actions for debt renegotiation of households and companies	50.3%
29	Encourage research related to medicines	50.2%
30	Preservation of culture and demarcation of indigenous and quilombolas lands	50.0%
31	"Specific health policies aimed at women, LGBTQIA+ population, disabled people and among other vulnerable groups"	49.2%
32	Maintain the value of 600 reais for Auxílio Brasil	49.0%
33	Meet the carbon gas reduction targets assumed by the country at the 2015 Paris Conference	48.9%
34	Improve and reduce the prices of transport services through the structuring of concessions and public-private partnerships	48.7%
35	Structuring the medical career in the SUS with mechanisms of attraction and recognition	48.6%
36	Regulation and protection of workers' labor rights by application	48.4%
37	Encouraging entrepreneurship through credit facilitation and debureaucratization	47.6%
38	Health facilities with good performance should monitor and assist those with lower performance	46.5%
39	Creation of policies to provide hybrid work and home office for women with children	45.7%
40	Actions to curb tax evasion	44.9%
41	Revocation of the spending ceiling	44.9%
42	Improve public job positions and salary plans with incentives related to goals	44.1%
43	Creation of policies that guarantee the inclusion and permanence of the LGBTQIA+ population in the labor market	44.0%
44	Encouraging female entrepreneurship through the facilitation of credit and microcredit	43.6%
45	"Combining face-to-face teaching with distance learning in basic education, analyzing regional peculiarities"	43.5%
46	Policy for valuing state-owned companies and those against privatization	43.5%
47	Implement a federal Public Integrity strategy	43.1%
48	"Establish the basic foundations of the subjects, removing ideological connotations and with a view to parents as the main actors in children's education"	43.0%
49	Actions to combat illegal mining	42.8%
50	New fuel pricing policy	42.3%
51	Actions to encourage the creative economy	41.7%
52	Expand the privatization of state-owned companies and national infrastructure concessions	41.7%
53	Reinforce the consolidation of the national cancer care support program	41.3%
54	Expand the articulation between the public and private health sectors	41.2%
55	Fostering agro-industry and national production of inputs	40.1%
56	Country's formal adherence to the OECD Council's Public Integrity Recommendation	38.1%
57	"Offer Green Bonds to finance investments considered sustainable in the areas of transport, energy and between others"	37.9%
58	"Expand, redesign, and improve the qualification programs of the police"	35.3%
59	Culture-focused policies through articulation with private sector institutions and companies and civil society organizations	34.4%
60	Maintain current labor legislation	32.9%
61	Implement national guidelines for the promotion and defense of police human rights	32.3%
62	Increase national production of fertilizers	32.3%
63	Continue programs related to encouraging physical activity for primary care	32.2%
64	Create a new National LGBTI+ Public Policy Committee	31.9%
65	Encouraging mining activity within a logic of environmental protection	31.6%
66	"Consolidate and expand land regularization actions, allied to the strengthening of legal institutions that ensure access to firearms"	27.4%
67	Investment in the Armed Forces and promotion of their international participation as in UN-sponsored missions	25.7%

Table 2.2: List of issues included in the Brazilian platform, ranked by win percentage.

2.4.2 Ranking of Agreements

Tables 2.1 and 2.2 present the ranking of agreements for France and Brazil, respectively. The top issues in France include the "use of 100% renewable energy by 2050," the "increase of personnel in public hospitals," and "increasing the minimum wage." In Brazil, the top-ranked issues were "Valorize the minimum salary to recuperate the purchasing power," "Create a pro-

gram that expands the guarantee of citizenship for the most vulnerable and brings a universal minimum income,” and “Invest on the management of the SUS (the public healthcare system).” Roughly, we observe that raising salaries and better conditions in healthcare systems are highly prioritized topics that seem to transcend the geopolitical context. We go deep into the main agreements by examining them in terms of political support.

Figure 2.3 a-d shows the political responsiveness in France and Brazil. In France, participants approved a higher percentage of issues associated with their self-reported political orientation. Participants self-reported from the political left approved 83.7% of issues from left-wing candidates but only 28.5% of proposals from right-wing candidates. Similarly, participants self-reported from the political right approved 66.1% of issues from right-wing candidates but only 56.7% of issues from left-wing candidates. This result remains consistent after targeting Macron on the Left or Right (Figure 2.3 b-c). By relying on self-reported intended presidential preferences in Brazil, we find that participants with an intended preference for Lula approved 95.8% of issues from Lula but only 76.5% of issues from Bolsonaro. Similarly, participants with an intended preference for Bolsonaro approved 89.6% of issues from Bolsonaro but only 73.7% of Lula’s issues. To contextualize the previous percentages, it should be noted that, on average, the approval rate in France was around 50% and 77.4% in Brazil.

2.4.3 Ranking of Divisiveness

Divisiveness from Socio-Demographic Variables

Next, we split the participants according to six dimensions: Political orientation (Left/Right), Location (Capital/Region), Zone (Urban/Rural), Sex (Male/Female), Age (Young/older people), and Educational Attainment (Less than undergraduate/Undergraduate or more) to study divisiveness between groups (See Methods). Figure 2.3 e-j compares the ranking of agreements when we split the population according to these groups in France. Taking into account that the $1 - R^2$ can be interpreted as a measure of polarization, the most divisive dimensions included the Political orientation ($R^2 = 30.4\%$, $P < 0.01$), Sex ($R^2 = 72.4\%$, $P < 0.01$), and Age ($R^2 = 73.2\%$, $P < 0.01$). By looking at divisiveness by political orientation (e), we find that “Restoration of border control by France leaving the Schengen agreements (ranked 36 overall),” ranks first among self-identified participants from the political right ($W = 65.2\%$), but has much fewer support among participants self-identified from the political left ($W = 26.9\%$). This high level of political divisiveness ($d = 38.4\%$) means the issue is 38.4 percentage points of distance from members of the political left and right. Similarly, we find a high level of divisiveness for “Expel foreigners whose behavior is part of radical Islamism and who are registered in the anti-terrorism files” (rank 93, $d = 36.8\%$, $W_{left} = 25.6\%$, $W_{right} = 62.5\%$), and for “Deport foreign offenders at the end of their sentence” (rank 77, $d = 32.8\%$, $W_{left} = 28.4\%$, $W_{right} = 61.2\%$). Conversely, “Create a citizen income,” is a proposal that ranks higher among self-identified participants from the political left but relatively low among those self-identified from the political right (rank 7, $d = 17.6\%$, $W_{left} = 66.2\%$, $W_{right} = 48.6\%$).

The analysis of Brazilian data shows similar associations to the ones observed in the French platform. In this case, the most divisive dimensions include Political orientation ($R^2 = 0\%$, $P = 0.96$), Age ($R^2 = 52.4\%$, $P < 0.01$), and Sex ($R^2 = 65.1\%$, $P < 0.01$). Analogously to our previous analysis, Figure 2.3 k-o show the divisiveness of issues in Brazil exploring differences in (k) political orientation, (l) geography (here, capital includes municipalities of Brasilia, Sao

Paulo, and Rio de Janeiro), (m) sex, (n) Age, and (o) Educational attainment. The issues to develop “specific health policies aimed at women, LGBTQIA+ population, disabled people and among other vulnerable groups” (rank 31, $d = 37.7\%$, $W_{left} = 58.4\%$, $W_{right} = 20.7\%$) ranks higher among self-identified participants from the political left, but is less prioritized by self-identified participants from the political right. Thus, “Expand the privatization of state-owned companies and national infrastructure concessions” (rank 52, $d = 62.9\%$, $W_{left} = 12\%$, $W_{right} = 74.9\%$) is the main priority for participants self-identified with the political right, but it is ranked on the bottom for self-identified participants with the political left.

Again, by splitting participants by other socio-demographic dimensions, such as sex (male versus female), geography (capital versus regions, urban versus rural), age (older and younger people), and education, we still spot some controversial issues that can be ascribed to a particular dimension. For instance, “Reserve social security assistance only for people of French nationality” (rank 39 overall) is more prevalent among older participants than among younger participants in France, or “Equal pay policy between men and women performing the same function” (rank 11 overall) is the second priority for women but the 22nd for men in Brazil. One reason why political divisions lead to Divisiveness could be due that our platforms were settled in a political episode, and people more engaged in political participation are also leaner to online participation. Nevertheless, the fact that other dimensions can be linked to Divisiveness suggests that our approach can encompass a multidimensional measurement. We will explore the last point in subsequent sections.

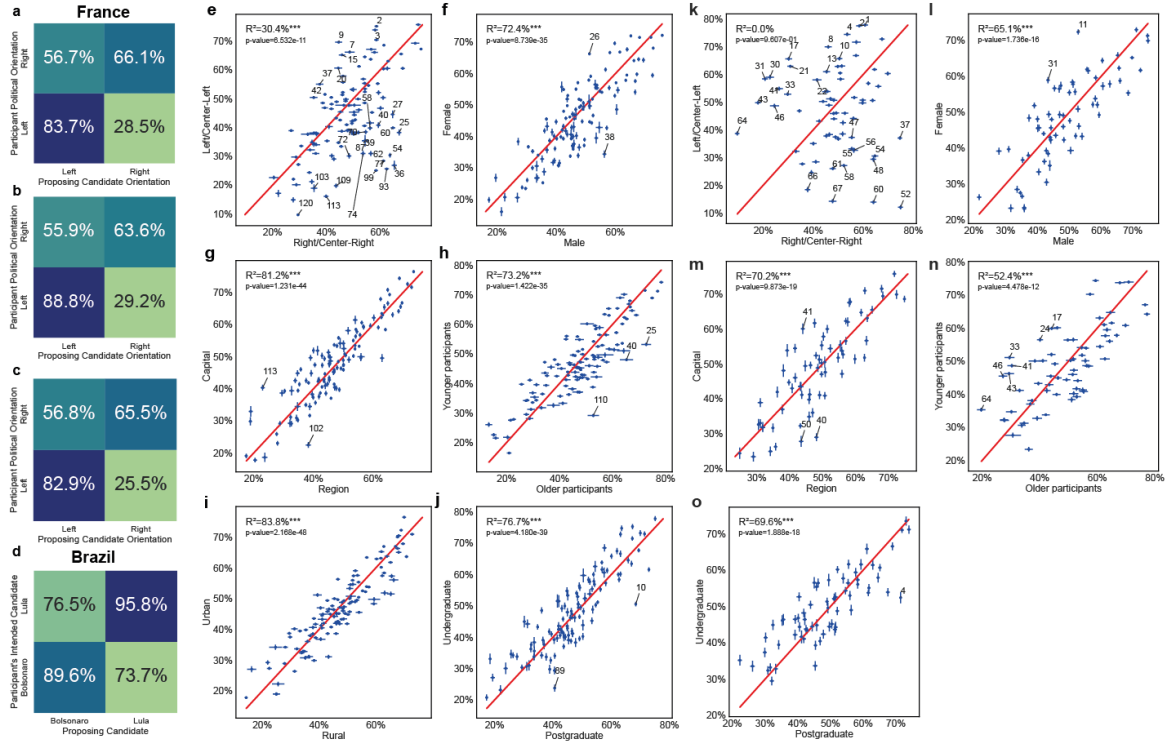


Figure 2.3: Understanding participants preferences. a-c Confusion matrix between participants self-reported preferences and the political orientation of the proposing candidates. (Participants_{Left}=690, Participants_{Right}=154). Emmanuel Macron was considered as part of the: a center (excluded) (Issues_{Left}=21, Issues_{Right}=9), b right (Issues_{Left}=19, Issues_{Right}=14), and c left (Issues_{Left}=8, Issues_{Right}=11). d Confusion matrix between participants intended to vote for Jair Bolsonaro/Lula and the proposals from Bolsonaro and Lula. (Participants_{Lula}=346, Participants_{Bolsonaro}=113; Issues_{Lula}=24, Issues_{Bolsonaro}=23). e-j comparison of proposal’s rankings when splitting the population according to e self-reported political orientation (N_{Left}=101,190, N_{Right}=22,043), f location (N_{Capital}=35,141, N_{Region}=118,067), g zone (N_{Urban}=131,252, N_{Rural}=26,792), h sex (N_{Female}=43,243, N_{Male}=116,697), i age (N_{Younger}=145,260, Participants_{Older}=20,490), and j education level (N_{LessthanUndergraduate}=18,011, N_{Undergraduateormore}=145,359) in France. k-n comparison of proposal’s rankings when splitting the population according to k self-reported political orientation (N_{Left}=24,910, N_{Right}=6,470), l location (N_{Capital}=6,676, N_{Region}=32,401), m sex (N_{Female}=14,437, N_{Male}=23,891), n age (N_{Younger}=33,685, N_{Older}=5,243), and o education level (N_{LessthanUndergraduate}=6859, N_{Undergraduateormore}=31,959) in Brazil. Error bars show 95% CIs computed using 30 bootstrap iterations (score of proposals) of half-size samples and are in some places thinner than the symbols in the figure. We label proposals showing a win percentage difference greater than 15% (IDs correspond to win percentage rank and can be obtained from tables 1 (France) and 2 (Brazil)). For e-o, we report the R² calculated as the square of Pearson’s correlation estimated from a two-sided alternative hypothesis, as determined by the SciPy library (v.1.9.3). The number of participants corresponds to the ones that responded to the self-report questionnaire, provided at least one preference (after excluding the “Equal” selection in Pairwise Comparison setup in Brazil), and were not targeted as suspicious by our bot detection system. *Note:* $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

2.4.4 Analysis of Divisiveness

In this vein, we use the definition of Divisiveness presented in eq. 2.3 to characterize divisive issues. Figure 2.4 a-b visualizes win percentage (x-axis) and divisiveness (y-axis) in France and Brazil. Overall, we report that both dimensions are uncorrelated (France $R^2 = 2\%$, $P < 0.1$; Brazil $R^2 = 12.5\%$, $P < 0.01$). The most divisive issues are “Create a Constituent Assembly to pass to the Sixth Republic” (11, $D = 53.6\%$), “Do not send French soldiers to Ukraine” (89, $D = 53.2\%$), “Engrave in the Constitution the superiority of French law over

international law” (38, $D = 51.8\%$), “Restore ENA” (the National School of Administration) (117, $D = 50.8\%$), “Acquisition of French nationality only by descent or by merit” (27, $D = 50.6\%$) in France. Similarly, the most divisive issues are “Revocation of the spending ceiling” (41, $D = 70.7\%$), “Expand the privatization of state-owned companies and national infrastructure concessions” (52, $D = 67.8\%$), “Investment in the Armed Forces and promotion of their international participation as in UN-sponsored missions” (67, $D = 67.3\%$), “Maintain current labor legislation” (60, $D = 61.0\%$), “Consolidate and expand land regularization actions, allied to the strengthening of legal institutions that ensure access to firearms” (66, $D = 60.5\%$) in Brazil.

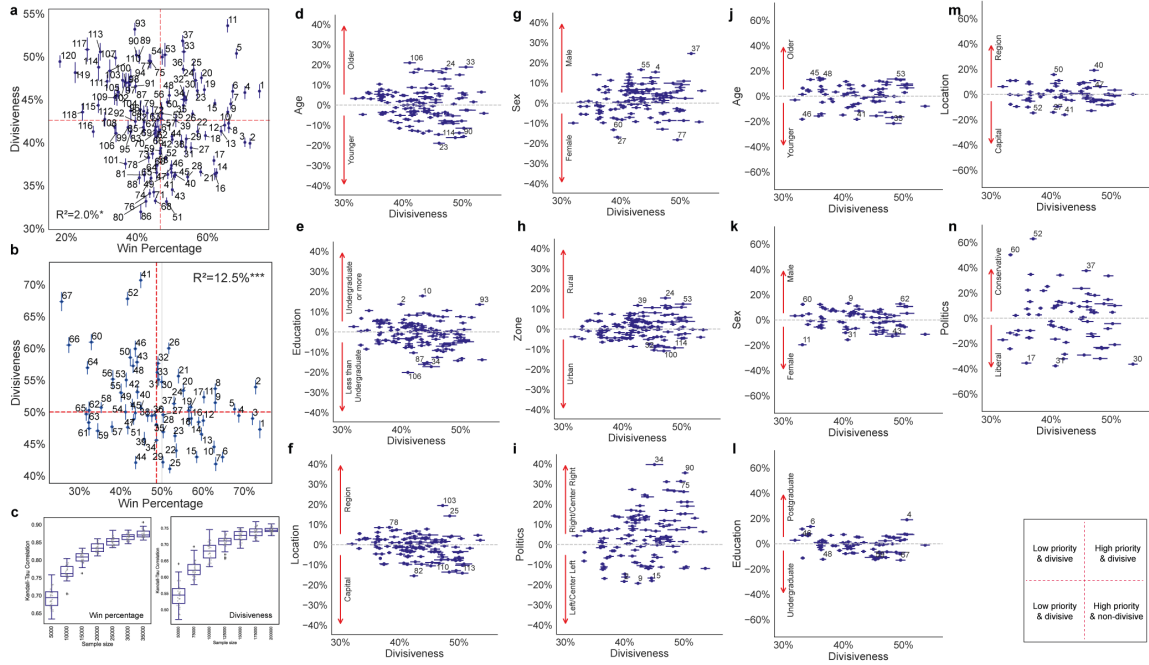


Figure 2.4: Divisiveness analysis. Win percentage (x-axis) versus divisiveness (y-axis) in a France ($N=217,034$ pairwise preferences) and b Brazil ($N=49,390$ pairwise preferences). For a-b, we report the R^2 calculated as the square of Pearson’s correlation estimated from a two-sided alternative hypothesis, as determined by the SciPy library (v.1.9.3). c Convergence in the rankings of agreements and divisiveness estimated using the Kendall-Tau correlation in France. Box plots show the Kendall-Tau correlation between the ranking obtained with the full sample and a random sample of the size indicated in the x-axis. Boxplot figures: center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by the seaborn library (v.0.12.1); whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, and circles represent individual data points. Multidimensional divisiveness for d-i France and j-n Brazil. Divisiveness as estimated in equation 2 (x-axis) compared to divisiveness estimated using self-reported (c (FR), j (BR)) age, (e (FR), l (BR)) education, (f (FR), m (BR)) location, (g (FR), k (BR)) sex, (h (FR)) zone, and (i (FR), n (BR)) political orientation. Each point in a-b and d-n represents the mean score of a proposal, and the error bars represent the 95% confidence interval of the proposal’s score and are, in some places, thinner than the symbols in the figure. Both values are calculated by bootstrapping half of the dataset 30 times. *Note:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Win Percentage	-0.14 (0.09)							-0.09 (0.09)
Age		0.25*** (0.09)						0.11 (0.09)
Education			0.18* (0.09)					0.13 (0.09)
Location				0.27*** (0.09)				0.20** (0.09)
Politics					0.30*** (0.09)			0.16* (0.09)
Sex						0.16* (0.09)		0.11 (0.09)
Zone							0.23*** (0.09)	0.14 (0.09)
Constant	0.00 (0.09)	0.00 (0.09)	0.00 (0.09)	0.00 (0.09)	0.00 (0.09)	0.00 (0.09)	0.00 (0.09)	0.00 (0.08)
R ²	0.02	0.06	0.03	0.07	0.09	0.03	0.05	0.22
Adj. R ²	0.01	0.06	0.02	0.07	0.08	0.02	0.05	0.17
Num. obs.	120	120	120	120	120	120	120	120

Table 2.3: Issues-level associations of Divisiveness in France. Each model represents a multivariate regression. We standardized both independent and dependent variables to have a mean of zero and a standard deviation of one. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Win Percentage	-0.35*** (0.12)							-0.20** (0.09)
Age		0.31** (0.12)						0.25*** (0.09)
Education			0.03 (0.12)					0.05 (0.09)
Location				0.15 (0.12)				0.13 (0.09)
Politics					0.62*** (0.10)			0.48*** (0.10)
Sex						0.36*** (0.12)		0.20** (0.10)
Zone							0.08 (0.12)	0.04 (0.09)
Constant	0.00 (0.12)	0.00 (0.12)	0.00 (0.12)	0.00 (0.12)	0.00 (0.10)	0.00 (0.12)	0.00 (0.12)	0.00 (0.09)
R ²	0.13	0.10	0.00	0.02	0.38	0.13	0.01	0.54
Adj. R ²	0.11	0.08	-0.01	0.01	0.37	0.11	-0.01	0.48
Num. obs.	67	67	67	67	67	67	67	67

Table 2.4: Issues-level associations of Divisiveness in Brazil. Each model represents a multivariate regression. We standardized both independent and dependent variables to have a mean of zero and a standard deviation of one. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Multidimensional Divisiveness Thus far, we have introduced and stressed our analysis of divisiveness. Nevertheless, the most divisive topics can be associated with diverse groups in society. For example, “Revocation of the spending ceiling” in Brazil is more divisive in terms of location and age. However, the divisions of “Expand the privatization of state-owned companies and national infrastructure concessions” are driven mainly by the political dimension.

To address this point, we first explore the Divisiveness calculated through our approach (x-axis) with the one estimated using self-reported information (Figure 2.4 d-i (France) and j-n (Brazil)). We observe that some of the most divisive issues for a specific group are also part of the ones mapped by our method. For instance, Figure 2.4 e compares Education in France. Here, the issue “Only process asylum applications from abroad” is by far more prioritized by participants with an undergraduate degree or more. But the issue “Engrave in the Constitution the superiority of French law over international law” is more divisive by participants who self-reported as male in comparison to female. We interpret that this aspect corresponds to preliminary evidence of multidimensional Divisiveness. Nevertheless, this seems just to be an anecdotal association. To challenge the previous statement, we second test the idea that our measure of divisiveness aggregates effects from diverse socio-demographic characteristics. Tables 2.3 (France) and 2.4 (Brazil) summarize eight regression models in France and Brazil, respectively. On the one hand, we regressed one model for each socio-demographic factor separately, and on the other hand, we controlled all the factors. We call Model 8 our “final model.”

For our final model, Location ($P < 0.05$) and Political orientation ($P < 0.1$) are positively associated with Divisiveness in France. An increase of 1 standard deviation in political orientation is associated with an increase of 0.16 standard deviations of Divisiveness (equivalent to $\frac{0.08-0.16}{0.05} = 25.6\%$ of increment). For Brazil, we find a positive relationship between Political orientation ($P < 0.01$), Age ($P < 0.01$), and Sex ($P < 0.05$) with Divisiveness. Interestingly, by looking at the dimensions separately in both countries, we observe that Political orientation is the dimension with the highest R^2 (Brazil: $R^2 = 0.38$; France: $R^2 = 0.09$), followed by Sex in Brazil ($R^2 = 0.13$) and Location in France ($R^2 = 0.07$). These findings align with previous conclusions of Cantador et al., suggesting that divisive issues can encompass effects from both ideological and socio-demographic characteristics. Again, non-causal explanations for this association of Divisiveness and Political orientation can be ascribed to the fact that participation in digital democracy systems could be associated with a pre-existent political interest and that both platforms were released in electoral periods. Even though we control by six dimensions often studied in terms of social cleavages, we call for further analysis in terms of including additional features.

Next, we compare some characteristics of the aggregation of preferences using divisiveness compared to its corresponding aggregation function. On the one hand, we compute the convergence ratio of our data to reach the collective ranking. On the other hand, we implemented a naïve algorithm to manipulate both rankings.

Convergence Firstly, we aim to determine the minimum number of pairwise comparison data rows needed to compute a ranking (both agreements and divisiveness) that correlates, at least in a 0.75 Kendall-Tau (KT), with the scores calculated from all data available. Figure 2.5 a-b show the convergence for our (c) ranking of agreements and (d) ranking of divisiveness in France. We focus on the French data for two reasons: First, the data set is 4.34

times larger than the Brazilian data. Second, the convergence analysis could be biased due to the sample size. The sample size in Brazil is around 50,000, a much lower value than the threshold determined for divisiveness from the analysis using the French data. We obtained a KT higher than 0.75 with at least 10,000 issue pairs for agreements and at least 200,000 issue pairs for divisiveness (i.e., around 20 times more than the data necessary to get convergence in the ranking of agreements.) The results suggest that a ranking of agreements using pairwise comparison data converges with a relatively small set of preferences ($\sim 4 \times 10^5$) with a stricter convergence criterion, and our approach to mapping divisiveness is more computationally expensive than agreement. Additionally, Figure 2.5 c-d includes the same analysis for synthetic data generated via the Urn model with a correlation of 10% using the library Preflib (we detail this in the next paragraph). Consistently with our previous finding, the ranking of agreements requires fewer data points to converge in comparison to Divisiveness.

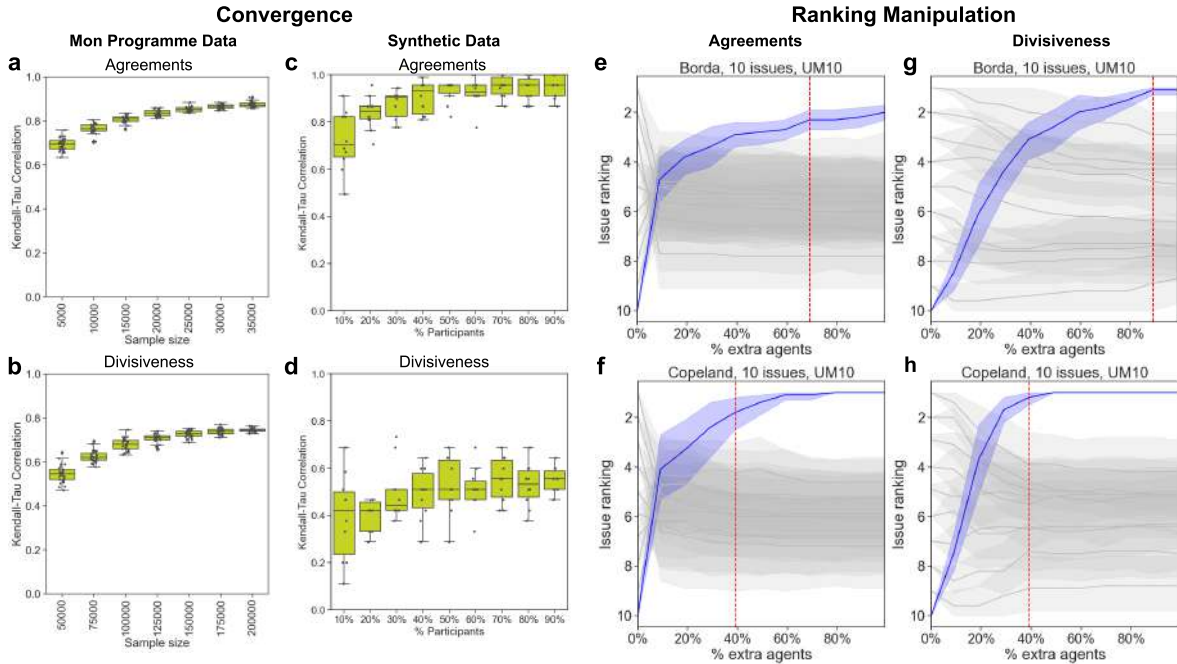


Figure 2.5: Robustness analysis for ranking of agreements/divisiveness in data (a-d) convergence and (e-h) manipulation. Convergence analysis using the French data for (a) agreements and (b) divisiveness, and synthetic data for (c) agreements and (d) divisiveness. Manipulation of ranking of agreements for ten issues using UM10 for (e) Borda and (f) Copeland and for ranking of divisiveness for (g) Borda and (h) Copeland. The vertical dashed red line shows the first computed threshold in which the least prioritized issue reaches the first position.

Manipulation Secondly, we examine how complex manipulating the ranking of Divisiveness is in comparison to its corresponding ranking of agreements. Specifically, we study the injection of “fake” ballots. Here, we use synthetic data generated via Impartial Culture (IC) and the Urn model utilizing a correlation of 10% and 50% (named UM10 and UM50, respectively). To simulate the profiles of preferences, we use the library PrefLib (Mattei & Walsh, 2013). We use this approach instead of data collected via our platforms since we aim to explore the manipulation in a complete set of preferences. To analyze the number of needed users to move the bottom-ranked issue to the top-ranked, we simulated preferences for 100 users with ten issues. The algorithm for manipulation of agreements consists of injecting new users, with profiles positioning the least ranked issue as her first priority and maintaining the order of all other issues. Conversely, the algorithm for manipulating Divisiveness consists of

injecting an odd profile that positions the less divisive issue as her first priority, followed by an even injection that sets the most divisive issue as her first priority. Figure 2.5 e-h show our results for the manipulation. We use a strict approach to evaluate its effectiveness, represented by a vertical dashed red line showing the first threshold in which the least prioritized issue reaches the first position, and the difference is statistically significant. While it takes around 75% extra participants to manipulate the ranking of agreements in favor of the last ranked issue for UM10, manipulating the ranking of Divisiveness requires roughly 90% of participants (Figure 2.5 e and g). We consistently observe the same pattern using Borda and Copeland and profiles generated through UM/IC. Our analysis yields that Divisiveness is not free of manipulation, though in this specific case, it takes more time for bots to distort their results than its ranking of agreements.

2.4.5 Empirical Boundaries of Divisiveness

Throughout this chapter, we have motivated the use of divisiveness for mapping divisive issues. We previously examined the nature of controversial issues in terms of their multidimensional dimension. Now, we explore how divisiveness behaves as an aggregation function. For this, we borrow two axiomatic properties studied in social choice, such as Pairwise Efficiency and Independence of Irrelevant Alternatives (IIA), to test some of its properties.

Pairwise Efficiency

While traditional research on social choice substantially relies on the impossibility of fairly aggregating individual preferences (Arrow, 1950), here we argue that under a large number of issues, a certain percentage of insatisfactions might be tolerable in order to reach a collective decision. For instance, we consider that violations occurred in situations where two issues are close to 50%-50% or whether one beats the other in 90%-10% can not be judged with the same severity. Returning to our notation, let x_{ij} be the total number of wins of issue i over j and let W_i be the win percentage of issue i , we define pairwise efficiency as the fraction of times in which $x_{ij} > x_{ji}$ given that $W_i > W_j$. Figure 2.6 a,j show the Pairwise Efficiency for the French and Brazilian data. We find that in both cases, the efficiency is around 80%, showing that overall, the collective ranking of agreements positions correctly in the higher and lower order. This result is interesting since it suggests that, trying to reach a collaborative ranking for 120 issues in France, around 96 are correctly positioned. It should be noted that most of the insatisfactions occur in the middle of the ranking.

Independence of Irrelevant Alternatives (IIA)

Continuing with the axiomatic analysis of divisiveness, inspired by the Arrow, 1950 axiom, we study the Independence of Irrelevant Alternatives (IIA): how much the ranking changes after the set of feasible issues is reduced. Let R be a ranking of issues (either by agreement or divisiveness). Consider removing an issue i from the data set and computing a new ranking of issues R' (that excluded i). For each issue j , we compute the Spearman's footrule d_{SF} between R and R' . It should be noted that to properly compare those rankings, we previously removed i from R and re-calculated the ranking of all remaining issues. This measure d helped us to compute the percentage of satisfaction of IIA by moving a threshold of acceptable distance in

both rankings. For example, suppose we have $R = \{a \succ b \succ c \succ d\}$, and we want to compare with $R' = \{a \succ d \succ c\}$ that is obtained by removing b from data. In this case, R is reduced to $\{a \succ c \succ d\}$. The distance in this example is $a = 0$, $c = 1$, $d = 1$.

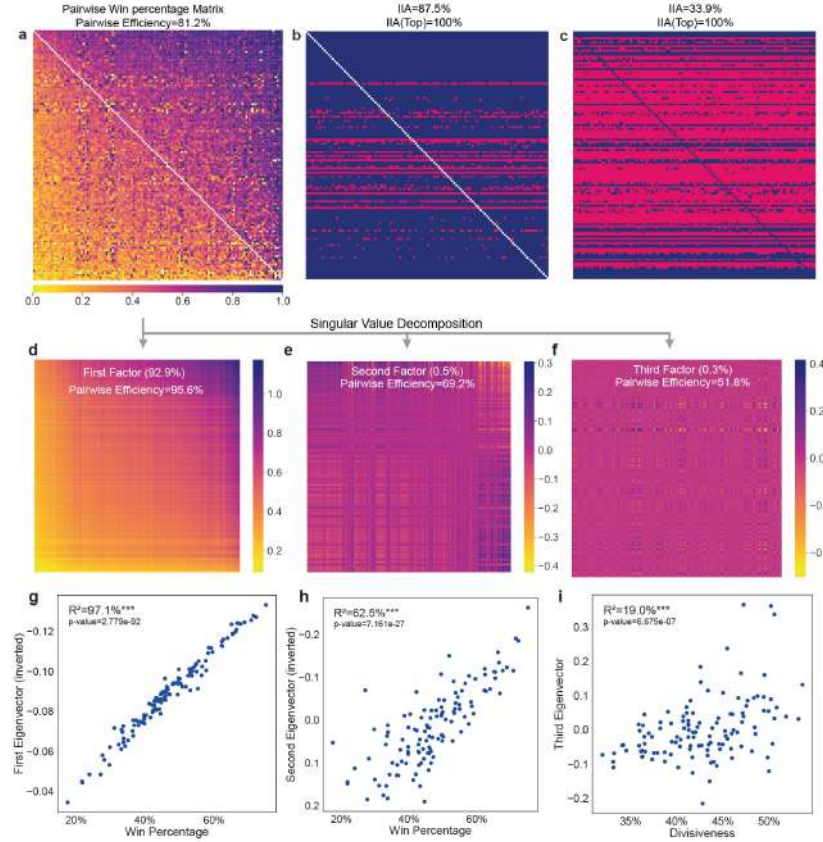


Figure 2.6: Axiomatic and matrixial properties of agreement and divisiveness for France. (a) Pairwise efficiency of the full matrix of preferences. Rows and columns represent proposals. Values indicate the win percentage of the proposal on the row when competing directly with the proposal on the column. Independence of Irrelevant Alternatives (IIA) for (b) the ranking of agreements (win percentage) and (c) divisiveness. In (c) and (d) we consider as robust changes involving less than four positions in the ranking. We compute the ranking of agreements with 30 bootstrap iterations and no bootstrapping was computed for the ranking of divisiveness. Singular Value Decomposition (SVD) of the matrix of pairwise preferences. Matrices corresponding to the first (d), second (e) and third (f) factors (eigenvectors). Correlation between first, second and third eigenvectors (unitary vectors of the Singular Value Decomposition) and (g-h) win percentages and (i) divisiveness. *Note:* $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Figure 2.6 b-c present the IIA violations by allowing a distance of four positions for the ranking of agreements and divisiveness in the French data (similarly, see Figure 2.6 k-l for Brazilian data). We observe that by being less strict in the allowed distance, we observe an increase in the percentage of IIA satisfaction. For instance, by allowing a maximum distance of 5 in the French data, 98.6% of the time, the data satisfies IIA. Similarly, by allowing a maximum distance of 5 in the Brazilian data, the data satisfies IIA 93.4% percent of the time.

