UFPE – Federal University of Pernambuco
CFCH – Philosophy and Human Sciences Centre
PPGCP – Graduate Program in Political Science

A systematic review of micro-level articles on second-order theory and European Parliament elections from 1980 to 2009

Victor Matheus de Santana Santos

Recife - PE
2022

UFPE – Federal University of Pernambuco
CFCH – Philosophy and Human Sciences Centre
PPGCP – Graduate Program in Political Science

A systematic review of micro-level articles on European Parliament elections and second-order theory from 1980 to 2009

Master's Thesis presented as a requirement to obtain the title of Master in Political Science by the Graduate Program in Political Science at the Federal University of Pernambuco, under the advisory of Prof. PhD Rodrigo Barros de Albuquerque.

Victor Matheus de Santana Santos

Recife - PE
2022

Catalogação na fonte
Bibliotecária Maria do Carmo de Paiva, CRB4-1291

S237s Santos, Victor Matheus de Santana.
A systematic review of micro-level articles on second-order theory and European Parliament elections from 1980 to 2009/ Victor Matheus de Santana Santos. – 2022.
108 f. : il. ; 30 cm.

Orientador: Prof. Dr. Rodrigo Barros de Albuquerque.
Dissertação (Mestrado) - Universidade Federal de Pernambuco, CFCH. Programa de Pós-Graduação em Ciência Política, Recife, 2022.
Inclui referências e apêndices.

1. Ciência Política. 2. Eleições. 3. Parlamento Europeu. 4. Partidos políticos. 5. Eleitores. 6. Voto. I. Albuquerque, Rodrigo Barros de (Orientador). II. Título.

320 CDD (22. ed.) (BCFCH2022-035)

UFPE – Universidade Federal de Pernambuco
CFCH – Centro de Filosofia e Ciências Humanas
PPGCP – Programa de Pós-Graduação em Ciência Política

Victor Matheus de Santana Santos

A systematic review of micro-level articles on second-order theory and European Parliament elections from 1980 to 2009

Recife, 18 de fevereiro de 2022.

BANCA EXAMINADORA:

_____

Prof. Dr. Rodrigo Barros de Albuquerque
Supervisor

_____

Profª Dra. Gabriela da Silva Tarouco
Internal Examiner

_____

Prof. Dr. Cairo Gabriel Borges Junqueira
External Examiner

A Deus, por mais esta benção.

Aos meus pais, por todo carinho, dedicação e investimento.

À Marina, por todo amor, companheirismo e estímulo. Obrigado por me dar forças mesmo diante do desânimo.

Ao Rodrigo, pela atenção e dedicação.

AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus por mais esta etapa alcançada. Diante de tudo que aconteceu nesses últimos dois anos, estar vivo e saudável é uma benção. Concluir a dissertação e o mestrado – feito inteiramente à distância e no meio de uma pandemia – já é uma vitória antes mesmo do veredito da defesa.

Agradeço também ao Professor e Orientador Rodrigo Albuquerque, que me acompanha desde a graduação. O senhor me inspira, como pessoa e como profissional. Sua supervisão foi fundamental para a conclusão deste trabalho no prazo e só tenho a agradecer por todo o tempo, atenção e cuidado que o senhor disponibilizou nesses dois anos; não só quanto à dissertação, mas também para comigo.

Aos Professores Gabriela Tarouco e Cairo Junqueira, por prontamente aceitarem o convite para a defesa. Mais especificamente, agradeço à Professora Gabriela, primeiramente pelas contribuições também dadas na qualificação desta dissertação, e segundamente por todo conhecimento que me foi repassado no decorrer do mestrado. Mais uma vez, agradeço novamente ao Professor Cairo, que coincidentemente, também esteve na minha banca de avaliação do TCC, ainda na UFS em 2019.

Ao Professor Dalson Figueiredo Filho, por todas as contribuições – desde a cadeira de Seminário de Dissertação, que me forneceu a base metodológica para a execução deste trabalho, até a qualificação – e por todo o estímulo que me foi dado para chegar até aqui.

Agradeço também, de forma especial, a Ian Batista, meu veterano da graduação e – caso tudo dê certo – meu futuro veterano do doutorado; por ter prontamente me ajudado na execução dos gráficos e dos scripts no Rstudio. Obrigado por todas as dicas e pelo estímulo em explorar essa linguagem, que ainda me deixa um pouco inseguro.

Ao colega Renato Brito, pelos momentos de desabafo e descontração, mesmo que à distância.

Finalmente, agradeço a toda minha família. Em especial, à minha mãe Lêda e ao meu pai Emerson, por ainda acreditarem e investirem no meu potencial, independentemente do que pensem sobre o ensino público superior brasileiro. Aos meus avós Arlette, Leal, Enock e Udjanir; pelos presentes que são em minha vida, independentemente de discordâncias políticas (e novamente, a Deus, pela benção que é ter os quatro ainda vivos e lúcidos para comemorarem mais este momento comigo). Aos meus irmãos Caio e Gabriel, especialmente

Gabriel, por acalmar o “loro” quando eu precisava de silêncio para estudar. Aos meus tios, em especial, meus padrinhos Elisângela e Luís Paulo, por serem referências para mim como pessoas e como brilhantes acadêmicos que são, ainda que em áreas do conhecimento diferentes.

Por último, mas não menos especialmente, à Marina, minha namorada, e por toda sua/nossa família. Você, mais do que todos, acreditou em mim – inclusive mais do que eu mesmo – e esteve comigo em todos os bons e maus momentos nesta jornada. Seja em Aracaju, Recife, São Paulo, ou onde for; sempre me incentivou a buscar o melhor para mim. Sou e serei eternamente grato a ti e ao nosso amor. Obrigado por aturar meus estresses, meu desespero, minha ansiedade e meus áudios de desabafo desnecessariamente longos e prolixos. Aos meus sogros Luciana e Gledson, ao meu irmão de outra mãe Júnior, minha concunhada Julianne e vó Marita. Obrigado por todos os momentos de apoio, amor, alegria e descontração. É uma benção ter vocês em minha vida, especialmente depois de dois anos tão difíceis. Muito obrigado!

ABSTRACT

This research consists of a systematic review of 54 articles published on the second-order elections theory and European Parliament (EP) elections between 1980 and 2009. The second-order elections theory was launched in 1980 by Reif and Schmitt (1980), one year after the first EP elections. This theory holds that second-order elections have lower turnouts, that governing parties usually perform worse, as do larger parties, while smaller or newer and/or radical parties are favoured. The theory assumes two main mechanisms to explain such voting patterns, both of which stem from the assumption that, for parties and voters, there is less at stake in these elections. Because there is less at stake, voters would either choose to vote sincerely – choosing smaller parties closer to their ideologies – or choose to punish or send a message to the governing parties – especially when second-order elections are held in the middle of the electoral cycle. Although these assumptions have been corroborated with aggregate-level data over the last 40 years, some authors have called attention to the scarcity and delay in the development of robust tests of the theory with individual-level data. This review found that tests with this type of data had been carried out since the 80s and 90s. However, most of these first articles with individual-level data did not perform robust statistical tests, as well as did not impact the literature, obtaining a low number of citations. Only at the turn of the 20th-21st century did this flaw begin gaining more prominence on the research agenda, especially between 2007 and 2009. Among other findings, this review also found that there is no publication bias on the studied topic - at least until 2009 - and that most of the authors on the subject are male and affiliated to German, Dutch, British and North American universities. Finally, it is expected that this review will help to better understand the second-order elections theory and to systematize the knowledge accumulated in these first 30 years of research on the subject. We also expect that this review will contribute as an incentive to carry out more systematic reviews in the larger Social Sciences field, demonstrating the importance of this method for the advancement of scientific knowledge and publishing in detail the step by step of our research design, inspired by Cooper (2016) and Figueiredo Filho *et al* (2014).

Keywords: second-order elections, European Parliament, systematic review, individual-level data.

RESUMO

A presente pesquisa consiste em uma revisão sistemática de 54 artigos publicados sobre a teoria de eleições de segunda-ordem e eleições para o Parlamento Europeu (PE) entre 1980 e 2009. A teoria de eleições de segunda-ordem foi lançada em 1980 por Reif e Schmitt (1980), um ano após as primeiras eleições para o PE. Esta teoria afirma que eleições de segunda-ordem têm *turnouts* mais baixos, que partidos governistas normalmente se saem pior, assim como partidos maiores, enquanto partidos menores ou mais novos e/ou radicais são favorecidos. A teoria pressupõe dois principais mecanismos para explicar tais padrões de voto, e ambos decorrem da hipótese de que, para partidos e eleitores, há menos em jogo nestas eleições. Por haver menos em jogo, eleitores optariam por votar de forma sincera - escolhendo partidos menores mais próximos de suas ideologias - ou optariam por punir ou dar um recado aos partidos governistas - especialmente quando as eleições de segunda-ordem são realizadas no meio do ciclo eleitoral. Apesar de tais postulados terem sido amplamente corroborados com dados ao nível agregado ao longo desses 40 anos, alguns autores têm chamado a atenção para a escassez e o atraso no desenvolvimento de testes robustos da teoria com dados a nível individual. Esta revisão descobriu que testes com este tipo de dado já eram realizados desde os anos 80 e 90. Contudo, a maior parte destes primeiros artigos com dados a nível individual não realizava testes estatísticos robustos, assim como também não impactaram a literatura, obtendo um baixo número de citações. Apenas na virada do século XX-XXI esta deficiência começou a ganhar mais destaque na agenda de pesquisa, sobretudo entre os anos de 2007 e 2009. Dentre outros achados, esta revisão também descobriu que não há viés de publicação no tópico estudado - pelo menos até 2009 - e que a maior parte dos autores sobre o tema são do gênero masculino e filiados a universidades alemãs, holandesas, britânicas e norte-americanas. Finalmente, espera-se que esta revisão ajude a melhor compreender a teoria de eleições de segunda-ordem e a sistematizar o conhecimento acumulado nestes primeiros 30 anos de pesquisas sobre o assunto. Esperamos também que esta revisão contribua como um incentivo para a maior realização de revisões sistemáticas na área de Ciências Sociais em geral, demonstrando a importância desse método para o avanço do conhecimento científico e publicizando em detalhes o passo a passo do nosso desenho de pesquisa, inspirado em Cooper (2016) e em Figueiredo Filho *et al* (2014).

Palavras-chave: eleições de segunda-ordem, Parlamento Europeu, revisão sistemática, dados a nível individual.

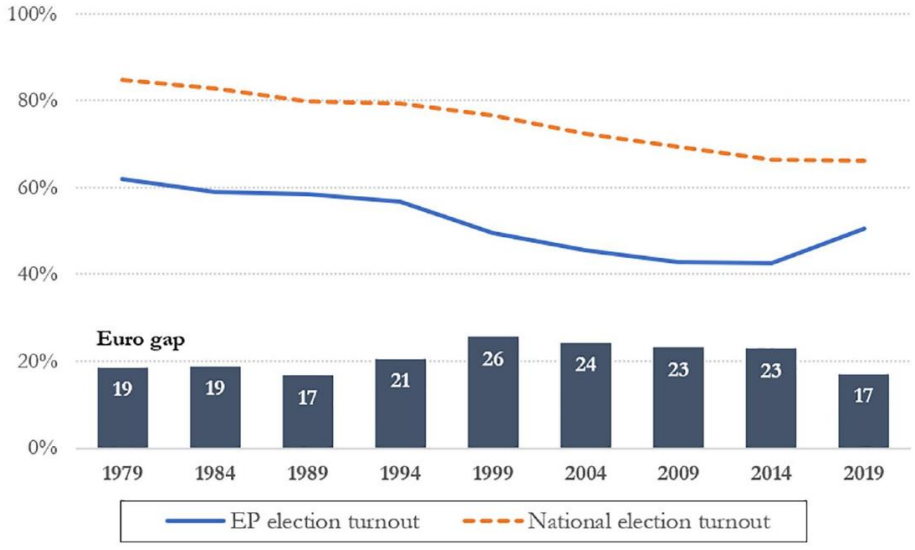
SUMMARY

INTRODUCTION

This research aims to analyse the formal, methodological and substantive profile of published articles that study elections for the European Parliament (EP) as second-order elections, based on a systematic review focused on three decades of research on the subject (1980-2009). The first article published on this topic dates back to 1980, a year after the first election, which took place in 1979. It was written by the Germans Karlheinz Reif and Hermann Schmitt and is entitled "Nine Second-Order National Elections – a conceptual framework for the analysis of European Election results". This article established the first postulates of what has become one of the most consolidated theories on EP elections.

According to Reif and Schimtt (1980) and Reif (1984b), EP elections are second-order ones because, among other factors, they have a low turnout rate, an increase in votes for small and/or opposition parties and are highly tied to the electoral cycle, which means that their results also depend on the distance between them and the first-order elections. First-order elections are the ones to the national parliaments, in the case of parliamentary regimes; and for president, in the case of presidential regimes. One of the most notable differences between these two types of election is the turnout rates. The graph below illustrates the difference between the average turnout rates of the national legislative elections and the EP elections. Schäfer (2021) calls this difference the "Euro gap" and states that it stems from the voters' perception that "there is less at stake", one of the main postulates of the second-order elections theory.

FIGURE 01 - Difference in participation between European and national parliamentary elections over time ('Euro gap').



Source: Schäfer (2021).

Another central proposition of this theory states that these elections are directly influenced by the electoral cycle: if they are held closer to the first-order ones, they tend to confirm the result of the latter; if they are held further away from first-order elections, second-order elections tend to favour the opposition (REIF, 1984b). This phenomenon is mainly due to the "opposition voting" logic. More specifically, the theory predicts that constituents, disappointed with the party they voted for in the previous first-order election, vote for the opposition in the second-order election, with the aim of punishing the governing party or coalition. Opposition voting tends to occur when EP elections are held in the middle of an electoral cycle between two first-order elections, when national government approval is commonly lower (REIF, 1984b). This effect is also stronger in countries with periodic changes in national governments (MARSH, 1998). Furthermore, second-order elections register high abstention rates and a greater number of votes for small and/or new parties. On its term, this other phenomenon also has its own logic, commonly named "sincere voting". Sincere voting tends to occur due to voters' perception that there is less at stake in EP elections, since the political arena considered to be the main or determining one is the national arena. Consequently, voters would not choose their votes based on strategic calculations such as which parties are most likely to win and form a government, which would benefit bigger parties. In this sense, these voters tend to choose smaller or more radical parties, which are closer to their opinions and interests.

These characteristics were present both in 1979 and on later elections, and several studies have already tested and confirmed these hypotheses, whether treating the elections as a whole (REIF, 1984b; EIJK, FRANKLIN & MARSH, 1996; REIF, 1997; NORRIS, 1997; MARSH, 1998; SCHMITT, 2005; HIX and MARSH, 2011; SCHMITT & TEPEROGLOU, 2015; SCHMITT & TOYGÜR, 2016), or by analysing some countries separately (FREIRE, 2004; SKRINIS & TEPEROGLOU, 2008; ERVIK, 2012; LAŞAN, 2020). Second-order election theory has also been tested with other types of elections, such as local (RALLINGS & THRASHER, 2005; MARIEN, DASSONNEVILLE & HOOGHE, 2015) and provincial/regional elections (FAUVELLE-AYMAR & LEWIS-BECK, 2011; MÜLLER, 2018), presidential elections in strong parliamentary systems (FORTES & MAGALHÃES, 2005; ELGIE & FAUVELLE-AYMAR, 2012), and referendums (ERKAN, 2010; GLENCROSS & TRECHSEL, 2011). Furthermore, it crossed geographical borders, as it was corroborated or refuted in several non-European cases such as Canada (CUTLER, 2008; THORLAKSON, 2015), Latin America (CARRERAS, 2017), Cabo Verde (DELGADO, 2017), Japan (IMAI & HINO, 2012) and Israel (NACHMIAS, ROSENTHAL & ZUBIDA,

2012). On the other hand, although the term “second-order elections” was coined by Reif and Schmitt in 1980, their work was inspired by previous studies of midterm elections in the United States (CAMPBELL, 1960; TUFTE, 1975; STIMSON, 1976) and provincial elections in Germany (DINKEL, 1977).

Despite having withstood the test of time, since more recent works corroborated once again the main postulates of the theory for the last elections held in 2019 (EHIN & TALVING, 2021), the second-order elections theory also has suffered criticisms and shortcomings. Hobolt and Wittrock (2011) question the distinction between issues relating to Europe and issues relating to each member countries’ national politics, since the European Union (EU) now deals with diverse policies and has a wide range of competencies. Schmitt, Sanz and Braun (2008) state that it was only a short time ago that researchers and academics started to analyse the theory with individual level data in a more consistent way.

Regarding this last issue, Schmitt *et al* (2020) claim that postulates at the aggregate level of the theory received more attention in the decades following the launch of the seminal article. However, key assumptions for the theory, such as the opposition voting and the high abstention rates due to the voters’ perception that there is little at stake, presuppose voters will behave in a certain way as political actors depending on the electoral arena which they will vote for. Although tests at the aggregate level indirectly corroborate these assumptions, according to Schmitt *et al* (2020), consistent tests at the individual level continued to be a gap in the research agenda until the first decade of the 21st century, as more comprehensive tests of these postulates are scarce and recent. Even at the end of the 1990s, Reif, (1997) and Norris (1997) already drew attention to this scenario and highlighted the importance of new studies and tests that used transnational surveys in their analyses. Despite not aiming to carry out an expressive literature review on the subject, the oldest work cited by Schmitt *et al* (2020) to perform tests at the individual level is a 2005 paper by Carruba and Timpone.

Other authors, such as Clark and Rohrschneider (2009) and Hobolt and Wittrock (2011), point out that the focus on tests at the aggregate level of previous works can run into problems such as the ecological fallacy. Briefly, the ecological fallacy occurs when data at the aggregate level of analysis are misinterpreted with respect to inferences at the individual level of analysis, leading to invalid inferences and correlations. This statistical paradox was first demonstrated by Robinson (1950) in an article referring to illiteracy among different social groups with United States (USA) census data. Robinson (1950) demonstrated that although data at the aggregate level indicated a strong correlation between Afro-descendance and

illiteracy (0.77), data at the individual level indicated that this correlation was considerably lower (0.20). With respect to the USA-born group, the difference was even greater: whereas data at the aggregate level indicated a negative correlation (-0.53), data at the individual level indicated a weak positive correlation (0.12). Although this statistical paradox was first described by Robinson (1950), it was Selvin (1958) who coined the term "ecological fallacy".

Despite Robinson's (1950) emphasis that data at the aggregate level should not be used to make inferences at the individual level, sometimes there is no data available at the individual level of analysis, or such data are fuzzy and biased. This problem has led to a series of studies on how to develop more robust ecological inferences, notably King, Rosen and Tanner (2004) and Liu (2007). Furthermore, some authors reaffirm the importance of studies that also take multilevel context into account, such as Subramanian *et al* (2009). In turn, Idrovo (2011) brings a more specific definition of ecological fallacy, listing three essential criteria for it to occur: data must be obtained at the aggregate level of analysis, these data must be used as a basis for inferences regarding individuals and the results obtained with other data at the individual level must be contradictory. Thus, Idrovo (2011) states that when the focus of the study is contextual or when data at the individual level do not exist, it is not possible to state that there is an ecological fallacy, being more appropriate to recognize that there is a possible relationship and that additional studies are needed.

As far as second-order election theory is concerned, some more recent works that use data at the individual level mention that too much focus on testing the theory with data at the aggregate level may incur in an ecological fallacy, such as Clark and Rohrschneider (2009), Hobolt and Wittrock (2011), Giebler and Wagner (2014) and Linek and Lyons (2007). However, in this small group of four works, only one presented considerably discrepancies when comparing the results obtained from individual level with the ones obtained from aggregate level data (LINEK and LYONS, 2007). On the other hand, Clark and Rohrschneider (2009), Hobolt and Wittrock (2011) and Giebler and Wagner (2014) got closer results to those obtained previously from aggregate level data. Given this scenario, it is not possible to state categorically that second-order elections theory is a clear case of ecological fallacy. From this perspective, it is necessary to compare the results of more studies with tests at the individual level. In addition, it is also important to check if there are works published before 2005 that have tested the theory with data at the individual level, one of the aims of this research.

There is also a shortage of comprehensive literature reviews on the subject. Between 1980 and 2010, only two major reviews were found. The first one was published in

1996 by Eijk, Franklin and Marsh, while the second one was published by Marsh and Mikhaylov in 2010. However, to the best of our knowledge, there are still no systematic reviews published on this topic, despite more than forty years of evolution and refinement of the theory.

Thus, this master thesis aims to systematically review previously published articles on EP elections as second-order ones since the Reif's and Schmitt's seminal article published in 1980 until 2009, one year before the release of Marsh's review, exposing the main research themes in each electoral round and how the agenda has evolved in the first thirty years of study. This work also aims to qualitatively analyse the research agenda dedicated to testing the theory with data at the individual level published until 2009, firstly seeking to verify the existence of older works dedicated to this type of test, as well as to demonstrate which were the limits that prevented or hindered the previous realization of these tests in a more robust way. Furthermore, this work also advocates for a greater execution of systematic and meta-analytic reviews in the Humanities and Social Sciences field. To this end, in addition to demonstrating step-by-step how this review was done and the benefits of conducting a review in a systematic way, we will also compare our results with what was written and published by Eijk, Franklin and Marsh (1996) and Marsh and Mikhaylov (2010), who opted for a narrative review of the theory.

This systematic review was designed according to Cooper's (2016) methodology, with a few adaptations due to the breadth of our object of study. The choice of the variables adopted here and their subsequent categorization was also inspired by Sampaio's and Figueiredo Filho's (2019) article on campaign financing for municipal elections in Brazil and adapted to our topic of interest.

In this way, we chose to detail the methodological procedures adopted here in the first chapter, for two main reasons. Firstly, Figueiredo Filho *et al* (2014) highlight that, despite their importance and usefulness for the accumulation of knowledge and for the progress of science in general, systematic and meta-analytical reviews are not usual methods in Human and Social Sciences. There is a preference for narrative-literary reviews, which leads to a low systematization of accumulated knowledge and low reliability of their results. In this sense, this master's thesis seeks to be part of this effort to raise awareness and propagate more reliable and replicable review methods, especially in the field of Social Sciences in Brazil, which is admittedly limited in the adoption of robust scientific methods, despite recent advances (SOARES, 2005; NEIVA, 2015).

The following three chapters will focus on exposing the results of the systematic review, which, as will be detailed in the first chapter, had its data divided into three main dimensions. The first dimension is the formal one, which includes variables such as the title of the article, authors, genre, year of publication, author's affiliation and country of origin, language, whether the article cites the seminal article, number of citations, etc. This dimension has a total of 44 variables. The data and results related to this dimension will be broken down in the second chapter.

In turn, the second dimension has 26 variables, and encompasses the methodological aspects of the literature, such as the selected articles' research questions, whether the article has an abstract or not, studied hypotheses, year of the studied elections, types and numbers of methods used and level of analysis. This last variable is of special interest, as it will also be analysed qualitatively. The data and results related to this dimension will be broken down in the third chapter, which also includes a brief summary of the development of the theory regarding the methods and data adopted, demonstrating how the research agenda evolved over time.

Subsequently, the third dimension is the substantive one, which comprises 24 variables. In this dimension, aspects such as dependent and independent variables adopted by the collected articles will be broken down, as well as their control variables, sample size, if these first works mention the issue of ecological fallacy, results, number and proportion of supported hypotheses, etc. This last variable is also one of the steps established by Cooper (2016) in the development of a systematic review. All these aspects will be reported in the fourth chapter. Finally, at the conclusion, a brief review of the work will be made, as well as additional comments on the limits of this master thesis and suggestions for future research.

It is clear, then, that this master's thesis aims to contribute to the research agenda on the theme of EP elections as second-order ones on several fronts. Firstly, as highlighted above, systematic and meta-analytic reviews are extremely important for the advancement of scientific knowledge, enabling more reliable and replicable results (FIGUEIREDO FILHO *et al*, 2014). In this sense, this research intends to point out the results already obtained in the area in the first thirty years of theoretical and methodological development, with due scientific rigor inherent to the systematic review method.

Secondly, it is recognized that there is some popular ignorance of the importance of the EU institutions and how they work (MARIANO; LUCIANO & BRESSAN, 2016). Even in academia, according to Costa and Brack (2013), the EP has been the most underrated institution compared to the other three main institutions of the EU. On the other hand, the

second-order elections theory provided explanations and interpretative keys for various contexts and facts experienced by Europe throughout integration, such as the low recurrent turnout, less investment in party campaigns and even the recent advance of Eurosceptic parties in the second decade of the 21st century. Following this logic, allied to other research agendas and relevant themes, such as the democratic deficit in the EU, this research is also interested in expanding the scope of the theory to a more diversified public, demonstrating how electoral dynamics work in the case of the EP, either for academics outside this research agenda or for a non-academic audience. In addition, understanding and knowing how these dynamics work in another integration process can help to understand and establish bases for further comparisons with other similar institutions in different regions, such as Parlasul of the Southern Common Market (MERCOSUR), bringing contributions to the regional integration agenda more generally.

Finally, this research has the potential to contribute to the EU studies agenda in Brazil. A search in the SciELO database with the terms "Parlamento Europeu" [European Parliament in Portuguese] found 25 articles, five of which correspond to works in Portuguese on various topics related to the EU published by Brazilian authors between 2009 and 2019.[1] A search in the same database with the terms "eleições segunda ordem" [second order elections in Portuguese] shows that among the only three articles found on the topic, all were published in Portuguese journals and were written by Portuguese authors.[2] This dissertation will be able to introduce this research agenda in Brazilian academic circles and also help to place Brazil on the map of potential internationally relevant studies on the subject.

[1] Search available at: https://search.scielo.org/?q=&lang=pt&count=15&from=1&output=site&sort=&format=summary&fb=&page=1&q=Parlamento+Europeu&lang=pt. Accessed in: December 3rd, 2021.
[2] Search available at: https://search.scielo.org/?q=&lang=pt&count=15&from=0&output=site&sort=&format=summary&fb=&page=1&q=elei%C3%A7%C3%B5es+segunda+ordem&lang=pt&page=1. Acessed in: December 3rd, 2021.

CHAPTER 01: Methodology and Research Design – a Systematic Review on EP Elections as Second-Order ones

1.1 On the Method: Systematic and Meta-Analytical Reviews and their Importance to Science

As stated in the previous section, despite their importance and usefulness for the development and accumulation of scientific knowledge, systematic and meta-analytic reviews are still little applied in Humanities and Social Sciences, fields traditionally more used to narrative-literary reviews. Despite giving the researcher greater freedom, this latter type of review does not adopt specific and standardized techniques. This can lead to a restricted or biased analysis of the subject of interest, either during sample collection or when publishing review results. Figueiredo Filho *et al* (2014) state that one of the most recurrent problems in non-systematic reviews is the convenient selection of the sample, usually selected considering the accessibility of the chosen texts or the author(s)' prior knowledge about them.

Another common problem in non-systematic reviews, also according to Figueiredo Filho *et al* (2014), is the difficulty – or even the impossibility – in replicating them. This in turn leads to reliability problems. Replication is an extremely important process for the advancement of science, being a fundamental component of the production of scientific knowledge (KING, 1995). Rocha *et al* (2013) describe three important benefits of replication. In addition to what has already been mentioned regarding the accumulation of knowledge, replication can also be used in a pedagogical way in the learning of new researchers, as well as protecting the scientific community from errors, fraud and manipulation. All these aspects reinforce the need for standardization of scientific review methods.

Two types of review that lend themselves to this role are the systematic review and the meta-analytic review. As for the former, Denyer and Tranfield (2009) define it in the following terms:

> Systematic reviews are a specific methodology that locates existing studies, selects and evaluates contributions, analyses and synthesises data, and reports the evidence in such a way that allows reasonably clear conclusions to be reached about what is known and what is not known (DENYER & TRANFIELD, 2009, p. 672).

As for the meta-analytical review, Glass (1976) defines it as:

> […] the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings. It connotes a rigorous alternative to the casual, narrative discussions of research studies

> which typify our attempts to make sense of the rapidly expanding research literature (GLASS, 1976, p. 03).

It is noticed, then, that the meta-analysis has a broader scope compared to a systematic review. In this sense, meta-analytic reviews aim not only to systematize the knowledge already obtained by previous studies, but also to synthesize its conclusions in a methodologically consistent way (ROSCOE & JENKINS, 2005), aiming to resolve disputes in the literature (IMBEAU *et al*, 2001) or estimate whether a particular method influences a particular pattern of results (FIGUEIREDO FILHO *et al*, 2014).