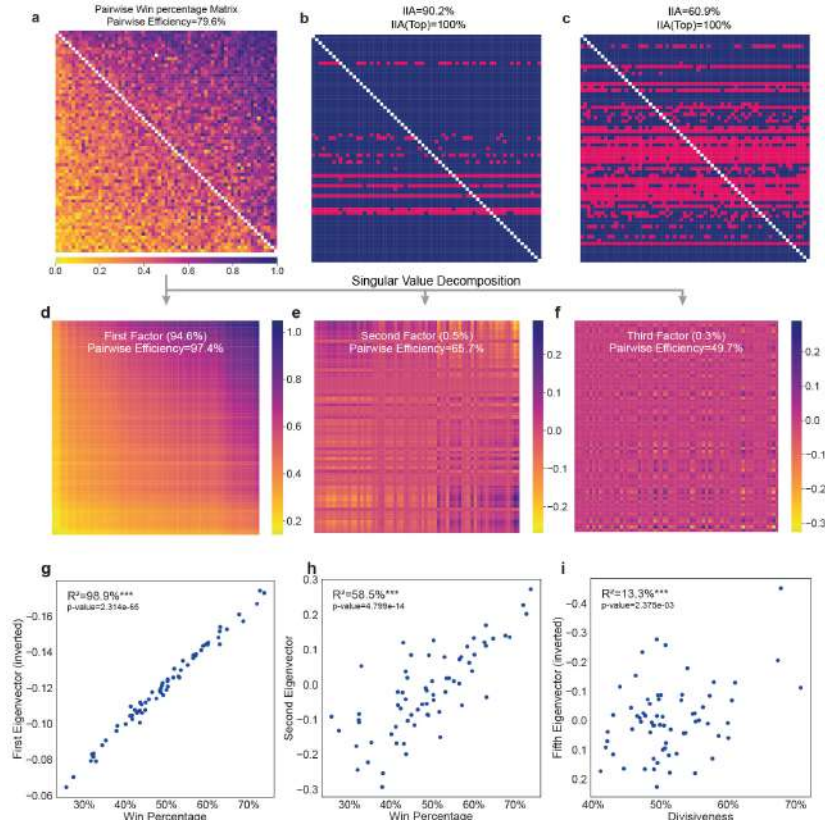


Figure 2.7: Axiomatic and matrixial properties of agreement and divisiveness for Brazil. (a) Pairwise efficiency of the full matrix of preferences. Rows and columns represent proposals. Values indicate the win percentage of the proposal on the row when competing directly with the proposal on the column. Independence of Irrelevant Alternatives (IIA) for (b) the ranking of agreements (win percentage) and (c) divisiveness. In (c) and (d) we consider as robust changes involving less than four positions in the ranking. We compute the ranking of agreements with 30 bootstrap iterations and no bootstrapping was computed for the ranking of divisiveness. Singular Value Decomposition (SVD) of the matrix of pairwise preferences. Matrices corresponding to the first (d), second (e) and third (f) factors (eigenvectors). Correlation vectors of the first, second and third eigenvectors (unitary vectors of the Singular Value Decomposition) and (g-h) win percentages and (i) divisiveness. *Note:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Singular Value Decomposition

Third, we complement these axiomatic properties by examining our pairwise matrix using linear algebra's Singular Value Decomposition (SVD) technique. Our view is that unitary vectors from this decomposition could potentially be associated with the ranking of agreements and divisiveness. Figures 2.6 and 2.7 show the matrixial properties for France (d-f) and Brazil (m-r). We observe that the first and second eigenvectors are positively associated with our measure of agreements (Figures 2.6 g-h France and 2.7 g-h Brazil). Interestingly, in the case of divisiveness, we observe a slight correlation with the third eigenvector in France and the fifth eigenvector in Brazil, suggesting that higher-order vectors could map divisiveness. It should be noted that techniques such as the Analytic Hierarchy Process use the eigenvector to compute its scores, suggesting that higher-order eigenvectors can be related to divisive issues. Nevertheless, the last statement is beyond the scope of this chapter.

2.4.6 Robustness of Key Results from Divisiveness

To conclude, we extend our analysis of Divisiveness beyond the data collected through our platforms or simulated profiles. We use benchmark data sets utilized in Computational Social Choice, deposited on [PrefLib.org](https://preflib.org). This repository is widely known in this community and is extensively used to study the aggregation of preferences. Here we focus on data sets labeled as *Strict Order Complete*, which represents the type of preferences in which voters provided a complete and strict ordering over the alternatives (thus, no ties are allowed between alternatives). This category offers the opportunity to study Divisiveness under a scenario of complete information about voters' preferences, unlike data collected *Mon Programme* or *Brazucracia*, which are incomplete preferences and in non-political contexts. Specifically, we used the sushi rank (Kamishima, 2003) and the skating ranking from Olympic games (Smith, 2000) data sets. The main advantage of these data sets is that these preferences are derived from people's voting/scores. Figure 2.8 reproduces the same key results in two Strict Order Complete data sets. Overall, we could reproduce both the uncorrelated pattern between Divisiveness and its corresponding aggregation function, the higher pairwise efficiency, and the associations of unitary eigenvectors from SVD with the measures of agreement and Divisiveness.

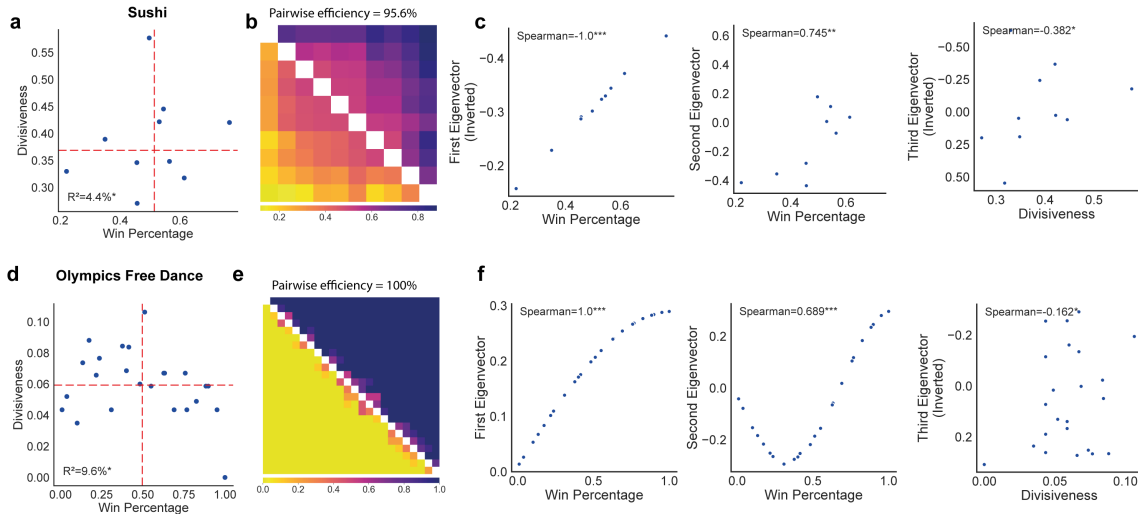


Figure 2.8: Reproducing key results in two Strict Order Complete data sets deposited on Preflib.org. Win percentage versus divisiveness for a Sushi preferences (Participants_{Sushi}=5,000, Proposals_{Sushi}=10), d Judge scoring of free ice-skating dance in the Olympics (Participants_{Ice-skating}=9, Proposals_{Ice-skating}=24), b-c e-f Singular Value Decomposition of these two datasets. For a, d, we report the R^2 calculated as the square of Pearson's correlation estimated from a two-sided alternative hypothesis, and for c, f, we report the Spearman correlation estimated from a two-sided alternative hypothesis. Both values are determined by the SciPy library (v.1.9.3). Note: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

The evidence presented throughout this chapter allows us to suggest that Divisiveness can be interpreted as the second central moment for any aggregation function. This seems to be supported by several reasons. First, given its conceptualization, it is a measure that can be applied to any aggregation function based on relative comparisons. Second, the evidence provided suggests that the issues identified using this approach can be framed as divisive or controversial. Third, the key results presented in this study are consistent with data collected in political and non-political episodes.

2.5 Discussion

This chapter used data collected from two real-world digital democracy systems released by our research group during electoral times in France and Brazil, aiming to study agreements and disagreements in these periods. Our results yielded that the main agreements provide little information about the campaign priorities of presidential candidates. In contrast, contentious issues seem to be most informative over political action, in line with previous findings reported by Ash et al., 2017 for the U.S. Senate election.

Then, we focus on understanding and systematically mapping divisive issues. First, we used self-reported data from participants to characterize divisive issues from six socio-demographic dimensions: sex, age, educational attainment, location, zone, and political orientation. We find that political orientation drives most of the divisions between participants in electoral periods, followed by sex and age. In fact, our platforms did not collect data from other characteristics that might be more relevant to explain citizen divisions, such as race or religious groups in Brazil. In order to tackle this drawback, we introduce an agnostic approach to identify divisive issues. We coined the term *Divisiveness* to this method that, inspired by pairwise comparison, can map the outcome dispersion of any aggregation function. Using data collected by our platforms, benchmark data sets from social choice, and synthetic data, we subsequently show that Divisiveness scores are uncorrelated to the ones obtained by its corresponding aggregation function, and this result seems to be robust both for political/non-political data and complete and incomplete preferences. We interpret that Divisiveness represents an understudied form of aggregation of preferences in social choice and can be interpreted as the second momentum of any aggregation function (that can be interpreted as the first momentum). In principle, this duality could be understood as the one between mean and variance. The fact that the same set of preferences already used in aggregation functions can also unveil divisive issues becomes relevant in the sense that we can move the boundaries over how we interpret the aggregation of preferences.

In addition, we mention a potential use case for our methodology. Anecdotally, the issue of “prohibiting the burkini at municipal swimming public pools” was the lowest-ranked issue on the French platform. In fact, an amendment from the Grenoble City Council allowing the use of burkinis in all public pools divided public opinion and generated an active and tense debate over this policy (LeMonde, 2022). Interestingly, based on data collected by *Mon Programme* one month before this debate, our approach labeled this issue as divisive (Ranked 15 out of 120). What is striking here is our ability to systematically identify controversial issues via online participation. This feature can contribute to better governance to detect “in time” potential conflicts and grievances at the grassroots level.

We suggest that the use of digital technologies can lead to mapping priorities and divisions in political episodes. Nevertheless, we are still far from fully understanding the nature of online participation. First, we point out an existent trade-off between the elicitation method and the quantity/quality of data collected. This chapter did not study how the content and user interface affected online involvement—specifically in terms of early dropouts and carelessness—calling for follow-up studies in the intersection of user interaction and data quality (For an extensive analysis of the nature of online involvement can be found at Chapter 4). Second, the skewness or self-selection toward “young, well-educated men living in urban areas” is a common pattern found in other digital democracy systems. This aspect is relevant whether we desire to map forms of societal polarization, opening a pending challenge for methods to

map citizen divisions or “polarization” using existent and representative instruments, not only relying on one-time non-representative data collected via surveys or online participation. One alternative could be examining open resources as election outcomes at the polling station level. Third, digital tools are the target of cyber-attacks and suspicious participation. Anecdotally, we identified server-side attempts to inject data into the French platform less than an hour after its release on Twitter. Even though our platform avoided these evident intentions to detriment legitimacy in online participation, we suggest that proposing proactive and reactive forms of removing suspicious participation is crucial to guarantee reliability in the results.

We expect that the study of Divisiveness will contribute to new research avenues from computational social science toward the intersection of agreements and Divisiveness. Specifically, we call for further analysis exploring how the identification of contentious issues can be conceptualized as a standard form of polarization.

Mapping Citizen Divisions Through Elections

Abstract

Elections can unveil citizens' enthusiasm and discomfort concerning political candidates, parties, and issues. While a substantial body of literature studies the election outcomes from the perspective of winners and losers, an under-explored condition to understand societal divisions emerges from citizen voting patterns. Here, we examine the concept of Election Polarization (EP) to address citizens' divisions on Election Day. We present an agnostic approach to map EP that relies exclusively on election data and considers the competitiveness of candidates (Between-EP) and their voting dispersion throughout a territory (Within-EP). We use both synthetic data and presidential election results from France, Chile, and the United States to show that our approach successfully identified theoretical expectations of "polarized" elections. Furthermore, we validate its robustness over the election type, aggregation scale, use of abstentions/spoilt votes, and the number of candidates. Finally, our analysis reveals that state-level Within-EP and Between-EP in the U.S. are positively associated with political polarization and political interest, respectively, shedding light that EP could potentially encompass a simple and reliable proxy of political polarization.

This chapter is based on the following articles:

Navarrete, C; Macedo, M; Stojkoski, V; Parada-Contzen, M. (2023). Election Polarization: Mapping Citizen Divisions Through Elections. *Manuscript under review*.

3.1 Introduction

Elections are a medium utilized by citizens to demonstrate their enthusiasm or discomfort concerning political candidates, parties, and issues (Makarenko, 2015). Recent unforeseen election outcomes—results that most pre-election polls can not anticipate—, including Brexit, the 2016 U.S. presidential election, and surprising performances of far-right candidates around the world have inspired a line of argument focused on the links between election results and polarization (Lamont et al., 2017; Layton et al., 2021) and how the places more neglected by the globalization “take revenge” of big cities by voting anti-establishment candidates (Rodrik, 2018; Rodriguez-Pose, 2018). While several techniques have been proposed both from economics, political science, and computer science to comprehend different facets of polarization (See, e.g., Fiorina and Abrams, 2008; Iyengar et al., 2012), reliable ideological polarization data, especially at sub-national level, are expensive and challenging to collect. Furthermore, most of these measures depend on one-time surveys collected from a small amount of the population. This issue is relevant since only a few countries have national election instruments to study this phenomenon, finding most work pertains to the United States, whose political system is not comparable to other realities and lacks a simple and generalized alternative across the world. Given that the main political consequence of an election is dividing the electorate into winners and losers (Wojtasik et al., 2013), we interpret that voting patterns can be analyzed from the lens of a measure of polarization.

This chapter examines the concept of Election Polarization (EP) to quantify how divided citizens are on Election Day. We argue that political candidates and parties represent a pool of different ideological perspectives on society. Citizens who vote for the same candidate or political party feel more similar to each other than those with a different preference (Esteban & Ray, 1994). To map the EP, we propose an agnostic approach that considers citizens’ voting patterns throughout a territory (Within-EP) and the closeness of candidates (Between-EP). Our analyses indicate that our method suits any aggregation scale (e.g., city, county, state, or country), is robust beyond a specific election type (e.g., presidential or senate), and is applicable for any number of candidates. Notably, we claim that EP can potentially encompass both “ideological” polarization and socio-demographic characteristics of geographies, providing evidence that it can be a measure of quasi-political polarization (i.e., a proxy for political polarization) and eventually become an index of electoral polarization. Additionally, we provide guidelines for a suitable data schema that can be applicable to any election data set, along with a collection of around ten million rows of curated election data sets publicly deposited on GitHub for further research. The latter is an advancement into the creation of standards for the use of election data and their applications in studies from Computational Social Science.

Chapter 3 is structured as follows. Section 2 presents an overview of polarization measurements and unexpected election outcomes. Section 3 methodologically introduces our approach to mapping Election Polarization. Section 4 presents our main findings, and Section 5 discusses the implications of our approach and provides a brief summary of its potential applications.

3.2 Literature Review

Electoral times are likely one of the most important political events for democracies. Although they should be considered routine for a healthy democracy, sometimes “unexpected” situations arise. We understand an unexpected election outcome as one that most pre-election polls could not anticipate, consequently shocking political experts and the population. The issue identified by this chapter concerning these unexpected outcomes is due that journalists often try to stress a simplified hypothesis that the “population is more polarized than in the past”; nevertheless, these statements are lacking in most cases from a quantitative basement. Here, we start exploring some contemporaneous examples from France, the United States, and Chile.

3.2.1 Unexpected election outcomes

Two decades ago in France, the far-right candidate Jean-Marie Le Pen from the *National Rally* (a far-right party) shocked political experts by surprisingly reaching the runoff of the 2002 Presidential Election, despite that most before-election polls affirmed that Lionel Jospin would face Jacques Chirac in the ballotage (Durand et al., 2004). Even though Chirac easily won the election with 82.21% of votes, the surprising electoral performance of Le Pen raised the feeling that France might become polarized (Mayer, 2013). Recently, this feeling climaxed with the *Yellow Vest Movement* in 2018, increasing support for the far-right candidates in the last elections, and massive protests against the raise of retirement age from 62 to 64 years old in 2023, revealing the social cleavages in the country (Valentin, 2022). Yet, France is not the only Western democracy in which we found surprisingly good performances for anti-establishment candidates in recent years. Likely, the most argued unexpected result is the 2016 United States Presidential Election. Donald Trump’s victory surprised both scholars, observers, and experts (Jacobson, 2017). His rhetoric focused on the white working-class concerns and exacerbating polarization in several dimensions (Abramowitz & McCoy, 2019; Lamont et al., 2017) did not leave any citizen indifferent, as shown by the opposite feelings of his supporters and opponents (Luttig et al., 2017). But, when we talk about a “polarized” election in the U.S., we also might refer to close outcomes, such as the 2000 Presidential Election between Al Gore and George W. Bush. For example, a margin of only 537 votes between them in Florida showed that the citizens were virtually divided in two on Election Day. Nonetheless, unexpected electoral outcomes are not a sole feature for developed countries. Chile, one of the most stable democracies in Latin America in the last decades, has seen flourishing conflicts and protests that challenge this image. In October 2019, a rise of 30 Chilean pesos (0.04 USD) in subway fares set off the most massive demonstrations in the last 30 years (The motto was “No son 30 pesos, son 30 años.”—*It’s not 30 pesos, it’s 30 years*). The political response was starting a new constitutional process (LaTercera, 2020). Scholars have considered this social eruption as a consequence of a gradual political polarization growth (Fábrega et al., 2018; Morales Quiroga, 2020) reported several years before. Thus, two years after the “October Revolution”—a movement that claimed profound transformations in society, a far-right candidate reached the runoff in the 2021 Presidential Election for the first time, and the proposed constitutional draft as a way out of the crisis was surprisingly rejected with more than 62% of votes in 2022, despite of expectations for a tight result.

3.2.2 The Footprint of Voting Patterns

Undoubtedly, elections play a central role in this chapter. From them, we also can learn more about citizens' preferences and how socio-economic and inequality conditions can shape the results. A line of argument has explored how socio-demographic and economic characteristics are associated with voting patterns. Depending on where electors live, either in urban areas or the countryside, their preferences can differ significantly, amplifying their differences (Rodrik, 2018; Rodriguez-Pose, 2018; Scala et al., 2015). Furthermore, socio-demographic characteristics of geographies seem to unravel voting patterns. In the United States, if the number of sedans encountered in a city is higher than the number of pickup trucks, the city is likely to vote for a Democrat during the next presidential election (88% chance); otherwise, it is expected to vote Republican (82%) (Geburu et al., 2017). Of course, candidates can take advantage of these cleavages in order to win an election. Evidence from the 2018 Brazilian presidential election suggests that Jair Bolsonaro took advantage of latent and unexploited demographic segregation in order to win the election (Layton et al., 2021). Similar findings were found in the 2016 United States presidential election, showing that unusually explicit appeals to racial and ethnic resentment and anti-establishment feelings attracted strong support from white working-class voters in favor of Donald Trump (Abramowitz & McCoy, 2019).

3.2.3 Polarization in Social Science

The preceding discussions concerning unexpected election outcomes and voting patterns lead us to think that all roads lead to polarization. Expressly, we understand that the study of polarization is tied by definition to elections. In a broad sense, polarization can be defined as the division into two or more contrasting groups in a society. The literature distinguishes polarization from several branches, such as **ideological**—opposite opinions about family issues such as same-sex marriage and abortion, **geographical**—spatial differences throughout a territory, **social**—segregation of population into low-income and high-income groups, **political**—opposite political platforms or **affective**—the difference between positive in-group bias towards the party someone supports and negative out-group bias towards other parties, to name a few (Dalton, 2008; Fiorina & Abrams, 2008; Iyengar et al., 2012). The salience of capturing its temporal and spatial dynamics relies on their effects on our society. For instance, polarization has been linked with inequality and conflicts (Esteban et al., 2012; Esteban & Ray, 2011). An increase in inequality at the local level can be associated with an increase in the probability of supporting an extreme ideologically political party (Winkler, 2019). Party polarization is associated with greater importance for substantive policy considerations for citizens' vote choice in competitive elections (Lachat, 2011). More recently, affective and elite polarization have been associated with lower attitudes towards COVID-19 (Charron et al., 2022). Nonetheless, not all associations for polarization have a bad connotation. Previous work has established that political polarization is positively associated with control of corruption perception (Brown et al., 2011; Testa, 2012) and Shi et al., 2019 found evidence that more ideologically polarized editorial teams produce articles of a higher quality than homogeneous teams.

Today, we find several techniques to map facets of polarization, both by using self-reported surveys (e.g., ANES, 2021), data from social media (e.g., Twitter) (Barberá, 2015), behavioral measures of trust or favoritism over political parties (Waugh et al., 2009), analyzing socio-demographic dimensions (e.g., rural population, income, or election outcomes) (Scala

et al., 2015), or measuring distance in a network (Hohmann et al., 2023). Similarly, we find methodological studies motivated by the “disappearing middle class” phenomenon (Esteban & Ray, 1994; Wang & Tsui, 2000; Wolfson, 1994; Zhang & Kanbur, 2001) that refers to the fact that a society is polarized whether the “middle class” tends to shrink over time. An example is the seminal contribution of Esteban and Ray, 1994, in which they introduce the identification-alienation framework; that is, a society is polarized whether there are individuals who identify themselves with those having a similar feature (e.g., income) and are alienated against other groups. Thus, a recent work from Faliszewski et al., 2023, based on concepts borrowed from elections, propose two measures that help to map polarization and diversity.

Elections in democratic countries represent a rich and reliable data source on society’s preferences. First, as transparency is crucial for guaranteeing reliability in electoral systems, election data is often openly shared at the precinct level or on a very disaggregated level. Second, some widely-known democracy indexes such as the “Democracy Index” (Kekic et al., 2007) uses election outcomes to provide measures of democracy wealth. Although there are studies on polarization using election data (See, e.g., Abramowitz and Saunders, 2008; Johnston et al., 2016), most of the work is motivated by the two-party political system from the United States, which is not a feature of multi-party systems such as the ones on France, Chile, or most Western European parliamentarism. Furthermore, this literature largely depends on one-time data collected via election studies, which only a few countries can afford. To our concern, most ideological polarization measures speak to the U.S. data, lacking a generalized alternative around the world. Hence, it seems reasonable to argue that the countries that could benefit from research on polarization using election data are lower- and middle-income ones.

3.3 Methods

This section outlines the steps pursued in conceptualizing our measure of election polarization. We start unpacking the antagonism of candidates, followed by their summation into polarization. To conclude, we cover some classical polarization measurements as alternative definitions for election polarization.

3.3.1 Definition of Antagonism

We ground the study of election polarization on the identification-alienation framework (Esteban & Ray, 1994), proposing that voters identify themselves as similar to those who choose the same candidate. At the same time, voters feel alienated from those with a different candidate preference. We will refer to this division level of a candidate i as **antagonism**. To quantify the antagonism, we consider three characteristics: i) The competitiveness between candidates, ii) the voting dispersion of a candidate within a geography, and iii) the relative relevance of candidates. Furthermore, we consider that this antagonism can be decomposed by an in-group (or **between-**) and out-group (or **within-**) components, as suggested by Palacios-González and Garcia-Fernández, 2012, that follows a similar decomposition to the one framed for the Analysis of Variance (ANOVA).

First, we introduce some notation. Let N be the number of candidates and let M be

the number of electoral units (e.g., district, precinct) in which electors cast their votes. For a candidate i , let $score_{i,k}$ and $votes_{i,k}$ be her percent and number of votes in k -unit, respectively, and let \overline{score}_i be the voting percentage of i in the election.

Between-Antagonism

This dimension relies on the competitiveness **between** candidates. In this dimension, we consider that a population divided 50-50 in an election of two candidates is more polarized than a 99-1 outcome since the voters are far away from a unanimous consensus over the winner. The same rule can be generalized for elections with more than two candidates (33.3% for three, 25% for four, etc.). Furthermore, we point out that a result of 50-49-1 between three candidates should be labeled as polarized since the third candidate is irrelevant to the outcome.

We formally define the **Between-Antagonism** of candidate i , as the distance (measured by the percent of votes) with respect to all the other candidates in an election as follows:

$$\text{Between-A}_i = \frac{\sum_{j=1}^N \sum_{k=1}^M votes_{i,k} (1 - |score_{i,k} - score_{j,k}|)}{N(N-1) \sum_{k=1}^M votes_{i,k}} \quad (3.1)$$

We introduced a variation of the alienation proposed by Esteban and Ray in order to capture the distance between groups. Here, we subtract the distance between two candidates from the unit. That is, whether the distance between two candidates is minimum (two candidates with the same percent of votes), the antagonism is maximum (or 1). Conversely, whether the distance between two candidates is maximal (one candidate reaches the unanimity of votes), the antagonism is minimal (or 0). The values range from 0 and $1/N$, when values closer to $1/N$ indicate that the candidate has a greater between-antagonism. Thus, **Between-Antagonism** of a candidate i represents how much this candidate was preferred over the others.

Within-Antagonism

This dimension explores the individual performance of a candidate i **within** a geography. That is, whether a candidate i concentrates the electoral stronghold in a region or not. Consider a city calling for candidacies to elect a future mayor. The city is divided into two districts: *Alpha* and *Beta*, composed of 300 voters each. In turn, the districts divide the electors into three precincts equitably. Let $\{A, B, C\}$ be three candidates running for the position. By exploring the outcome per district (*Alpha*: $A=30, 29, 31$; $B=20, 19, 21$; $C=50, 52, 48$ and *Beta*: $A=5, 0, 85$; $B=5, 40, 15$; $C=90, 60, 0$ —Each number represents a precinct within the district), we observe that candidates obtained the same votes ($C = 150, A = 90, B = 60$). However, we observe that the voting dispersion for each candidate at the precinct level was almost null in *Alpha* in comparison to *Beta*, where candidate C even obtained 0 votes in a precinct of *Beta*. This idea leads to the definition of **Within-Antagonism** as a form to capture the voting dispersion of a candidate within a geography.

$$\text{Within-}A_i = \frac{\sum_{k=1}^M \text{votes}_{i,k} |\text{score}_{i,k} - \overline{\text{score}_i}|}{(N-1) \sum_{k=1}^M \text{votes}_{i,k}} \quad (3.2)$$

The **Within-Antagonism** ranges from 0 to $1/N$, where values closer to $1/N$ indicate that the distribution of the votes on candidate i is less balanced. As this dimension of antagonism relies on the performance of the same candidate throughout a territory, we argue that it could capture rural-urban, north-south, or poor-rich groups' divisions, to name a few. In order to establish this link, we subsequently explore this statement. It should also be noted that this definition speaks to previous work using dispersion as a measure of polarization.

Antagonism

Taken together, we define the total **Antagonism** of a candidate i , A_i , as the summation of the within- and between- antagonism. This is inspired by the within-group and between-group variation in ANOVA.

$$A_i = \text{Between-}A_i + \text{Within-}A_i \quad (3.3)$$

3.3.2 Definition of Election Polarization (EP)

Election polarization is based on the antagonism of political candidates. Formally, the antagonism A of a political candidate i is the level of division it generates on Election Day. Again, the antagonism has two dimensions: between-antagonism and within-antagonism. The between-antagonism represents the competitiveness of a candidate in an election, whereas the within-antagonism describes her voting dispersion within a territory (See Methods 3.3.1).

We then define Election Polarization (EP) as the sum of the antagonism across all candidates on Election Day, i.e,

$$\text{EP} = \sum_{i=1}^N A_i. \quad (3.4)$$

This measure is generalized to elections with more than two candidates and for any number of precincts/aggregation units. We motivated this definition by the idea that the antagonism of each candidate represents a fraction of information over the level of polarization in society, and the sum of all their antagonisms must represent a measure of polarization (Esteban & Ray, 1994).

Analogously to the antagonism, the EP can also be decomposed into the Within-EP and Between-EP:

$$\text{Between-EP} = \sum_{i=1}^N \text{Between-A}_i, \quad (3.5)$$

$$\text{Within-EP} = \sum_{i=1}^N \text{Within-A}_i. \quad (3.6)$$

We show subsequently that this distinction helps us to differentiate the dimensions of election polarization.

3.3.3 Numerical Examples

Let A and B be two candidates, and let 1 and 2 be two locations, in which each location covers 100 voters. In the following examples, we show how our approach behaves.

Example 1 Geographies 1 and 2 have opposite preferences. Whereas 1 fully support candidate A , 2 does for candidate B .

Location	Candidate	Votes	Percentage
1	A	100	100%
1	B	0	0%
2	A	0	0%
2	B	100	100%

In this case, the Within-A for A and B is equal to $1/2$, whereas the Between-A for both candidates is equal to 0. Consequently, the Within-EP for this election is 1 (maximal), and the Between-EP is 0 (minimal).

Example 2 Both 1 and 2 are completely divided, where the voters in each are divided in support of candidates A and B .

Location	Candidate	Votes	Percentage
1	A	50	50%
1	B	50	50%
2	A	50	50%
2	B	50	50%

In this case, the Between-A for A and B is equal to $1/2$, whereas the Within-A for both candidates is equal to 0. Consequently, the Between-EP for this election is 1 (maximal), and the Within-EP is 0 (minimal).

Example 3 Candidates A and B reach the same number of votes, though each one wins with small voting variations both in locations 1 and 2. Whereas candidate A won in location 1 with 55% of votes, candidate B won in location 2 with 55% of votes as well.

Location	Candidate	Votes	Percentage
1	A	55	55%
1	B	45	45%
2	A	45	45%
2	B	55	55%

In this case, the Between-A for each candidate is 0.45, and the Within-A is 0.05. Consequently, the Between-EP is equal to 0.9, and Within-EP is equal to 0.1. As the overall result divides the electorate 50-50, the EP is equal to the unit.

Example 4 Unlike preceding examples, cities 1 and 2 agree that candidate A is the majoritarian winner, showing consistent support of 80% in each location.

Location	Candidate	Votes	Percentage
1	A	80	80%
1	B	20	20%
2	A	80	80%
2	B	20	20%

In this case, the Between-A for each candidate is 0.2, and the Within-A is 0. Consequently, the Between-EP is equal to 0.4, and Within-EP is equal to 0. Overall, the EP is equal to 0.4.

To sum up, figure 3.1 illustrates four fictional elections of two and three candidates with six precincts each. These examples show the boundaries of elections with maximal and minimal Within/Between-EP. Then, we suggest in Figure 3.3 b a form of classifying elections according to four quadrants split by competitiveness and voting dispersion. In other words, greater voting dispersion presents greater Within-EP, and closer elections present greater Between-EP.

Numerical properties of Election Polarization

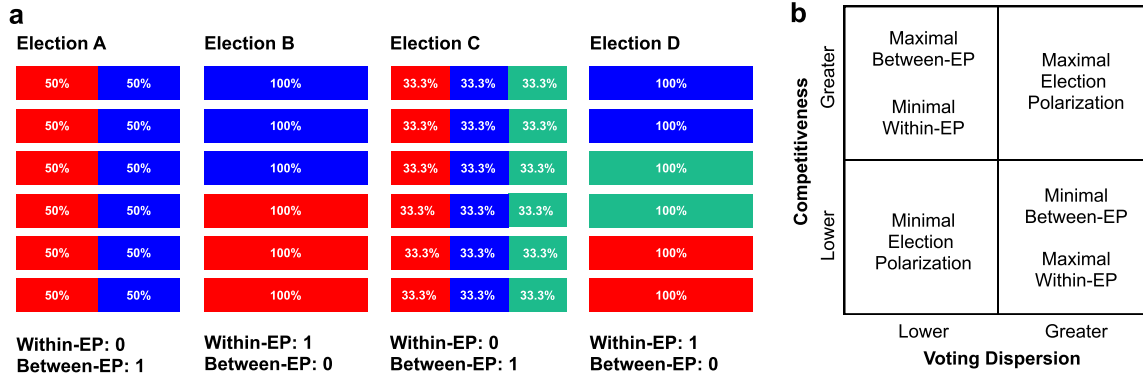


Figure 3.1: An illustrative example of our approach for two (Elections A-B) and three (Elections C-D) candidates. Each row represents a precinct, and each color represents a candidate. (b) Magic quadrant to summarize the criteria considered to map EP. The x-axis represents the voting dispersion (associated with Within-EP), and the y-axis represents the closeness (associated with Between-EP).