Although some authors treat systematic and meta-analytic reviews as synonyms (FIGUEIREDO FILHO *et al*, 2014), other authors consider that they both are separate parts of the same process. Although this distinction is well established in Medical and Health Sciences, a field in which systematic review methods are already consolidated, some authors in Social Sciences and Humanities have also started to differentiate the two types of reviews, such as Petticrew and Roberts (2006) and Dacombe (2018).

According to Dacombe (2018), meta-analysis is a proceeding done after systematic review. It is not possible to carry out a meta-analysis without a prior systematic review of the literature. However, not every systematic review is accompanied by a meta-analysis, either because of the authors' research interests or because of the research object itself. Since meta-analyses consist of the quantitative synthesis of results from different parameterized studies, not every research object favours or allows this type of analysis.

This may also help to explain the resistance of academics and researchers in the fields of Humanities and Social Sciences to adopt meta-analysis as a recurrent method in their research. As we deal with social phenomena, in which it is often difficult or even impossible to quantify and measure them in a scientifically adequate way, meta-analysis is not always perceived as a viable methodological alternative. Sometimes, there are even quantitative studies on a certain phenomenon, but the great difference in the methods adopted, in the cases analysed or even in the way of measuring the object of study prevents a good parameterization of the results. This is precisely the case of this research, which aims to collect and analyse the results of 30 years of published articles on a theory.

However, it is emphasized here that meta-analysis is only the last step in a systematic review process. Although quantitative synthesis is not always feasible, a rigorous standard for the collection and selection of studies, as well as the analysis of their results, is always possible, even in the face of complex and abstract objects of study such as comprehensive social phenomena and theories of human behaviour. This is precisely why we

believe that there is still a lot of room in the academic literature in Humanities and Social Sciences for the dissemination and advancement of systematic review methods. To the best of our knowledge, for instance, there is no systematic review on second-order elections theory so far, despite more than four decades of study and refinement of the theory.

Even so, it is not in our interest to present a historical overview of the evolution of systematic and meta-analytic review methods in the Social Sciences. This work has already been done by Dacombe (2018). Therefore, in the next subsection, the step-by-step of a systematic review and how this step-by-step was applied in this research are followed. The practical problems we faced while carrying out the research will also be reported, as well as its limitations.

1.2. A Step-by-Step Guide on Systematic Reviews and our Adapted Research Design

As stated earlier, we used the step-by-step set by Cooper (2016) in carrying out a systematic review, set out in seven main steps. Cooper (2016) summarize these steps in a table partially reproduced below.

The first step established by Cooper (2016) in carrying out a systematic review is the identification of the research problem, which, in fact, is the first step of any scientific method. Figueiredo Filho *et al* (2014) emphasize that, if the objective is a meta-analysis, it is necessary that the results of the works on the research object can be quantitatively synthesized, as described in the subsection above. It is also important to highlight what the research objectives are. Resolve any dispute in the literature? Summarize the knowledge already obtained on the subject so far? Detail methodological aspects or report the historical evolution of the research agenda on the topic of interest? Such aspects should be considered in this first phase, since the research design may need to be adapted according to pre-established objectives. As far as this work is concerned, such aspects have already been reported in the introduction.

The second step established by Cooper (2016), which is actually the first practical step of this method, is the collection of literature. Are you going to work with the population or with a sample of it? What will be the sampling criteria? Where do the collected studies come from? What types of sources are being researched, and why? Such questions must be answered at this stage, in order to guarantee the transparency, standardization and replicability of the review to be carried out.

Table 1 – Research Synthesis Conceptualized as a Research Project

Step	Research question asked at this stage of the synthesis	Primary function Served in the synthesis
1 Formulating the problem	What research evidence will be relevant to the problem or hypothesis of interest in the synthesis?	Define the (a) variables and (b) relationships of interest so that relevant and irrelevant studies can be distinguished.
2 Searching the literature	What procedures should be used to find relevant research?	Identitfy (a) sources (e.g., reference, databases, journal(s) and (b) terms used to search for relevant research.
3 Gathering information from studies	What information about each study is relevant to the problem or hypothesis of interest?	Collect relevant information about studies in a reliable manner.
4 Evaluating the quality of the studies	What research should be included in the synthesis based on (a) the suitability of the methods for studying the synthesis question, and/or (b) problems in research implementation?	Identify and apply criteria that separate studies conducted in ways that correspond with the research question from studies that do not.
5 Analyzing and integrating the outcomes of studies	What procedures should be used to summarize and combine the research results?	Identitfy and apply procedures for (a) combining results across studies and (b) testing for differences in results between studies.
6 Interpreting the evidence	What conclusions can be drawn about the cumulative state of the research evidence?	Summarize the cumulative research evidence with regard to its strength, generality, and limitations.
7 Presenting the results	What information should be included in the report of the synthesis?	Identify and apply editorial guidelines and judgement to determine aspects of methods and results readers of the report will need to know.

Source: Partial reproduction of table 1.2. by Cooper (2016, p. 42-45).

In the case of this research, it was restricted to articles published between 1980 and 2009 in English, Spanish, French and Portuguese. This time frame chosen covers the entire period of academic production on second-order elections theory – since the seminal article by Reif and Schmitt, released in 1980 – until the publication of the second major review of the subject, published by Marsh and Mikhaylov in 2010. We also chose to collect articles in languages other than English for two main reasons. First, it is possible that specific research agendas on the subject have emerged among authors and research groups restricted to one of these languages. Second, it is also possible that there are interesting case studies of EP elections as second-order elections restricted to those languages.

As for the types of publications analysed, it was decided to restrict the collection to articles published in academic journals. Despite the notorious relevance and impact of

some books on the topic of interest (REIF, 1984a; OPPENHUIS, 1995; EIJK & FRANKLIN, 1996; BLONDEL, SINNOTT & SVENSSON, 1998; PERRINEAU, GRUNBERG & YSMAL, 2002; LODGE, 2005; MAIER, 2006; EIJK & BRUG, 2007; MARSH, MIKHAYLOV & SCHMITT, 2007; NIELSEN & FRANKLIN, 2016; VIOLA, 2016; KRITZINGER *et al*, 2020), this type of material proved unfeasible for collection due to its low availability online in a complete and free form, especially in the case of books from the 1980s and 1990s, whose availability in full online format for free is even rarer. Therefore, it was decided to exclude this type of material before data collection, even knowing that important contributions would not be included in the final sample.

In turn, monographs, master theses and dissertations were included in the collection stage. However, one problem was noticed while carrying out this stage: a somewhat considerable part of these works was indexed in restricted access repositories, making the equalized collection of this type of data unfeasible. Therefore, as was done with books, these types of publications were also excluded from the final sample, but were coded in a preliminary exclusion sample.

With regard to working papers, this type of publication was also codified in a preliminary sample, but did not enter the final sample. Despite the notorious importance and impact of some working papers with theory tests with data at the individual level (SCHMITT & EIJK, 2002; MANOW, 2005; CLARK & HUSLEY, 2007; SCHMITT, SANZ & BRAUN, 2008; WEBER, 2009); it was not possible to include such works in the final sample, firstly, due to the large amount of texts and data to be read, collected and analysed. As will be described later, the final sample consists of 54 published articles, from which 96 variables were collected. Such articles were also read in full for data collection. If working papers were included in the final sample, at least 28 more working papers would be added, something that unfortunately was not possible to do in time for the conclusion of this research. Another five working papers were subsequently published as articles until 2009, and entered the final sample as published articles. A second minor problem was that a little number of working papers did not inform where they were presented. This made it impossible to collect some relevant variables for the analysis, in addition to not making it clear whether the work was a working paper presented at an academic event or a mere draft of a work to be published as an article or book chapter. Therefore, we chose to not code these working papers, to avoid biasing the sample of works that did not make it into the final sample and to prevent honest coding errors.

Finally, regarding the means of collection, two databases were chosen. The main one for carrying out the research was the Scopus online repository, which contains mostly published articles and some indexed books and book chapters. In Scopus, files were collected in two different ways. First, the list of indexed works that cited the seminal article by Reif and Schmitt (1980) was followed. This article had an impressive number of 1376 citations on the Scopus database as of December 18th, 2021 (including books and book chapters), of which 313 were published until 2009. Secondly, searches were carried out with the terms "second-order", "second-order elections", "European elections" and "European Parliament" in different positions, with quotation marks and with the Boolean operators AND and OR. After some tests, we opted for the combination of the following terms: "'second-order' AND 'second-order elections' OR 'European Parliament' OR 'European elections'", which provided 286 works as of December 18th, 2021, of which 56 were published until 2009. It was noticed that most of them were already indexed in the first list, even though a few important articles did not cite Reif and Schmitt (1980). These first results demonstrate right away that the research agenda on second-order elections has expanded and flourished enormously in the last 12 years, which account for more than three-fourths of all works indexed in Scopus that cite the seminal article or mention the topic in their title or in their abstract. A more detailed analysis of this finding will be made on the conclusion.

As there is the possibility that some articles are not indexed in Scopus, it was also decided to search in Google Scholar. On this database other types of work that did not enter the final sample were also found, such as working papers, master theses and dissertations. In this second platform, the terms "second-order elections" and "European Parliament" were searched with quotation marks but without Boolean operators in between them. Searches were also carried out with such terms translated into Spanish, French and Portuguese. Due to the immense amount of works found, in this database the searches were filtered by year of release. Our search obtained 3579 entries (between articles, working papers, drafts, monographies, master theses and dissertations) published since 1980, as of December 18th, 2021. A more detailed analysis of those first search results will follow on the next sub-section.

Due to the broad search pattern and the large temporal scope – even with the restriction until 2009 –, some exclusion criteria were adopted at the stage of literature collection. First, articles from journals that do not occupy the first and second percentiles of the Scimago Journal Rank (SJR) in their scientific area for at least half of the years in which the journal was evaluated were excluded; in addition to journals that eventually do not adopt

the blind peer review procedure by the year 2020. This criterion was adopted as a way to preserve the quality of the works already at the time of collection. The SJR was chosen since it is a journal classification list of free and open access, which includes a greater number of European journals than the Journal Citation Report list in the area of Political Science and International Relations.

A second important exclusion criterion refers to the fact that not all articles that cite the seminal article by Reif and Schmitt (1980) have the second-order elections theory as a central or highlighted aspect in their analyses. In turn, other works analyse this theory, but do not focus on EP elections, analysing other types of elections and/or non-European cases, as shown in the introduction. Thus, articles that only cite the seminal article *en passant* and that do not bring more substantial discussions or analyses about the theory were not included. Likewise, articles that do not contain at least one case of EP election were also not included in the final sample.

Following the execution of the systematic review, the third step according to Cooper (2016) consists of collecting information about the studies. This collection of information takes place from a list of variables of interest to the author, which must be specified in advance. According to Figueiredo Filho *et al* (2014), the collection of as many variables as possible is recommended, as the monetary and logistical cost of a new collection is generally prohibitive.

In fact, in our personal experience, while collecting information about the texts, it was noted that some relevant variables had not been previously listed. As the collection of such information took place from the reading of the articles and manually coding in Excel coding sheets, the addition of new variables delayed the expected progress of the research, since, for each variable added, it was necessary to return to the articles already catalogued in search of the new data of interest. This problem, however, did not affect the progress of the research as a whole, although this step took longer than expected to finish. Thus, the list of the variables of interest must be well designed and well thought out before this step.

As for this study, the choice of the selected variables, as highlighted in the introduction, was inspired by Sampaio and Figueiredo Filho's (2019) article on local election campaign financing in Brazil and adapted to the topic of interest. 95 variables were collected, divided into three dimensions (formal, with 44 variables[3]; methodological, with 26 variables;

[3] The large number of variables in the formal dimension is justified by practical and counting issues. In works with more than one author, it became impossible to code the two or more authors in the same cell. The same

and substantive, with 24 variables), already briefly described in the introduction. For reasons of space and textual fluidity, the tables with the detailed variables for each dimension and the way they were coded are available in appendices A, B and C after the bibliographical references. More details about these variables will also be described in the following chapters, dedicated to each one of the dimensions. In addition to these 94 variables, a last nominal variable dedicated to observations or comments on the selected articles and/or some variations in their coding was also included in the database.[4] Finally, the collection of data was done in a chronological way, following a year-by-year basis, to facilitate the chronological analysis of the collected information and results.

The fourth and next step, according to Cooper (2016), is the assessment of the quality of the selected material. This step is important as it prevents studies with questionable methods or unreliable results from entering the final sample. Cooper (2016) describes in detail different methods of evaluating the quality of selected works. It is noteworthy that such methods are more easily applicable when the topic of interest studied in the analysed works is measurable or when the analysed researchers use methods that allow interventions such as experiments, something that is not always possible in Humanities and Social Sciences studies.

In the specific case of this work, another potential problem is that we intend to analyse the evolution of the research agenda on European elections as second-order elections. For this, it is necessary to take into account debates and criticisms between authors, as well as initial works whose methodologies were improved later. Therefore, this step was adapted for the purpose of the research. The adoption of a selection criterion based on the impact classification of academic journals during the data collection also served as a means of indirect control of the quality of the selected material.

The fifth step is the analysis and synthesis of studies in the case of a meta-analysis. Because this work is a systematic review, and also because of the broad scope of the research object, which encompasses studies that are not strictly empirical, it was decided at this stage to calculate the frequency of findings in the expected direction of the selected articles, based on the null-hypothesis rejection parameter. For this, the proportion of supported hypotheses in each work were verified to check, for example, if there is publication bias.

---

procedure was followed for the universities of origin of each of the authors, with their respective cities and countries coded in different variables for each one of them, which already add up to 20 different variables.

[4] For example, in the case of Rabier (1984), the author of the text was an employee of the European Commission, and this information was entered in the variable "University of the first author"; despite the fact that the European Commission is not a university. Thus, an observation was also left in the variable "Observations", to clarify this and other possible variations in data coding.

In turn, the sixth step consists of interpreting the collected data. This stage is central, as it is in this stage that the researcher must make inferences from the data and results aggregated in previous works. As highlighted earlier, in this master's thesis, the evolution of studies on the topic of interest will take place from a methodological standpoint, focused on the variable "level of analysis".

Finally, the seventh and last step according to Cooper (2016) is the presentation of the research results, which, in this case, took place in the format of a master's thesis. Although the Meta-Analysis Reporting Standards (MARS) model is recommended (FIGUEIREDO FILHO *et al*, 2014; COOPER, 2016), it was decided to divide the disclosure of the research results into three different parts, according to the categorization of the variables into each of its dimensions. We chose to not follow MARS firstly because a master's thesis is traditionally divided into chapters, and secondly, because of the comprehensive amount of data collected.

Finally, it was possible to identify three gaps in our research design beforehand, even before its execution. Two of them have already been described. The first one concerns the availability of materials such as books, monographs, dissertations and theses online. Consequently, important books that impacted the research agenda, some of them with tests with data at the individual level, had to be excluded. Secondly, due to limited research execution time, working papers also had to be excluded from the final sample. Finally, the second-order elections theory was initially named and tested by two Germans, and Germany remains one of the central countries in the European integration process. Thus, it is possible that there is an important research agenda on this topic developed in the German language, which will not be covered in this review.

In the next sub-section, it will be shown how the final *n* of 54 articles was reached. In addition, a brief descriptive analysis of the works that did not enter the final sample will be carried out.

1.3. Collection of Literature: Descriptive Analysis of Exclusion Procedures and Excluded Works

Conducting a systematic review involves a rigorous control, not only of the texts that will be part of the final sample, but also of those that did not pass the exclusion criteria. For transparency and bias testing purposes of the exclusion criteria adopted in this review, a list of works that did not enter the final sample was drawn up, as well as a few variables of interest about these works were also collected.

In more detail, the group of excluded works can be divided into two main subgroups. The first group consists of those works that were promptly excluded, and whose data were not collected even for the excluded works sample list. Due to the immense number of entries found in both databases and due to the manual collection of data, it became impracticable to collect data from so many works, even for the small amount of variables restricted to the excluded works sample list, which from now on, will be called "second list". Thus, the following cases did not enter the second list and, therefore, did not have any data collected:

- Books or book chapters;
- Book reviews;
- Articles that eventually did not mention the term "second-order elections" in the text, or that only mentioned it in the footnotes;
- Works that were published in the same year as a working paper and as an article, to avoid duplication (in this case, data were collected only from the published article);
- Works that were not available in PDF or html format;
- Drafts that did not indicate they were presented at academic events;
- Due to an unknown error in the Google Scholar Search engine, some works published outside the time frame we searched in appeared in the results, which we did not consider in our population (for example, when the search engine was restricted to the year 2000, but one of the works found by Google Scholar was published or presented in 2007).
- Entries that were not written in English, French, Portuguese or Spanish.

In turn, all other cases of articles that did not make it into the final sample were catalogued in the second list; together with working papers, monographs, master's theses and doctoral theses. For this list, two variables related to the exclusion criteria violated by each of these studies were elaborated. One variable encodes the exclusion criteria in categorical form (1-14) and another encodes them in nominal form (in words). The 14 exclusion categories adopted for the second list were as follows:

- (1) The article does not have the second-order elections theory as its main or central aspect, despite citing the seminal article;
- (2) The article does not have the second-order elections theory as its main or central aspect, not citing the seminal article;
- (3) The article does not deal with any case of EP elections in detail, despite citing the seminal article;

- (4) The article does not deal with any case of EP elections in detail, not citing the seminal article;
- (5) The article cites the seminal article, but does not have the theory as a central or main aspect, nor does it study any case of EP elections in detail (1 and 3 combined);
- (6) The article does not cite the seminal article, and does not have the theory as a central or main aspect, nor does it study any case of EP elections in detail (2 and 4 combined);
- (7) Monograph;
- (8) Master's thesis;
- (9) Doctoral thesis;
- (10) The journal in which the article was published did not reach the first two percentiles of the Scimago Journal Ranking (SJR) (quality proxy);
- (11) The journal in which the article was published is not registered in the SJR, making quality assessment impossible;
- (12) Working paper that entered the final sample as a published article;
- (13) Article is a translation of an original work;
- (14) Working papers in general.

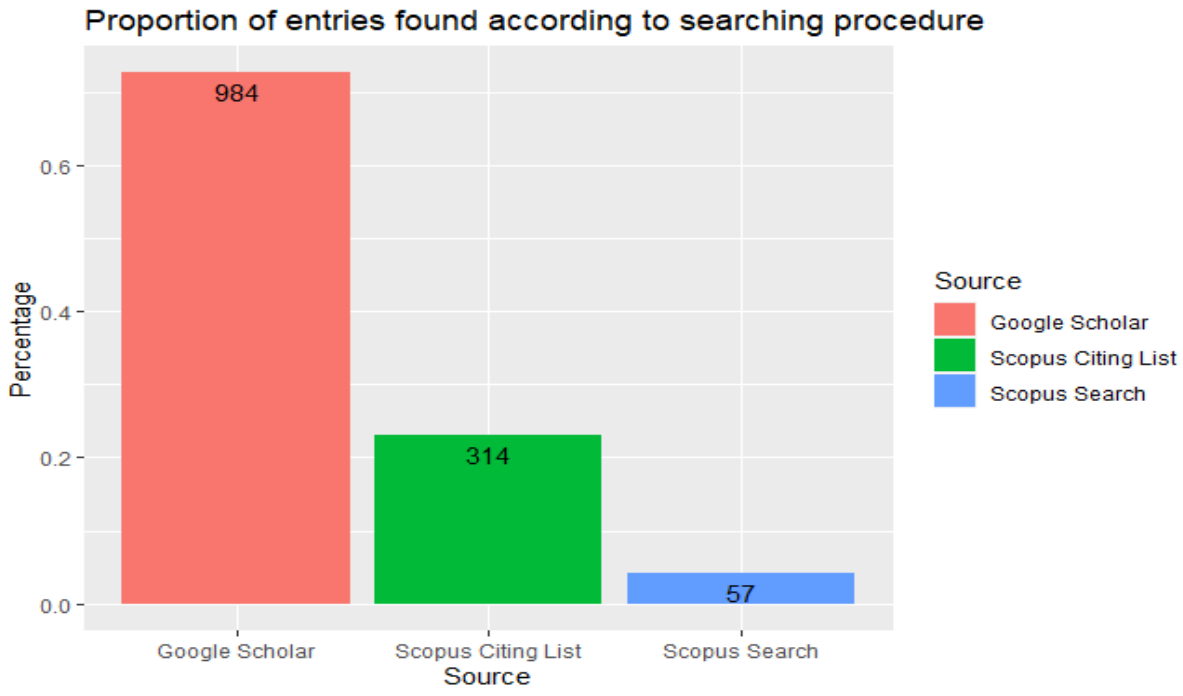
Thus, as of December 18th, 2021, 1355 entries were obtained in the three searches performed in the two databases considering the time frame of 1980 until 2009. Of these, 314 were obtained from the list of works citing the seminal article registered in Scopus; 57 were obtained from the second search in Scopus and another 984 works were obtained from the searches carried out in Google Scholar in the four different languages selected previously. However, it is noteworthy that a great part of these entries are duplications. Several duplications were verified: between the two Scopus lists, between the Scopus lists and the Google Scholar search results, and even within Google Scholar's own searches. A good part of these duplications was expected, as – as will be demonstrated in the following chapters – most of the relevant literature cites the seminal article, in addition to mentioning the term "second-order elections" recurrently, which makes these works stand out in Google Scholar's search algorithms.

However, there were cases where we found the same work mentioned more than once in Google Scholar's results, and we also found works in other languages in searches which the key terms were in English. Works were found in several languages, ranging from German to Greek, Russian, Turkish and even Japanese. If, on the one hand, the existence of works in such different languages shows the impressive geographic reach of this theory, on the other hand, Google Scholar proved to be a search tool of dubious precision, which should

be used with great care and attention. Consequently, some works in French, Spanish and Portuguese were found twice on Google Scholar, firstly in the search with the key terms in English, and again in the search with the key terms in their respective languages. Some of them were also found previously on Reif and Schmitt's (1980) citing articles list on Scopus, appearing at least three times on all searches performed.

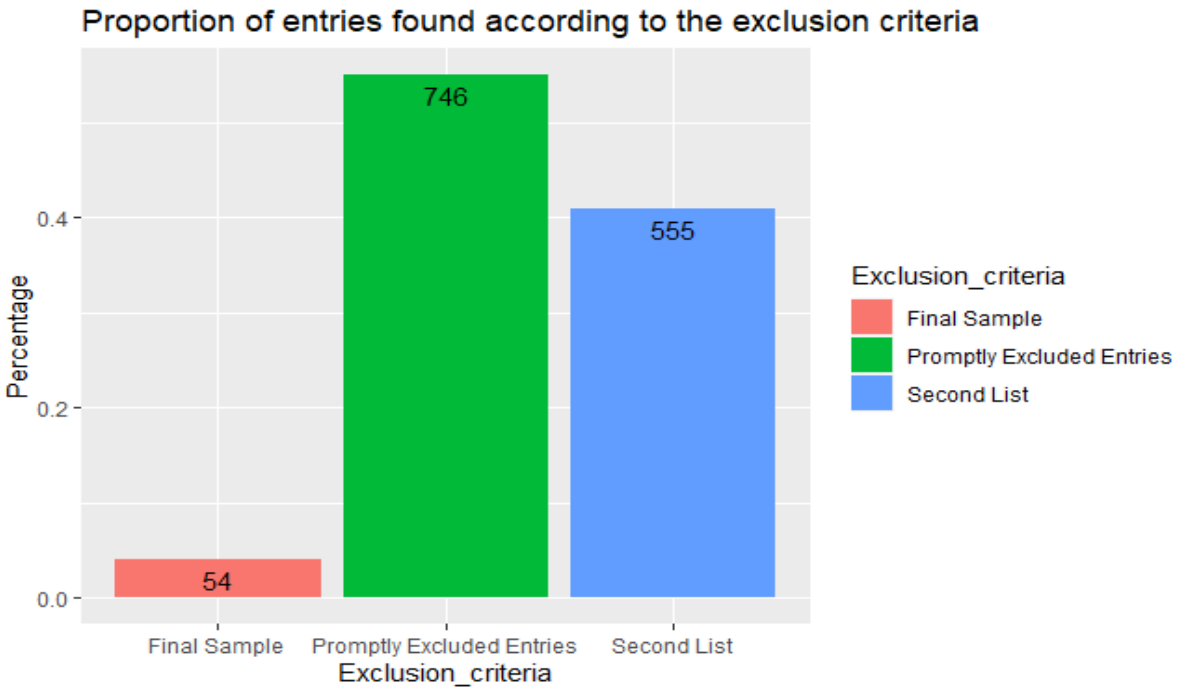
Unfortunately, as works in other languages and duplications were promptly excluded, it is not possible to determine their share of the 1355 entries obtained. What can be determined is that, out of this maximum total value, 54 works entered the final sample (all published articles), while another 555 entered the second list. Thus, "only" 609 out of 1355 entries were catalogued somehow. Charts 01 and 02 illustrate the proportion of works promptly excluded compared to the works that entered the final sample or second list and the proportion of entries found according to each searching procedure. From now on, all graphics in this dissertation were produced by the author in Rstudio 4.0.5 software.[5]

CHART 01 – Proportion of entries found according to searching procedure



[5] The entire script, as well as the database and individual graphics images are available at the following link on Github: https://github.com/Victor-Santana-Santos/SOE_database. To obtain the personal access token to such documents, please send an email to victor.santanasantos@ufpe.br.

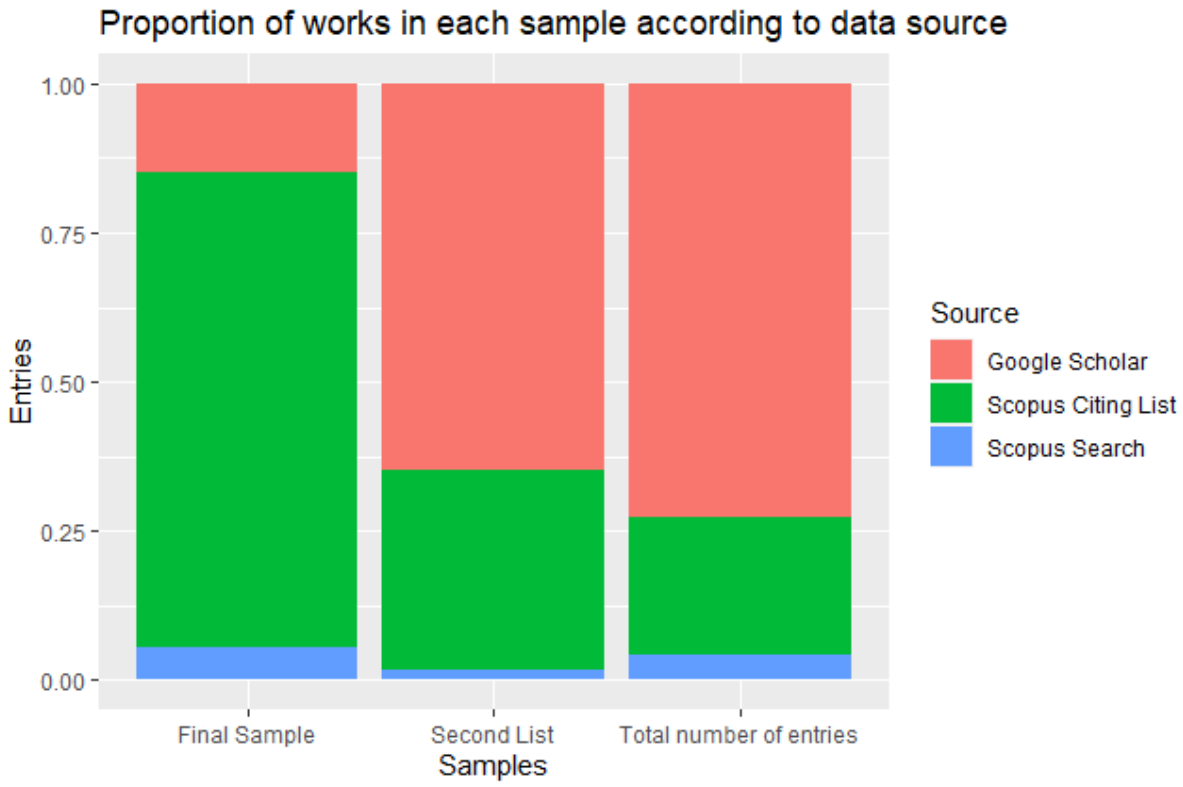
CHART 02 – Proportion of entries found according to exclusion criteria



Thus, it is noted that at least more than half of the entries found were promptly excluded, for reasons already listed above. Another 40.91% were later excluded, according to the exclusion criteria listed in the research design. It can be seen, therefore, that the selection criteria were quite precise and rigorous, since only 3.98% of the total entries found reached the final sample.