3.3.4 Election Data Schema

From a data science perspective, the main issue regarding the use of election data in research is caused by they are shared in unstandardized formats. Essentially, no guidelines are adopted by countries to share electoral administrative records. Here, we suggest a straightforward guideline inspired by tidy data schema and following the work of Baltz et al., 2022 with election data in the U.S. We consider that an appropriate form to share election data should be agnostic to language barriers, consider the properties of elections, and guarantee comparison at any aggregation scale. Roughly, the features that we consider relevant to accurately define an explainable election data set are the following:

Candidate Candidate name.

Party Political party associated with the candidate. In case that candidate is not defined, this feature is equivalent to her.

Political Coalition In the case of multi-party systems and coalitions allowed, this feature includes the political pact name.

Polling ID We point out that we create a unique identifier for the minimal aggregation unit called *polling_id* (e.g., the precinct in the United States for the general election). This feature concatenates hierarchical aggregate levels according to the data origin. For example, the *polling_id* joins state and county FIPS in the United States (2000), and the *polling_id* joins region, province, commune, and polling station name in Chile (2021).

Flag Candidates There are characteristics of elections that are not considered in the final results. This can be exemplified by abstentions, blank, or spoilt votes. For these cases, we

consider that these options can be included as additional rows representing “candidates”, but they should be flagged in order to avoid computing final results with them. In this vein, this feature is a boolean variable that establishes the differentiation between candidates and votes that are not part of the final result, but it can be used for research purposes.

Value Number of votes for a candidate in a specific Polling ID.

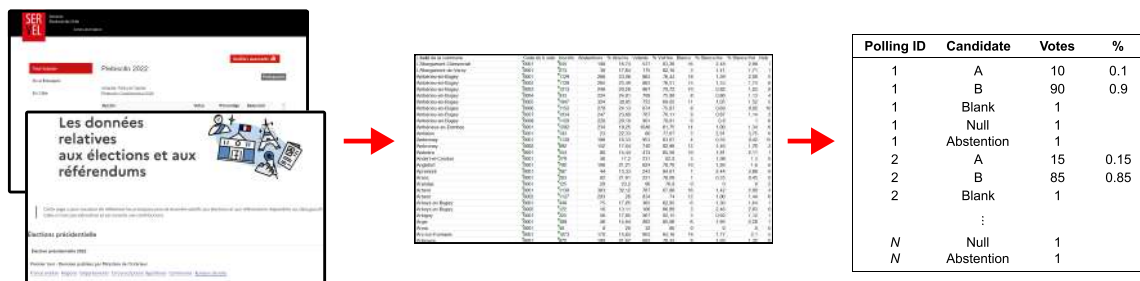
Rate Voting percentage for a candidate in a specific Polling ID. By default, this value does not consider the candidates that are flagged.

3.3.5 Election Data

For this study, we downloaded presidential election data at the polling station level from Chile (2013-2021) (Servel, 2023) and France (2002-2020) (InterieurGouv, 2023). For the United States, we use county-level presidential election data (2000-2020) curated by the MIT Election Lab (Baltz et al., 2022). For robustness and validation analysis performed in following sections, we use the precinct-level data for the 2016/2020 House, Senate, and Presidential elections (MIT Election Data and Science Lab, 2018a, 2018b, 2022a, 2022b, 2022c) in the United States.

We focus on these countries in order to study our measure of polarization in presidential systems. These data sets were curated in order to allow comparisons over time and across geographies. As suggested before, we created a unique identifier associated with the most granular geographical level here. For instance, this ID corresponds to the polling station in Chile but county in the United States.

Figure 3.2 illustrates the data collection and curation to perform our study. Roughly, we downloaded the official files from the government institutions in charge of elections (Electoral Service in Chile and Ministry of Internal Affairs in France), then we structured the data according to four features: candidates’ names, polling identifier, number of votes, and voting percentage.



We downloaded election data published by official institutions in each country.

Specifically, we use presidential elections in France (2002-2022), the United States (2000-2020), Chile (2013-2021) and general elections in Spain (2019), Belgium (2019), Germany (2021), and Italy (2022)

Despite its transcendence in our democracy, election data is not under a standard framework around the world. Thus each country decides independently on an appropriate form to share this data. For instance, candidates’ votes are shared in columns in France, whereas they correspond to data rows in Chile.

In order to tackle this issue, we propose a standardized format that considers tidy data guidelines and allows including of both abstentions/spoil votes.

Figure 3.2: An illustrative example of data collection and curation in France and Chile.

The data curated during the preparation of this study are deposited on Harvard DataVerse and GitHub (<https://github.com/cnavarreteliz/election-polarization>). Additionally, we created a didactic web application that allows the audience to interact with the results of this chapter and access the main findings of EP. This resource is available online at <https://electionpolarization.com>, and its repository is under an MIT license.

We create an additional file named `{country}_{year}_{round}_location.csv.gz`, that contains metadata regarding the voting unit. Both files can be merged using `polling_id` feature. Additional documentation can be found on GitHub. We exclude the data sets from the United States as part of our contribution since they were previously curated by Baltz et al., 2022.

Country	Year	Election Round	Nro. candidates	Nro. rows	Abstentions	Blank	Nulls
France	2002	First round	16	1,026,256			
France	2002	Runoff	2	128,282			
France	2007	First round	12	787,296			
France	2007	Runoff	2	131,234			
France	2012	First round	10	679,259			
France	2012	Runoff	2	135,864			
France	2017	First round	11	600,635	✓	✓	✓
France	2017	Runoff	2	138,484			
France	2022	First round	12	1,045,230	✓	✓	✓
France	2022	Runoff	2	139,364			
Chile	2013	First round	9	372,141			
Chile	2013	Runoff	2	81,572			
Chile	2017	First round	8	471,790	✓	✓	✓
Chile	2017	Runoff	2	85,780			
Chile	2021	First round	7	466,390	✓	✓	✓
Chile	2021	Runoff	2	93,278			

Table 3.1: Description of data sets curated and publicly available as part of our contribution. All aggregation levels are polling stations.

3.3.6 Polarization Measurements

This section includes alternative definitions of polarization utilized throughout this chapter. First, we adapt classical polarization measurements, such as Esteban-Ray, Wang-Tsui, Dispersion, and Reynal-Querol, in order to compute the antagonism for candidate i and, consequently, the polarization level. We created a version of these algorithms for Python, which is part of our scientific contribution and is deposited on GitHub. Second, we provide a brief explanation of our measure of political polarization, which is used subsequently in the results section.

Esteban and Ray

$$ER_i = K \sum_{i=1}^N \sum_{k=1}^M \sum_{j=1}^N votes_{i,k}^{1+\alpha} votes_{j,k} |score_{i,k} - score_{j,k}| \quad (3.7)$$

Where $votes_{i,k}$ is the number of votes of i in k , $score_i$ the percent of votes of i in k , and K and α represents freedom degrees introduced by the authors. We set $\alpha = 0.25$ and $\alpha = 1$.

In all cases, we set $K = \frac{1}{\left(\sum_{k=1}^M votes_{i,k}\right)^{(2+\alpha)}$.

Wang and Tsui

$$WT_i = \sum_{i=1}^N \frac{K}{P} \sum_{k=1}^M votes_{i,k} \left| \frac{score_{i,k} - m_i}{m_i} \right|^\gamma \quad (3.8)$$

Where P is the number of voters $P = \sum_{k=1}^M votes_{i,k}$, m is the median percent of votes of i , γ is a freedom degree in which $\gamma \in \{0, 1\}$, and K represents a freedom degree. We set $K = 1$ and test with $\gamma = 0.5$ and $\gamma = 0.75$.

Dispersion

$$Dispersion_i = \sum_{i=1}^N \sqrt{\frac{\sum_{k=1}^M (score_{i,k} - \mu_i)^2}{M - 1}} \quad (3.9)$$

Where M is the number of units, $score_{i,k}$ the percent of votes of i in k , and μ_i is the voting percentage of i in the election.

Reynal-Querol

$$RQ = 1 - \sum_{i=1}^N \left(\frac{1/2 - \mu_i}{1/2} \right)^2 \mu_i \quad (3.10)$$

Where N is the number of candidates and μ_i is the voting percentage of i in the election. In contrast to the latter polarization measurements, we do not split the calculation by voting unit here since Reynal-Querol should be conceptualized rather as a measure of closeness between candidates. Furthermore, we point out that this definition is only suitable for elections between two candidates.

Mapping Political Polarization

We subsequently verify whether Election Polarization can be a suitable proxy for political polarization. Here we follow the methodology proposed by Baker et al., 2020 to compute political polarization using a self-reported ideological scale collected from the Cooperative Election Study (CES) (Kuriwaki, 2022). To our knowledge, this is the only survey suitable to compute state-level political polarization. Let a be a participant, and let A be the total number of participants in the CES survey. First, we map the ideological score of party p for a given geography as follows:

$$\text{Ideology}_p = \sum_{a=1}^A S_{ap} \cdot w_a \quad (3.11)$$

The parameters are: S_a the ideological score of participant a (1=Lean, 2=Not very strong, 3=Strong), w_{ap} is the weight of a . Second, we define Political Polarization (PP) as the ideological distance between the Democratic ($p = 1$) and Republican ($p = 2$) parties.

$$\text{PP} = |\text{Ideology}_1 - \text{Ideology}_2| \quad (3.12)$$

We use this measure to compare its relationship with respect to Election Polarization.

3.4 Results

Here, we present our method to estimate election polarization and use this approach to explore its boundaries by using simulations and data from the presidential elections in the United States, France, and Chile.

3.4.1 Numerical Properties of Election Polarization

We explore the properties of EP using numerical simulations and compare them to the properties of other measures of polarization.

First, we simulate twelve elections under different levels of competitiveness and dispersion for a territory divided into 1,000 units with 100 voters each. The candidates' voting percentage per unit was drawn and sampled using a Gaussian distribution. Figure 3.3 a-b confirms that competitiveness is associated with greater Between-EP, and voting dispersion is positively associated with Within-EP. For example, the simulated elections (5) and (6) from figure 3.3 a shows electors divided 50-50, but a greater voting dispersion is observed in (6). We find that the Within-EP in (6) is almost five times higher than in (5), and the Between-EP in (5) is around 1.5 times higher than in (6). Nonetheless, the EP in both scenarios is equal to the unit.

To generalize this observation, in elections between N candidates, greater EP values are reached when the outcome of each candidate is closer to $1/N$. We numerically observe lower Within-EP and higher Between-EP values as the number of candidates is greater (e.g., $N > 10$). This fact suggests that the comparison of EP across geographies must be performed by using the same number of candidates. Conversely, no significant differences are observed whether the number of electoral units is greater.

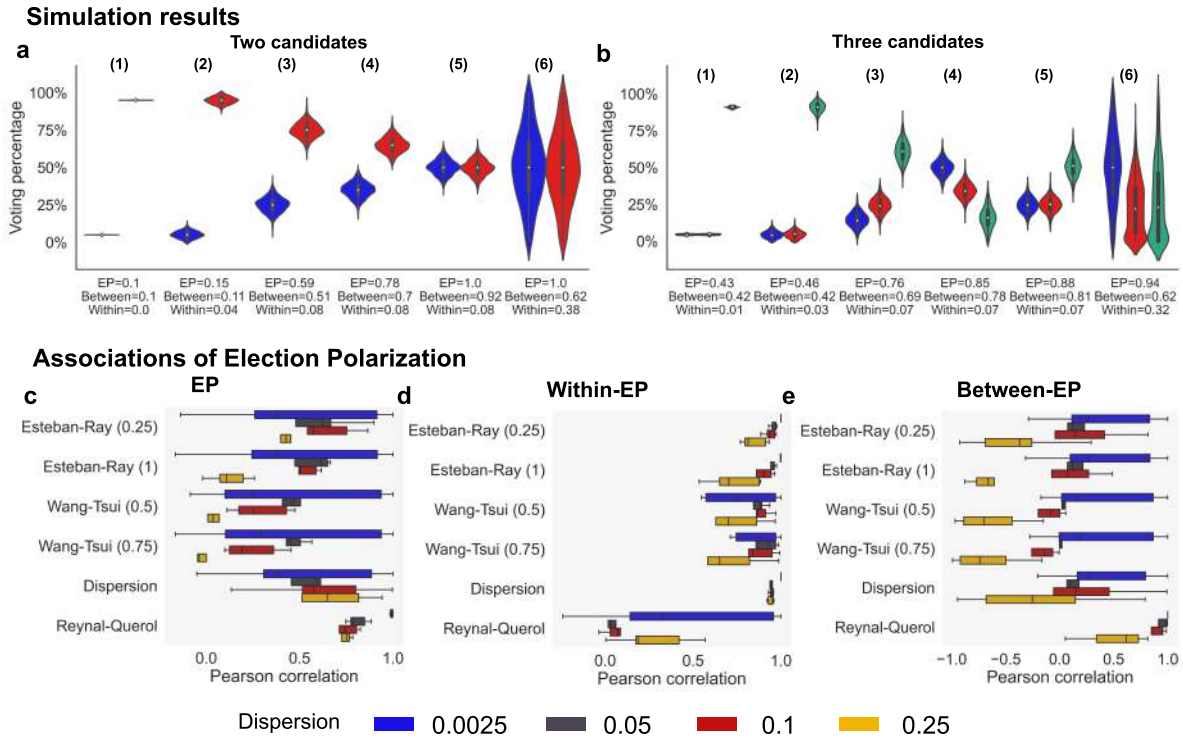


Figure 3.3: Numerical properties of Election Polarization. Violin plot distributions for synthetic election data for (a) two candidates and (b) three candidates. Each color represents a candidate. **Comparison of classical polarization measures for synthetic elections between two candidates.** (c-e) Correlation of polarization measures with the (c) EP, (d) Within-EP, and (e) Between-EP. Values in parentheses for Esteban-Ray and Wang-Tsui represent the degree of freedom used. We simulated 100 elections with 1000 units each, with an equal number of votes per unit (in our case, 100).

Second, we investigate whether classical measurements of polarization—Esteban-Ray (ER), Wang-Tsui (WT), Dispersion (DP), and Reynal-Querol (RQ)—can already capture EP. Our analysis indicates that these approaches are not suitable in all scenarios for accurately labeling polarization that emerges from election data. We interpret that the reason is due that these approaches are motivated by different phenomena. For instance, Reynal-Querol introduces her method for religious polarization, and both Esteban-Ray and Wang-Tsui drove their contributions by the “disappearing middle class” literature. Even so, we find that the Within-EP positively correlates with ER , DP , and WT , but these measures tend to fail as the election gets closer to a unanimous outcome (a condition of expected minimal polarization) and under higher voting dispersion levels (Figure 3.3 d). Conversely, Between-EP positively correlates with RQ (Figure 3.3 e); nevertheless, the latter—by definition—does not capture the heterogeneity of voting patterns throughout units.

Overall, we find that our approach consistently reaches the maximal polarization for close elections (e.g., 50-50) regardless of voting dispersion and reaches a minimal value for unanimous elections. Furthermore, by comparing two elections with the same voting percentages for candidates, the more dispersed election will always reach a greater value (See Figure 3.4) both in EP and Within-EP. We interpret that this condition is satisfied because of a complementary effect among the Within and Between components. For instance, Between-EP could be associated with elections in which candidates’ ideological positions are not distinguished for the population, and Within-EP can be interpreted as a fragmentation measure. Even so, both dimensions provide valuable information in understanding polarization.

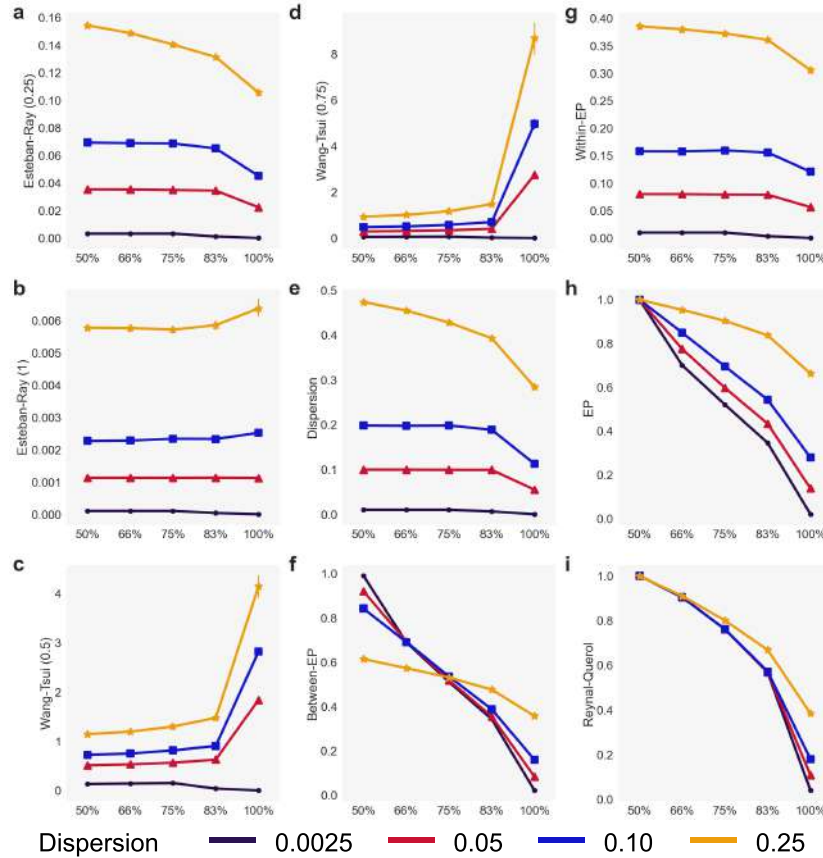


Figure 3.4: Comparison of classical polarization measures and EP by using synthetic data. The x-axis represents the sampled mean, and each bar represents a dispersion level. Error bars are 95% and are, in some places, thinner than the symbols in the figure

3.4.2 EP in Real Data

Next, we conducted an empirical study to explore whether we could quantify and identify theoretical expectations of politically “polarized” elections reported in the literature for Chile, France, and the United States. To our knowledge, this is a first-time attempt towards a comparable measure of polarization across geographies. Additionally, we investigate whether EP is associated with political dimensions, such as political polarization and political interest in an election. To this end, we used presidential election data from Chile (precinct level, 2013-2021), France (precinct level, 2002-2022), and the United States (county level, 2000-2020).

Figure 3.5 a-c describes the spatial and temporal dynamics of state-level EP for the United States from 2000 to 2020. Roughly, we observe greater Within-EP in the South-east, while the Between-EP is greater in the Midwest. At the country level, the fact that U.S. elections usually splits the electorate into two halves in the popular vote—reaching consequently EP values closer to the unit in the study period—highlights the need for decomposing our measure to observe the main differences. According to Between-EP, the most polarized election occurred in 2000, whereas according to Within-EP, the most polarized election occurred in 2016 (Figure 3.5 d-f). On the one hand, the competitiveness in the presidential election between Al Gore and George W. Bush explains the Between-EP in 2000. On the other hand, it is documented that Donald Trump’s discourse—focused on the white working-class Rural

America–raised political polarization, being consistent this trend with the one observed in the Within-EP. Then, we challenge our approach to find whether the EP distinguishes polarization at the sub-national level, performing a comparison of EP from Swing states with respect to Partisan states. Due to that, Swing states should be more electorally polarized—since we expect them to face more competitive elections (Ahn, 2022; Hoban, 2016), this is a good natural experiment to compare EP at sub-national level. Figure 3.5 g-i shows that Swing states presented greater polarization than Partisan states. Nevertheless, the difference is uniquely significant for Between-EP. This result is relevant since it provides evidence that our method identified that Swing states are more polarized than partisan ones, but these differences are mainly explained by competitiveness rather than geographical cleavages.

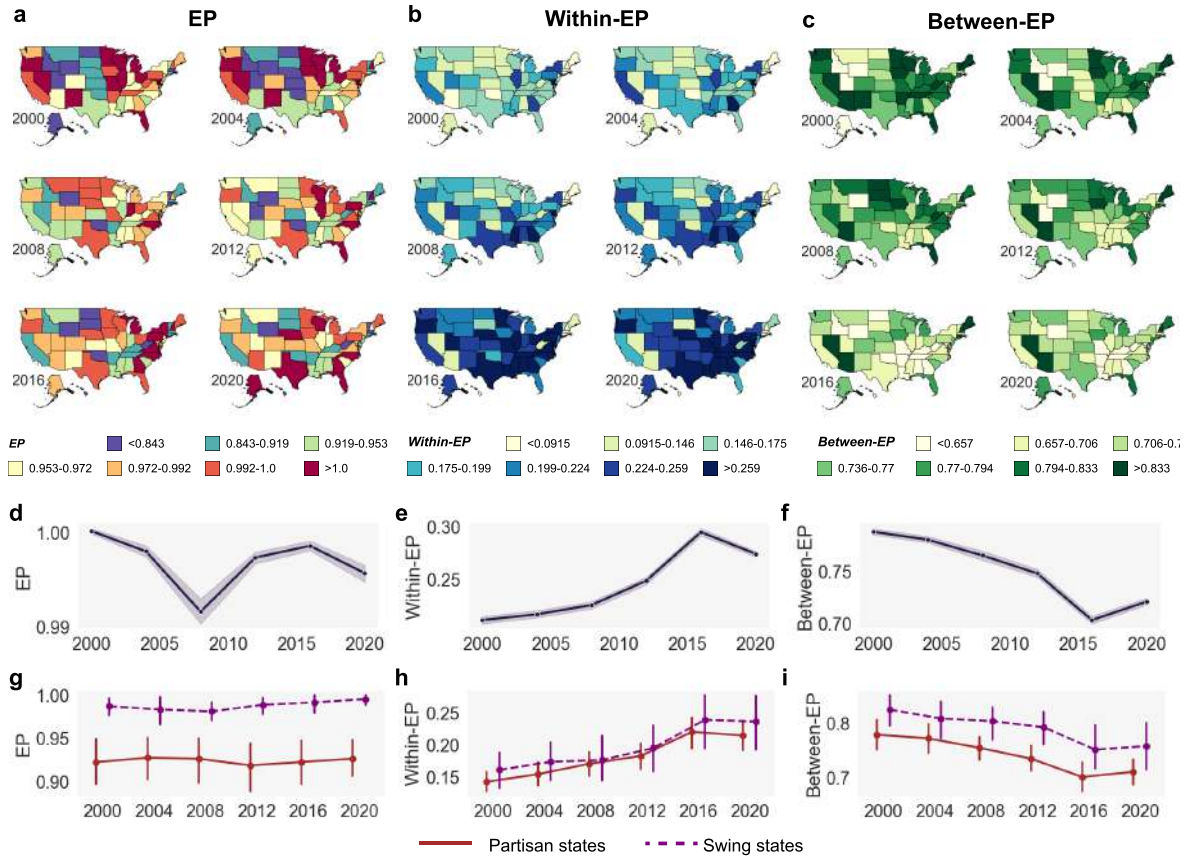


Figure 3.5: The geography of Election Polarization in the United States in the period 2000–2020. Temporal dynamics of state-level (a) EP, (b) Within-EP, and (c) Between-EP. Temporal dynamics of country-level (d) EP, (e) Within-EP, and (f) Between-EP. We bootstrapped 100 times the county-level data by sampling half of the counties in each iteration. Temporal dynamics of (g) EP, (h) Within-EP, and (i) Between-EP by Democrat, Republican, and Swing states. Error bars represent 95% CI.

We complement the previous findings with data from France and Chile (see Figure 3.6). In France, the most polarized elections occurred in 2002 and 2022, with a 26% of increase in Within-EP in the last five years. There are reasons to believe that these elections were polarized. On the one hand, the 2002 election was the first time a candidate from the *Rassemblement national* (a far-right party) made it into the runoff. On the other hand, the 2017–2022 period (between two presidential elections) witnessed massive protests (*Yellow Vest Movement*), climaxing the feeling that the French population was more polarized than in the past. From a political perspective, Within-EP growth is worrying since seven candidates and the

three majorities (Macron, Le Pen, and Mélenchon) were the same in both elections. For Chile, we observe an increment of 6.3% of Within-EP and 8.1% in Between-EP from 2013 to 2021, which is in line with theoretical expectations.

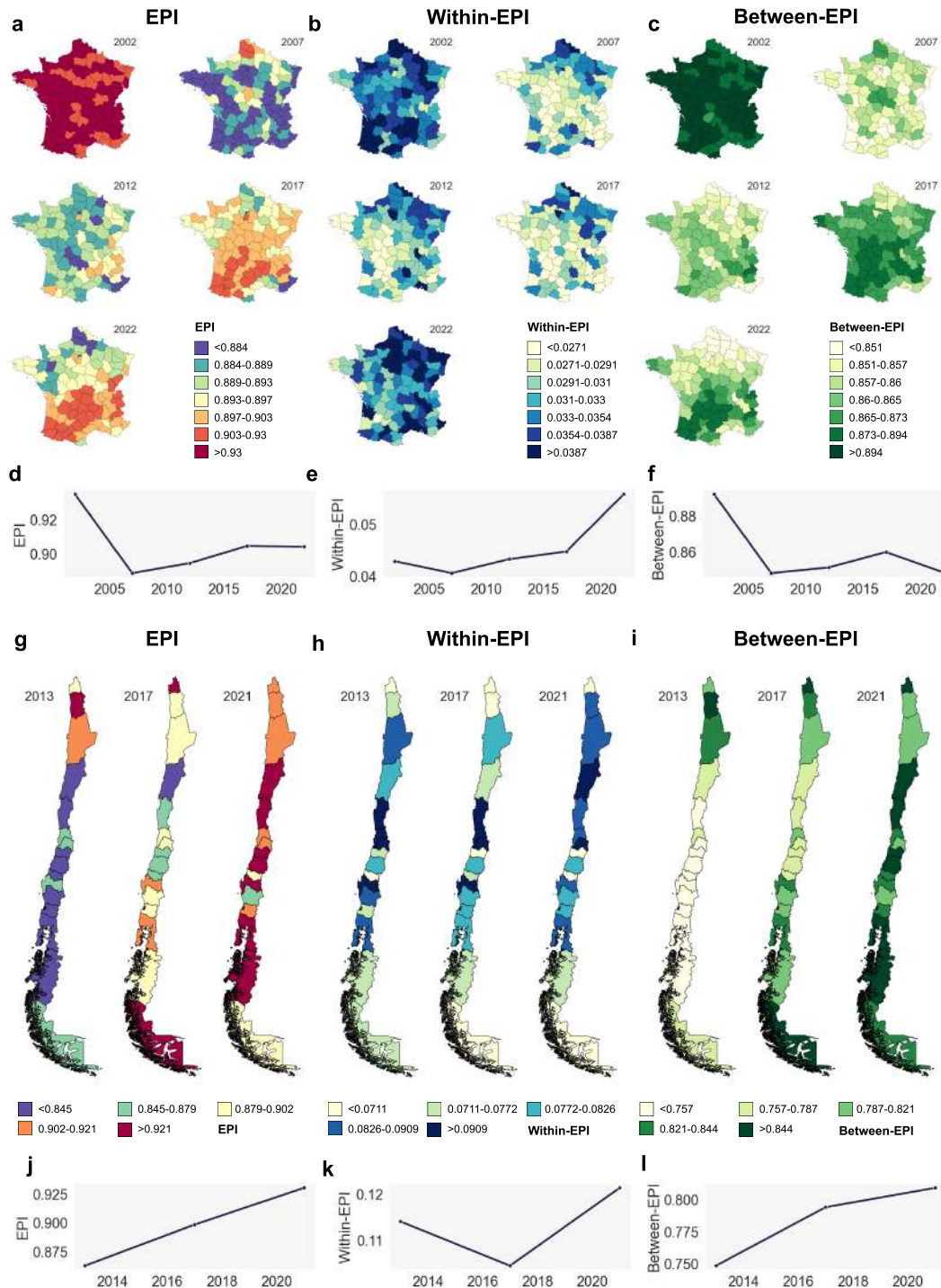


Figure 3.6: The geography of Election Polarization in France (2002-2022) and Chile (2013-2021). Temporal dynamics of department-level (a) EP, (b) Within-EP, and (c) Between-EP in France. Temporal dynamics of country-level (d) EP, (e) Within-EP, and (f) Between-EP in France. Temporal dynamics of region-level (g) EP, (h) Within-EP, and (i) Between-EP in Chile. Temporal dynamics of country-level (j) EP, (k) Within-EP, and (l) Between-EP in Chile. We bootstrapped 100 times the county-level data by sampling half of the counties in each iteration. Error bars represent 95% CI.

Overall, the previous analysis yielded that EP can be suitable for measuring polarization. Specifically, our findings suggest that Within-EP is the dimension that aligns with theoretical expectations, and instead, Between-EP could be interpreted as a complementary measure of election outcomes.

3.4.3 Political Polarization, Political Interest, and Election Polarization

In addition to empirically measuring EP, we use political polarization and political interest to analyze their relationship with Within/Between EP. Our assumption is that the political climate surrounding an election should be associated with election polarization.

Here, we focus on a state-level analysis in the United States since we rely on fine-grained data for a longer study period. To the best of our knowledge, the U.S. is the only country with sub-national level data to compute a form of “ideological” polarization. Using data from the Cooperative Election Study (CES) and the U.S. Election Project (McDonald, 2023), we investigate the combined characteristics that encompass political polarization—the ideological distance between Democrats and Republicans, defined in eq. 3.12—, and political interest—proxied by the voting turnout on Election Day. Our analysis does not target causation; rather, we focus on understanding whether the dimensions of EP are associated with political dimensions or whether its effects are already captured by socio-demographic features. To ensure that socio-demographic characteristics and not other factors cause any increase in political polarization or turnout in the U.S., we control our analysis by population density, GDP per capita, personal income per capita, educational attainment, and a dummy variable for swing states. We also hold those unobserved factors in an Election Year by using year-fixed effects since an increase in polarization could result from multiple factors other than socio-demographic variables. For instance, the Cambridge Analytica scandal might play a role in the polarization identified in the 2016 Election.

We set up panel regressions as follows:

$$y_{g,t} = \beta_0^T \text{EP}_{g,t} + \beta_1^T \text{EP}_{-g,t} + \beta_2^T \mathbf{X}_{g,t} + \beta_3 \mathbf{S}_g + \mu_t + \epsilon_{g,t}$$

Where $y_{g,t}$ is the dependent variable for the geography g in year t (in our case, the political polarization and political interest), $\text{EP}_{g,t}$ is a vector of EP components (Between-EP and Within-EP) and $\text{EP}_{-g,t}$ is a vector of spatial EP components. $\mathbf{X}_{g,t}$ is a vector of control variables that account for other factors addressed in the literature. Here, we use an inequality level (Gini coefficient), population density, education attainment, annual unemployment rate, personal income per capita, and GDP per capita. \mathbf{S}_g is a dummy to control by swing states, μ_t is the period fixed effect, and $\epsilon_{g,t}$ is the error term. The variables have been normalized by subtracting the average and dividing by the standard deviation before running the regressions. Standard errors are clustered at the state level.

We present the models for state-level associations of political polarization in Table 3.2 and political interest in Table 3.3. From each table, we consider Model 6 as our “final model”.

	(1)	(2)	(3)	(4)	(5)	(6)
Inequality level (t)	0.12 (0.08)				0.04 (0.09)	-0.03 (0.08)
Unemployment (t)		0.15 (0.10)			0.07 (0.10)	0.12 (0.09)
Within-EP (t)			0.32*** (0.09)		0.39*** (0.08)	0.27*** (0.10)
Between-EP (t)				-0.10 (0.10)	0.05 (0.11)	0.01 (0.12)
Inequality * Unemployment (t)					0.03 (0.04)	0.03 (0.04)
Spatial Within-EP (t)						0.30** (0.12)
Spatial Between-EP (t)						-0.03 (0.11)
Ed. Attainment (t)					-0.44*** (0.16)	-0.31** (0.15)
Personal Income PC (t)					0.17 (0.24)	0.22 (0.22)
GDP PC (t)					-0.28 (0.28)	-0.19 (0.22)
Pop. Density (t)					0.56*** (0.20)	0.27 (0.22)
Num. obs.	204	204	204	204	204	204
R ² (full model)	0.37	0.37	0.46	0.37	0.57	0.61
R ² (proj model)	0.04	0.05	0.18	0.04	0.35	0.41
Adj. R ² (full model)	0.35	0.36	0.45	0.35	0.54	0.58
Adj. R ² (proj model)	0.02	0.02	0.16	0.01	0.31	0.36
Num. groups: year	4	4	4	4	4	4

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 3.2: State-level association of Political Polarization in the United States (2008-2020). Each model represents a multivariate regression, with year-fixed effects. Standard errors are in brackets and are robustly clustered by state. We standardized both independent and dependent variables to have a mean of zero and a standard deviation of one. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Political Polarization We find a positive relationship between Within-EP and Spatial Within-EP with political polarization (0.27, $P < 0.01$ and 0.29, $P < 0.05$, respectively), being slightly less significant than educational attainment (-0.3, $P < 0.1$) (Table 3.2)—after controlling by socio-demographic conditions. No significant associations are observed for Between-EP. This result is interesting for several reasons. First, it suggests that Within-EP might be interpreted as a form of quasi-political polarization of geographies. Second, it extends the idea that election data can be a reliable proxy for capturing political polarization—especially for lower and middle-income countries without national surveys to characterize this phenomenon. Third, opens a window to study political polarization at multiple and disaggregated scales, such as counties or districts.