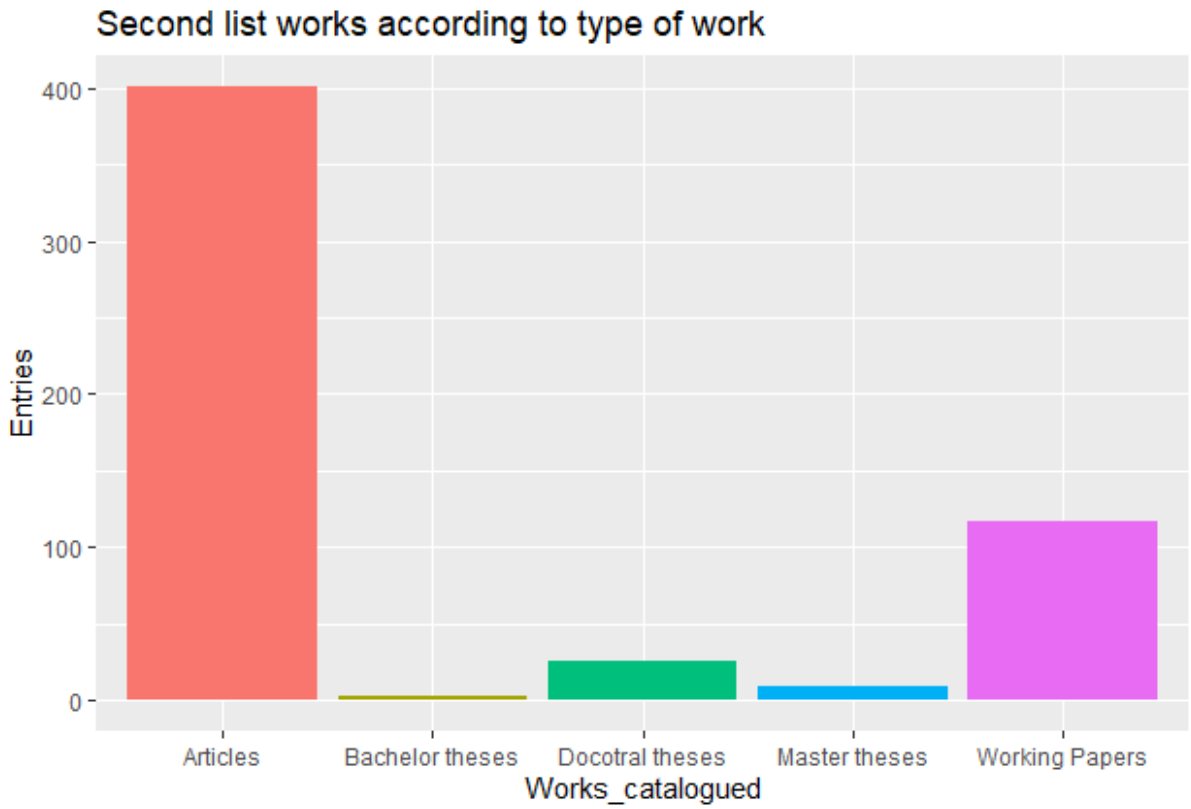
Another finding is the greater proportion of texts found in Scopus in the final sample and in the second list compared to the total number of entries found. On the one hand, this was to be expected, as Scopus was given priority in the order of searches. More specifically, our main search base was Scopus. First, the list of indexed works that cited the seminal article was checked. Subsequently, the Scopus search engine was checked with the keywords described in the previous subsection. Most of the texts obtained at this stage had already been found in the first list: of the 57 works found in the Scopus search engine, only 12 did not mention the seminal article; among these 12, only three entered the final sample. Google Scholar searches were, therefore, the final step, prioritizing key terms in English and, later, searching with terms in French, Portuguese, and Spanish; in that order. Chart 03 illustrates the difference in proportion between the three lists according to the source of the texts.

CHART 03 – Proportion of works in each sample according to data source



Regarding the second list only, most of the works catalogued were articles, reaching 401 out of 555 works, around 72.25% of this sample. 117 working papers, 25 doctoral theses, nine master theses and only three monographs were also found, these last three were all from the year 2009. This may demonstrate that second-order election theory is a topic more discussed by established academics with more expertise on the subject than by academics in training, at least until the year 2009. Chart 04 illustrates these proportions.

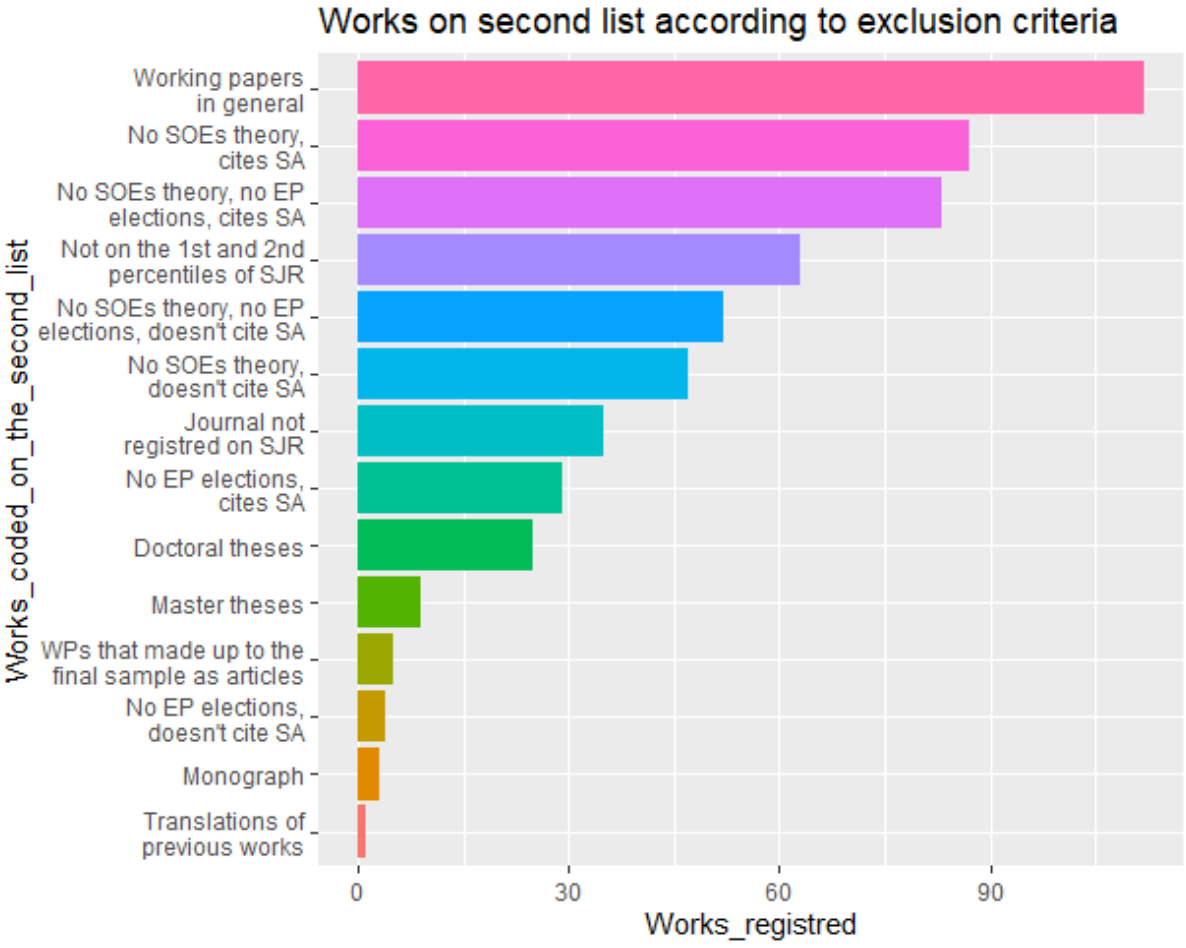
CHART 04 – Second list works according to type of work



Taking into account the exclusion criteria registered only for the works catalogued in the second list, it can be seen that criteria 1 (the work does not have the second-order elections theory as its main or central aspect, despite citing the seminal article), 5 (the work cites the seminal article, but does not have the theory as a central or main aspect, nor does it study any case of EP elections in detail) and 6 (the work does not cite the seminal article, and does not have the theory as a central or main aspect, nor does it study any case of EP elections in detail) were the most common among published articles, with 87 (15.67% of the sample, 21.69% of the published articles on the second list), 83 (14.95% of the sample, 20.69% of the articles) and 52 (9.36% of the sample, 12,96% of the articles) occurrences each. As previously stated, only three monographs were found, while nine master's theses and 25 doctoral theses were registered. Regarding the criteria related to the quality of the journals, 63 articles (11.93%) were excluded because their respective journals were not classified in the first and second percentiles of the SJR. In turn, another 34 articles (6.43%) were also excluded because their respective journals were not indexed in the SJR. Only one translation was registered. Finally, regarding working papers, 117 were catalogued (21.08%), of which five entered the final sample as published articles. Chart 05 illustrates these trends.[6]

[6] SOEs stands for “second-order elections”. SA stands for “Seminal article”. EP stands for “European Parliament”. SJR stands for “Scimago Journal Ranking”. WPs stands for “working papers”.

CHART 05 – Works on second list according to exclusion criteria

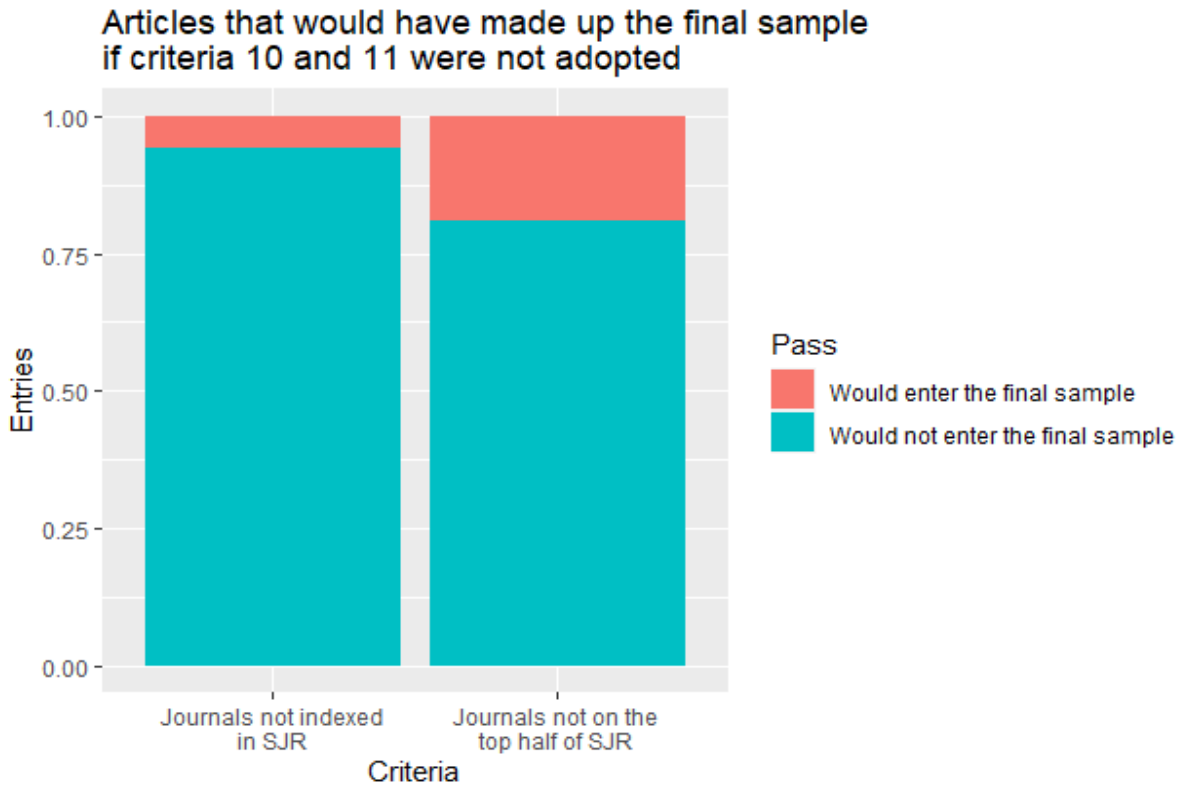


Excluding the distinction between articles that cite the seminal article and those that do not cite it; 134 articles (24.14% of the sample, 33.41% of the articles) were excluded for not addressing the theory as the central or main topic of study, while 33 works (5.94% of the sample, 8.22% of the articles) were excluded for not addressing EP elections directly. Another 135 (24.32% of the sample, 33,66% of the articles) were excluded for not meeting these two criteria at the same time. It can be seen then that 302 out of 401 articles were excluded by these criteria (about 75.31% of the total number of articles in the second list and about 54.51% of the entire second list sample).

This demonstrates that, quantitatively, the criteria related to the quality of the articles found make up only 18.37% of the total sample in the second list. Considering only the 401 catalogued articles, this share rises to 24.43%. However, it is known that the English language has become the *lingua franca* of the scientific community today. Consequently, articles published in English end up having a greater reach and impact, to the detriment of articles published in other languages. It is possible that the adoption of journals' impact

factors as a proxy for the quality of publications has affected non-English-language journals in a non-random way. Therefore, it is necessary to conduct a bias test of the criteria relating to the quality of the journals found. For this, all articles that did not enter the final sample due to criteria 10 and 11 were read, to verify whether they, disregarding such criteria, would enter the final sample or not. Among the 63 articles whose journals did not reach the first two percentiles in their respective areas of study in the SJR, 12 of them would have entered the final sample; whereas only two out of 35 articles would have entered, in the case of articles published in journals not indexed in this ranking. Chart 06 illustrates the results obtained.

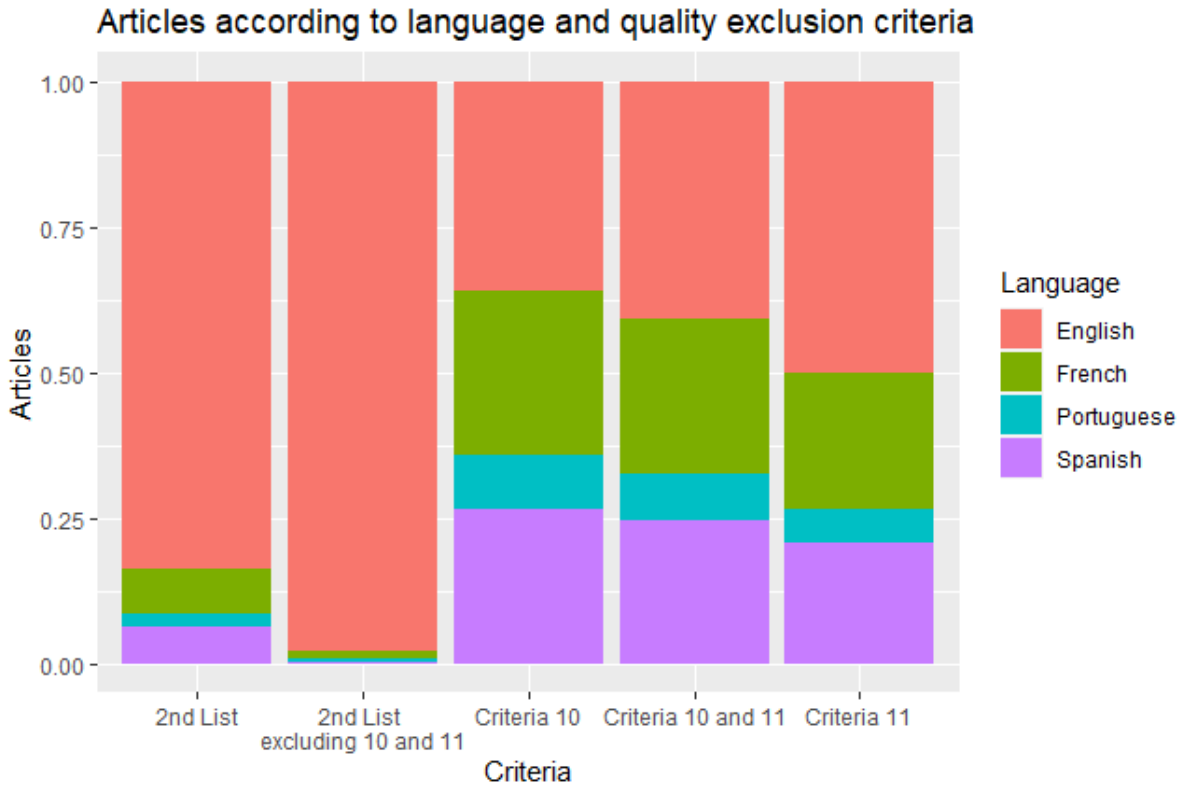
CHART 06 – Articles that would have made up the final sample if criteria 10 and 11 were not adopted



Although such articles add up a small part of these samples and, mainly, an even smaller part of the total sample of the second list; it remains to be verified whether this distribution according to the language of the articles followed a random pattern or not. For this, two comparisons are needed. First, we compared the linguistic distribution of articles that did not enter by other criteria with those that did not enter by criteria 10 and 11. Chart 07 demonstrates that there is indeed a linguistic bias both in relation to the lower impact of non-English language journals and the fact that a greater number of non-English language journals are not even registered in the SJR. While articles in English add up to more than 80% of all

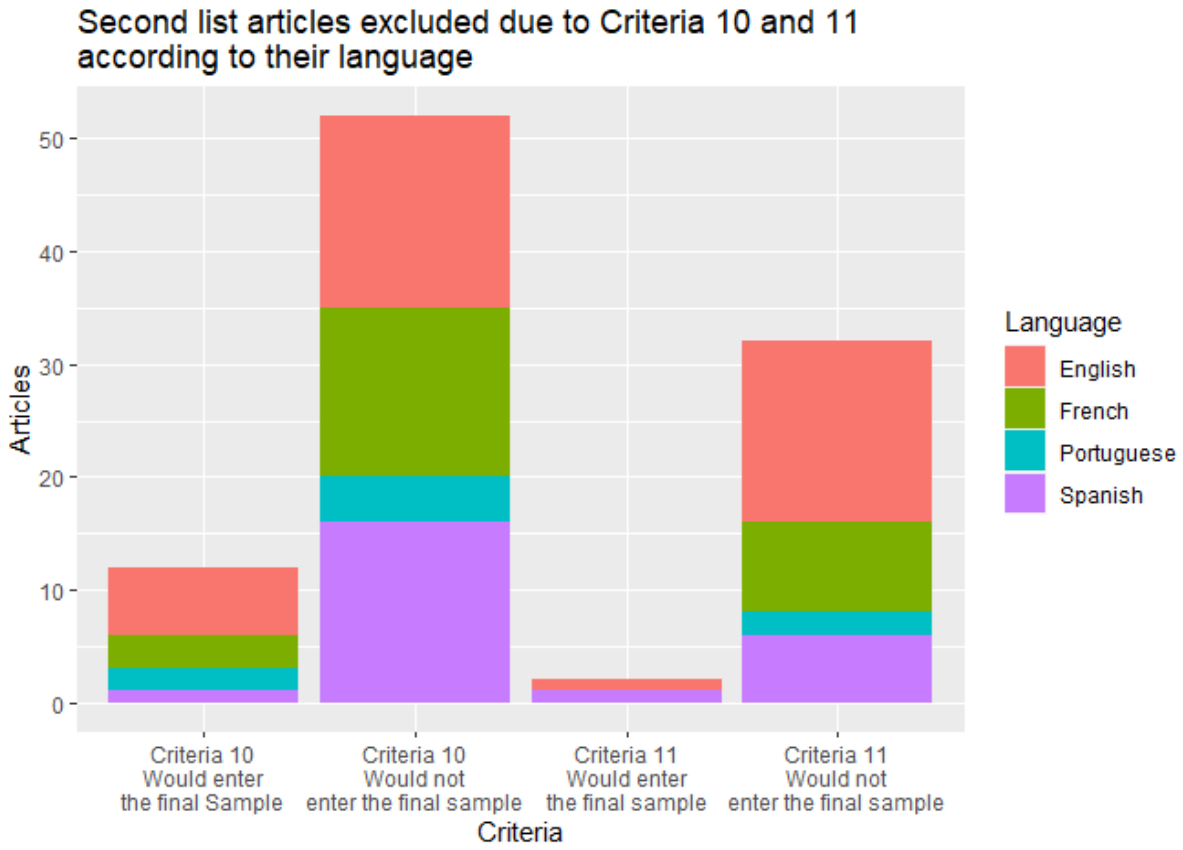
articles catalogued in the second list; they only reach a little more than 40% of the subsample of articles excluded by criteria 10 and 11.

CHART 07 – Articles according to language and quality exclusion criteria



Still, this does not mean that articles in French, Portuguese and Spanish excluded by criteria 10 and 11 would not be excluded by other criteria, if criteria 10 and 11 were not adopted. In other words, it is likely that a large part of this literature is not relevant for the purposes of this review and would end up being excluded anyway. As said before, 12 of 63 articles excluded by criteria 10 would have entered the sample, as well as other two of 35 articles excluded by criteria 11. At least half of these 14 articles were written in English, while three were written in French, two in Portuguese and two in Spanish. Chart 08 illustrates this distribution.

CHART 08 – Second list articles excluded due to Criteria 10 and 11 according to their language



Although they make up a very small part of the second list, it is clear that the adoption of exclusion criteria that use the impact of academic journals as a proxy for the quality of publications carried a bias that limited the entry of articles not written in English in the final sample. This bias mainly affected articles published in Spanish and Portuguese. As will be described in the following chapter, no articles in Portuguese or Spanish reached the final sample, while only two articles in French entered this sample. A small search in the SJR showed that no Portuguese-language journal published in Portugal dedicated to Political Science and International Relations reached the first two percentiles of the ranking for the year 2019, which clearly illustrates the linguistic bias present in the evaluations and formulas that make up the impact factor of academic journals.

Thus, it is understood from the execution of our research design that using the impact factor as a proxy for quality may be a non-optimal criterion in the evaluation of journals in different languages. This criterion can be useful in the case of reviews restricted to articles in a single language, especially in English, since there is also a greater number of academic journals that publish partially or exclusively in English. However, in the case of

multilingual reviews, the differences related to this criterion must be considered and parameterized before the text collection stage.

However, this does not automatically invalidate our research design or the results obtained by this systematic review. Since one of the main components of the impact factor is the articles' number of citations, this indicates that such articles were probably little cited by other authors, which reduces the relevance of these works as a whole, minimizing their potential contributions for the literature. Even so, some of these works bring potentially important discussions and results for the interests of this research. Five of these 14 articles bring tests with data at the individual level.

Among them, the article by Jadot (2002), published in French, stands out. This article provides a very interesting multilevel analysis of turnout and voter abstention in the 1999 elections, using data from the 1999 European Election Studies. It performs regression tests both at the individual and aggregate level, in order to determine which variables have the greatest impact on the turnout. Jadot (2002) finds that numerous variables impact abstention at both micro and macro levels. What was surprising to her was that variables such as the electoral cycle and support for integration were not significant. This led the author to criticize the theory of second-order elections, even bringing up the problem of ecological fallacy. She reinforces the hypothesis of habitual voting, especially with the results obtained with individual-level data. One of the shortcomings of the article, however, is that it does not publicise the test results in table, graphics or appendix format.

The criteria adopted in this review were considerably rigid and precise, resulting in a final sample of 54 articles, of which only two were written in French and all others in English. The following chapter begins the exposition of the data and results collected from these 54 articles, starting with the variables of the formal dimension of analysis. Such variables are the simplest to analyse, as they consist of objective criteria, such as the journal in which the article was published and its number of citations.

CHAPTER 02: Formal Dimension – Analysis and Results

This chapter aims to present the variables related to the formal dimension of the systematic review, as well as their respective results obtained. These variables are related to aspects such as title, number of pages, source of the work, genre and authors' university of origin, number of citations, among others; in general, variables that do not require reading the texts in full.

After a review of all exclusion criteria, a total of 54 articles were classified for the final sample. Regarding the journals in which the articles were published, two journals stood out: European Journal of Political Research (with eight articles, 14.81% of the sample) and Electoral Studies (with eleven articles, 20.37%). There is also a predominance of English publishers: 44 out of 54 articles were from British-published periodicals (81.48%). Another six articles were from North American publishers, while another four were from journals from Ireland, Germany, France and Austria. Charts 09 and 10 illustrate these trends.