	(1)	(2)	(3)	(4)	(5)	(6)
Inequality level (t)	-0.13 (0.10)				-0.12 (0.09)	-0.15* (0.09)
Unemployment (t)		-0.17* (0.10)			-0.06 (0.08)	-0.00 (0.10)
Within-EP (t)			-0.01 (0.13)		0.11 (0.16)	0.07 (0.11)
Between-EP (t)				0.20* (0.11)	0.31* (0.17)	0.23** (0.11)
Inequality * Unemployment (t)					-0.01 (0.04)	-0.03 (0.04)
Spatial Within-EP (t)						-0.04 (0.14)
Spatial Between-EP (t)						0.22 (0.16)
Ed. Attainment (t)					0.53*** (0.15)	0.51*** (0.15)
Personal Income PC (t)					0.08 (0.25)	0.05 (0.25)
GDP PC (t)					0.11 (0.19)	0.26 (0.23)
Pop. Density (t)					-0.16 (0.25)	-0.33 (0.21)
Swing State	0.24 (0.27)	0.30 (0.26)	0.27 (0.27)	0.16 (0.27)	0.30 (0.26)	0.30 (0.25)
Num. obs.	303	303	303	303	303	303
R ² (full model)	0.27	0.27	0.26	0.30	0.51	0.54
R ² (proj model)	0.03	0.03	0.02	0.07	0.35	0.39
Adj. R ² (full model)	0.26	0.26	0.24	0.28	0.48	0.52
Adj. R ² (proj model)	0.01	0.01	-0.01	0.04	0.31	0.36
Num. groups: year	6	6	6	6	6	6

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 3.3: State-level association of Voting Turnout in the United States (2008-2020). Each model represents a multivariate regression, with year-fixed effects. Standard errors are in brackets and are robustly clustered by state. We standardized both independent and dependent variables to have a mean of zero and a standard deviation of one. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Political Interest Next, Table 3.3 shows that Between-EP positively correlates with the voting turnout (0.23, $P < 0.05$) and it is the second most important variable in terms of size effect in the final model. Here, we do not observe significant associations for Within-EP. Since we have voting turnout data for France and Chile, we find similar univariate correlation effects in both countries (France (department-level): $\rho_{2022} = 0.46$, Chile (region-level): $\rho_{2021} = 0.19$). A potential explanation suggests that as competitive elections attract participation, Between-EP might capture this pattern. Notably, this condition might point out that polarization for Swing states is mainly driven by competitiveness and not by ideology.

In summary, we interpret that we can label an election as polarized by simply knowing whether two conditions are satisfied: it attracts total participation of the electorate (or close to

that), and voters' preferences are highly clustered geographically. If we set up Within-EP and Between-EP as our dependent variables (See Appendix), the socio-demographic characteristics are better predictors of Within-EP, providing evidence that macroeconomic conditions in Election Year impact election outcomes, and it can be used as a proxy of political polarization.

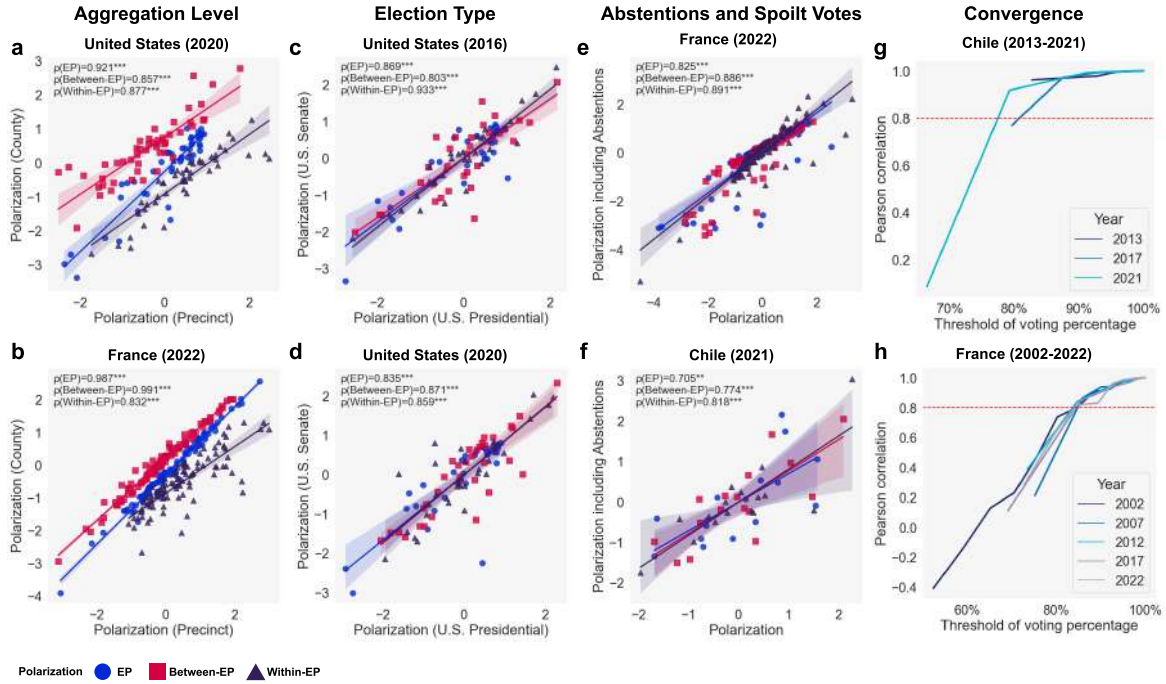


Figure 3.7: Robustness across aggregation levels using real data. Robustness by aggregation level in the (a) United States (2020) and (b) France (2022). Robustness by election type–State-level EP of senate vs. presidential–in the United States in (c) 2016 and (d) 2020. Robustness by including abstentions and spoilt votes as candidates in (e) France (2022) and (f) Chile (2021). Convergence of EP values in (g) Chile (2013-2021) and (h) France (2002-2022). The value in the top-left corner is the Pearson correlation. The dashed line in (g-h) represents the correlation equal to 0.8. The variables have been normalized by subtracting the average and dividing by the standard deviation before running the regressions. *Note:* $*p < 0.1$; $**p < 0.05$; $***p < 0.01$

3.4.4 Robustness across Aggregation Levels

To conclude, we perform a robustness analysis to validate and extend the reliability of our method by testing five approaches that can bias the result: a specific election type (e.g., presidential or senate), the use of a particular aggregation scale, the role of abstentions and spoilt votes, the removal of protest candidacies, and the differences between first round and runoff.

Firstly, we compute the regional-level EP in France (department-level, 2022) and the United States (state-level, 2020) by aggregating preferences at precinct and county scales. Figure 3.7 a-b shows in both countries a strong positive correlation for all EP components (France: EP=0.987, $P < 0.01$, Within-EP=0.991, $P < 0.01$, Between-EP=0.832, $P < 0.01$; United States: EP=0.921, $P < 0.01$, Within-EP=0.877, $P < 0.01$, Between-EP=0.857, $P < 0.01$), discarding that a specific aggregation level could alter the main results.

Secondly, we compare the state-level EP of two elections held on the same day using data from general and senate elections in the United States in 2016 and 2020 (precinct level).

Figure 3.7 c-d shows a strong positive correlation between the EP computed for both elections (2016: EP= 0.869, $P < 0.01$, Within-EP= 0.933, $P < 0.01$, Between-EP= 0.803, $P < 0.01$; 2020: EP= 0.835, $P < 0.01$, Within-EP= 0.859, $P < 0.01$, Between-EP= 0.871, $P < 0.01$). Since both elections are celebrated simultaneously, a potential non-causal explanation for this correlation level would be that as citizens are influenced by similar events (e.g., discussions with friends, ads campaigns, TV, and social media), they mirror a similar polarization level through their votes.

Thirdly, we analyze whether abstentions, blanks, and null votes distort the findings. Figure 3.7 e-f shows a positive correlation between the EP computed using all the candidates with the one computed including abstentions, null and blank votes both in France (2022) (EP= 0.825, $P < 0.01$, Between-EP= 0.886, $P < 0.01$, Within-EP= 0.891, $P < 0.01$) and in Chile (2021) (EP= 0.705, $P < 0.05$, Between-EP= 0.774, $P < 0.01$, Within-EP= 0.818, $P < 0.01$). Nevertheless, we do not observe a strong correlation between EP and Between-EP in Chile (2017) (EP= 0.304, Between-EP= 0.276, Within-EP= 0.945, $P < 0.01$ in Appendix). We conclude that abstentions do not bias the Within-EP; however, additional analyses should be required to interpret Between-EP under scenarios with high abstention and spoilt votes.

Fourthly, we test whether EP computed on a subset of candidates is feasible. This is relevant to the extent of comparing elections with different numbers of candidates. Using elections from France and Chile (since they both have the characteristic of having more than two competitive candidates per election), we compute the EP by relying on the top-2 candidates, then top-3, and so on, until completing all of them (See Appendix). The excluded candidates were grouped in the category “other”. Figure 3.7 g-h shows the Pearson correlation between the EP computed from all candidates and the one calculated using the top- N candidates. We observe a strong positive correlation (greater than 0.8) for any number of candidates that represent at least 80 % of votes, being robust to this finding for all elections analyzed. For instance, this paper relied on the top-8 candidates in France to perform the main analysis since they consistently represented more than 80% preferences across time.

Fifthly, we examine whether EP computed from runoffs is equivalent to the one computed using the first round in presidential systems. Since both rounds are usually close in time (two weeks in France and four weeks in Chile), it seems reasonable to think that the changes in population polarization must be marginal. Nevertheless, we know that candidates moderate their political positions in the second rounds (Bordignon et al., 2016) and citizens might be involved in voting by the “lesser evil” in some elections. Figure 3.8 shows the correlation of EP for France (2002-2022) and Chile (2013-2021). We systematically observe that the EP from the first and second rounds are uncorrelated or exhibit a slight correlation. For instance, in the 2002 election in France, where it is documented that citizens voted by the lesser evil, the correlation for Between-EP was 0.270 ($P < 0.05$), and the one for Within-EP was 0.363 ($P < 0.05$). This leads us to discard the use of runoffs for mapping EP.

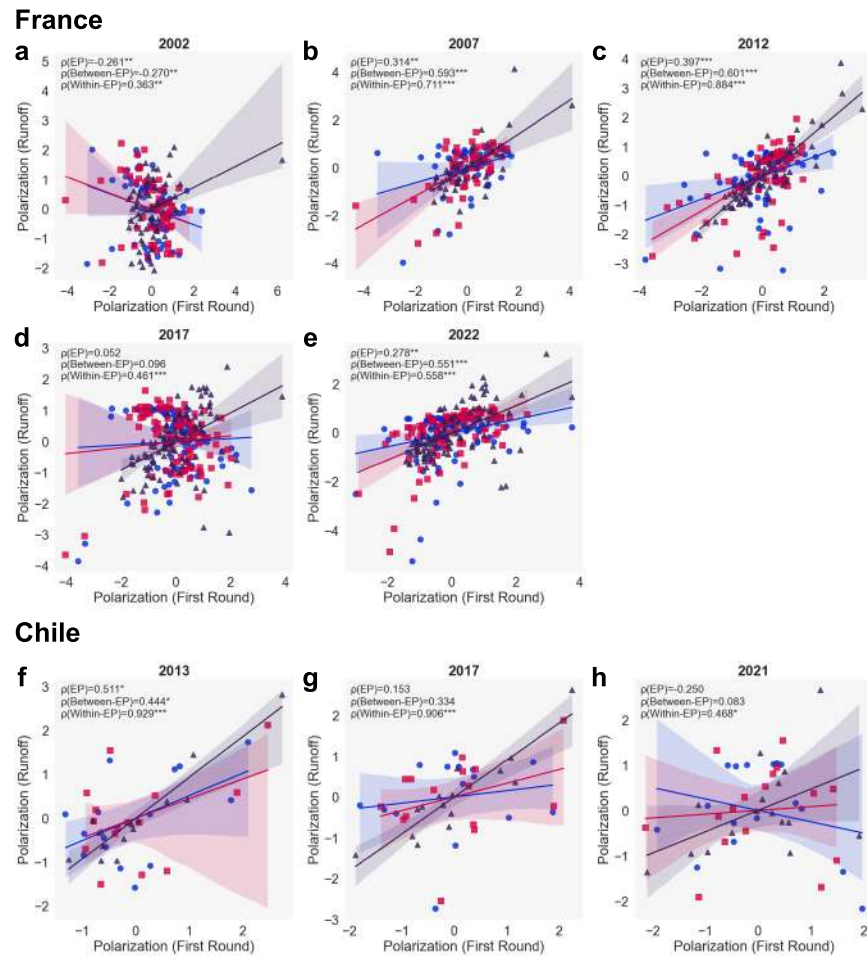


Figure 3.8: Comparison of Election Polarization and its components computed using First Round and Runoff data in France (a) 2002, (b) 2007, (c) 2012, (d) 2017, (e) 2022, and Chile (f) 2013, (g) 2017, and (h) 2022. Note: $*p < 0.1$; $**p < 0.05$; $***p < 0.01$

These results computationally validate the usefulness of applying our metrics using election data under different assumptions (e.g., temporally, spatially) and relying on incomplete and noisy data, contributing to the reliability of this approach.

3.5 Discussion

In a broad sense, polarization can be defined as the division of contrasting groups in a society. This phenomenon is addressed from several branches, such as ideological, geographical, social, political, or affective, to name a few. The temporal and spatial dynamics of polarization are related to effects seen in society, such as inequality and conflicts. These dynamics corroborate little by little to extreme ideological political parties, and even lower attitudes towards health problems such as COVID-19.

Previous work has established that candidates take advantage of polarization levels. Evidence from the 2018 Brazilian presidential election suggests that Jair Bolsonaro exploited demographic segregation in order to win the election (Layton et al., 2021). Similar findings

were observed in the 2016 United States presidential election, showing that unusually explicit appeals to racial and ethnic resentment and anti-establishment feelings attracted strong support from white working-class voters in favor of Donald Trump (Abramowitz & McCoy, 2019). Our view is that measuring and distinguishing polarization from election outcomes is essential to studying their impacts on our society quantitatively.

This chapter examines the concept of Election Polarization (EP) as a measure of citizen divisions on Election Day. In particular, we present an agnostic method to compute EP by distinguishing it into geographical (Within-EP) and competitiveness (Between-EP) components. Using synthetic and empirical data, we demonstrated that our approach successfully labeled polarizing elections and growing trends presented in the literature for both France, the United States, and Chile. Additionally, we find a positive relationship of Between-EP and political interest in France, Chile, and the United States and Within-EP with political polarization in the United States, respectively, yielding that we can define a polarized election as satisfying two conditions: it attracts total participation of the electorate (or close to that), and voters' preferences are highly clustered geographically. However, we caution that our findings do not claim causation, raising additional analysis to confirm that the Within-EP could be defined as a proxy for political polarization.

Even though this study is limited to three countries with majoritarian systems, which are not able to represent the diversity of electoral systems and political contexts around the world, we believe that election polarization is a straightforward concept in understanding and comprehending polarization from a geographical perspective. Furthermore, the fact that the approach presented in this study has the potential to become an indicator of polarization in geographies opens the possibility of extending the study of polarization beyond U.S. borders. Nevertheless, we call for further focus on causation mechanisms between social interactions and the voting patterns revealed on Election Day. A second stage could be linking the ideological position of candidates with the antagonism level and exploring the before and after election polls and their relationship with the EP.

To conclude, we set up some considerations. First, the geographical hierarchy of an election represents the largest aggregation level to calculate the EP. The reason is that local candidates' quality could bias the party's support in a location (Roy & Alcantara, 2015). For example, the House of Representatives election outcomes in the United States do not work for computing state-level EP since citizens in two congressional districts from the same state cast different ballots. Second, our method is not valid for non-free-and-fair elections—a referendum with 99% approval in a dictatorship does not mean that society is not polarized at all. Finally, run-off data sets are unsuitable for mapping EP because the values computed from the first and second rounds are uncorrelated in some cases.

Online Involvement in Two Real-World Digital Democracy Systems

Abstract

The consensus-oriented approach pursued by recent digital democracy efforts has shown promising results in decreasing polarization, increasing citizen involvement, and reaching collective decisions. Yet, there is little evidence of how participation inside these systems can be affected by their endogenous characteristics. Inspired by the fact that citizens prefer to avoid divisive issues in order to abstain from pressures and related literature in behavioral studies, this chapter examines how the content and user interface condition anonymous online involvement. Specifically, here we use data collected from two real-world digital democracy systems presented in Chapter 2. Our analysis yields that participants did not drop out of the platforms by merely reading a contentious issue nor abstaining from them in greater value. Furthermore, we report that careless participation is associated with demographic features and the time spent on the platforms. Our concluding discussion suggests a set of insights and recommendations for platform planners and practitioners of digital democracy.

4.1 Introduction

Digital technologies have contributed to enhancing citizen involvement in decision-making processes. Nevertheless, transitioning from sporadic to more routinary participation instances represents a triple challenge—both for citizens, public institutions, and platform planners (the ones that design and release online participation tools)—for the creation of user-friendly digital democracy tools. At the same time, the spread of these systems has shown promising results in reducing partisan polarization (Fishkin et al., 2021), creating policy regulations (Small et al., 2021), and unraveling citizen demands of social movements (González-Bailón et al., 2011). However, the digital sphere also shows the worst of our society. The proliferation of trolls and the role played by anonymity spreading hate speech or unpolite behavior have eroded confidence in the transformative potential of digital technologies. Despite this pessimistic view, we still consider that anonymous participation, under controlled conditions, can have potential benefits. Specifically, a line of argument in political science argues that people avoid face-to-face discussions over divisive issues in order to abstain from social pressures and confrontations (Gerber et al., 2013). This leads us to think that including these issues in an anonymous setup could contribute to better revealing the extent of disagreements in society. This assumption implicitly suggests the interface significantly impacts the form we grasp citizen preferences. Related literature in behavioral studies sheds light on data quality in online experiments (Meade & Craig, 2012). This is worthy for platform planners, allowing them to make strategies to reduce the cognitive burden and, at the same time, collect reliable preferences from participants.

This chapter extends the analysis of data collected through the two real-world digital democracy systems introduced in Chapter 2 to investigate whether the content, user interface, and demographic characteristics affect online involvement. Our contribution can be split two-fold. On the one hand, we examine whether including contentious issues in an anonymous environment conditions early dropouts. On the other hand, we assess how the elicitation method used to collect preferences impacts the data quantity and to what extent the fatigue of long participation periods could be associated with careless participants. Our study fills a gap in the literature on the interplay of behavioral studies from psychology, digital democracy, and hypotheses from political science.

Chapter 4 is structured as follows. Section 2 retrieves a review of contentious issues in political discussions, the role of anonymity on the Internet, and inputs from behavioral studies in psychology. Section 3 overviews our methodology. Section 4 presents the main findings, and Section 5 discusses our final remarks.

4.2 Literature Review and Hypothesis Formulation

Our study investigates whether contentious issues could potentially affect participation in digital democracy systems. We relate our work with literature about divisive issues in political episodes, the role played by consensus-oriented digital democracy systems in enhancing citizen involvement, and anonymous participation on the Internet.

4.2.1 Contentious issues and Anonymity

Discussing contentious topics is an act that we even keep away from doing with relatives and friends. Related literature has established that people avoid talking about contentious issues to abstain from confrontations or social pressures (Chen & Berger, 2013; Gerber et al., 2013; Simons & Green, 2018). Nevertheless, the fact that these issues can better highlight political stances than the main agreements incentivizes candidates to focus on them in their political campaigns (Ansolabehere, Puy, et al., 2015; Ash et al., 2017). Of course, a topic can be controversial both in political and non-political contexts; however, when they are of a political nature, there is evidence that social media reinforces the avoidance of being exposed to them through filter bubbles (Barberá et al., 2015). Perhaps citizens feel more comfortable talking about these topics in anonymous environments. According to Perbawani et al., 2018, people prefer anonymity in online political participation to protect from online threats, avoid conflicts, and even legal consequences, and Kang et al., 2013 described that participants seek anonymity for multiple reasons, such as protecting their jobs or because previous bad experiences. We then understand that people can feel less restrained in giving their opinion under an anonymous context, showing a dissociation from their in-person behaviors (Suler, 2004). Unfortunately, feeling that your comments can not be tracked back to you also leads people to show their worst version. Online anonymous involvement is likely to be more uncivil and impolite than when individuals reveal their identity (Rowe, 2015), and removing the anonymity can nearly eliminate negative comments (Kilner & Hoadley, 2017).

Here we address that digital democracy systems have a role to play in terms of deliberation. There are good reasons to believe that online tools can foster fructiferous conversations over controversial issues. Notable examples include Pol.is (Small et al., 2021) establishing regulation for Uber drivers in Taiwan, the crowd-sourcing approach for writing a new constitution for Iceland (Landemore, 2015), and the Stanford Deliberative Program (Fishkin et al., 2018; Fishkin et al., 2021) decreasing polarization of opinion over contesting issues between Republicans and Democrats. These examples show that an adequate use of digital technologies can be helpful for democracy. In fact, Helbing et al., 2023 address the potential of these tools to establish more inclusive and representative democracies, emphasizing the benefits of consensus-oriented setups. Nevertheless, focusing on consensus does not mean the total exclusion of divisions. While little work has explored to what extent the inclusion of contentious issues threatens online participation, we find a few recent examples in the literature. According to Cantador et al., 2020, divisive issues reach a lower audience than consensual ones, and Navarrete et al., 2022 suggest that divisive issues derived from online participation can better suit political interest in electoral times (See Chapter 2).

While the development of consensus-oriented frameworks for digital democracy is seen as a guideline today, we identify a gap in the literature on how the content, specifically the mere inclusion of contentious issues, can hamper online involvement. Going back to our review, we hypothesize that face-to-face avoidance of contentious issues discussions should not be mirrored in anonymous online participation, arguing that participants should feel disinhibited from giving their opinions since they are not challenged for the reasons behind them. This leads to our first hypothesis:

Hypothesis 1: Asking for contentious issues should not condition the likelihood of early dropout in an anonymous digital democracy system.

4.2.2 Collecting preferences in online participation

Having hypothesized that including divisive topics should not be detrimental to anonymous online participation, we turn our focus to examining whether the characteristics of elicitation methods themselves affect participation and collective decisions. From social choice, we know that the outcome of a decision using the same data set can vary dramatically by aggregating the preferences using a particular voting rule (Klamler, 2005). Nevertheless, the voting rule is not the only feature that determines a collective decision. The user interface design can condition our selections (e.g., Cambre et al., 2018).

This study focuses on pairwise comparison technique. This approach, traced back to Thurstone, 1927, considers that data collected via relative comparisons of issues (i.e., which issue do you prefer? A or B) can be aggregated into a reasonable collective ranking. The success and the wide adoption of pairwise comparison as a mechanism for crowdsourcing in online setups could be treated by two main headings: the reduction of the cognitive burden from participants and its simple interface. To date, this technique has shown impressive results, both scaling up data collection (Salganik & Levy, 2015), mapping urban perception (Saleses et al., 2013), studying moral dilemmas of self-driving cars (Awad et al., 2018), giving accurate recommendations from recommender systems (Kalloori et al., 2018), but, most notably for our purposes, is that in the political arena, it contributed to mapping citizen preferences in political episodes (Hidalgo, 2019). Nonetheless, asking participants for a large number of comparisons is a task that increases the cognitive burden of participants exponentially over time. The reason is simple: the number of comparisons for N issues is given by $\frac{N(N-1)}{2}$. Considering a web application that displays one comparison per screen, asking for 100 issues quickly involves around 5000 screens, making it nearly impossible to achieve from a participant. This leads to the proposition of alternative methods that could potentially collect the same data structure as a pairwise comparison setup. For instance, Lee et al., 2014 show that using a small percentage of issue pairs can approximate the rank expected for complete preferences. In terms of elicitation methods, Brams and Sanver, 2009 propose Fallback Voting, which considers two stages: first, asking users for their approvals and second, asking them to rank the issues they approved. From a theoretical perspective, the data structure from Fallback Voting should be equivalent to pairwise comparison (when ties are not allowed). The latter leads to our second hypothesis, i.e.:

Hypothesis 2: Elicitation methods that aim to collect the same data structure from the same population might yield variations in terms of data quantity, though not data quality.

Assuming that the elicitation method could affect the data collection in terms of quantity but not quality, we revise related literature in the study of data quality. To frame our assessment in digital democracy systems, we borrow some definitions from behavioral studies in psychology, specifically from the strand of literature that explores participation in online behavioral research on recruitment platforms such as Amazon Mechanical Turk or Prolific and online survey services. According to Peer et al., 2022, quality can be framed from several perspectives, which the most important for researchers in the discipline are attention, comprehension, reliability, and dishonesty. There is a main agreement that the attention of users is not an endless feature, especially in long questionnaires (Bowling et al., 2021). Likewise, participants reduce their attention at the end of online experiments (Berry et al., 1992) and

can be careless responders (Meade & Craig, 2012). A line of argument suggests that carelessness could be associated with personality traits (Bowling et al., 2016) and demographic characteristics of participants (Nichols & Edlund, 2020).

Perhaps, we can not really distinguish whether a participant gives accurate or reliable preference in online involvement. Therefore, we believe it is crucial to examine the boundaries of carelessness in digital democracy in order to advance our understanding of online participation in political episodes. Related literature addresses the relevance of computing indices and the identification of outliers to guarantee high-quality data, such as including attention checks to verify whether the individual is attentive to the questions (Meade & Craig, 2012; Newman et al., 2021) or the use of *Mahalanobis Distance* as the extent of measuring the distance from the average response pattern to identify outliers. Ward and Meade, 2018 suggests that more engaging instructions (e.g., applying cognitive dissonance) can increase respondent motivation. Even though this literature considers the removal of careless participants from the main analysis, the potential threat of removing noisy responders can lead to non-representative samples and be detrimental to external validity. This is a touchy discussion for digital democracy systems. For instance, a minoritarian preference, by definition, will be far from the group decision and does not mean that should be removed by using Mahalanobis Distance.

In light of the discussion, it is reasonable to think that the preferences provided by a participant during the first two or three minutes of participation can be more careful than those provided after 15 minutes inside the platform, and this increase in carelessness can detriment collective decisions. This assumption drives our third and fourth hypotheses, i.e.:

Hypothesis 3: The quality of responses provided by participants decreases over time.

Hypothesis 4: Careless responses significantly affect collective decisions that emerged from digital democracy systems.

4.3 Methods

This section outlines our methods. We start by describing the platforms and the reasons behind the user interface chosen. Then, we describe our criteria to map divisive issues, data quantity, and data quality. To conclude, we introduce three logistic regression models performed to understand the demographic characteristics of carelessness.

4.3.1 The Platforms

We released two real-world digital democracy systems preceding the 2022 presidential elections in France and Brazil. Both platforms aimed to help in the creation of a collaborative government program by asking participants their opinions on a large number of issues, 120 in France and 67 in Brazil, curated from the presidential candidates' government programs. These platforms were conceptualized, designed, and released integrally by

our research team. Specifically, this dissertation contributed to coding a standard framework for both tools, which is deposited on GitHub with the name of Opencracia (<https://github.com/CenterForCollectiveLearning/opencracia>).

Participation in these platforms was anonymous and voluntary. The only incentive to continue participating was to receive a personalized government program and contribute to a collaborative one, publicly shared on the results page. As mentioned in Chapter 2, we did not claim to be a representative population sample. Instead, our focus was creating a controlled environment for mapping citizen preferences.

From a platform planner side, we attempted to balance asking a large number of questions and the cognitive burden of participants. To understand how the interface could potentially affect the data collection, we tested a total of five scenarios. We coded a JavaScript routine to assign participants to these scenarios randomly, not being aware of the existence of other ones.

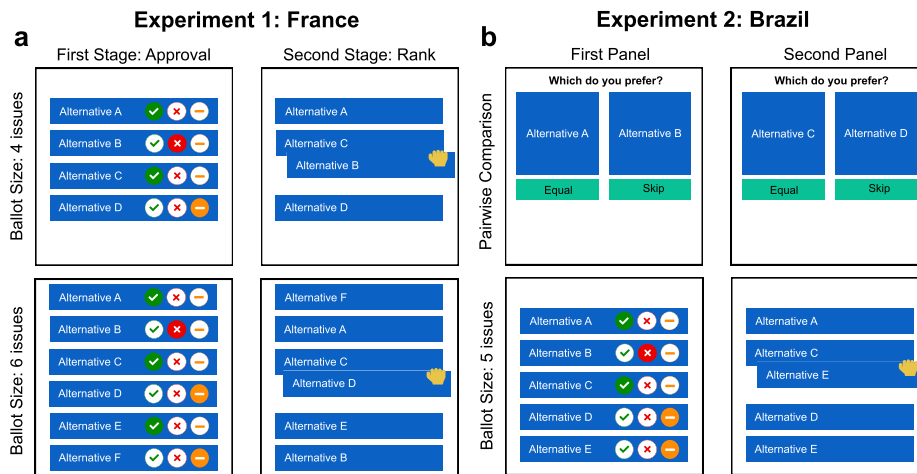


Figure 4.1: Unpacking the user interface of Rank screen. Illustration of experimental systems released in (a) France and (b) Brazil.

France: *Mon Programme 2022*

The French platform (<https://monprogramme2022.org>) is integrally an online adaptation of Fallback Voting (Brams & Sanver, 2009). The first stage asked participants to approve/disapprove/abstain a set of issues, and the second stage asked them to rank the ones they approved. While it is evident that asking for 120 issues is unfeasible in a web application, especially on mobile screens, we overcome this drawback by creating universes: a scenario that samples k issues per screen, where $k \in \{4, 5, 6\}$. The immediate effect is dividing the participation into more than one screen. For example, it took 40 screens to complete the experiment for a participant assigned to Universe 6, but 60 screens for one assigned to Universe 4. Here we alternate between approval and rank screens until a participant responds to all of them. Figure 4.1 a roughly illustrates the differences between universes 4 and 6.

Brazil: *Brazucracia*

The Brazilian platform (<https://brazucracia.org>) was launched seven months after its French twin. This tool included technical upgrades derived from the French experience. The most significant change concerns running an A/B test by comparing the fallback voting setup and the classical framework of pairwise comparison (Figure 4.1 b). It should be noted that the fallback voting here only included screens with five issues each, and we first displayed all approval screens and then the rank ones.

Both platforms randomly assigned the issues order displayed to each participant. This means that each one has a pseudo-unique experience. Even though we could consider edge cases where two participants read the issues in the same order, this is a very low-probability event. Given that issues were divided by screens in fallback voting, we included a progress bar in the experiment that ranges from 0 to 100%, which 100% means participation provided preferences in all issues. The progress bar was calculated with the number of screens completed.

Unpacking user-interface

This section describes the composition of each screen in terms of design and expected user interaction.

Approval screen The first step of fallback voting can be framed as approval ballots. By platform design, participants were limited to three alternatives: approve, disapprove, or abstain. They were not allowed to skip any preference. Only after answering all issues, participants could move to the next screen (Figure 4.2). We hypothesize that the inclusion of the abstain button contributes as a buffer to avoid eliciting unpleasing preferences and, eventually, increase engagement with the platform. Unfortunately, our experiments did not control approval screens by excluding abstention buttons.

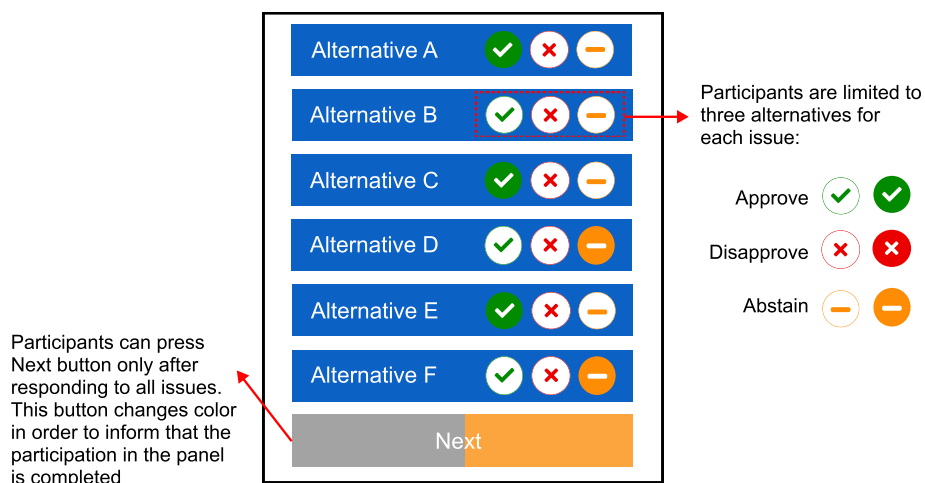


Figure 4.2: Illustrative example of Approval screen for six issues.

Rank screen The second step of fallback voting can be framed as a ranking ballot. Participants receive a randomly sampled order of k issues. Due to the fact that this screen only includes issues previously approved by the participant, we first include issues from the preceding approval screen, completing the k issues with approved issues from all screens. Then, they could either accept the proposal and continue their participation or use the interactive interface (drag-and-drop and up/down buttons) to modify the proposal (Figure 4.3).