CHART 09 – Most recurrent journals in the final sample

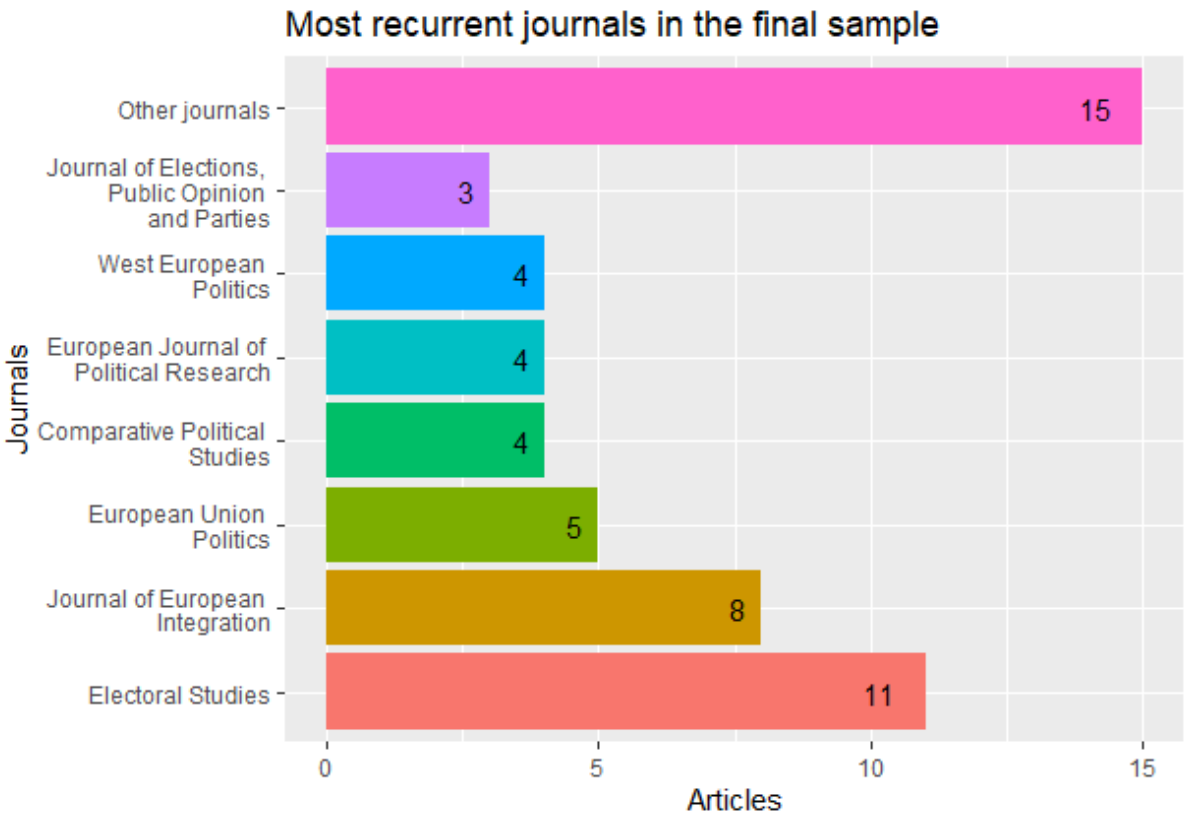
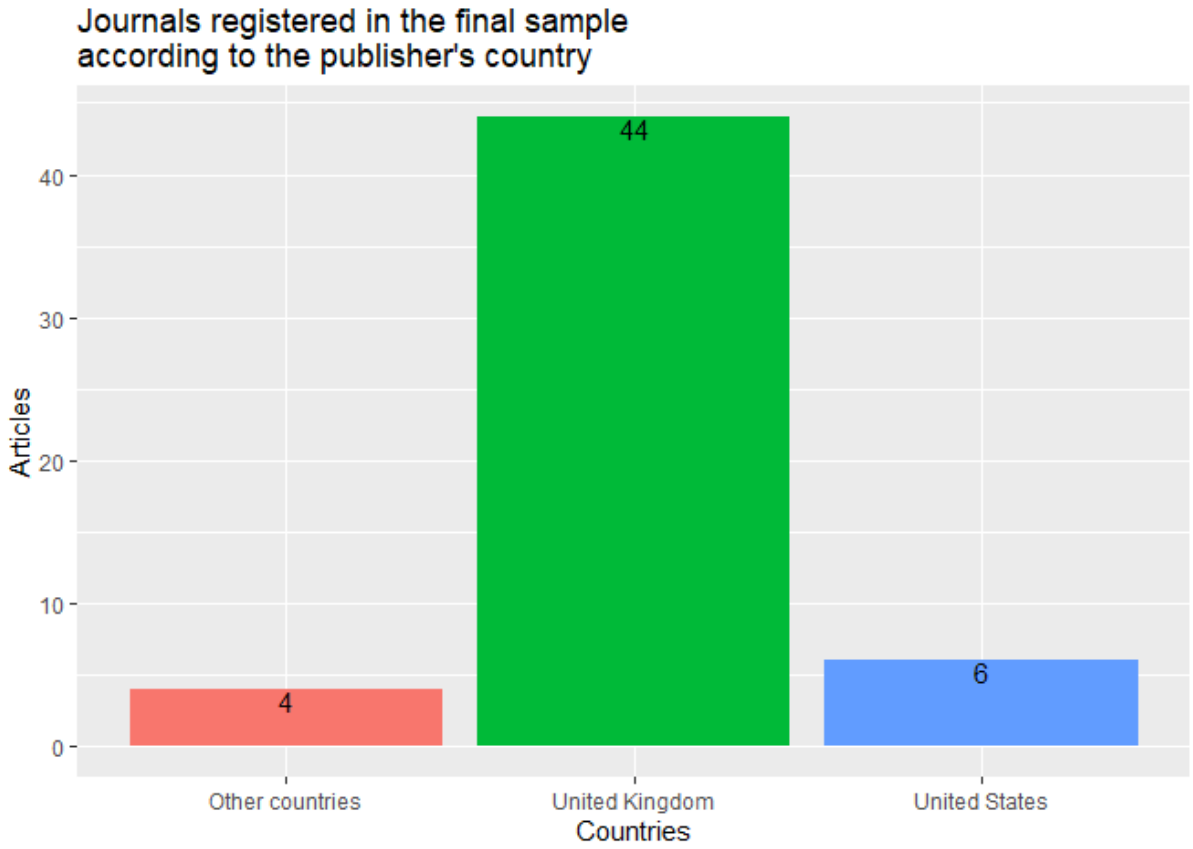
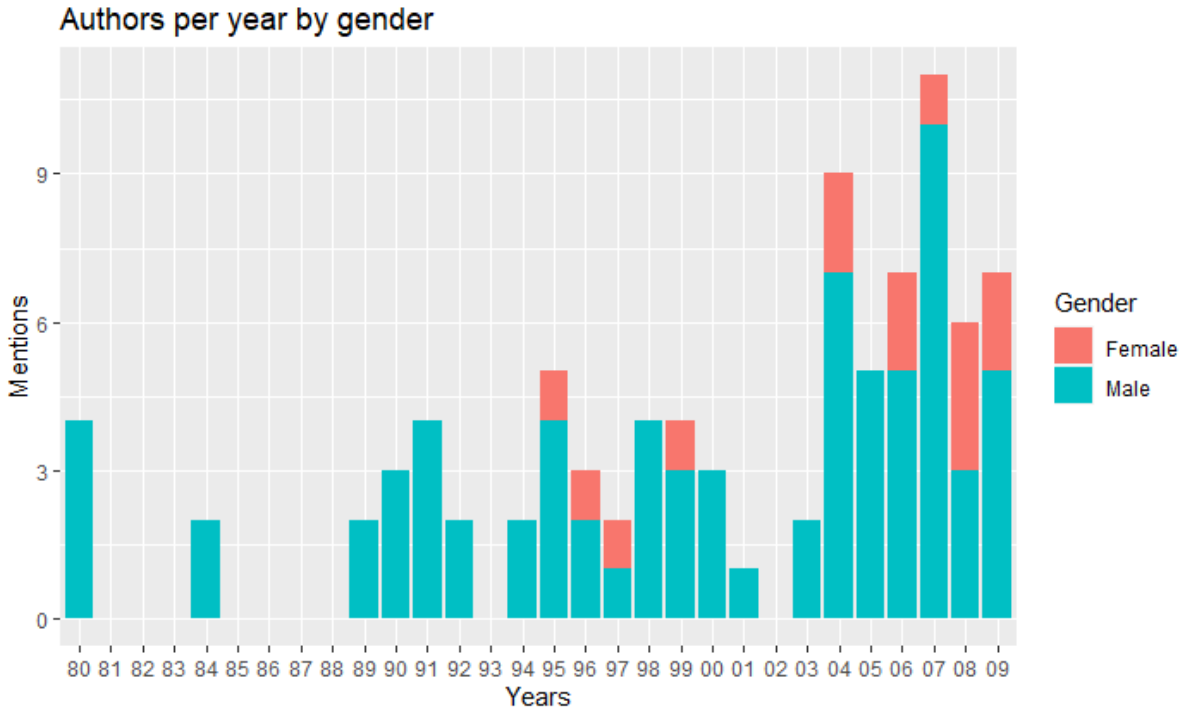


CHART 10 – Journals registered in the final sample according to the publisher's country



Moreover, 24 articles were of single authorship (44.44%). Among the authors, four names stood out: the German Karlheinz Reif, with three mentions; Michael Marsh, with five mentions; the Dutch Cees van der Eijk, with five mentions; and the German Hermann Schmitt, with six mentions. All other authors had either one or two mentions. Still with regard to authors, a large discrepancy was noted with regard to gender: in a total of 96 author-year records, 82 were from male authors (85.41% of the total). Among female authors, Bridget Taylor stood out, the only female author registered twice on the final sample. On her turn, the first female author was registered only in 1995 (SMITH, 1995). These results clearly illustrate the consequences of structural machismo in society, which delayed the entry of a significant number of women into academia, causing a great gender inequality in the production of scientific knowledge, with negative repercussions to this day. However, the trend is towards greater gender equalization. In fact, chart 11 demonstrates that the number of women who have published articles on the theory of second-order elections and EP elections has increased over the years. This inequality is expected to have fallen a little further after 2010, but this has yet to be corroborated in future work. The chart also illustrates the trend towards a greater academic production on the topic over the years.

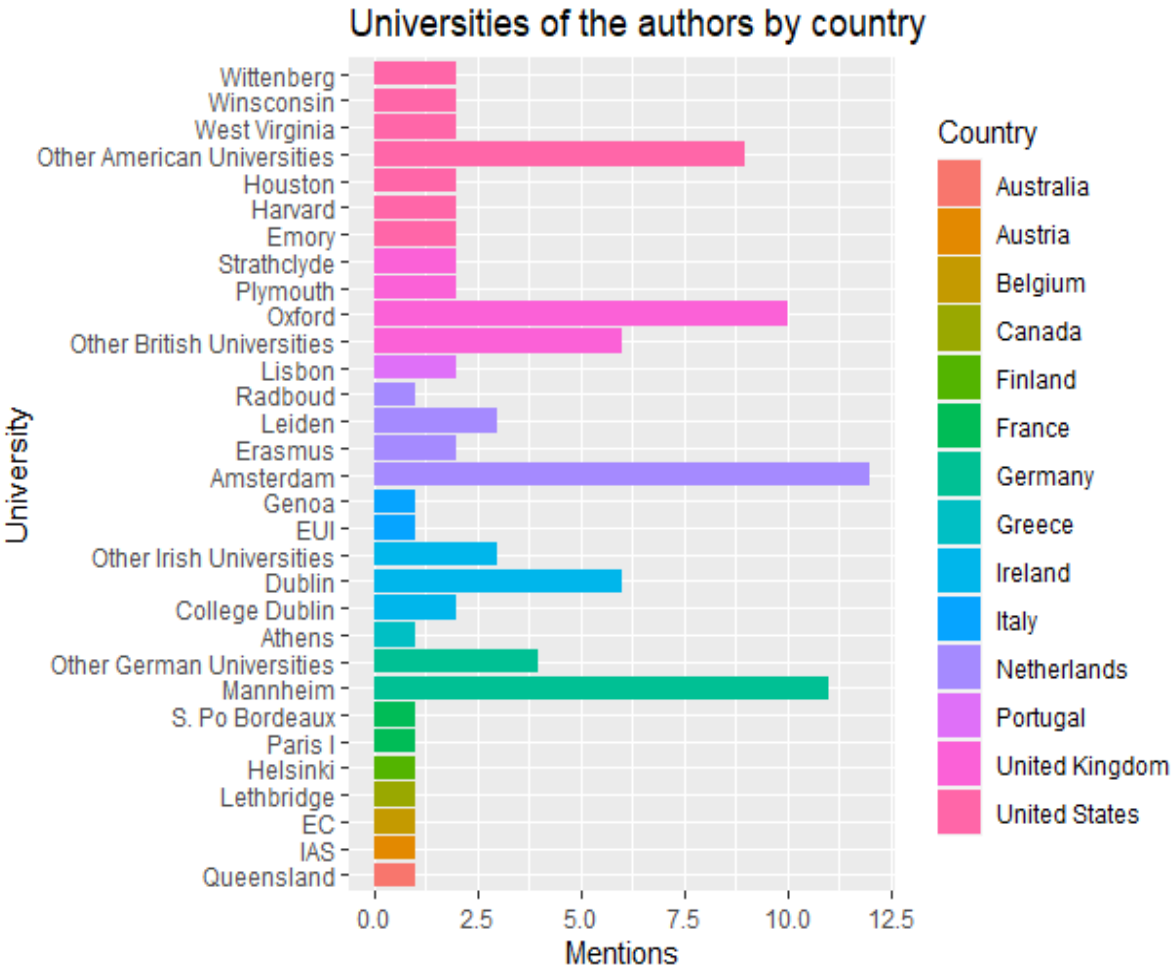
CHART 11 – Authors per year by gender



Regarding the authors' universities of origin, three of them stood out: the University of Oxford, with nine mentions (9.37% of the total); the University of Mannheim, with eleven mentions (11.45% of the total), and the University of Amsterdam, with twelve mentions (12.50% of the total). Not coincidentally, the University of Mannheim is the home university of Reif and Schmitt, authors of the seminal article. Similarly, Cees van der Eijk, the second most recorded author, is affiliated with the University of Amsterdam.

At the country level, authors from North American and British universities stood out, with 21 and 20 records respectively (21.87% and 20.83%). Next, Dutch and German universities stood out, with 18 and 15 records respectively (18.75% and 15.62%). These four countries alone account for 77.08% of all works coded in the final sample in relation to the countries of the authors' home universities. A total of 49 different universities or other affiliations were registered, from 14 countries. Chart 12 illustrates such trends. It was decided to remove the "University of" from the names to facilitate the illustration and not to pollute the graphic visually. Other names were also abbreviated.

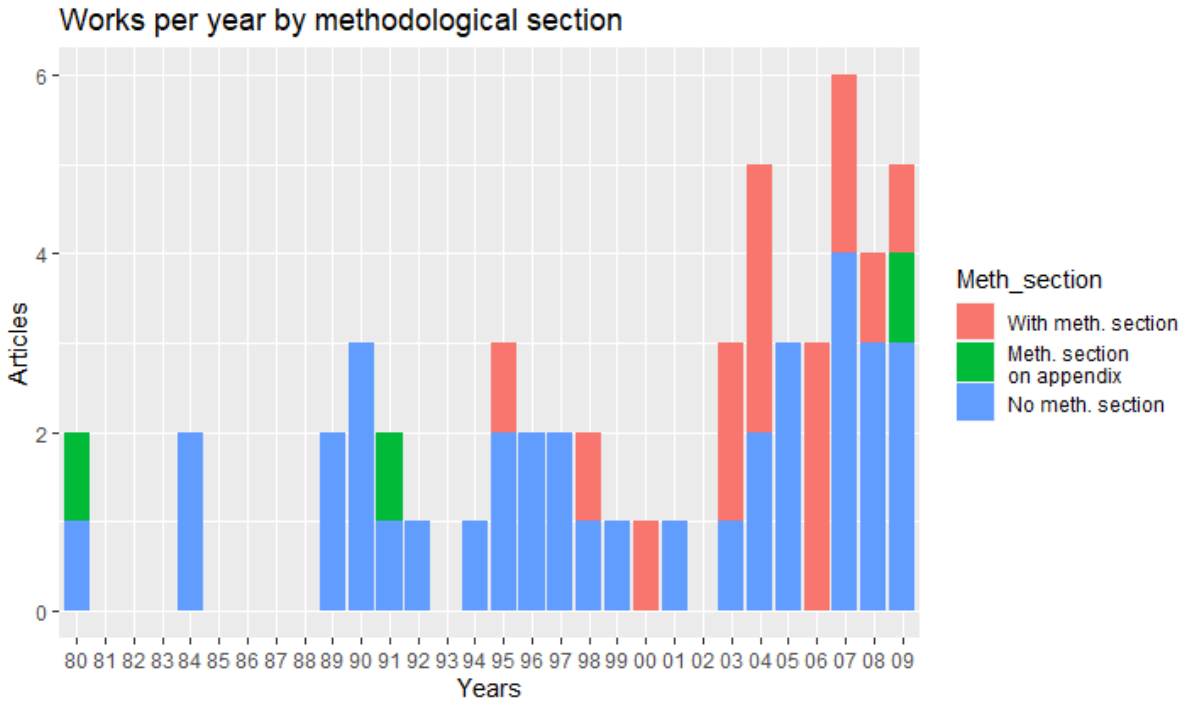
CHART 12 – Universities of the authors by country



Acronyms: EUI - European University Institute; EC - European Commission; IAS - Institute for Advanced Studies; S. Po - Science Po.

Still regarding the formal dimension, it was noted that 14 articles had sections dedicated exclusively to methodology (25.92% of the final sample), with another three having a methodological section in the appendix. It was noticed that the existence of sections dedicated to the methodology became more common over time, as shown in chart 13 below.