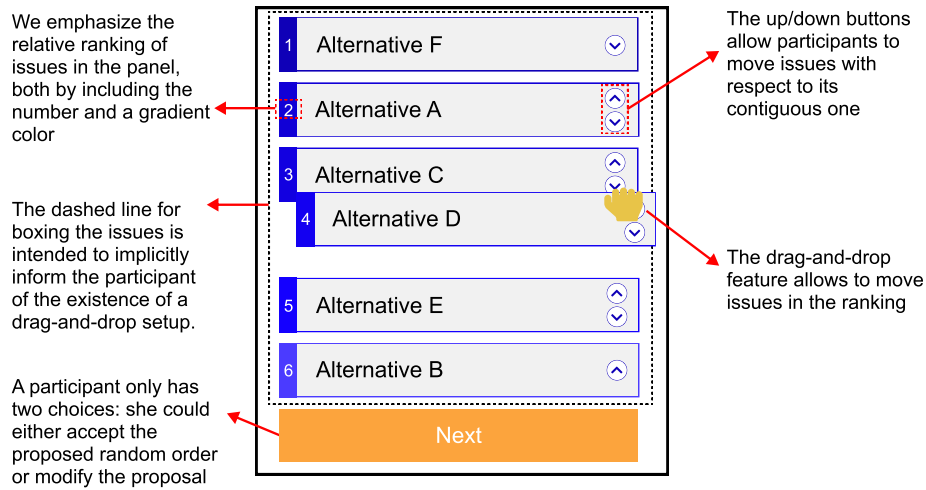


Figure 4.3: Illustrative example of Rank screen for 6 issues.

Pairwise comparison screen The classical pairwise comparison framework asked participants their relative preference over two issues emphasized in the user interface with greater boxes. In contrast with fallback voting, here we allowed participants to skip questions or allow ties (both preferences are equally relevant). The selection of an option implicitly means moving to the next screen (Figure 4.4).

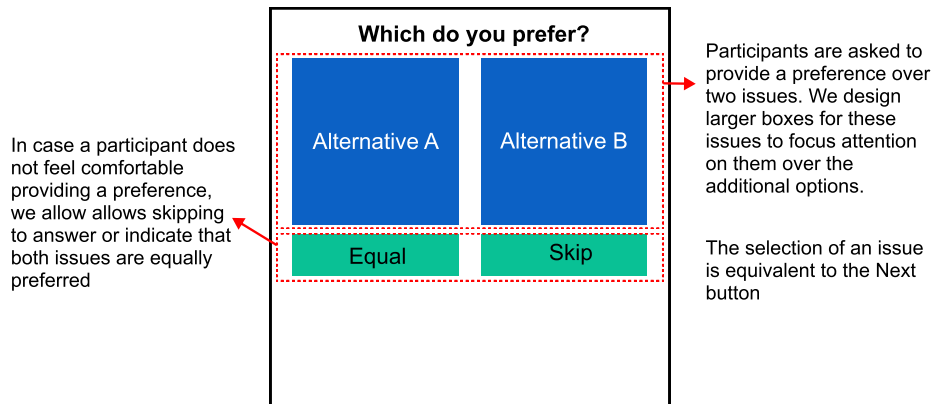


Figure 4.4: Illustrative example of Pairwise Comparison.

4.3.2 Mapping contentious issues in ballots

Our first hypothesis calls us to investigate whether early dropouts are conditioned by merely reading a contentious issue. Here we focus on data collected in fallback voting setups.

We define a contentious issue as one that is ranked over 85 percent of all issues in a ranking of divisive issues. Out of a total of 120 issues in France, 18 were classified as divisive. Similarly, out of a total of 67 issues in Brazil, 10 were classified as divisive. We determine the issues' divisiveness by using two measures borrowed from Chapter 2: On the one hand, **political divisiveness**: the absolute score difference between the data aggregated from participants who self-identified as the political left and individuals who self-identified as the political right. On the other hand, the agnostic **divisiveness** approach, introduced in Chapter 2, aggregates preferences using a given aggregation function (in our case, weighted Borda count).

Then, we conduct a quasi-experimental approach to distinguish our treatment and control conditions. Since we randomly assigned the issues displayed per screen, we define our conditions as follows:

Treatment condition: Participants who read at least one contentious issue in the first approval screen.

Control condition: Participants who did not read any collective contentious issue in the first approval screen.

We focus on the first approval screen for two reasons. On the one hand, this screen gives participants a first glance at the nature of topics asked on the platform. On the other hand, we rely on a larger sample of preferences since almost all participants who previously accepted the consented form provided at least one preference. Out of a total of 585 participants in Brazil, 304 were classified in the treatment condition (52%) and 281 in the control condition (48%). Out of a total of 1529 participants in France, 779 were classified in the treatment condition (50.9%) and 750 in the control condition (49.1%).

4.3.3 Measuring data quantity and quality

We start by describing our definition of data quantity and quality; then, we outline the methods addressed to measure data quality and finish with an interplay between quantity and quality.

Data quantity

This study defines quantity as the number of issue pairs gathered per minute. The first part of hypothesis 2 leads us to investigate whether data collected via Fallback Voting is more efficient in terms of quantity than the classical pairwise comparison framework. Here, we focus on the Brazilian data. Specifically, we define our treatment and control conditions as follows:

Treatment condition: Participants assigned to fallback voting framework.

Control condition: Participants assigned to classical pairwise comparison framework.

Out of a total of 860 participants, 437 were in the treatment condition (50.8%), whereas 423 were in the control condition (49.2%).

Data Conversion

To compare data collected both from our treatment and control conditions, we convert the fallback voting data into issue pairs. The following part details the conversion and the assumptions considered.

Approval screen. For each participant, we assume that an approved issue will always be preferred over a disapproved one. We point out that this condition does not depend on displaying both issues on the same screen. For example, let A and B be two issues; whether the user approved issue A on the first screen and disapproved issue B on screen 1 or 20, we always consider that $A > B$. As a side note, we do not include abstentions since we can not make any inference from these preferences.

Ranking screen. Here, we follow a more strict rule. We assume that only the sorted issues belonging to the same screen can be converted into pairwise comparison data. That is, each screen of k issues is converted to $\frac{k(k-1)}{2}$ issue pairs. Even though this assumption is ambitious, given that our framework did not allow ties between issues, we address this point in a subsequent section over consistency and transitivity.

Data quality

Next, we evaluate the responses in terms of data quality. While defining quality is ambitious in online participation, since we can not judge the level of knowledge over issues or beliefs, for our purposes, we define quality as the extent to which reliable and real can be interpreted the preferences of a participant in an online system. We borrow some concepts from behavioral studies to the digital democracy arena. Specifically, we included the following metrics i.e.:

Interaction index As mentioned above, the platform randomly assigned the order of issues displayed in the rank screen. Participants then can make two possible decisions: accepting the proposal of sorted ranking or updating this screen through drag-and-drop (or up/down buttons). We define the interaction index as a boolean that represents whether the participant modified the rank proposed (1) or only accepted it (0).

Abstention ratio This measure represents the percentage of abstentions provided by a participant in an approval screen. Values range from 0 to 1, in which 0 represents approval/disapproval responses in all issues, whereas 1 represents abstention in all of them.

Response speed We measure the time that participants spend on each screen. First, we concatenated all milestones involved in the life cycle of participation. That is, we join data

rows from the consented form, approval screen, rank screen, and self-reported information popup. Then, we measure the time difference (in seconds) between each timestamp. While we expect a decay in response time, we consider unusually fast responders or extremely long response times to be proxies of inattention or, eventually, suspicious participation. We follow recommendations from the literature, setting 2 seconds per response as the cut-off for inattentive responses in the approval screen. This is converted into a cut-off of $2k$ seconds at screen level.

Longstring index This measure represents the number of identical contiguous responses over the same option provided in an approval screen. The values range from 1 to k , where k is the maximum number of issues per screen. We consider greater longstring values to be a signal of a lack of interest.

Due to the experiment design, participants were sometimes asked to order the same two issues in different rank screens. For validity purposes, we tested two measures, namely consistency and transitivity.

Consistency We measure the percentage of times participants consistently prefer the same issue over another in repeated rank screens. That is, from the rank $A > B > C$, we use $A > B$, $B > C$, and $A > C$ data points. It should be noted that this index is calculated uniquely for issue pairs compared more than once by the same participant.

Transitivity We measure the percentage of non-cyclic triplets derived from rank screens. To avoid over-representing transitive triplets derived from the same screen, we only use contiguous pairs. Going back to our example, from the rank $A > B > C$, we use $A > B$ and $B > C$ data points. The procedure we use to quantify the transitivity is as follows:

- We compute all the possible triplets. In the case of our French platform, the number of triplets for 120 issues is 280,840 (i.e., $\binom{120}{3}$). In the Brazilian platform, the number of triplets for 67 issues is 47,905.
- For each triplet and participant, we assess the percentage of cycles. To exemplify this, consider a situation where a participant has selected $A > B$ 5 times, $B > C$ 5 times, $A > C$ once, and $C > A$ once. Here, a total of 50 triplets are formed ($A > B > C$ 25 times and $B > C > A$ 25 times), leading to a transitivity of 50%.

4.3.4 Econometric models

We estimate whether the likelihood of early dropouts and the decrease in data quality can be associated with the nature of the content, the platform setup, and the demographic characteristics of participants. All models presented in this chapter correspond to logistic regressions.

Analysis 1: Divisive issues in approval screen

First, we assess the likelihood of early dropouts after responding to the first approval screen. Let p be a participant. Our dependent variable is a dummy that represents whether p dropped the platform (1) or continued her participation (0). Additionally, we control by characteristics associated with the screen that could lead to dropouts, such as the wording of issues (proxied by the number of characters), the inclusion of a divisive issue (as a dummy), the response speed, and the Longstring index. Our logistic regression model is set up as follows:

$$\text{logit}(y_p) = \beta_0 + \beta_1 \text{length}_p + \beta_2 \text{divisive}_p + \beta_3 \text{time}_p + \beta_4 \text{longstring}_p + \epsilon_p \quad (4.1)$$

Where y_p is our dependent variable, length_p is the average length of issues, divisive_p is a dummy variable to represent the inclusion of a divisive issue, time_p is the response time, longstring_p is the longstring index, and ϵ_p is the error term. Both the longstring index and time were standardized with a mean of zero and a standard deviation of one.

Unfortunately, this regression analysis can not take into account the socio-demographic characteristics of participants since, by platform design, we only have self-reported data from participants who continued responding.

Analysis 2: Likelihood of modifying proposed rank

Second, we assess the likelihood that a participant modifies the randomly assigned order of issues in the rank screen. Specifically, we aim to predict the interaction index. Let p be a participant and let s be the s -rank screen seen by a participant. The universe-fixed effect logistic regression model takes on the following form:

$$\text{logit}(y_{ps}) = \beta_1 \text{progress}_{ps} + \beta_2 X_p^T + \beta_3 \text{progress}_{ps} \cdot X_p^T + \gamma_u \quad (4.2)$$

Where y_{ps} refers to the likelihood that the rank screen is updated by p (1, the randomly assigned order of issues was modified; 0, otherwise), progress_{ps} refers to the percentage of completed progress by p , X refers to a vector that accounts for individual characteristics (Sex: 1, female; 0, male. Education: 1 less than undergraduate, 0 undergraduate or more. Age: 1, younger than the median; 0, older than median), $\text{progress}_{ps} \cdot X_p^T$ represents the interaction term between demographic features and the progress, and γ_u refers to the universe fixed-effect.

Analysis 3: Predicting abstentions in approval screen

Third, we predict the abstention over an issue i in approval screens. Let p be a participant and let s be the s -approval screen. Similarly to our second analysis, we set up an issue-fixed effect logistic regression as follows:

$$\text{logit}(y_{ips}) = \beta_1 \text{progress}_{ps} + \beta_2 X_p^T + \beta_3 \text{progress}_{ps} \cdot X_p^T + \gamma_i \quad (4.3)$$

Where y_{ips} is our dependent variable (1, p abstained in issue i in the s screen; 0, otherwise), progress_{ps} refers to the percentage of completed progress by p , X is a vector that accounts for the demographic characteristics of participants, $\text{progress}_{ps} \cdot X_p^T$ is the interaction term between demographic features and the progress, and γ_i represents the issue-fixed effect.

Final remarks

The interaction term is included in analyses 2 and 3 to estimate the relationship between demographic characteristics and progress in participation in terms of assessing whether these features could potentially accentuate careless behaviors.

4.3.5 Interplay of data quantity and quality

Our concluding analysis assesses the second part of Hypothesis 2 and Hypothesis 4 by computing the ranking of agreements using data collected via fallback voting and the classical pairwise comparison framework. We assume that whether the first elicitation method is suitable for increasing participant data, it should reach equivalent collective decisions to the ones captured by the classical framework. We introduce some notation. Let x_{ij} the number of total wins of issue i over j , m the number of issues. Next, we use adapted versions of Copeland and Borda for incomplete preferences, i.e.:

Borda score

The Borda score of an issue i for incomplete preferences represents the fraction of times that i was selected in the entire data set.

$$\text{Borda}_i = \frac{\sum_{j=1}^m x_{ij}}{\sum_{j=1}^m (x_{ij} + x_{ji})} \quad (4.4)$$

Borda_i ranges from 0 to 1. Values closer to 0 mean that i is not prioritized. Conversely, values closer to 1 mean that i reaches a unanimous agreement.

Copeland

This version of Copeland scores the issues according to the number of pairwise majority contests an issue wins, loses, and ties with respect to the other issues (Copeland, 1951). First, let d_{ij} be a dummy variable to represent whether i beats j .

$$d_{ij} = \begin{cases} 1 & \text{if } x_{ij} > x_{ji} \\ 1/2 & \text{if } x_{ij} = x_{ji} \\ 0 & \text{if } x_{ji} > x_{ij} \end{cases} \quad (4.5)$$

Next, the Copeland score of an issue i is given by:

$$\text{Copeland}_i = \frac{\sum_{j=1, i \neq j}^m d_{ij}}{m-1} \quad (4.6)$$

Copeland_i ranges from 0 to 1. Consider a data set without ties, if an issue has $\text{Copeland} = 0.6$, it indicates that it beats 60% of other individual issues. If there is a Condorcet winner (a given issue who beats each other in individual comparisons), the Copeland score for that issue will be 1.

4.4 Results

This section outlines our main findings. We split the analysis into the study of divisive issues and data quality/quantity.

4.4.1 Divisive issues

We start examining whether contentious issues hamper the life cycle of participants. Figure 4.5 a-d show the percentage of early dropouts from the French and Brazilian platforms in the first (a,c) and second approval screens (b,d). Using two measures for classifying divisive issues, we do not observe that participants who read a contentious issue have greater dropouts in comparison to the ones who did not read them. What is striking from these figures is that after the second approval screen in Brazil, this difference seems to be significant for one of our divisiveness measures, coinciding with the fact that this platform popped up a questionnaire asking for self-reported socio-demographic information. Perhaps at that moment, participants might realize that they were no longer completely anonymous, and they could tend to moderate their responses. To mitigate concerns with respect to this argument, we compare the issues' approval ratio of participants who filled out the questionnaire and the ones who did not fill it out, along with comparing these ratios before and after the popup. The results remain consistent, observing that none of the issues have a significant difference in their approval ratio (See Figure 4.5 e-f for the ten most significant differences in France and Brazil, respectively).

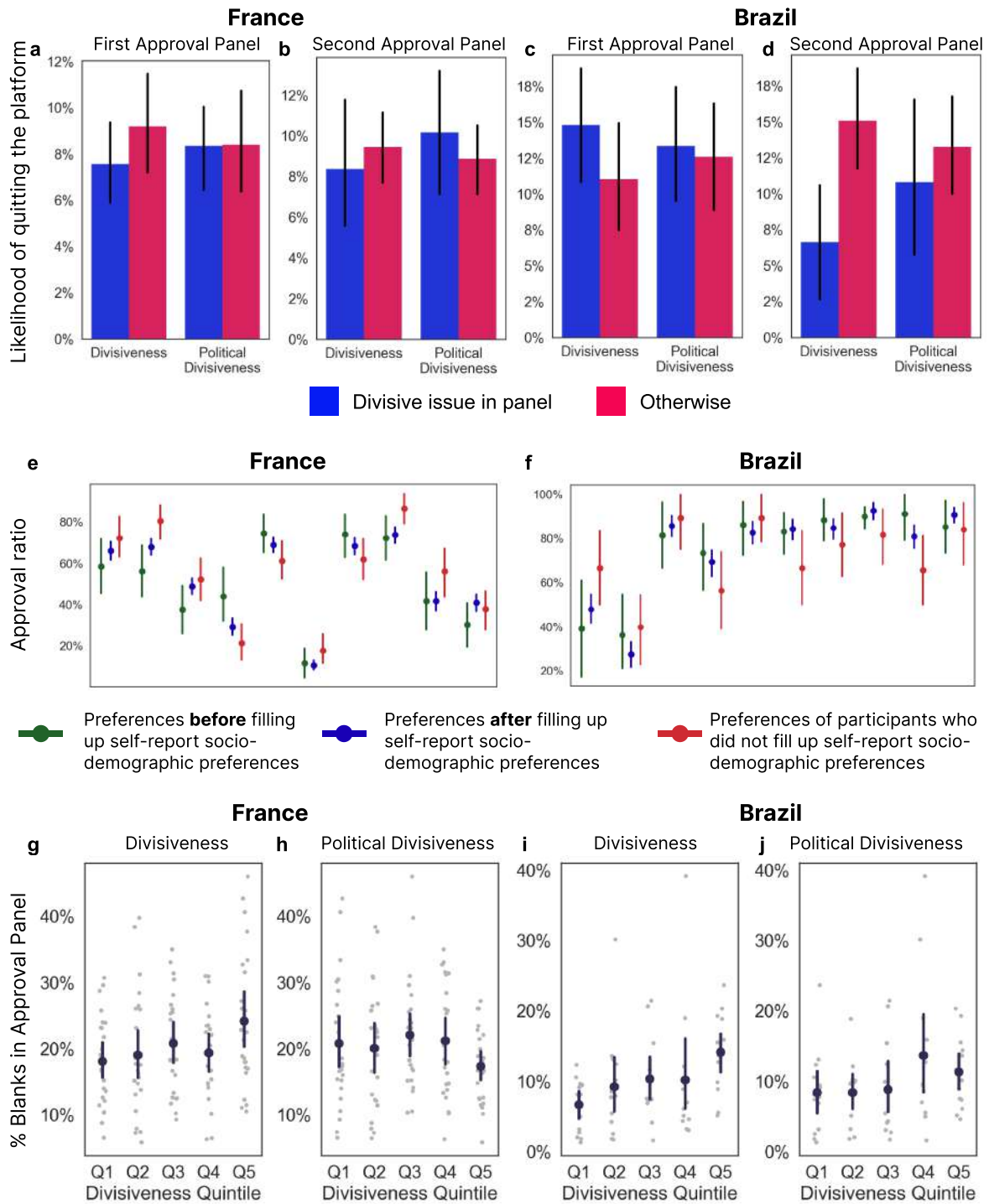


Figure 4.5: Impact of divisive issues in online participation. Percentage of dropouts after responding to the (a,c) first and second (b,d) approval screens in France and Brazil. The ten most significant differences in approval ratios before/after filling up self-reported socio-demographic preferences observed in (e) France and (f) Brazil. Percentage of blanks in approval screens using (g) Divisiveness and (h) Political Divisiveness in France (i-j for Brazil, respectively).

Nonetheless, it is still not clear whether endogenous characteristics of screens, such as the ballot size or issue length, can be associated with early dropouts. Then, we use our logistic regression defined in eq. 4.1. Table 4.1 summarizes the econometric analyses for

France (models 1-2) and Brazil (models 3-4). The results confirm our previous observation: divisive issues are not associated with the likelihood of early dropouts in any of the platforms. Surprisingly, while 91.6% of participants who accepted the consented form in Brazil continued to the second screen, we observe that an increase in one standard deviation of the average issue length reduces the odds of dropout by 83%. A non-causal explanation might be due that reading longer issues could lead to feeling more involved with the tool; nevertheless, this finding could not be reproduced for the French data.

Next, we provide evidence suggesting that the longstring index can be effectively interpreted as a synonym for careless participation. In the French platform, an increase of one standard deviation in this index is associated with a rise of 25% in the dropout odds. The reason why the Longstring index is not explanatory in Brazil could be related to its greater approval ratio (75%), being these effects indistinguishable from a genuine interest in providing reliable preferences with carelessness.

	Model 1	Model 2	Model 3	Model 4
Divisive Issue	-0.21 (0.19)	-0.22 (0.19)	0.34 (0.25)	0.38 (0.26)
Issue length		-0.08 (0.09)		-0.37*** (0.13)
Response speed		-0.22 (0.22)		-0.54 (1.10)
Longstring index		0.23*** (0.09)		0.16 (0.13)
Constant	-2.29*** (0.13)	-2.33*** (0.13)	-2.09*** (0.19)	-2.20*** (0.22)
McFadden R2	0.00	0.01	0.00	0.03
AIC	882.62	873.98	454.03	449.18
BIC	893.29	900.63	462.78	471.04
Log Likelihood	-439.31	-431.99	-225.02	-219.59
Deviance	878.62	863.98	450.03	439.18
Num. obs.	1529	1526	585	585

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 4.1: Logistic regression analysis for early dropouts (1, participant abandoned platform after first approval panel; 0, otherwise). Each model represents a multivariate regression. We standardized all non-boolean variables to have a mean of zero and a standard deviation of one.

To provide more validity to our previous findings, we examine another approach focused on abstention in approval screens. Whether participants feel discomfort giving their opinion over divisive issues, they should abstain from them at a greater rate in comparison to non-divisive ones. For this, we classify the issues into quartiles, where Q1 includes the 20% less divisive issues, and Q5 includes the 20% most divisive ones. Figure 4.5 g-j show that divisive issues have a slightly higher percentage of abstentions than non-divisive ones, though these differences are not statistically significant. Whereas Q5-issues were abstained in 22%, the Q1-issues were abstained 19% in France (Figure 4.5 e). In Brazil, we observe a marginal difference, showing that Q1-issues were abstained 7% in comparison with the 12% for Q5-issues (Figure 4.5 g). This finding is consistent with all our measures of divisiveness. We interpret that being an anonymous experiment and allowing abstention are necessary conditions to grasp preferences over agreements and disagreements. Furthermore, the fact that contentious issues

did not receive a significantly higher abstention suggests that face-to-face social pressures could not influence anonymous online participation. Specifically, we provide a sort of validation to our platform planning assumption that the “abstain” button could act as a buffer whether participants did not feel comfortable giving their preference. Regrettably, this study could not assess the latter beyond anecdotal results, calling for follow-up studies evaluating these metrics.

Overall, the results yield that anonymity under controlled conditions can be beneficial in terms of including content on contentious issues. Unfortunately, since our platforms did not collect data on participants who decided not to respond to the first approval screen after accepting the consent form, we can not claim any causation. Nonetheless, in light of the previous evidence, these results lead us to support hypothesis 1.

4.4.2 Data quantity

Having established that contentious issues did not condition early dropouts, we now examine the participation in terms of data quantity. As the Brazilian platform was limited to 30 pairwise comparison screens, we assess the time required to collect 30 issue pairs from a traditional pairwise comparison setup compared to fallback voting.

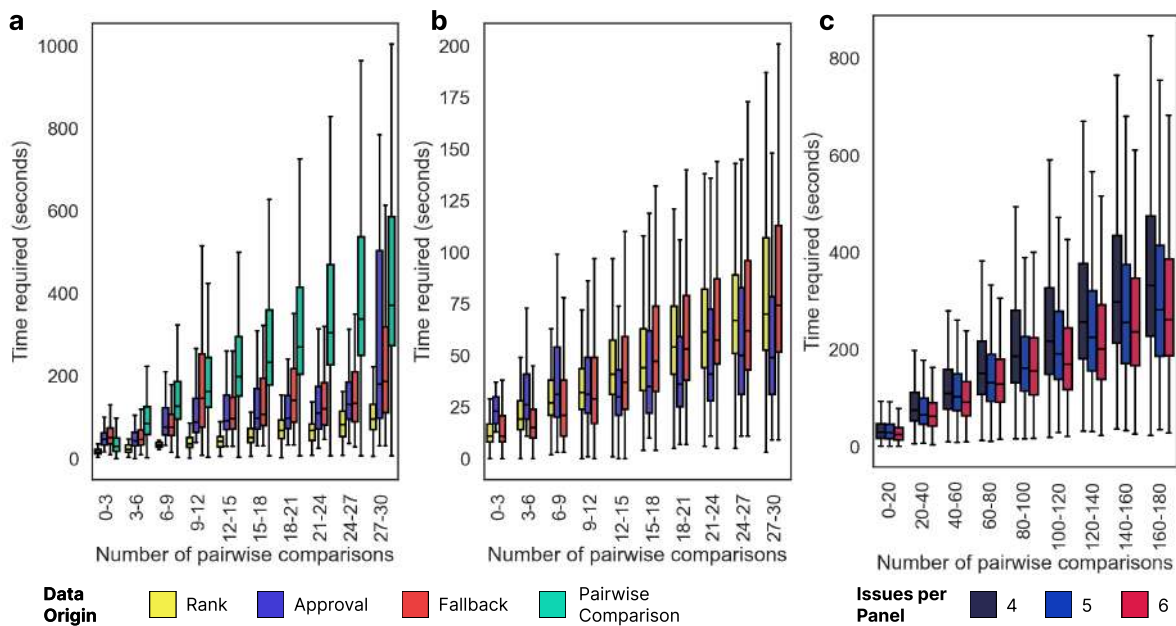


Figure 4.6: Comparison of data collected in (a) Brazil and (b) France. Comparison of data collected in France in terms of the number of issues included (c).

From Figure 4.6 a, we learn that pairwise comparison behaves similarly to fallback voting at early stages of participation, showing that for collecting 9-12 pairs, the first spend a median time of 163 seconds in comparison to 143 seconds for the latter. Nevertheless, as the participation extends for a longer period, we see significant differences that favor fallback voting. For example, collecting 27-30 issue pairs in pairwise comparison took a median time of 372 seconds in comparison to 187 seconds of fallback voting. Nonetheless, the rank screen was the most efficient method, which took a median of 95 seconds to collect 27-30 issue pairs. We interpret that the high approval ratio in Brazil (around 75%) conditions the issue

pairs collected by approval screens and, consequently, fallback ones. This view is grounded in the results observed in France (a platform with around 50% of approvals), showing that approval screens collected 27-30 pairs in a median time of 49 seconds in comparison to the ones collected by rank screens in a median of 70 seconds. What is striking here is that despite our pre-experiment beliefs, fallback voting does not seem to be as efficient as simply using rank screens, finding evidence that the most important condition for data collection with this framework is related to the expected approval ratio. Whether the platform planners expect a high percentage of approvals, our analysis suggests that traditional frameworks such as rank or pairwise comparison can perform better. However, in cases with a more balanced approval ratio, fallback voting and approval screens can better lead to the data collection. To discard any potential impact due to the screen size, we compare the data collected by Universe in the French platform. Figure 4.6 c shows that the number of issues per screen does not affect the collection of issue pairs. Of course, longer ballots could be challenging to follow online; nevertheless, we tackle these potential concerns in terms of data quality in the following section.

The results presented lead us to support the first part of hypothesis 2. We demonstrated that fallback voting successfully increased the number of issue pairs collected in a time slot. Furthermore, we point out that implementing this elicitation method is not enough to obtain more data from participants. As shown in Figure 4.6, an unbalanced ratio of approvals/disapprovals in approval screens can condition the amount of data collected—as in the case of the Brazilian platform.

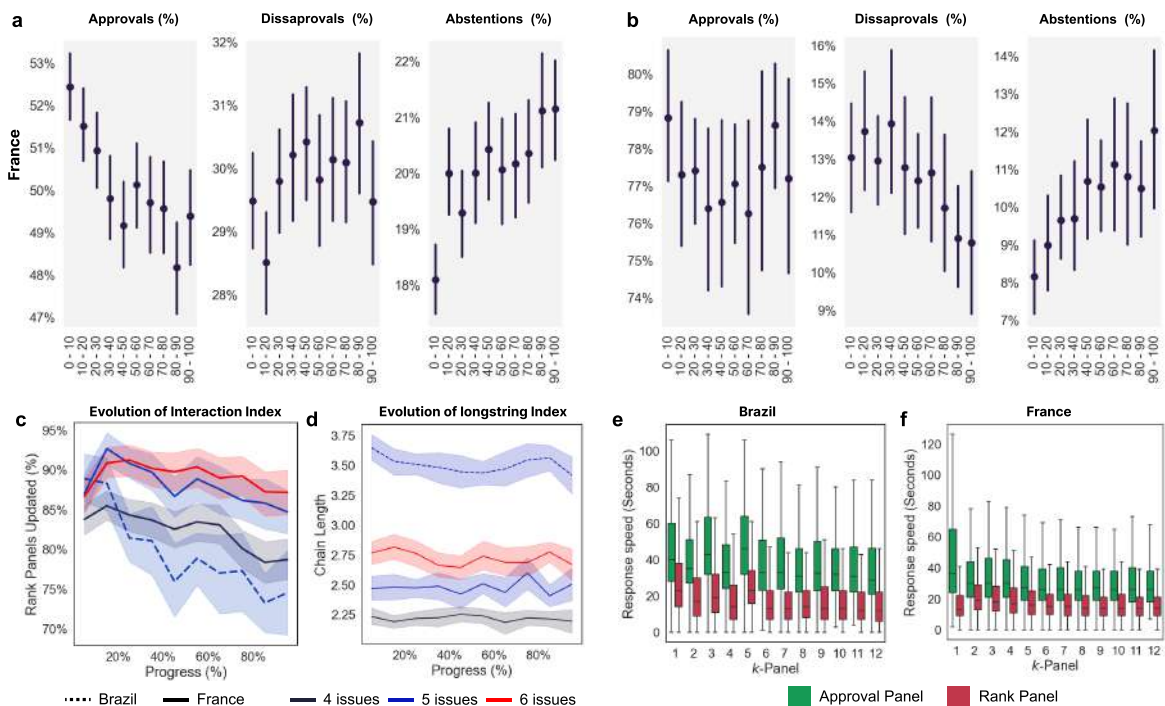


Figure 4.7: The nature of online participation.

4.4.3 Data quality

Next, we evaluate the nature of preferences in terms of data quality. To assess hypothesis 3, we use four approaches: the percentage of abstentions in approval screens, the average interaction index, the average longstring index, and response speed throughout the life cycle of participants.

Abstentions in approval screens Figure 4.7 a-b present the main findings in approval screens for France and Brazil. We measure the ratio of approvals, disapproval, and abstentions per screen as its percentage of selection over the total issues. Our baseline using all data is an approval ratio of 50% in France and 75% in Brazil. In France, we find that the approval ratio decreased from 53% on the first screen to 50% on the last one. Conversely, in Brazil, we observe a decrease in disapprovals, from 13% on the first screen to 11% on the last screen. Nevertheless, all these variations do not show significant differences between the experiment’s beginning and end (Figure 4.7 c). What is striking here is that both platforms showed an increase in the abstention ratio over time, finding an average increase of 2.6% of abstention rate every ten minutes of participation.

	Model 1	Model 2	Model 3	Model 4
Progress Bar	0.05*	0.17**	0.27***	0.60**
	(0.03)	(0.09)	(0.09)	(0.24)
Sex (1, female)	0.22***	0.23***	0.11**	0.07
	(0.02)	(0.04)	(0.05)	(0.10)
Education (1, less than undergraduate)	-0.17***	-0.19***	0.21***	0.02
	(0.03)	(0.05)	(0.06)	(0.11)
Age (1, younger)	0.05*	0.11**	-0.07	0.21
	(0.03)	(0.05)	(0.07)	(0.15)
Progress * Sex		-0.02		0.08
		(0.06)		(0.17)
Progress * Education		0.04		0.37*
		(0.09)		(0.20)
Progress * Age		-0.13		-0.53**
		(0.09)		(0.24)
Log Likelihood	-46977.84	-46976.62	-6252.23	-6247.92
Deviance	93955.69	93953.24	12504.45	12495.84
Num. obs.	98876	98876	21293	21293

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 4.2: Logistic regression analysis for the likelihood of abstentions (1, participant abstained; 0, otherwise). Each model represents a multivariate regression with issue-fixed effects. We standardized all non-boolean variables to have a mean of zero and a standard deviation of one.

To complement this analysis, Table 4.2 summarizes four logistic regression models to understand whether characteristics of participants condition abstentions. Models 1-2 correspond to French data, and models 3-4 correspond to Brazilian data. First, we find a positive relationship between the time participants spent on the platforms and the likelihood of abstaining from a preference. An increase of one percent in the progress bar is associated with an increase of 18.5% and 82.2% in the odds of abstaining from a preference in France ($P < 0.05$) and Brazil ($P < 0.05$), respectively. Second, we learn that self-reported female participants abstained

1.3 times more than male participants by looking at socio-demographic characteristics. We report that the educational attainment effect goes in different directions. Participants self-labeled as less than undergraduate abstained their opinion 15.6% less than those from more than undergraduate in the French platform. Conversely, this effect is not significant in the Brazilian data. Surprisingly, we found that demographic characteristics and the completion time have significant interaction effects on the Brazilian platform. We report that younger participants have a 41.1% greater likelihood of abstaining than older participants ($P < 0.05$), and participants with less than an undergraduate abstained 1.45 times more than those with at least an undergraduate degree over time.

Interaction Index From Figure 4.7 c, we learn that the average interaction with the rank screen decreases over time in all scenarios. While 93.2% of participants modified the proposed rank in the second screen in France (universe 5), this percentage fell almost ten points to 83.4% at the 24th rank screen. This finding is consistent in all our scenarios. Even though we observe that rank screens with 4 issues are less updated than the ones with 5-6, we interpret that having fewer issues could coincide with a higher probability of proposing the correct order rather than inattention or carelessness. To assess whether this decay in interaction is associated with the characteristics of participants, we perform the econometric analysis introduced in eq. 4.2.

	Model 1	Model 2	Model 3	Model 4
Progress Bar	-0.55*** (0.07)	-0.86*** (0.19)	-1.15*** (0.16)	-0.88** (0.43)
Sex (1, female)	0.28*** (0.05)	0.17* (0.10)	-0.19** (0.09)	-0.15 (0.19)
Age (1, young)	0.56*** (0.06)	0.41*** (0.12)	0.27** (0.12)	0.35 (0.26)
Education (1, less than undergraduate)	-0.22*** (0.06)	-0.32** (0.12)	-0.89*** (0.10)	-0.68*** (0.21)
Progress * Sex		0.23 (0.17)		-0.08 (0.32)
Progress * Education		0.19 (0.21)		-0.40 (0.35)
Progress * Age		0.28 (0.20)		-0.15 (0.43)
Log Likelihood	-7790.19	-7787.97	-1660.12	-1659.38
Deviance	15580.38	15575.93	3320.24	3318.75
Num. obs.	20177	20177	3423	3423

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 4.3: Logistic regression analysis for the likelihood of modifying proposed rank (1, rank updated; 0, otherwise). Each model represents a multivariate regression with universe-fixed effects.

Table 4.3 presents the findings for France (Models 1-2) and Brazil (Models 3-4). We describe the even models as our “final models”. We observe that an increase of one percent of completion is associated with a 57.7% of reduction in the odds of modifying the rank proposal in France and 58.5% in Brazil. In France, self-reported young participants were 1.75 times more likely than self-reported older participants to update the rank screen (1.43 times in Brazil; nonetheless, this value is not significant in the final model). Associations in the same

direction were observed for education, in which participants who self-reported studies level less than undergraduate was associated with a 19.7% reduction in the odds of modifying the proposed rank in France and 49.3% in Brazil. One reason why older participants are less prone to modify the proposed rank might be because the platform design is not as clear and intuitive as we believed. The use of drag-and-drop, without additional explanations than the dashed box, and the addition of up/down buttons for each issue seem not to be appropriate when we expect to reach older audiences. Perhaps, not modifying the proposed rank could not be addressed in terms of data quality, but just the lack of intuitiveness in the design from our side.

Longstring Index Figure 4.7 e shows the longstring index over time. We observe that the average longstring remains constant throughout the participation cycle, not finding evidence that supports our hypothesis 3 for this measure. Again, the fact that the Brazilian platform reported a majoritarian approval ratio can explain why this index is not suitable for measuring carelessness since almost everyone has a longstring value of 3 out of a maximum of 4 in each panel. In this sense, we call for studying this index in further studies focused on approval ballots.

Response speed Figure 4.7 e-f show the response time per screen. Here we report the findings for universes of 5 issues. Overall, we observe that rank screens require less time to be responded in comparison to approval screens. Participants took a median time of 36.5 seconds to respond to the first approval screen compared to a median time of 14 seconds for the first rank screen in France. Conversely, in Brazil, participants took a median time of 40 seconds to respond to the first approval screen compared to a median time of 32 seconds for the first rank screen. At the end of the experiment in France, on the 24th screen, participants took a median time of 24 seconds in comparison to a median time of 15 seconds.

While the French platform reported a clear difference with respect to the Brazilian one in the response speed of approval vs. rank screens, we associate this difference with the intertwined of both screens. Again, the Brazilian platform first asked for all approval screens and then for rank screens, contrasting with the approach adopted in France. In light of the evidence, it seems reasonable to think that intertwining screens contribute to reducing the cognitive burden of participants. One possible interpretation considers that participants “save” time reading the issues in approval screens and then retain this information on the next screen, which was not the case in the Brazilian platform.

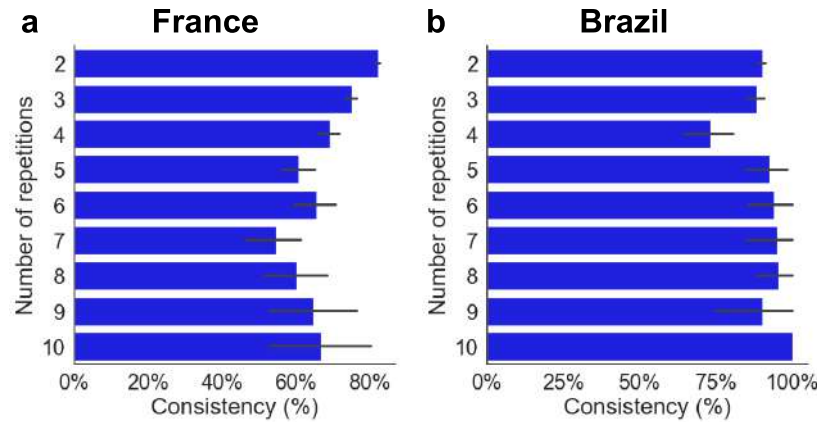


Figure 4.8: Evaluation of consistency in (a) France and (b) Brazil.

Consistency In France, around 8 out of 10 participants consistently ranked two issues in the same order when they had to respond twice (82.2%), and this value decreased to an average of 79.4% by assessing all data (Figure 4.8). In Brazil, the consistency is greater than the one observed in France. Here we find an average consistency of 89.9%. Greater levels of consistency for participants responding in the same order even after asking 8 or 10 times observed in the figure seems to be more suspicious rather than a clear consistency in beliefs. Again, participants only ranked issues that they previously approved on. We interpret that asking for a larger number of issues in France (and consequently a larger number of screens) could partially explain the lower consistency.

By evaluating consistency at participant-level, 87.9% of participants in France and 97.5% in Brazil showed a consistency higher than 95%. In other words, a reduced number of participants provided most of the inconsistencies in our data.

Transitivity We report a transitivity of 74.6% for the Brazilian data and 82.6% for the French one. By uniquely looking at the triplets that include at least one divisive issue, we find that the transitivity in Brazil decreases to 71.3% and increases to 83.7% in France.

To sum up, some of our indices show an increased carelessness from participants over time. For example, the fact that participants decreased by 10% the interaction in rank screens could be seen as a promising finding in terms of carelessness. Furthermore, the values for consistency and transitivity show that despite following strict assumptions for measuring them, we still obtained relatively high levels of these metrics. Nevertheless, we can not discard that a non-clear design for older participants and the fatigue of staying over 15 minutes on the same website could cause this decrease. In this sense, we find mixed findings for our hypothesis 3.

4.4.4 Interaction data quantity/quality

Our previous analysis yields that some quality metrics for responses could decrease over time and that the elicitation method dramatically increases the data collection. Nevertheless, none of our analyses assessed the interplay of elicitation methods, data quantity, and carelessness. Specifically, how does this interaction affect collective decisions? To test this question, we

compared the ranking of preferences using an adapted version of Copeland and Borda for incomplete data. Since we previously demonstrated that we could decompose fallback voting data into issue pairs, we set the classical pairwise comparison framework as our “final ranking”. Here, we describe the findings from the aggregation using Copeland.

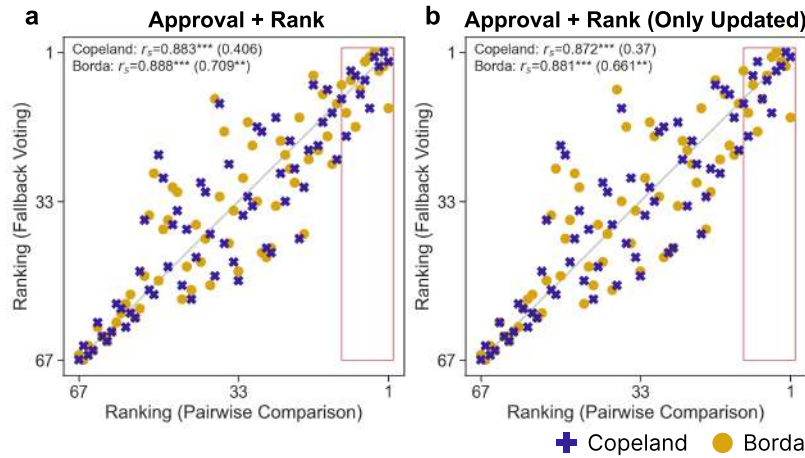


Figure 4.9: The nature of pairwise comparison data. Association Ranking of agreements using pairwise comparison data and (a) rank, (b) rank, only updated, (c), approval and rank, and (d) approval and rank, only updated. Parentheses values represent the Spearman correlation for the top-10 issues from the ranking using pairwise comparison data.

We report the Spearman correlation between the ranking of all issues (r_s) and the top-10 issues from the pairwise comparison scenario ($r_{s,top}$). Figure 4.9 a shows that uniquely using data from rank screens is strongly correlated to our final ranking ($r_s = 0.883$, $P < 0.01$) though the top-10 ranked issues are slightly correlated ($r_{s,top} = 0.406$). Nevertheless, by simply removing non-updated rank screens (a measure for data quality), figure 4.9 b shows that the correlation marginally decreased to 0.872 ($P < 0.01$), but we continue seeing slightly correlated values for the top-ranked ($r_{s,top} = 0.37$).

We previously demonstrated that fallback voting dramatically increased data collection (See hypothesis 2); here, we show this efficiency is not accompanied by reaching the same collective decisions and the mild correlations using samples of the data from rank screen. Even though the exclusion of potential careless participation—proxied by the interaction index—slightly increases the correlation, the evidence does not seem enough to reach collective decisions similar to the ones obtained with the traditional framework. In this sense, we support hypothesis 4.

4.5 Discussion

Our analysis of data collected through two real-world digital democracy systems offers an outlook on how the content and user interface condition online involvement in terms of early dropouts, data quantity, and careless participation. Participants in anonymous environments were not likelier to drop out after reading about a divisive issue than those who did not read them, nor did they omit their opinion on them in a greater percentage. Overall, both platforms showed that participants slightly increased their odds of abstaining in approval screens and decreased interaction odds in rank screens over time, raising concerns about potential fatigue or carelessness.

We subsequently show that the decrease in carelessness is especially relevant for rank screens, showing that participants reduce by around 10 percent the interaction with the proposed random order of issues over time. We found that participants tend to be consistent throughout screens. Even though the response speed slightly decreased over the experiment, we assume that this could be caused due that individuals “learned” about the platform usage rather than being exclusively associated with carelessness. This is reaffirmed by the fact that only 10 participants (0.45 %) could be categorized as fast-speed responders in both platforms, giving evidence that response time is not enough to argue detriment in data quality.

Our results raise that although all scenarios tested aimed to collect an equivalent data set to pairwise comparison, the outcome derived both in terms of data collection and collective decisions can vary dramatically. We demonstrated that fallback voting showed impressive results in maximizing the amount of data collected. Nevertheless, the ranking of preferences (a metric of quality) using fallback voting reaches mild correlations compared to the one with data obtained by the traditional pairwise comparison framework. Altogether, we suggest that the inclusion of guidelines from behavioral studies is necessary for the study and creation of digital democracy systems. For instance, phrasing engaging introductions, including attention checks, and implementing heuristics to strategically display the order of preferences can contribute to increasing online involvement.

Nevertheless, the debate about removing careless participants from the data sample in digital democracy systems must not be seen only from a statistical perspective. Eliminating preferences from these participants can lead to non-representative samples and detriment external validity. For instance, we report carelessness associated with certain demographic characteristics, such as age and educational attainment. Can we consider excluding participants merely flagged as suspicious as an appropriate method to examine citizen participation? To further exemplify this point, the fact that female participants abstained more than male participants is not a reason for carelessness; instead, it could be simply related to not feeling comfortable giving an opinion about the topic. Therefore, removing abstentions would lead to the detriment of female representation, which, as mentioned in Chapter 2, is already minoritarian.

We hope that this study contributes to expanding our understanding of digital participation and how the medium could shape collective decisions. We consider that our findings are relevant for platform planners interested in creating consensus-oriented approaches.

Conclusions and Perspectives

Has the electorate become more polarized than in the past? This naïve question has been a driving force during the preparation of this thesis. While we confess that this question holds more complexity than it looked at three years ago, this dissertation aims to shed some light on how recent computational advancements can improve our understanding of citizen division and polarization. We hope to have convinced the readers that our methodological contributions are robust and useful for the literature to map divisive issues from online participation, measure the polarization that emerges from voting patterns, and explore carelessness in digital involvement.

Chapter 2 presented the main findings of two real-world digital democracy systems released in France and Brazil preceding their respective 2022 presidential elections. Motivated by related efforts on digital platforms, our platforms aimed to create a collaborative government program between participants. We used data collected by our systems as an opportunity to study agreements and disagreements in electoral times. We found that the main agreements did not seem to align with the topics most discussed by political candidates. Grounded in the literature, this led us to examine whether divisive issues are better predictors of political action than agreements. To this end, we calculated a ranking of divisive issues according to differences in issues' scores from six socio-demographic characteristics (political orientation, sex, age, location, zone, and educational attainment). Here we learned that political orientation is the one that explains the most divisions, followed by sex and age. Nevertheless, this approach is not generalizable and is biased toward the dimensions included in the study. This is relevant since we did not cover other potentially significant dimensions in Brazil, such as religious belief or race. This led to conceptualizing an agnostic approach that we call *Divisiveness*, which aims to map divisive issues for any aggregation function based on pairwise comparison. We found that the issues identified by our method could be traced back to several dimensions. For instance, a divisive topic identified in France was “Restoration of border control by France leaving the Schengen agreements,” which is mostly explained by political orientation and educational attainment. Finally, by using data collected by our platforms, synthetic data, and benchmark data sets from social choice, we subsequently showed that the ranking of Divisiveness is uncorrelated to its corresponding aggregation function, it took more issue pairs than the ranking of agreements to converge and could be studied following axioms from social choice.

Chapter 3 assessed how voting patterns on Election Days can serve to unravel layers of polarization. Inspired by theoretical measurements from economics and the active debate from political science, we formalize the concept of Election Polarization as a form to quantify citizen divisions on Election Day. This measure can be decomposed by competitiveness between candidates (Between-EP) and voting dispersion throughout a geography (Within-EP). To compute EP, we propose a data-driven and agnostic approach that relies exclusively on election data, and it can be generalized for any aggregation level, number of candidates, and election

type. Using synthetic elections and presidential election data from France (2002-2022), Chile (2013-2021), and the United States (2000-2020), we show that our approach successfully identified theoretical expectations of “polarized” elections, indicating that these effects mostly speak to Within-EP. Next, we confronted our findings with a measure of political polarization and political interest using data collected by electoral studies in the United States, showing that state-level Within-EP was positively associated with political polarization and Between-EP was positively associated with political interest (proxied by voting turnout). Altogether, we can define that an election is polarized by simply knowing whether two conditions are satisfied: high voting turnout and preferences highly geographically clustered.

In **Chapter 4** we empirically examined the data collected by the platforms presented in Chapter 2 in terms of how the content and interface conditioned user participation. Our assumption is that an anonymous environment under controlled conditions is beneficial to understanding citizen disagreements. We reported that participants did not drop out of the platforms after reading a contentious issue compared to the ones who did not read any of them. Then, borrowing some definitions from behavioral studies in psychology and social choice, we observed that careless participation tended to increase over time. For example, the abstentions increased at a rate of around 3% per 10 minutes, and the interaction with rank screens decreased by 10% over time. While we reported that some demographic characteristics of participants were significant in explaining this carelessness, we stated that this might not be a synonym for careless participation. For instance, female participants were more likely to abstain in comparison to male participants, and older participants were more likely to accept the random proposition in rank screens. Removing data associated with these features would lead to under-representing these groups from the collective decisions. Finally, we demonstrated that the data collection using Fallback voting could dramatically increase the number of issue pairs collected; nevertheless, this is not enough to reach the same decisions derived from the classical pairwise comparison setup.

While we claim that our main contributions are mainly empirical, we recognize that one of the most significant issues in the literature concerns access to reliable and standardized data sets to study citizen preferences. Hence, as part of our contribution, we deposited around two million pairwise comparison preferences collected by our platforms and over 10 million rows of curated election outcomes from Chile (2013-2021) and France (2002-2020). These data sets are deposited on Harvard Dataverse and GitHub. Furthermore, we released on GitHub two foundational libraries for the aggregation of preferences in Python (Comchoice: <https://github.com/CenterForCollectiveLearning/comchoice>) and the releasement of customized digital democracy systems (Opencracia: <https://github.com/CenterForCollectiveLearning/opencracia>).

Perspectives This dissertation aimed at the interplay of both computational and social perspectives. On the one hand, the computational perspective refers to our challenges in creating and deploying online tools and analyzing massive data sets. Of course, releasing a digital democracy system is not just coding a front-end with Javascript. It involved hundreds of hours iterating designs, preferring using GDPR-compliant libraries, configuring web servers to mitigate cyber attacks, hashing identifiable features from users (e.g., IP address), enabling server-side cache, and hosting our code in servers from the University, to name a few. Studying these steps separately makes no sense; however, we guarantee that all of them were crucial for successfully launching our platforms in France and Brazil. On the other hand, the social perspective refers to our decisions based on ethics and social dilemmas intertwined with

the computational extent. For example, from past experience releasing digital democracy systems, this dissertation decided not to allow participants to include their own issues in our platforms, avoid collecting sensitive information, and not recommending a presidential candidate as an output. Even though we recognize that “recommending a match” with a presidential candidate could significantly increase participation and data collection, we were unsure whether our recommendations could potentially bias the results. In this case, we consider that not having proper studies of recommender systems could alter participation, especially sensitive in contexts of political polarization reported in France and Brazil.

Ethical challenges for practitioners Our work presented a few examples of how digital democracy systems can contribute to a better understanding of our society. Despite this, we flag some ethical concerns that platform planners and digital democracy practitioners should consider. The fact that real-world digital democracy systems are released and maintained by privates raises a potential conflict of interest that can be detrimental to people’s trust in these technologies. We recall the Italian political party Five Star Movement (5-S) to exemplify this concern. This party, born with the promise to empower the population by making collective decisions through a digital platform maintained by the party (Treré & Barassi, 2015), has shown that facts are far from fulfilling this altruistic promise. Whether 5-S leaders want to push an idea in a certain direction, e.g., anti-immigration policies, they manipulate the wording of questions to obtain the desired results. Furthermore, the data processing is not transparent or auditable, noting that the company—owned by the party’s founder—manages the server and database. This dissertation indicates that Ethical Boards are crucial in establishing boundaries to conduct research using online platforms. Nevertheless, private organizations are not under this regulatory process. The lessons learned warn us of potential misbehavior that can be mitigated by adopting some safeguards from academia, such as avoiding collecting personal information, offering transparent approaches, establishing explicit purposes of the platforms, and following consensus-oriented frameworks, to name a few.

Future Perspectives Here, we address some potential future steps underlying our findings. First, we consider the study of regional-level election polarization worldwide and its spatial and temporal dynamics, both for presidential and parliamentary systems, to be relevant and timely. Preliminary analyses using data from European countries not included in this dissertation suggest that the election polarization spread might follow a virus pattern; nevertheless, this still needs further validation and follow-up studies. Another potential path is performing a behavioral study for Divisiveness. Specifically, an interesting question not addressed by our work is the evaluation of the “second momentum” postulate. Whether Divisiveness is reversible to its aggregation function, data collected from an experiment asking “Which issue do you consider more divisive?” by being aggregated by our metric of Divisiveness should be equivalent to the ranking of agreements of an experiment asking for “Which issue do you prefer?”.

Undoubtedly, our final remarks address that our methodological contributions from Chapters 2 and 3 can be enriched with a theoretical perspective. For instance, our following-up work from Divisiveness provides a theoretical formalization that considers the population size and studies some of its properties, such as manipulation and robustness (Colley et al., 2023). Thus, a recent and interesting work from Faliszewski et al., 2023 designs theoretical measurements to map polarization using elections, following a similar motivation to the one introduced in Chapter 3. Taken together, these recent works coming out from different

disciplines beyond political science show a promising interest in characterizing polarization. Finally, we hope to continue advancing our understanding of the dynamics of polarization by relying on the results presented in this dissertation.

Bibliography

- Abramowitz, A., & McCoy, J. (2019). United States: Racial resentment, negative partisanship, and polarization in Trump's America. *The ANNALS of the American Academy of Political and Social Science*, *681*(1), 137–156.
- Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, *70*(2), 542–555.
- Abrams, S. J., & Fiorina, M. P. (2012). “The big sort” that wasn't: A skeptical reexamination. *PS: Political Science & Politics*, *45*(2), 203–210.
- ACS. (2023). American Community Survey (ACS) [(Accessed on 05/10/2023)].
- Ahn, H. (2022). *Three essays on political polarization in the United States* (Doctoral dissertation). University of Missouri–Columbia.
- Alcalde-Unzu, J., & Vorsatz, M. (2013). Measuring the cohesiveness of preferences: an axiomatic analysis. *Social Choice and Welfare*, *41*, 965–988.
- ANES. (2021). ANES 2020 Time Series Study Full Release [dataset and documentation]. July 19, 2021 version [(Accessed on 04/17/2023)].
- Ansolabehere, S., Puy, M. S., et al. (2015). *Issue-salience, Issue-divisiveness and Voting Decisions* (tech. rep.).
- Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of political economy*, *58*(4), 328–346.
- Ash, E., Morelli, M., & Van Weelden, R. (2017). Elections and divisiveness: Theory and evidence. *The Journal of Politics*, *79*(4), 1268–1285.
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64.
- Baker, S. R., Baksy, A., Bloom, N., Davis, S. J., & Rodden, J. A. (2020). *Elections, political polarization, and economic uncertainty* (tech. rep.). National Bureau of Economic Research.
- Baltz, S., Agadjanian, A., Chin, D., Curiel, J., DeLuca, K., Dunham, J., Miranda, J., Phillips, C. H., Uhlman, A., Wimpy, C., et al. (2022). American election results at the precinct level. *Scientific Data*, *9*(1), 651.
- Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political analysis*, *23*(1), 76–91.
- Barberá, P., Casas, A., Nagler, J., Egan, P. J., Bonneau, R., Jost, J. T., & Tucker, J. A. (2019). Who leads? Who follows? Measuring issue attention and agenda setting by legislators and the mass public using social media data. *American Political Science Review*, *113*(4), 883–901.
- Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, *26*(10), 1531–1542.
- BEA. (2023). Interactive Data Application — U.S. Bureau of Economic Analysis (BEA) [(Accessed on 03/30/2023)].
- Benade, G., Itzhak, N., Shah, N., Procaccia, A. D., & Gal, Y. (2018). Efficiency and usability of participatory budgeting methods.
- Benade, G., Nath, S., Procaccia, A. D., & Shah, N. (2021). Preference elicitation for participatory budgeting. *Management Science*, *67*(5), 2813–2827.
- Berry, D. T., Wetter, M. W., Baer, R. A., Larsen, L., Clark, C., & Monroe, K. (1992). MMPI-2 random responding indices: Validation using a self-report methodology. *Psychological assessment*, *4*(3), 340.

- Best, S. J., & Krueger, B. S. (2005). Analyzing the representativeness of Internet political participation. *Political Behavior*, *27*, 183–216.
- Bishop, B. (2008). *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt.
- Borda, J. d. (1784). Mémoire sur les élections au scrutin. *Histoire de l'Académie Royale des Sciences pour 1781 (Paris, 1784)*.
- Bordignon, M., Nannicini, T., & Tabellini, G. (2016). Moderating political extremism: single round versus runoff elections under plurality rule. *American Economic Review*, *106*(8), 2349–2370.
- Boudou, J., Colley, R., & Grandi, U. (2022). Itero: An Online Iterative Voting Application (demo track). *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI 2022)*, 5896–5899.
- Bowling, N. A., Gibson, A. M., Houpt, J. W., & Brower, C. K. (2021). Will the questions ever end? Person-level increases in careless responding during questionnaire completion. *Organizational Research Methods*, *24*(4), 718–738.
- Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. *Journal of Personality and Social Psychology*, *111*(2), 218.
- Brams, S. J., & Sanver, M. R. (2009). Voting systems that combine approval and preference. *The mathematics of preference, choice and order: essays in honor of Peter C. Fishburn*, 215–237.
- Brandt, F., Conitzer, V., & Endriss, U. (2012). Computational social choice. *Multiagent systems*, *2*, 213–284.
- Brill, M. (2021). From Computational Social Choice to Digital Democracy. *30th International Joint Conference on Artificial Intelligence (IJCAI 2021)*, 4937–4939.
- Brown, D. S., Touchton, M., & Whitford, A. (2011). Political polarization as a constraint on corruption: A cross-national comparison. *World Development*, *39*(9), 1516–1529.
- Cabannes, Y. (2004). Participatory budgeting: a significant contribution to participatory democracy. *Environment and urbanization*, *16*(1), 27–46.
- Cambre, J., Klemmer, S., & Kulkarni, C. (2018). Juxtapeer: Comparative peer review yields higher quality feedback and promotes deeper reflection. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13.
- Can, B., Ozkes, A. I., & Storcken, T. (2015). Measuring polarization in preferences. *Mathematical Social Sciences*, *78*, 76–79.
- Cantador, I., Cortés-Cediel, M. E., & Fernández, M. (2020). Exploiting Open Data to analyze discussion and controversy in online citizen participation. *Information Processing & Management*, *57*(5), 102301.
- Casal Bértoa, F., & Rama, J. (2021). Polarization: what do we know and what can we do about it? *Frontiers in Political Science*, 56.
- Census. (2023). State area measurements and internal point coordinates [(Accessed on 05/02/2023)].
- Census.gov. (2023). Historical Income Tables: States [(Accessed on 03/30/2023)].
- Charron, N., Lapuente, V., & Rodríguez-Pose, A. (2022). Uncooperative society, uncooperative politics or both? Trust, polarization, populism and COVID-19 deaths across European regions. *European Journal of Political Research*.
- Chen, Z., & Berger, J. (2013). When, why, and how controversy causes conversation. *Journal of Consumer Research*, *40*(3), 580–593.
- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022). Social capital I: measurement and associations with economic mobility. *Nature*, *608*(7921), 108–121.

- Chomel, V., Cuvellé-Magar, N., Panahi, M., & Chavalarias, D. (2022). Polarization identification on multiple timescale using representation learning on temporal graphs in Eulerian description. *NeurIPS 2022 Temporal Graph Learning Workshop*.
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9).
- Colley, R., Grandi, U., Hidalgo, C., Macedo, M., & Navarrete, C. (2023). Measuring and Controlling Divisiveness in Rank Aggregation. *arXiv preprint arXiv:2306.08511*.
- Condorcet, N. d. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. de l'Imprimerie Royale.
- Congge, U., Guillamón, M.-D., Nurmandi, A., Sihidi, I. T., et al. (2023). Digital democracy: A systematic literature review. *Frontiers in Political Science*, 5, 4.
- Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J. .-, Sanchez, A., et al. (2012). Manifesto of computational social science. *The European Physical Journal Special Topics*, 214, 325–346.
- Copeland, A. H. (1951). *A reasonable social welfare function* (tech. rep.). Mimeo, University of Michigan USA.
- Dalton, R. J. (2008). The quantity and the quality of party systems: Party system polarization, its measurement, and its consequences. *Comparative political studies*, 41(7), 899–920.
- De Condorcet, N. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Cambridge University Press.
- DiMaggio, P., Evans, J., & Bryson, B. (1996). Have American's social attitudes become more polarized? *American journal of Sociology*, 102(3), 690–755.
- Duclos, J.-Y., Esteban, J., & Ray, D. (2004). Polarization: concepts, measurement, estimation. *Econometrica*, 72(6), 1737–1772.
- Durand, C., Blais, A., & Larochelle, M. (2004). The polls in the 2002 French presidential election: An autopsy. *Public Opinion Quarterly*, 68(4), 602–622.
- Edelmann, A., Wolff, T., Montagne, D., & Bail, C. A. (2020). Computational social science and sociology. *Annual Review of Sociology*, 46, 61–81.
- Esteban, J., Mayoral, L., & Ray, D. (2012). Ethnicity and conflict: Theory and facts. *Science*, 336(6083), 858–865.
- Esteban, J., & Ray, D. (2011). Linking conflict to inequality and polarization. *American Economic Review*, 101(4), 1345–1374.
- Esteban, J.-M., & Ray, D. (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, 819–851.
- Fábrega, J., González, J., & Lindh, J. (2018). Polarization and electoral incentives: The end of the Chilean consensus democracy, 1990–2014. *Latin American Politics and Society*, 60(4), 49–68.
- Faliszewski, P., Kaczmarczyk, A., Sornat, K., Szufa, S., & Was, T. (2023). Diversity, Agreement, and Polarization in Elections. *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023)*.
- Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annu. Rev. Polit. Sci.*, 11, 563–588.
- Fishkin, J., Garg, N., Gelauff, L., Goel, A., Munagala, K., Sakshuwong, S., Siu, A., & Yandamuri, S. (2018). Deliberative democracy with the Online Deliberation platform. *The 7th AAAI Conference on Human Computation and Crowdsourcing*.
- Fishkin, J., Siu, A., Diamond, L., & Bradburn, N. (2021). Is deliberation an antidote to extreme partisan polarization? Reflections on “America in one room”. *American Political Science Review*, 115(4), 1464–1481.

- Frank, M. W. (2023). The U.S. Income Inequality Page of Mark W. Frank [(Accessed on 03/30/2023)].
- Fukuyama, F. (2021). Making the Internet Safe for Democracy. *Journal of Democracy*, 32(2), 37–44.
- Garzia, D., & Marschall, S. (2012). Voting advice applications under review: the state of research. *International Journal of Electronic Governance*, 5(3-4), 203–222.
- Garzia, D., & Marschall, S. (2016). Research on voting advice applications: state of the art and future directions. *Policy & internet*, 8(4), 376–390.
- Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences*, 114(50), 13108–13113.
- Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., & Hill, S. J. (2013). Who wants to discuss vote choices with others? Polarization in preferences for deliberation. *Public opinion quarterly*, 77(2), 474–496.
- Gibbard, A. (1973). Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, 587–601.
- González-Bailón, S., Borge-Holthoefer, J., Rivero, A., & Moreno, Y. (2011). The dynamics of protest recruitment through an online network. *Scientific reports*, 1(1), 1–7.
- Google. (2023). reCAPTCHA v3 — Google for Developers [(Accessed on 06/22/2023)].
- Hacker, K. L., & van Dijk, J. (2000). *Digital democracy: Issues of theory and practice*. Sage.
- Hanna, A., Wells, C., Maurer, P., Friedland, L., Shah, D., & Matthes, J. (2013). Partisan alignments and political polarization online: A computational approach to understanding the French and US presidential elections. *Proceedings of the 2nd Workshop on Politics, Elections and Data*, 15–22.
- Hashemi, V., & Endriss, U. (2014). Measuring diversity of preferences in a group. *ECAI*, 423–428.
- Helbing, D., Mahajan, S., Fricker, R. H., Musso, A., Hausladen, C. I., Carissimo, C., Carpentras, D., Stockinger, E., Sanchez-Vaquerizo, J. A., Yang, J. C., et al. (2023). Democracy by Design: Perspectives for digitally assisted, participatory upgrades of society. *Journal of Computational Science*, 102061.
- Helbing, D., & Pournaras, E. (2015). Society: Build digital democracy. *Nature*, 527(7576), 33–34.
- Hidalgo, C. A. (2019). Chile: web poll sifts policies amid riot, rallies and curfews. *Nature*, 575(7783), 443–444.
- Hoban, B. (2016). Why are swing states important? [(Accessed on 04/19/2023)].
- Hoehe, M. R., & Thibaut, F. (2022). Going digital: how technology use may influence human brains and behavior.
- Hoffmann, J. P., & Miller, A. S. (1998). Denominational influences on socially divisive issues: Polarization or continuity? *Journal for the Scientific Study of Religion*, 528–546.
- Hohmann, M., Devriendt, K., & Coscia, M. (2023). Quantifying ideological polarization on a network using generalized Euclidean distance. *Science Advances*, 9(9), eabq2044.
- Hsiao, Y.-T., Lin, S.-Y., Tang, A., Narayanan, D., & Sarahe, C. (2018). vTaiwan: An empirical study of open consultation process in Taiwan.
- Huckfeldt, R., Johnson, P. E., & Sprague, J. (2002). Political environments, political dynamics, and the survival of disagreement. *Journal of Politics*, 64(1), 1–21.
- Insee. (2023a). Niveau de vie et pauvreté par région Revenus et patrimoine des ménages — Insee [(Accessed on 03/30/2023)].
- Insee. (2023b). Structure et distribution des revenus, inégalité des niveaux de vie en 2020 — Insee [(Accessed on 03/30/2023)].

- Insée. (2023c). Toujours plus d'habitants dans les unités urbaines - Insee Focus - 210 [(Accessed on 03/30/2023)].
- InterieurGouv. (2023). Les résultats de toutes les élections — Les élections en France [(Accessed on 04/16/2023)].
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology social identity perspective on polarization. *Public opinion quarterly*, 76(3), 405–431.
- Jacobson, G. C. (2017). The Triumph of polarized partisanship in 2016: Donald Trump's improbable victory. *Political Science Quarterly*, 132(1), 9–41.
- Jerit, J., & Barabas, J. (2012). Partisan perceptual bias and the information environment. *The Journal of Politics*, 74(3), 672–684.
- Johnston, R., Manley, D., & Jones, K. (2016). Spatial polarization of presidential voting in the United States, 1992–2012: the “big sort” revisited. *Annals of the American Association of Geographers*, 106(5), 1047–1062.
- Kalloori, S., Ricci, F., & Gennari, R. (2018). Eliciting pairwise preferences in recommender systems. *Proceedings of the 12th ACM Conference on Recommender Systems*, 329–337.
- Kamishima, T. (2003). Nantonac collaborative filtering: recommendation based on order responses. *Proceedings of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 583–588.
- Kang, R., Brown, S., & Kiesler, S. (2013). Why do people seek anonymity on the internet? Informing policy and design. *Proceedings of the SIGCHI conference on human factors in computing systems*, 2657–2666.
- Kekic, L., et al. (2007). The Economist Intelligence Unit's index of democracy. *The Economist*, 21, 1–11.
- Kilner, P. G., & Hoadley, C. M. (2017). Anonymity options and professional participation in an online community of practice. In *Computer Supported Collaborative Learning 2005* (pp. 272–280). Routledge.
- Kim, Y. M., Hsu, J., Neiman, D., Kou, C., Bankston, L., Kim, S. Y., Heinrich, R., Baragwanath, R., & Raskutti, G. (2018). The stealth media? Groups and targets behind divisive issue campaigns on Facebook. *Political Communication*, 35(4), 515–541.
- Klamler, C. (2005). On the closeness aspect of three voting rules: Borda–Copeland–Maximin. *Group Decision and Negotiation*, 14, 233–240.
- Klüver, H., & Bäck, H. (2019). Coalition agreements, issue attention, and cabinet governance. *Comparative Political Studies*, 52(13-14), 1995–2031.
- Kuriwaki, S. (2022). *Cumulative CCEs Common Content*. <https://doi.org/10.7910/DVN/I2DB6>
- Lachat, R. (2011). Electoral competitiveness and issue voting. *Political Behavior*, 33(4), 645–663.
- Lamont, M., Park, B. Y., & Ayala-Hurtado, E. (2017). Trump's electoral speeches and his appeal to the American white working class. *The British journal of sociology*, 68, S153–S180.
- Landemore, H. (2015). Inclusive constitution-making: The Icelandic experiment. *Journal of Political Philosophy*, 23(2), 166–191.
- LaTercera. (2020). Chile inicia histórico proceso para reemplazar su Constitución: Congreso acuerda plebiscito para abril de 2020 - La Tercera [(Accessed on 04/16/2023)].
- Layton, M. L., Smith, A. E., Moseley, M. W., & Cohen, M. J. (2021). Demographic polarization and the rise of the far right: Brazil's 2018 presidential election. *Research & Politics*, 8(1).
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., et al. (2009). Computational social science. *Science*, 323(5915), 721–723.

- Lazer, D. M., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., et al. (2020). Computational social science: Obstacles and opportunities. *Science*, *369*(6507), 1060–1062.
- Lee, D., Goel, A., Aitamurto, T., & Landemore, H. (2014). Crowdsourcing for participatory democracies: Efficient elicitation of social choice functions. *Proceedings of the Second AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2014)*, 133–142.
- Lees-Marshment, J., Dufresne, Y., Eady, G., Osborne, D., van der Linden, C., & Vowles, J. (2015). Vote Compass in the 2014 New Zealand election: Hearing the voice of New Zealand voters. *Political Science*, *67*(2), 94–124.
- Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*(S1), 392–410.
- LeMonde. (2022). French court suspends authorization for burkinis in Grenoble’s swimming pools [(Accessed on 06/27/2023)].
- Levendusky, M. S. (2009). The microfoundations of mass polarization. *Political Analysis*, *17*(2), 162–176.
- Luttig, M. D., Federico, C. M., & Lavine, H. (2017). Supporters and opponents of Donald Trump respond differently to racial cues: An experimental analysis. *Research & politics*, *4*(4), 2053168017737411.
- Makarenko, B. (2015). The Role of Elections in Democracy. In A. Przeworski (Ed.), *Democracy in a Russian Mirror* (pp. 130–146). Cambridge University Press.
- Mann, A. (2016). Computational social science. *Proceedings of the National Academy of Sciences*, *113*(3), 468–470.
- Martin, L. W. (2004). The government agenda in parliamentary democracies. *American Journal of Political Science*, *48*(3), 445–461.
- Mattei, N., & Walsh, T. (2013). Preflib: A library for preferences <http://www.preflib.org>. *Algorithmic Decision Theory: Third International Conference, ADT 2013, Bruxelles, Belgium, November 12-14, 2013, Proceedings 3*, 259–270.
- Mayer, N. (2013). From Jean-Marie to Marine Le Pen: electoral change on the far right. *Parliamentary Affairs*, *66*(1), 160–178.
- McDonald, M. P. (2023). US Elections Project - Election Data [(Accessed on 05/06/2023)].
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological methods*, *17*(3), 437.
- Mellon, J., Peixoto, T., & Sjoberg, F. M. (2017). Does online voting change the outcome? Evidence from a multi-mode public policy referendum. *Electoral Studies*, *47*, 13–24.
- Mieres Brevis, M. (2020). La dinámica de la desigualdad en Chile: Una mirada regional. *Revista de análisis económico*, *35*(2), 91–133.
- MIT Election Data and Science Lab. (2018a). *U.S. President Precinct-Level Returns 2016*. <https://doi.org/10.7910/DVN/LYWX3D>
- MIT Election Data and Science Lab. (2018b). *U.S. Senate Precinct-Level Returns 2016*. <https://doi.org/10.7910/DVN/NLTQAD>
- MIT Election Data and Science Lab. (2022a). *U.S. House of Representatives Precinct-Level Returns 2020*. <https://doi.org/10.7910/DVN/VLGF2M>
- MIT Election Data and Science Lab. (2022b). *U.S. President Precinct-Level Returns 2020*. <https://doi.org/10.7910/DVN/JXPRES>
- MIT Election Data and Science Lab. (2022c). *U.S. Senate Precinct-Level Returns 2020*. <https://doi.org/10.7910/DVN/ER9XTV>
- Morales Quiroga, M. (2020). Estallido social en Chile 2019: participación, representación, confianza institucional y escándalos públicos. *Análisis político*, *33*(98), 3–25.

- Navarrete, C., Ferrada, N., Macedo, M., Colley, R., Zhang, J., Grandi, U., Lang, J., & Hidalgo, C. A. (2022). Understanding Political Agreements and Disagreements: Evidence from the 2022 French Presidential Election. *arXiv preprint arXiv:2211.04577*.
- Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology, 70*(3), 1380–1402.
- Nichols, A. L., & Edlund, J. E. (2020). Why don't we care more about carelessness? Understanding the causes and consequences of careless participants. *International Journal of Social Research Methodology, 23*(6), 625–638.
- Palacios-González, F., & Garcia-Fernández, R. M. (2012). Interpretation of the coefficient of determination of an ANOVA model as a measure of polarization. *Journal of Applied Statistics, 39*(7), 1543–1555.
- Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 1*.
- Perbawani, P. S., Rahayu, R., & Anshari, I. N. (2018). Online political participation and Netizen anonymity in Indonesia's digital democracy. *PCD Journal, 6*(2), 185–212.
- Persily, N. (2017). The 2016 US Election: Can democracy survive the internet? *Journal of democracy, 28*(2), 63–76.
- Peters, D., Pierczyński, G., & Skowron, P. (2021). Proportional participatory budgeting with additive utilities. *Advances in Neural Information Processing Systems, 34*, 12726–12737.
- Prothro, J. W., & Grigg, C. M. (1960). Fundamental principles of democracy: Bases of agreement and disagreement. *The Journal of Politics, 22*(2), 276–294.
- Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior, 58*, 354–360.
- RePass, D. E. (1971). Issue salience and party choice. *American Political Science Review, 65*(2), 389–400.
- Reynal-Querol, M. (2002). Ethnicity, political systems, and civil wars. *Journal of conflict resolution, 46*(1), 29–54.
- Rodrik, D. (2018). Populism and the economics of globalization. *Journal of international business policy, 1*, 12–33.
- Rodriguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge journal of regions, economy and society, 11*(1), 189–209.
- Rohla, R., Johnston, R., Jones, K., & Manley, D. (2018). Spatial scale and the geographical polarization of the American electorate. *Political Geography, 65*, 117–122.
- Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, communication & society, 18*(2), 121–138.
- Roy, J., & Alcantara, C. (2015). The candidate effect: Does the local candidate matter? *Journal of Elections, Public Opinion & Parties, 25*(2), 195–214.
- Saaty, T. L. (1988). What is the analytic hierarchy process? In *Mathematical models for decision support* (pp. 109–121). Springer.
- Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: mapping the inequality of urban perception. *PloS one, 8*(7), e68400.
- Salganik, M. J., & Levy, K. E. (2015). Wiki surveys: Open and quantifiable social data collection. *PloS one, 10*(5), e0123483.
- Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory, 10*(2), 187–217.

- Scala, D. J., Johnson, K. M., & Rogers, L. T. (2015). Red rural, blue rural? Presidential voting patterns in a changing rural America. *Political Geography*, *48*, 108–118.
- Servel. (2023). *Servel — servicio electoral de chile* [(Accessed on 04/27/2023)].
- Shadac. (2023). Bulk Data Download — State Health Access Data Assistance Center [(Accessed on 03/30/2023)].
- Shi, F., Teplitskiy, M., Duede, E., & Evans, J. A. (2019). The wisdom of polarized crowds. *Nature Human Behaviour*, *3*(4), 329–336.
- Simons, J. J., & Green, M. C. (2018). Divisive topics as social threats. *Communication Research*, *45*(2), 165–187.
- Small, C., Bjorkegren, M., Erkkilä, T., Shaw, L., & Megill, C. (2021). Polis: Scaling deliberation by mapping high dimensional opinion spaces. *Recerca: revista de pensament i anàlisi*, *26*(2).
- Smith, W. D. (2000). Range voting. *Working Paper*.
- Su, C. (2017). From Porto Alegre to New York City: Participatory Budgeting and Democracy. *New Political Science*, *39*(1), 67–75.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior*, *7*(3), 321–326.
- Sussell, J. (2013). New support for the big sort hypothesis: An assessment of partisan geographic sorting in California, 1992–2010. *PS: Political Science & Politics*, *46*(4), 768–773.
- Testa, C. (2012). Is polarization bad? *European Economic Review*, *56*(6), 1104–1118.
- Thurstone, L. L. (1927). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, *21*(4), 384.
- Treré, E., & Barassi, V. (2015). Net-authoritarianism? How web ideologies reinforce political hierarchies in the Italian 5 Star Movement. *Journal of Italian Cinema & Media Studies*, *3*(3), 287–304.
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*.
- Valentin, J. V. (2022). *The Polarization of the French society: a study of the Yellow Vests movement* (Master's thesis). UiT Norges arktiske universitet.
- Wang, Y.-Q., & Tsui, K.-Y. (2000). Polarization orderings and new classes of polarization indices. *Journal of Public Economic Theory*, *2*(3), 349–363.
- Ward, M., & Meade, A. W. (2018). Applying social psychology to prevent careless responding during online surveys. *Applied Psychology*, *67*(2), 231–263.
- Waugh, A. S., Pei, L., Fowler, J. H., Mucha, P. J., & Porter, M. A. (2009). Party polarization in congress: A network science approach. *arXiv preprint arXiv:0907.3509*.
- Weber, I., Popescu, A.-M., & Pennacchiotti, M. (2013). Data-driven political science. *Proceedings of the sixth ACM international conference on Web search and data mining*, 777–778.
- Winkler, H. (2019). The effect of income inequality on political polarization: Evidence from European regions, 2002–2014. *Economics & Politics*, *31*(2), 137–162.
- Wojtasik, W., et al. (2013). Functions of elections in democratic systems. *Political Preferences*, (4), 25–38.
- Wolfson, M. C. (1994). When inequalities diverge. *The American Economic Review*, *84*(2), 353–358.
- Zhang, X., & Kanbur, R. (2001). What difference do polarisation measures make? An application to China. *Journal of development studies*, *37*(3), 85–98.
- Zhang, X., & Kanbur, R. (2005). Spatial inequality in education and health care in China. *China economic review*, *16*(2), 189–204.

List of Figures

1.1	Interplay between our research questions and methodology	21
1.2	Milestones from the official candidacies and the Election Day.	22
1.3	Conceptualizing the Election Polarization.	25
2.1	Illustrative examples for Fallback Voting and Pairwise Comparison	37
2.2	Participation platform and pairwise comparison data	43
2.3	Understanding participants preferences	50
2.4	Divisiveness analysis	51
2.5	Robustness analysis for ranking of agreements and divisiveness	54
2.6	Axiomatic and matrixial properties of agreement and divisiveness for France . .	56
2.7	Axiomatic and matrixial properties of agreement and divisiveness for Brazil . .	57
2.8	Key results of Divisiveness in datasets from Preflib.org	58
3.1	Illustrative example of Election Polarization	70
3.2	An illustrative example of data collection and curation in France and Chile. . .	71
3.3	Numerical properties of Election Polarization	75
3.4	Comparison of classical polarization measures and EP by using synthetic data .	76
3.5	The geography of Election Polarization in the United States in the period 2000- 2020.	77
3.6	The geography of Election Polarization in France (2002-2022) and Chile (2013- 2021).	78
3.7	Robustness across aggregation levels using real data	82
3.8	Comparison of Election Polarization and its components computed using First Round and Runoff data in France and Chile	84
4.1	Unpacking the user interface of Rank screen	92
4.2	Illustrative example of Approval screen for six issues.	93

4.3	Illustrative example of Rank screen for 6 issues.	94
4.4	Illustrative example of Pairwise Comparison.	94
4.5	Impact of divisive issues in online participation.	101
4.6	Comparison of data collected in terms of issue pairs	103
4.7	The nature of online participation.	104
4.8	Evaluation of consistency in (a) France and (b) Brazil.	108
4.9	The nature of pairwise comparison data.	109

List of Tables

2.1	List of issues included in the French platform, ranked by win percentage. . . .	46
2.2	List of issues included in the Brazilian platform, ranked by win percentage. . .	47
2.3	Issues-level associations of Divisiveness in France	52
2.4	Issues-level associations of Divisiveness in Brazil	52
3.1	Description of data sets curated and publicly available as part of our contribution	72
3.2	State-level association of Political Polarization in the United States (2008-2020)	80
3.3	State-level association of Voting Turnout in the United States (2008-2020) . . .	81
4.1	Logistic regression analysis for early dropouts	102
4.2	Logistic regression analysis for the likelihood of abstentions	105
4.3	Logistic regression analysis for the likelihood of modifying proposed rank . . .	106

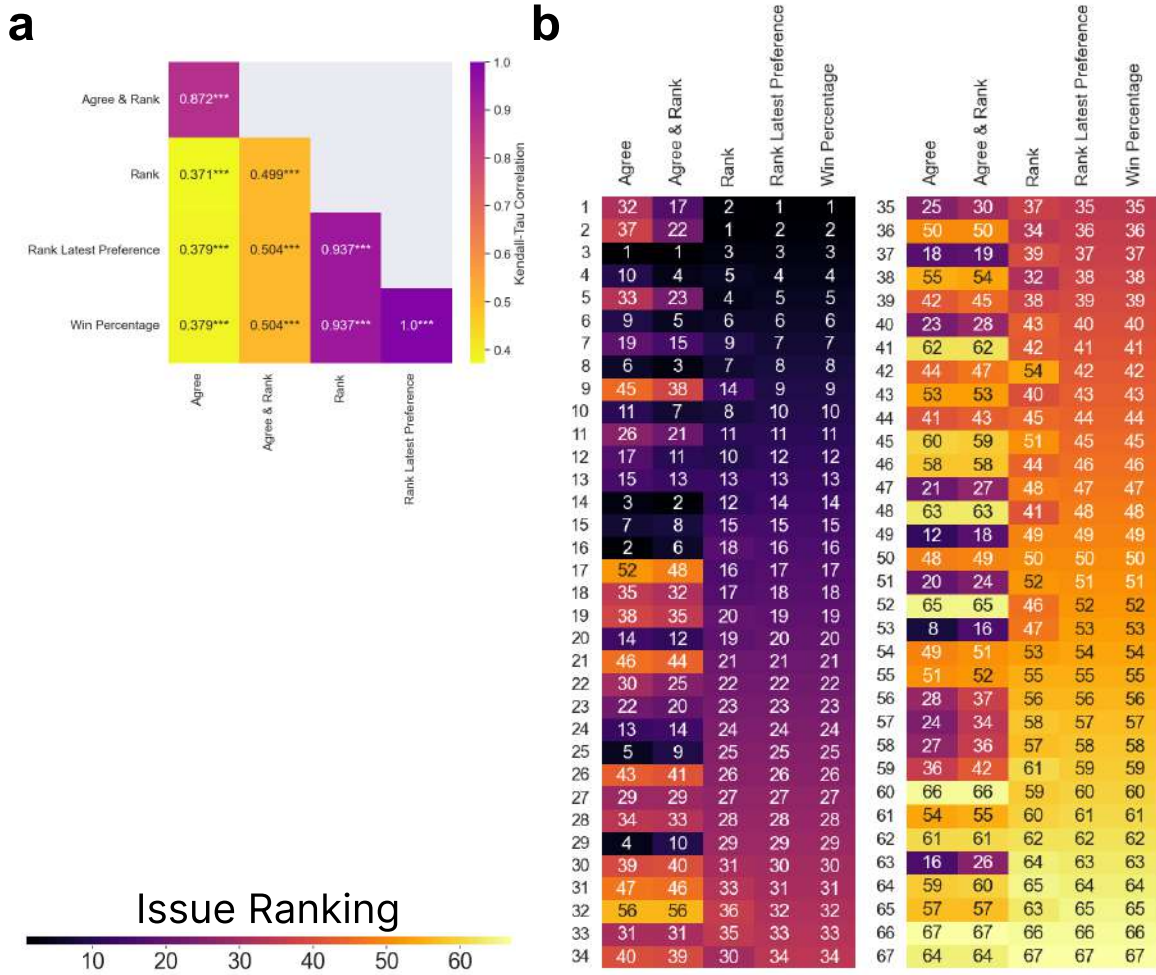


Figure A.2: Robustness between rankings of agreements by aggregation functions (b) and the Kendall-Tau correlation matrix (a) in Brazil. Win Percentage represents the ranking of agreements presented in the manuscript.

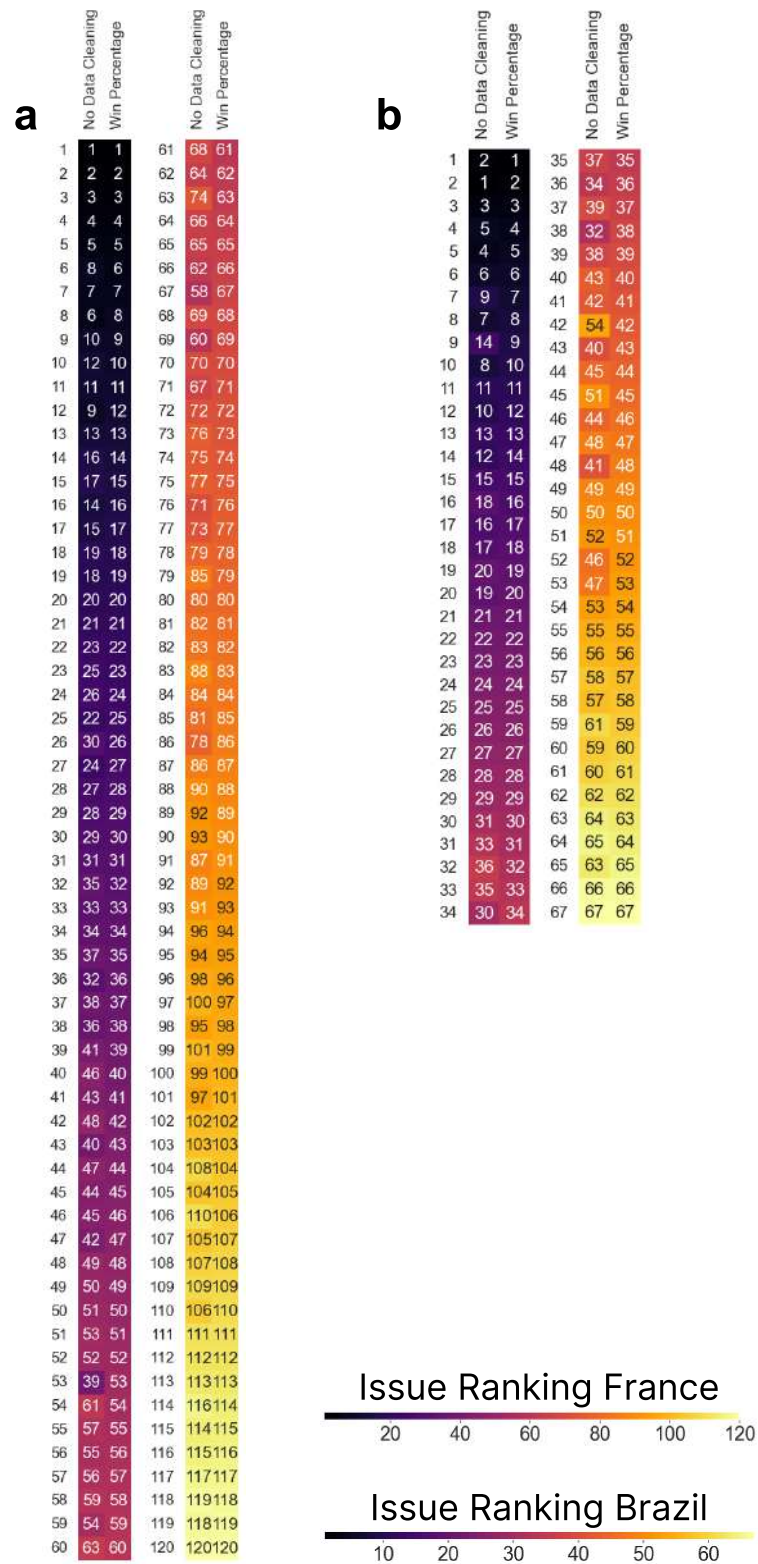


Figure A.3: Robustness between rankings of agreements including the suspicious/bot participation in (a) France and (b) Brazil. Win Percentage represents the ranking of agreements presented in the manuscript.

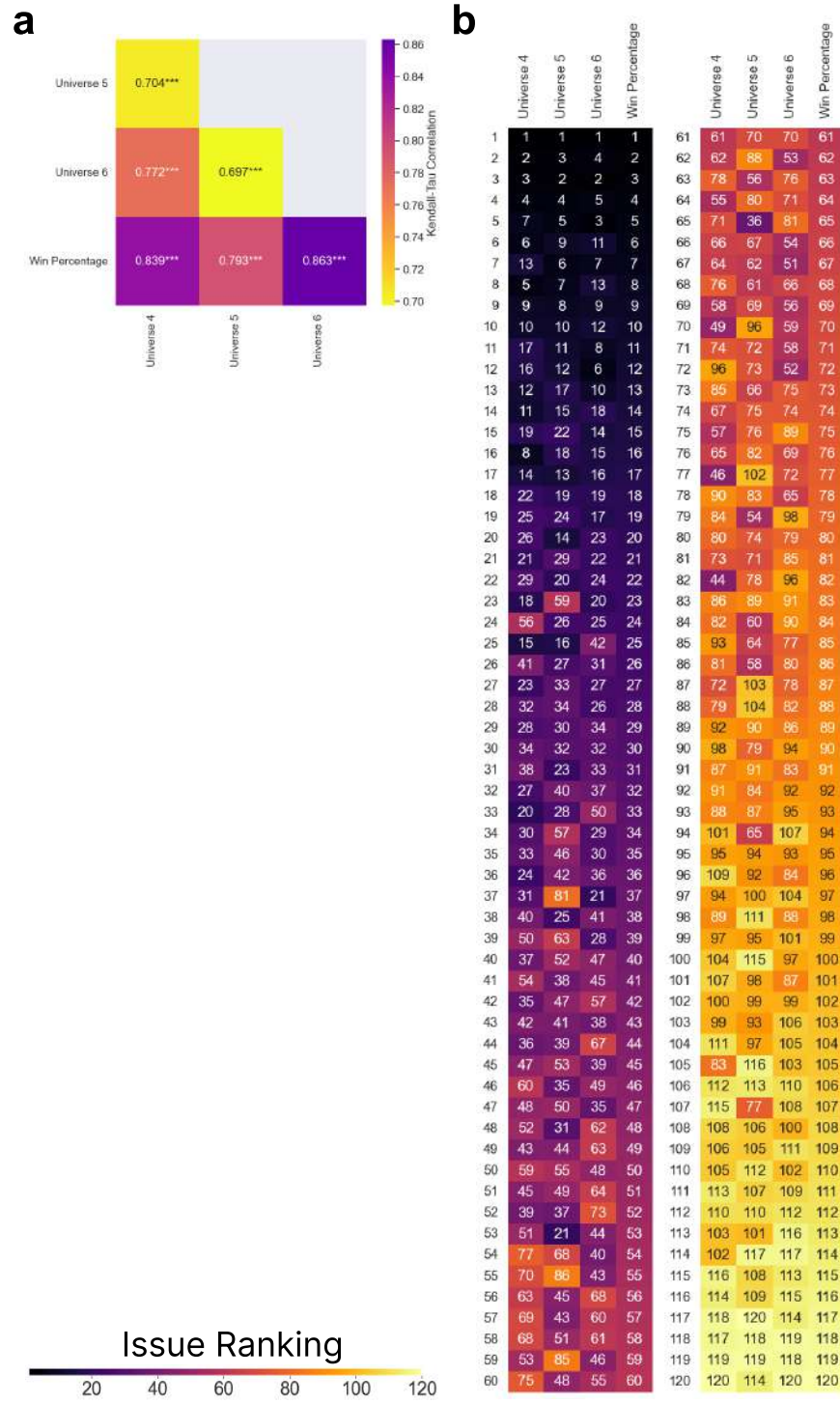


Figure A.4: Robustness between the rankings of agreements across universes (b) and the Kendall-Tau correlation matrix (a) in France. Win Percentage represents the ranking of agreements presented in the manuscript.

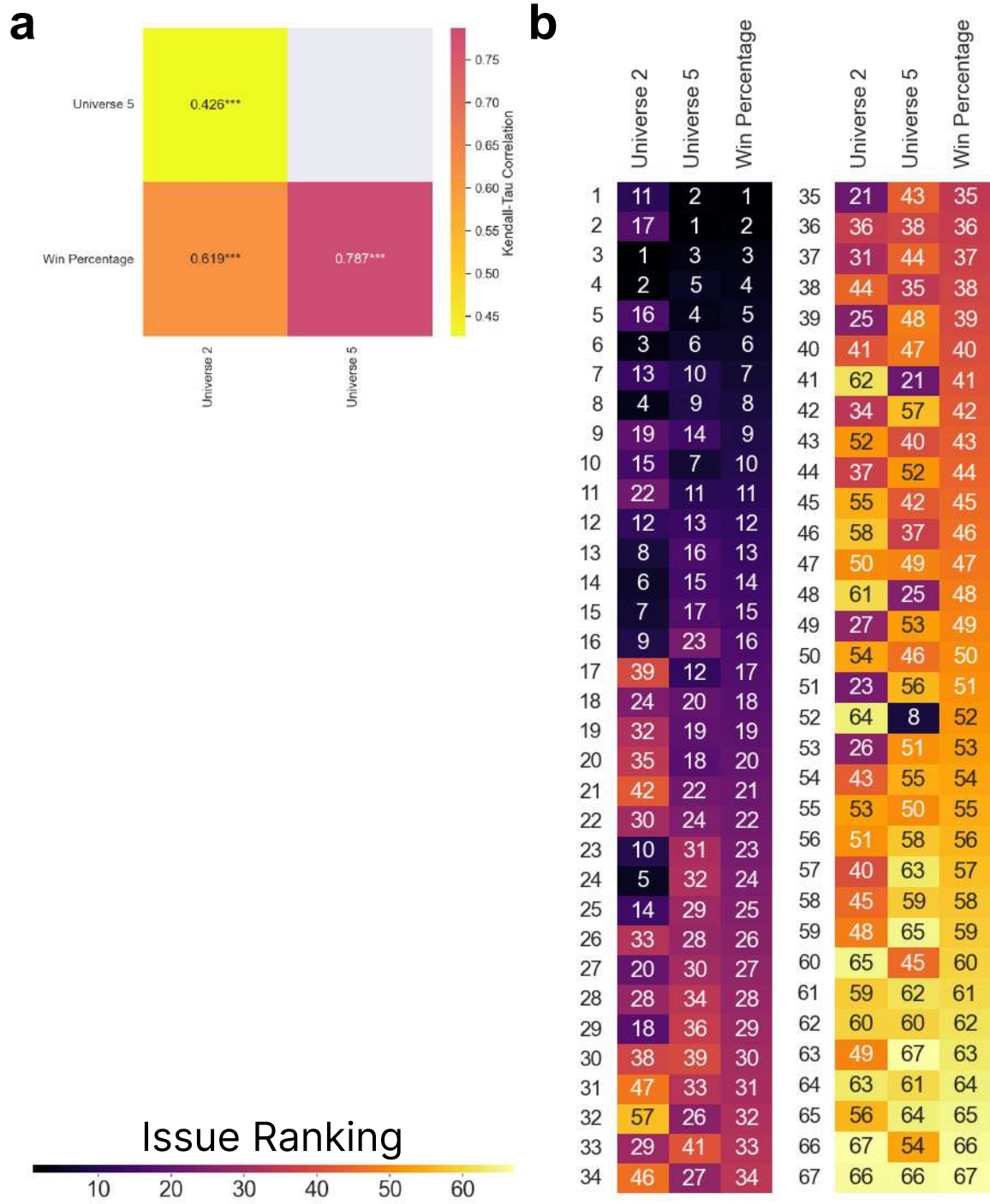


Figure A.5: Robustness between the rankings of agreements across universes (b) and the Kendall-Tau correlation matrix (a) in Brazil. Win Percentage represents the ranking of agreements presented in the manuscript.

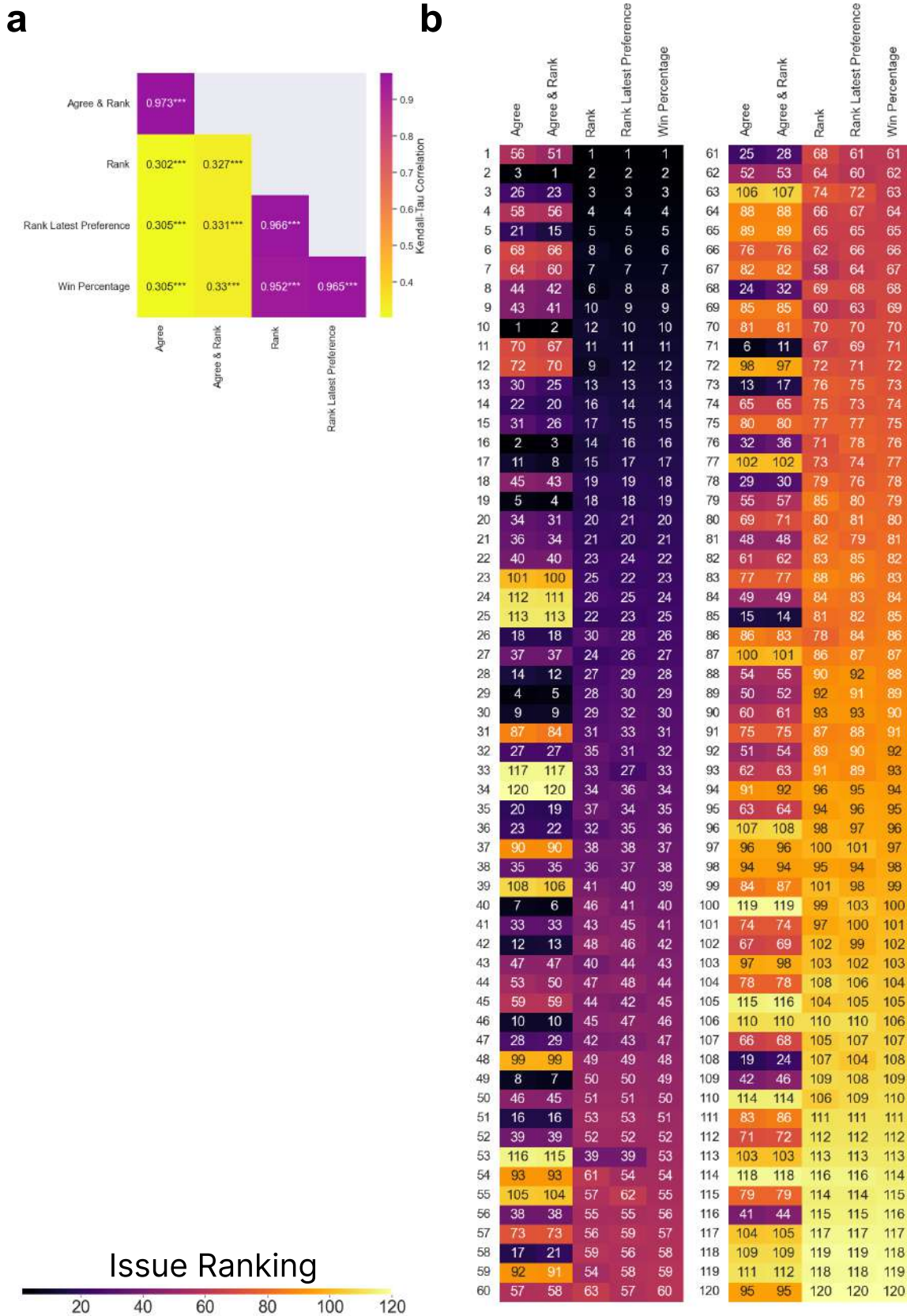


Figure A.6: Robustness between the rankings of agreements using different datasets (a) and the Kendall-Tau correlation matrix (b). The manuscript uses the Win Percentage ranking of agreements.

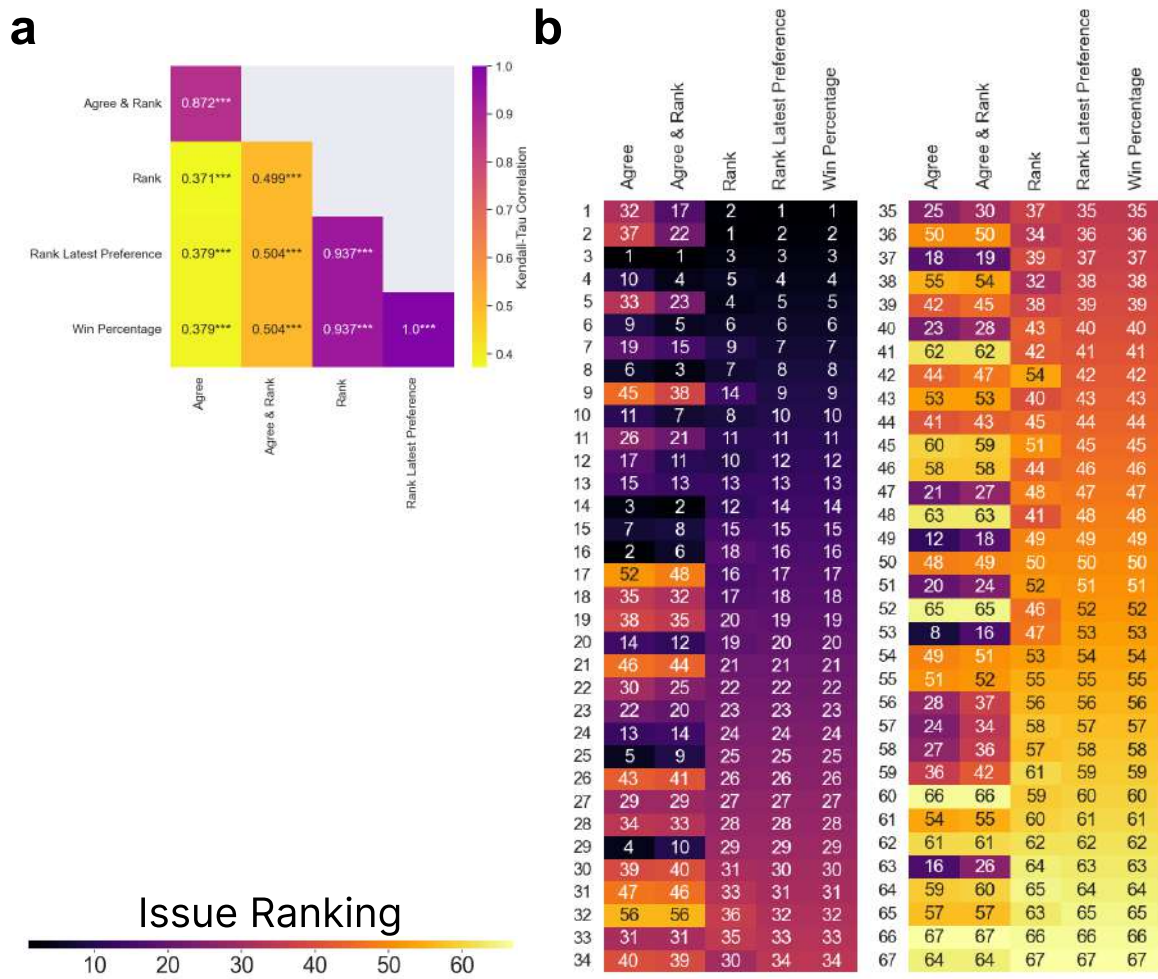


Figure A.7: Robustness between the rankings of agreements using different datasets (a) and the Kendall-Tau correlation matrix (b) in Brazil. The manuscript uses the Win Percentage ranking of agreements.

A.2 Robustness ranking of divisiveness

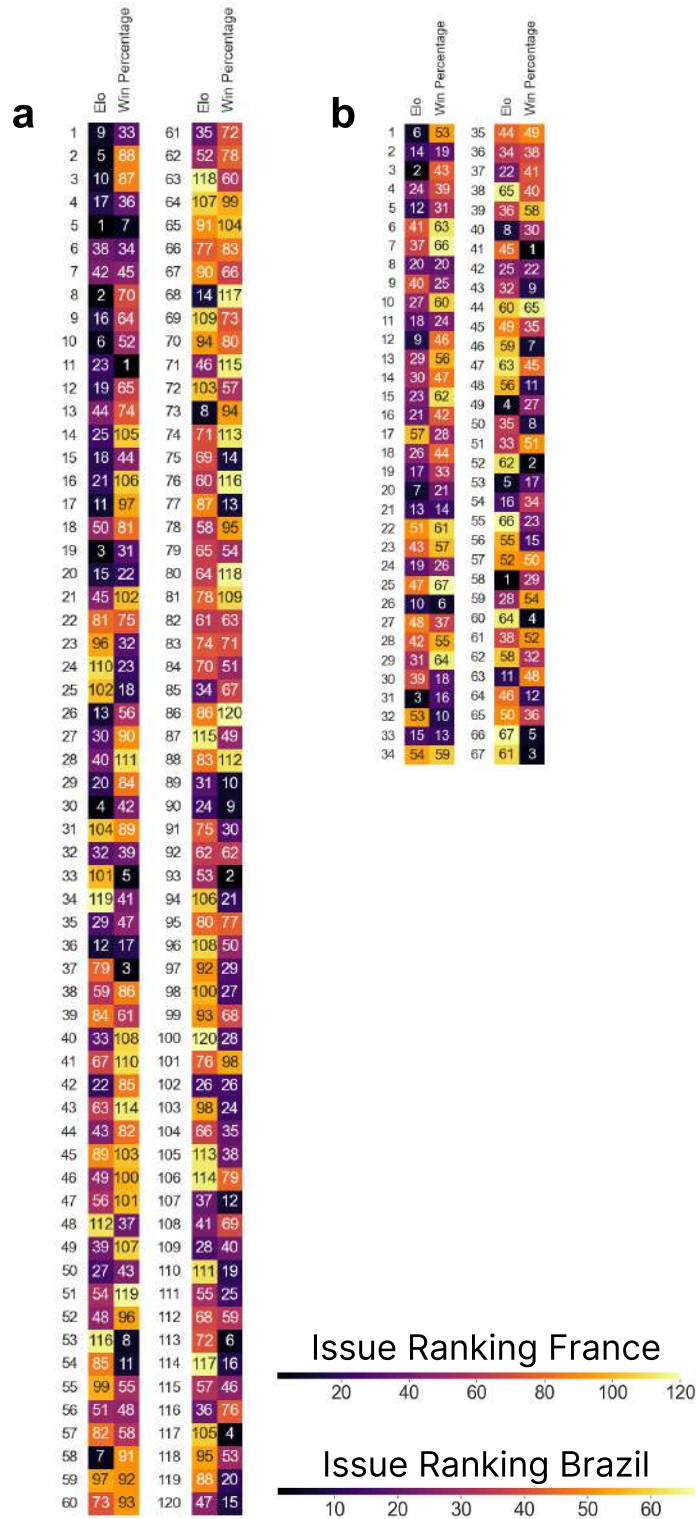


Figure A.8: Robustness ranking of divisiveness comparing the results using Elo and Win Percentage in (a) France and (b) Brazil. The manuscript presents the results using the divisiveness ranking of Win Percentage.

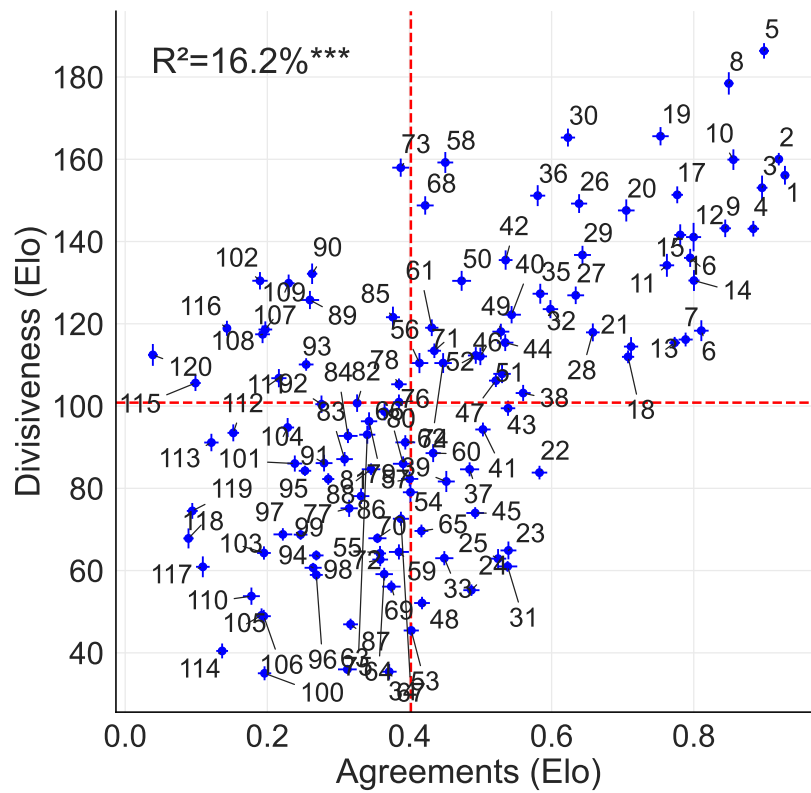


Figure A.9: Agreements and divisiveness using Elo score method in France.

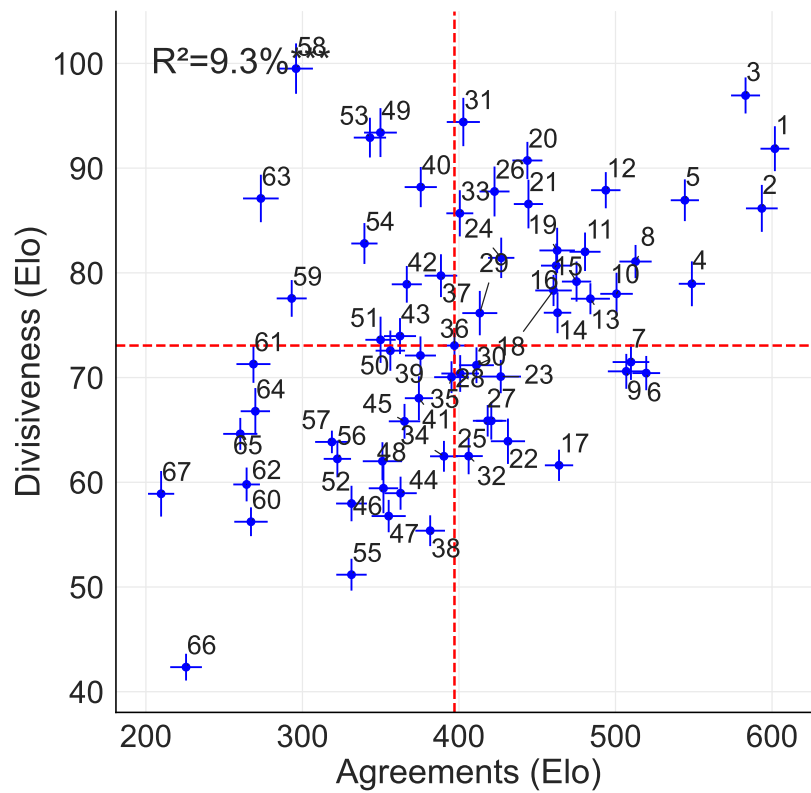


Figure A.10: Agreements and divisiveness using Elo score method in Brazil.

Appendix for Chapter 3

B.1 Socio-demographic characteristics

Socio-demographic characteristics, such as inequality level, annual unemployment rate, or median household income, were downloaded from official public institutions (e.g., Census Bureau) or repositories maintained by scholars or private institutions. Here we briefly describe the data origin and curation procedure.

B.1.1 United States

Census Bureau

- Land area in Sq. Km. from State Area Measurements (Census, [2023](#)).
- Population (Estimation) (2000-2020)

American Community Survey (ACS) This survey is the premier source for detailed population and housing information in the U.S. (ACS, [2023](#)). The Census Bureau releases every year a version of ACS, providing detailed household and population information.

- Income inequality (Gini coefficient) (2010-2020) (Census.gov, [2023](#))

Frank-Sommeiller-Price Series Data collection of inequality measures maintained by Professor Mark W. Frank, Ph.D. Professor of Economics at SHSU (Frank, [2023](#)). He offers a comprehensive panel of annual state-level income inequality measures constructed from individual tax filing data available from the Internal Revenue Service. We downloaded the Income inequality (Gini coefficient) in the period 2000-2018.

State Health Access Data Assistance Center (SHADAC) This institution at the University of Minnesota provides a collection of statistics in order to help bridge the gap between policy and health issues (Shadac, [2023](#)). From SHADAC, we downloaded the following data sets:

- Annual unemployment rate (2000-2020)

Bureau of Economic Analysis (BEA) The Bureau of Economic Analysis (BEA) is the federal agency responsible for producing statistics and data about the U.S. economy (BEA, 2023). We downloaded the following state-level data sets from BEA:

- Personal Income per capita (in Dollars) (2000-2020).
- Real Gross Domestic Product (Millions of chained 2012 dollars) (2000-2020).
- Personal Expenditure per capita (2000-2020).

Special considerations in the U.S.

Income inequality As seen, there is no data source that includes the state-level Gini coefficient for all our period of study. We then followed two approaches: Firstly, we input the Frank-series Gini coefficient for 2018 as the one in 2020. Secondly, we tested our models in a shorter period using the Gini from the Census and imputing 2010 as the value for 2008.

B.1.2 France

L’Institut national de la statistique et des études économiques (Insee) The bureau of statistics and economic studies (Insee) collects, produces, analyze, and diffuse information over French economy and society. The data was collected at the department level. Since we do not rely on recent data for all variables, we used the most recent value available for each characteristic:

- Rural population (2017) (Insée, 2023c).
- Poverty rate (2018) (Insée, 2023a).
- Income inequality (Gini coefficient) (2020) (Insée, 2023b).
- Median household income (in Euros). (2020) (Insée, 2023b).

B.1.3 Chile

Biblioteca del Congreso Nacional (BCN) The *Congress’ Library* is the institution in the hold of preserving documents and valuable information about the legislative process in Chile. They offer a collection of standardized data sets from national statistics products, such as the Census and the National Socioeconomic Characterization Survey (Encuesta de Caracterización Socioeconómica Nacional (Casen) in Spanish). Similarly to France, we used the most recent year.

- Rural population (Census, 2017).
- Poverty rate (Casen, 2020).
- Household monetary income (in CLP) (Casen, 2020).

Income inequality We relied on the region-level Gini coefficient (using 2017 data) computed by Brevis (2020) (Mieres Brevis, 2020).

B.2 Mapping Political Interest

We proxy political interest as the turnout level on Election Day. From France (2022) and Chile (2017, 2021), election data already includes the number of abstentions per polling station. In the United States, we downloaded voting statistics from the U.S. Elections Project (2000-2020) (McDonald, 2023), and used the Voting Eligible Population (VEP) turnout rate.

B.3 Data sampling

We explored in Chapter 3 the numerical properties of EP using synthetic data. Algorithms 1 and 2 summarize our approach for elections of two and three candidates, respectively. Here we sampled $\mu \in \{0.5, 0.66, 0.75, 0.8333, 1\}$, $\sigma \in \{0.0025, 0.05, 0.10, 0.25\}$, and we set the number of units M in 100. It should be noted that we consider as constant the number of votes per unit, in our case, 100.

Algorithm 1 Data sampling algorithm for a given election between two candidates.

Require: $\mu \in [0, 1], \sigma \in [0, 0.25], m > 0, N \leftarrow 2$

Ensure: $\sum_{k=0}^1 r_{i,k} = 1$ & $\sum_{i=0}^{m-1} \sum_{j=0}^1 v_{i,k} = 100m, \forall i \in [0, m-1]$

$a \leftarrow \mu$

▷ Mean

$s \leftarrow \sigma$

▷ Standard deviation

$M \leftarrow m$

▷ Number of voting units

for $i \leftarrow 0$ to $M - 1$ **do**

$r_{0,i} \leftarrow \mathcal{N}(a, s)$

if $r_{0,i} > 1$ **then**

$r_{0,i} \leftarrow 1$

else if $r_{0,i} < 0$ **then**

$r_{0,i} \leftarrow 0$

end if

$r_{0,i} \leftarrow r_{0,i} \bmod 0.01$

$r_{1,i} \leftarrow 1 - r_{0,i}$

end for

for $i \leftarrow 0$ to $N - 1$ **do**

for $j \leftarrow 0$ to $M - 1$ **do**

$v_{i,k} \leftarrow 100r_{i,k}$

end for

end for

Algorithm 2 Data sampling algorithm for a given election between three candidates.

Require: $\mu_1, \mu_2 \in [0, 1], \sigma_1, \sigma_2 \in [0, 0.25], m > 0, N \leftarrow 3$

Ensure: $\sum_{j=0}^2 r_{i,k} = 1$ & $\sum_{i=0}^{m-1} \sum_{j=0}^2 v_{i,k} = 100m, \forall i \in [0, m-1]$

$a_1 \leftarrow \mu_1$ ▷ Mean (1)
 $a_2 \leftarrow \mu_2$ ▷ Mean (2)
 $s_1 \leftarrow \sigma_1$ ▷ Standard deviation (1)
 $s_2 \leftarrow \sigma_2$ ▷ Standard deviation (2)
 $M \leftarrow m$ ▷ Number of voting units

for $i \leftarrow 0$ to $M - 1$ **do**

$r_{0,i} \leftarrow \mathcal{N}(a_1, s_1)$
 $r_{1,i} \leftarrow \mathcal{N}(a_2, s_2)$

for $j \leftarrow 0$ to 1 **do**

if $r_{i,k} > 1$ **then**
 $r_{i,k} \leftarrow 1$
else if $r_{0,i} < 0$ **then**
 $r_{i,k} \leftarrow 0$
end if

end for

if $(r_{0,i} + r_{1,i}) > 1$ **then**
 $r_{1,i} \leftarrow 1 - r_{0,i}$
end if

$r_{2,i} \leftarrow 1 - (r_{0,i} + r_{1,i})$

end for

for $i \leftarrow 0$ to $M - 1$ **do**

for $j \leftarrow 0$ to $N - 1$ **do**

$v_{i,k} \leftarrow 100r_{i,k}$

end for

end for

B.4 Additional Regression Models

Here we focus on analyzing the relationship between control variables and EP to explore whether our measure of polarization encompasses socio-demographic characteristics of territories and discard that just mirroring inequality or macro-economic conditions. We set up the regression model as follows:

$$y_{g,t} = \beta_0 + \beta_1 y_{g,t-1} + \beta_2 \text{EP}_{-g,t-1} + \beta_3^T \mathbf{X}_{g,t-1} + \beta_4^T \mathbf{X}_{g,t} + \mu_t + \gamma_g + \epsilon_{g,t} \quad (\text{B.1})$$

Where $y_{g,t}$ is the dependent variable for the geography g in year t (in our case, the EP, Within-EP, and Between-EP), $\text{EP}_{-g,t-1}$ is the average EP of neighbors, $y_{g,t-1}$ is the lag of the dependent variable, $\mathbf{X}_{g,t}$ is a vector of control variables in t , $\mathbf{X}_{g,t-1}$ is a vector of control variables in $t - 1$, μ_t is the period fixed effect, and $\epsilon_{g,t}$ is the error term. The lag period $t - 1$ represents the previous election. Control variables were normalized by calculating their logarithm before running the regressions. Standard errors are clustered at the state level.

EP as dependent variable: Tables [B.1](#), [B.4](#), and [B.7](#).

Within-EP as dependent variable: Tables [B.2](#), [B.5](#), and [B.8](#).

Between-EP as dependent variable: Tables [B.3](#), [B.6](#), and [B.9](#).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
EP (t - 1)	0.64*** (0.06)							0.61*** (0.08)	0.61*** (0.07)
Inequality level (t - 1)		-0.38* (0.22)						-0.14 (0.10)	-0.11 (0.10)
Unemployment (t - 1)			0.71 (0.56)					0.84* (0.43)	0.89*** (0.44)
Spatial EP (t - 1)				0.07 (0.07)				0.03 (0.04)	0.03 (0.04)
Spatial EP (t - 1)				0.07 (0.07)				0.03 (0.04)	0.03 (0.04)
Ed. Attainment (t - 1)					0.01 (0.15)				0.25* (0.13)
Log of personal Income PC (t - 1)						-0.05 (0.06)		0.02 (0.04)	-0.04 (0.04)
Log of GDP PC (t - 1)	0.02*** (0.01)	0.05*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.05*** (0.01)	0.01** (0.01)	0.01** (0.01)
Log of Pop. Density (t - 1)	-0.01** (0.00)	-0.02** (0.01)	-0.01 (0.01)	-0.01* (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01* (0.01)	-0.01** (0.01)	-0.01** (0.00)
Num. obs.	250	250	250	250	250	250	250	250	250
R ² (full model)	0.57	0.26	0.25	0.27	0.24	0.25	0.27	0.59	0.60
R ² (proj model)	0.57	0.26	0.25	0.27	0.24	0.25	0.27	0.59	0.60
Adj. R ² (full model)	0.56	0.24	0.23	0.25	0.22	0.22	0.24	0.57	0.58
Adj. R ² (proj model)	0.56	0.24	0.23	0.25	0.21	0.22	0.24	0.57	0.57
Num. groups: year	5	5	5	5	5	5	5	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.1: State-level regression analyses of EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Within-EP (t - 1)	1.01*** (0.02)							0.99*** (0.02)	0.99*** (0.02)
Inequality level (t - 1)		-0.21 (0.21)					-0.16 (0.21)	-0.06* (0.03)	-0.06* (0.03)
Unemployment (t - 1)			0.14 (0.53)					-0.21 (0.13)	-0.20 (0.13)
Spatial Within-EP (t - 1)				0.34*** (0.12)				0.04 (0.03)	0.05 (0.03)
Ed. Attainment (t - 1)					-0.15 (0.17)				0.02 (0.04)
Log of personal Income PC (t - 1)							-0.10* (0.05)	-0.01 (0.01)	-0.01 (0.01)
Log of GDP PC (t - 1)	0.00 (0.00)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.00* (0.00)	0.00* (0.00)
Log of Pop. Density (t - 1)	-0.00** (0.00)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.00** (0.00)	-0.00** (0.00)
Num. obs.	250	250	250	250	250	250	250	250	250
R ² (full model)	0.92	0.38	0.37	0.44	0.38	0.41	0.42	0.93	0.93
R ² (proj model)	0.91	0.28	0.28	0.35	0.28	0.32	0.32	0.91	0.91
Adj. R ² (full model)	0.92	0.36	0.35	0.42	0.36	0.39	0.40	0.92	0.92
Adj. R ² (proj model)	0.91	0.26	0.25	0.33	0.26	0.30	0.30	0.91	0.91
Num. groups: year	5	5	5	5	5	5	5	5	5

***p < 0.01; **p < 0.05; *p < 0.1

Table B.2: State-level regression analyses of Within-EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: ***p < 0.01, **p < 0.05, *p < 0.1.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Between-EP (t - 1)	0.74*** (0.07)							0.74*** (0.08)	0.74*** (0.08)
Inequality level (t - 1)		-0.17 (0.31)						-0.19 (0.30)	0.02 (0.09)
Unemployment (t - 1)			0.57 (0.65)					1.02*** (0.31)	1.04*** (0.10)
Spatial Between-EP (t - 1)				0.06 (0.05)				0.01 (0.02)	0.01 (0.02)
Ed. Attainment (t - 1)					0.16 (0.18)				0.21 (0.13)
Log of personal Income PC (t - 1)						0.05 (0.07)		0.05 (0.03)	0.00 (0.05)
Log of GDP PC (t - 1)	0.00 (0.00)	0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.00)	-0.00 (0.00)
Log of Pop. Density (t - 1)	-0.00 (0.00)	0.00 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.00)	-0.00 (0.00)
Num. obs.	250	250	250	250	250	250	250	250	250
R ² (full model)	0.62	0.12	0.13	0.13	0.13	0.13	0.13	0.64	0.64
R ² (proj model)	0.57	0.02	0.02	0.03	0.02	0.02	0.03	0.59	0.60
Adj. R ² (full model)	0.60	0.10	0.10	0.11	0.10	0.10	0.10	0.62	0.62
Adj. R ² (proj model)	0.56	-0.01	-0.01	0.00	-0.01	-0.01	-0.00	0.57	0.58
Num. groups: year	5	5	5	5	5	5	5	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.3: State-level regression analyses of Between-EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
EP (t - 1)	0.72*** (0.05)							0.70*** (0.07)	0.69*** (0.06)
Inequality level (t)		-0.07 (0.28)					-0.05 (0.28)	0.00 (0.12)	0.02 (0.12)
Unemployment (t)			0.36 (0.43)					0.03 (0.29)	0.05 (0.28)
Spatial EP (t)				0.12 (0.07)				0.04 (0.04)	0.04 (0.04)
Ed. Attainment (t)					-0.05 (0.25)				0.12 (0.12)
Log of personal Income PC (t)						-0.09 (0.10)	-0.09 (0.10)	-0.05 (0.05)	-0.09* (0.05)
Log of GDP PC (t)	0.01 (0.02)	-0.02 (0.04)	-0.01 (0.05)	0.02 (0.06)	-0.01 (0.07)	0.04 (0.08)	0.04 (0.08)	0.05 (0.04)	0.06 (0.04)
Log of Pop. Density (t)	-0.00 (0.00)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.00 (0.00)	-0.00 (0.00)
Num. obs.	250	300	300	300	300	300	300	250	250
R ² (full model)	0.55	0.03	0.03	0.10	0.03	0.03	0.03	0.55	0.56
R ² (proj model)	0.55	0.03	0.03	0.10	0.03	0.03	0.03	0.55	0.56
Adj. R ² (full model)	0.53	0.00	0.00	0.08	-0.00	0.01	0.00	0.53	0.53
Adj. R ² (proj model)	0.53	-0.00	0.00	0.07	-0.00	0.01	0.00	0.53	0.53
Num. groups: year	5	6	6	6	6	6	6	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.4: State-level regression analyses of EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Within-EP (t - 1)	1.02*** (0.02)							1.00*** (0.02)	0.99*** (0.02)
Inequality level (t)		0.16 (0.24)						-0.06 (0.04)	-0.06 (0.04)
Unemployment (t)			0.29 (0.52)					-0.12 (0.10)	-0.11 (0.10)
Spatial Within-EP (t)				0.52*** (0.13)				0.06** (0.03)	0.07** (0.03)
Ed. Attainment (t)					-0.20 (0.22)				0.02 (0.04)
Log of personal Income PC (t)						-0.35*** (0.12)	-0.37*** (0.12)	-0.02 (0.02)	-0.03 (0.02)
Log of GDP PC (t)	-0.00 (0.01)	-0.01 (0.04)	-0.00 (0.04)	0.06 (0.04)	0.02 (0.04)	0.22*** (0.08)	0.23*** (0.08)	0.02 (0.01)	0.02 (0.01)
Log of Pop. Density (t)	-0.00*** (0.00)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.01 (0.01)	0.01* (0.01)	-0.00 (0.00)	-0.00 (0.00)
Num. obs.	250	300	300	300	300	300	300	250	250
R ² (full model)	0.92	0.18	0.18	0.33	0.19	0.28	0.30	0.93	0.93
R ² (proj model)	0.91	0.01	0.00	0.19	0.01	0.13	0.15	0.91	0.91
Adj. R ² (full model)	0.92	0.16	0.16	0.31	0.16	0.26	0.28	0.92	0.92
Adj. R ² (proj model)	0.91	-0.02	-0.02	0.17	-0.01	0.11	0.12	0.91	0.91
Num. groups: year	5	6	6	6	6	6	6	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.5: State-level regression analyses of Within-EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Between-EP (t - 1)	0.74*** (0.07)							0.71*** (0.08)	0.72*** (0.08)
Inequality level (t)		-0.23 (0.29)						-0.01 (0.12)	-0.01 (0.12)
Unemployment (t)			0.07 (0.61)					0.18 (0.23)	0.18 (0.23)
Spatial Between-EP (t)				0.08* (0.05)				0.02 (0.02)	0.02 (0.03)
Ed. Attainment (t)					0.15 (0.26)				0.02 (0.12)
Log of personal Income PC (t)						0.26* (0.14)	0.27** (0.14)	0.09 (0.06)	0.09 (0.07)
Log of GDP PC (t)	0.01 (0.02)	-0.01 (0.04)	-0.01 (0.05)	0.01 (0.05)	-0.03 (0.07)	-0.18* (0.10)	-0.19* (0.10)	-0.04 (0.04)	-0.04 (0.04)
Log of Pop. Density (t)	-0.00 (0.00)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.00)	-0.00 (0.00)
Num. obs.	250	300	300	300	300	300	300	250	250
R ² (full model)	0.62	0.15	0.15	0.17	0.15	0.19	0.20	0.62	0.62
R ² (proj model)	0.57	0.03	0.02	0.05	0.03	0.07	0.09	0.58	0.58
Adj. R ² (full model)	0.61	0.13	0.12	0.15	0.13	0.17	0.18	0.61	0.60
Adj. R ² (proj model)	0.56	0.00	-0.01	0.02	-0.00	0.04	0.06	0.56	0.56
Num. groups: year	5	6	6	6	6	6	6	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.6: State-level regression analyses of Between-EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Inequality level (t - 1)	-0.38* (0.22)					-0.36* (0.21)	-0.49** (0.21)	-0.46** (0.21)
Unemployment (t - 1)		0.71 (0.56)					1.08 (0.66)	1.12* (0.66)
Spatial EP (t - 1)			0.07 (0.07)				0.09 (0.06)	0.09 (0.06)
Spatial EP (t - 1)			0.07 (0.07)				0.09 (0.06)	0.09 (0.06)
Ed. Attainment (t - 1)				0.01 (0.15)				0.24 (0.25)
Log of personal Income PC (t - 1)					-0.05 (0.06)	-0.04 (0.05)	0.03 (0.07)	-0.02 (0.10)
Log of GDP PC (t - 1)	0.05*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.05*** (0.01)	0.04*** (0.01)	0.04*** (0.01)
Log of Pop. Density (t - 1)	-0.02** (0.01)	-0.01 (0.01)	-0.01* (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01* (0.01)	-0.02** (0.01)	-0.02** (0.01)
Num. obs.	250	250	250	250	250	250	250	250
R ² (full model)	0.26	0.25	0.27	0.24	0.25	0.27	0.32	0.33
R ² (proj model)	0.26	0.25	0.27	0.24	0.25	0.27	0.32	0.33
Adj. R ² (full model)	0.24	0.23	0.25	0.22	0.22	0.24	0.29	0.30
Adj. R ² (proj model)	0.24	0.23	0.25	0.21	0.22	0.24	0.29	0.29
Num. groups: year	5	5	5	5	5	5	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.7: State-level regression analyses of EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Inequality level (t - 1)	-0.21 (0.21)					-0.16 (0.21)	-0.24 (0.21)	-0.22 (0.21)
Unemployment (t - 1)		0.14 (0.53)					-0.02 (0.52)	0.06 (0.49)
Spatial Within-EP (t - 1)			0.34*** (0.12)				0.31** (0.14)	0.36*** (0.14)
Ed. Attainment (t - 1)				-0.15 (0.17)				0.27 (0.20)
Log of personal Income PC (t - 1)					-0.10* (0.05)	-0.10* (0.05)	-0.04 (0.07)	-0.09 (0.08)
Log of GDP PC (t - 1)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)
Log of Pop. Density (t - 1)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02*** (0.01)	-0.02*** (0.01)
Num. obs.	250	250	250	250	250	250	250	250
R ² (full model)	0.38	0.37	0.44	0.38	0.41	0.42	0.45	0.47
R ² (proj model)	0.28	0.28	0.35	0.28	0.32	0.32	0.37	0.38
Adj. R ² (full model)	0.36	0.35	0.42	0.36	0.39	0.40	0.43	0.44
Adj. R ² (proj model)	0.26	0.25	0.33	0.26	0.30	0.30	0.34	0.35
Num. groups: year	5	5	5	5	5	5	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B-8: State-level regression analyses of Within-EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Inequality level (t - 1)	-0.17 (0.31)					-0.19 (0.30)	-0.32 (0.29)	-0.31 (0.27)
Unemployment (t - 1)		0.57 (0.65)					1.30** (0.58)	1.31** (0.57)
Spatial Between-EP (t - 1)			0.06 (0.05)				0.10** (0.04)	0.10** (0.04)
Ed. Attainment (t - 1)				0.16 (0.18)				0.06 (0.32)
Log of personal Income PC (t - 1)					0.05 (0.07)	0.05 (0.07)	0.12 (0.08)	0.11 (0.14)
Log of GDP PC (t - 1)	0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)
Log of Pop. Density (t - 1)	0.00 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)
Num. obs.	250	250	250	250	250	250	250	250
R ² (full model)	0.12	0.13	0.13	0.13	0.13	0.13	0.19	0.19
R ² (proj model)	0.02	0.02	0.03	0.02	0.02	0.03	0.09	0.09
Adj. R ² (full model)	0.10	0.10	0.11	0.10	0.10	0.10	0.15	0.15
Adj. R ² (proj model)	-0.01	-0.01	0.00	-0.01	-0.01	-0.00	0.05	0.05
Num. groups: year	5	5	5	5	5	5	5	5

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table B.9: State-level regression analyses of Between-EP as dependent variable. Each regression includes period-fixed effects. Standard errors in brackets. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.