CHART 13 – Works per year by methodological section

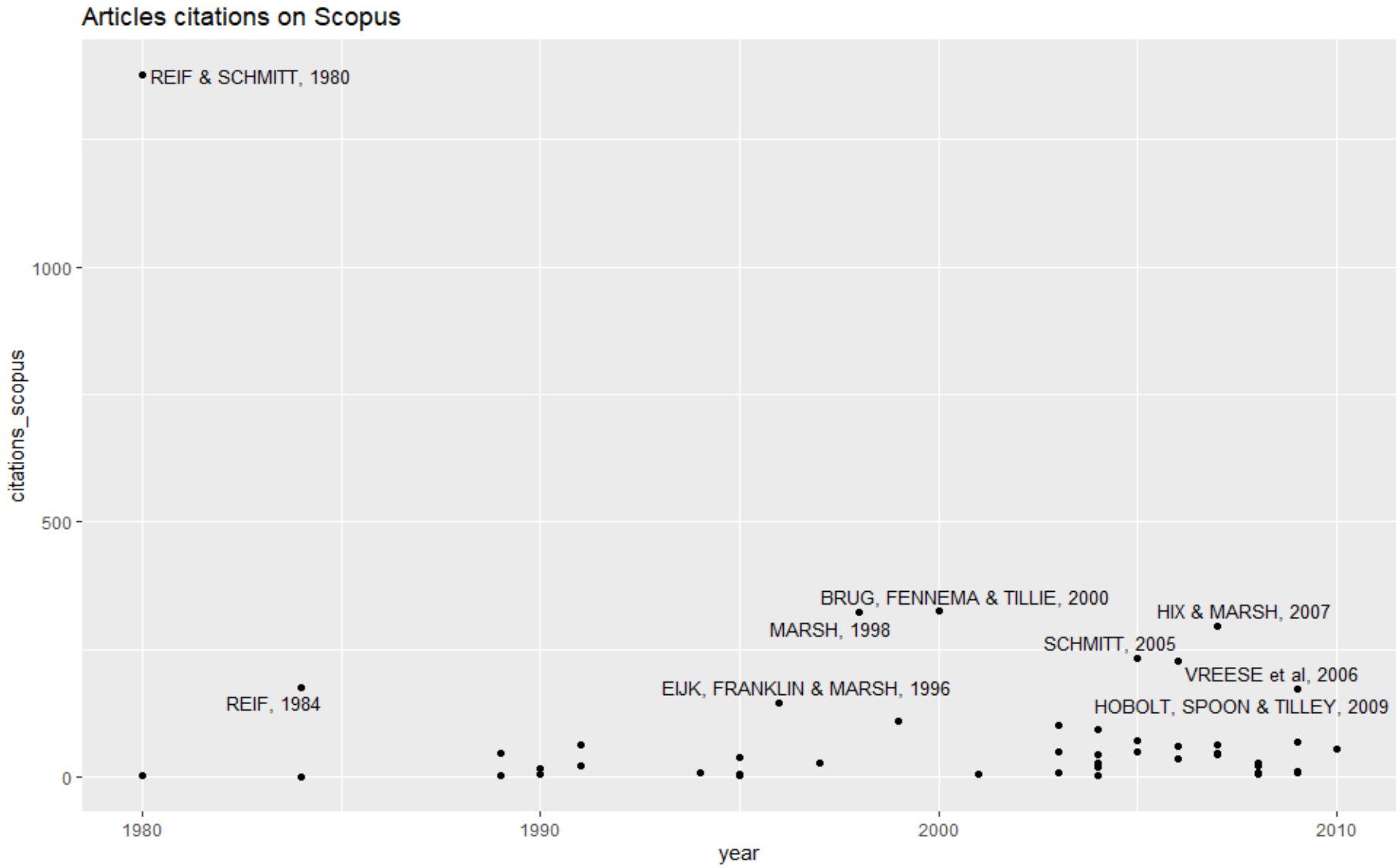


Regarding the language, the predominant language was English, with 52 articles (96.29% of the total), with another two works written in French. As highlighted in chapter 01, the introduction of exclusion criteria based on the impact factor as a proxy for the quality of publications may have hindered the inclusion of articles in languages other than English in the final sample. Not coincidentally, such articles have a low number of citations. Rabier (1984) is the least cited article in the sample, with a total of zero citations up to December 18th, 2021.[7] In his turn, Magni-Berton (2008) has five citations up to the same date.

Still on the formal dimension, about 48 works cite the seminal article in the bibliography (88.88%). Finally, among the 48 articles indexed in Scopus, the number of citations varied between 1376 (the seminal article) and 0 (RABIER, 1984), having as reference the date of December 18th, 2021. The average of citations was of 94.75; however, this value is skewed by the high number of citations from the seminal article. Excluding this article, the average drops to 67.59. The median of the sample excluding the seminal article was of 39 citations, indicating that most works have a low to average number of citations. Chart 14 illustrates the articles' number of citations and the year in which they were published. It is noticed that the seminal article is an outlier compared to the other works. Works not indexed in Scopus were excluded from the graphic representation.

[7] However, when proceeding with the reading of the texts for the systematic review, it was noted that one of them cited Rabier (1984) in detail. Schmitt and Mannheimer (1991) highlight Rabier (1984) as one of the pioneering articles in the use of survey data to analyse the turnout in European elections.

CHART 14 – Articles citations on Scopus



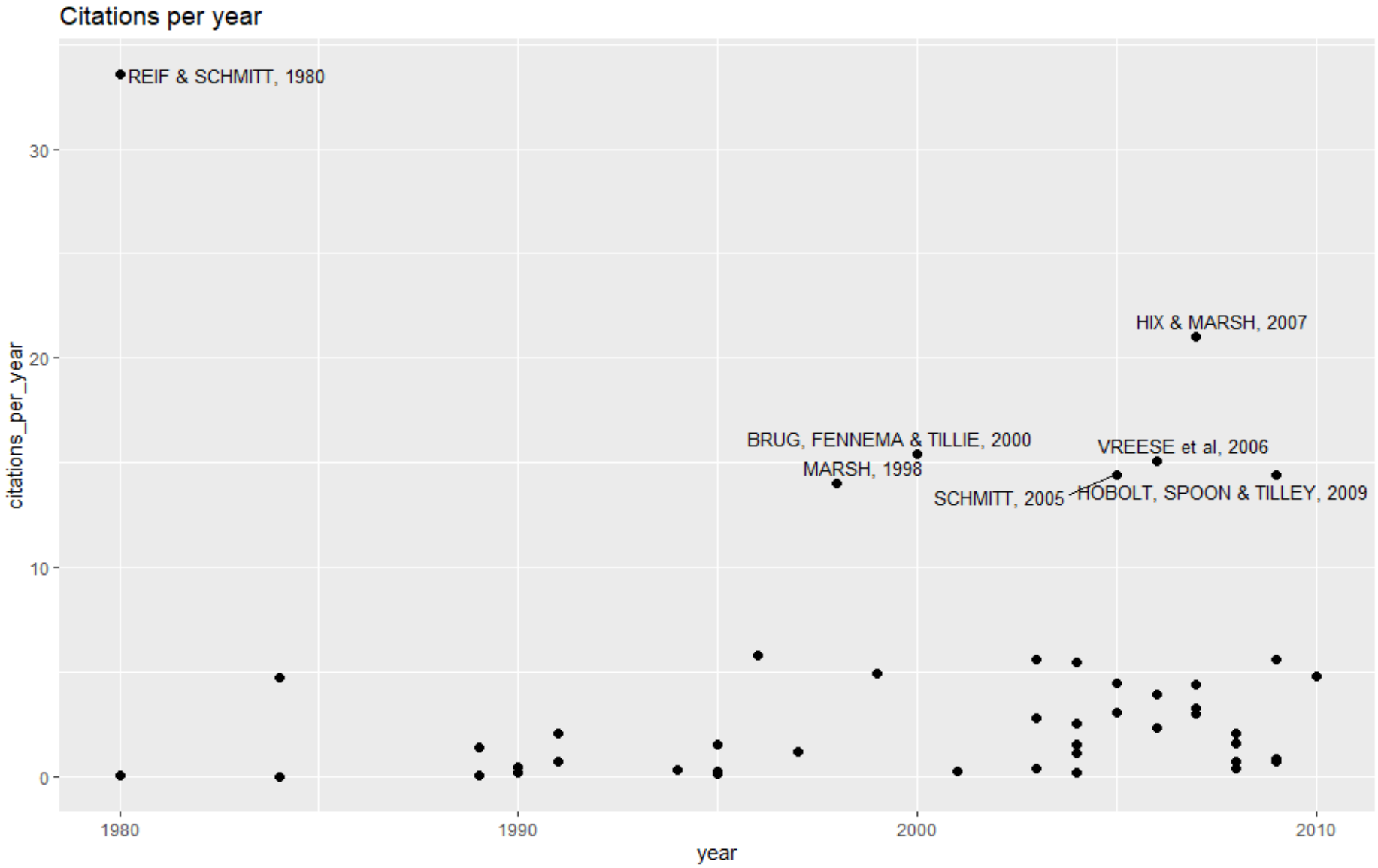
As much as newer works tend to have a higher number of citations than older works, as shown in the previous graph; the age of an article can influence its respective number of citations, since newer articles need to have a big impact to get the same number of citations as older works that have a stable number of citations per year. Thus, we also chose to divide the articles' number of citations by their age in years to observe the average number of citations per year.

The average number of citations per year of all articles indexed in Scopus selected for the final sample was of 4.45 citations per year. Excluding the seminal article, this value drops to 3.83. The median was of 2.06, excluding the seminal article, indicating that most articles in the final sample have a relatively low number of citations per year. Among the 48 articles indexed, seven of them stood out, obtaining at least an average of 14 citations per year. Chart 15 illustrates the disposition of the average citations per year for each article.

The article with the most citations per year is the seminal article by Reif and Schmitt (1980), with an average of 33.56. The second article with the highest average number of citations per year was Hix and Marsh (2007), with 21 citations per year. The other five highlighted works had between 14 (MARSH, 1998) and 15.42 (BRUG, FENNEMA & TILLIE, 2000) citations per year. Among the six articles with the most citations per year – excluding the seminal article – three performed theory tests with aggregate-level data (MARSH, 1998; SCHMITT, 2005 and HIX & MARSH, 2007) and two performed tests with individual-level data (BRUG, FENNEMA & TILLIE, 2000 and HOBOLT, SPOON & TILLEY, 2009). However, it is noteworthy that Brug, Fennema and Tillie (2000) focuses on the theme of the attractiveness of anti-immigration parties. This theme, despite being correlated with the second-order elections theory, does not cover the entire theoretical framework studied, meaning that more robust tests of the theory were not carried out by these authors. Finally, the sixth remaining work, Vreese *et al* (2006), focuses on the analysis of news content during EP election campaigns.

The data show that most of the authors who published articles about second-order elections and elections to the EP are men and come from German, British, Dutch and North American universities. There is also a predominance of British and American publishers. It was noticed that more recent articles have more frequently dedicated an exclusive section to the topic of methodology. In addition, the number of citations tends to be close to five citations per year on average.

CHART 15 – Citations per year



The following chapter continues exploring the data and results collected from the 54 articles in the final sample. The variables of the methodological dimension will be described, alongside with an overview of the articles that studied the theory with individual-level data and their results.

CHAPTER 03: Methodological Dimension – Analysis and Results

This chapter aims to present the variables related to the methodological dimension of the systematic review, as well as their respective results obtained. These variables are related to aspects such as if the article has an abstract, if its hypotheses are clearly stated, as well as the research questions, year of the EP elections analysed, level of analysis, etc.

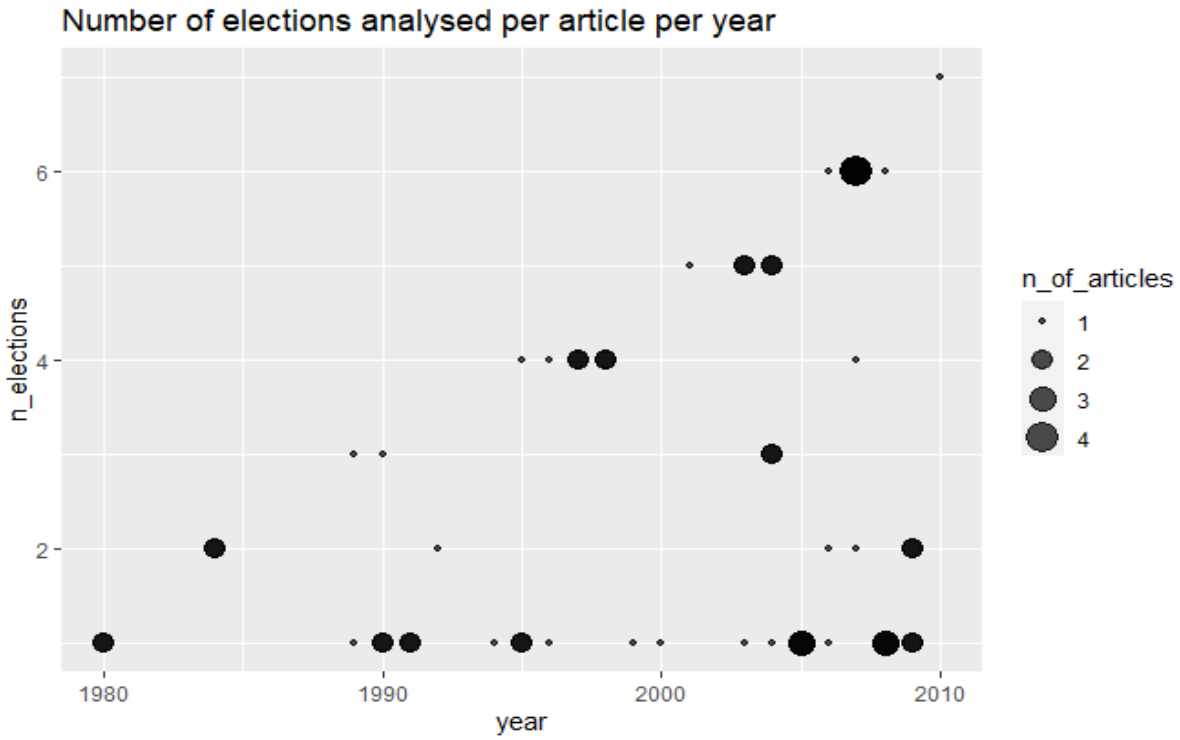
Regarding the methodological dimension, 42 of the 54 articles had an abstract (77.78% of the total). 23 works explicitly mentioned the research question in the abstract or in the introduction (42.59% of the total), while another ten mentioned it in some other section (18.51% of the total). As for the hypotheses, the pattern was different: only five studies had the hypotheses well defined and specified in the abstract or in the introduction, while another 21 did so in later sections. It is then noticed that most articles (about 51.85%) did not clearly specify which were the hypotheses adopted in the research. This made it difficult to construct the variable "Description of hypotheses", since, in these 28 works, the hypotheses had to be inferred from the reading. This variable is extremely important for the fourth stage of the review, which is the comparison of the proportion of findings in the expected direction, based on the parameter of corroborated and refuted hypotheses, which will be carried out in the fourth chapter. As careful as the reading is, the possibility of bias or misinterpretation of texts in which the hypotheses are not explicit cannot be excluded, which can impact the reliability of the results to be obtained in the fourth stage.

In any case, the number of hypotheses presented in each work varied between 0 – two short literature reviews (NORRIS, 1997; REIF, 1997), a case report (STEFANOVA, 2008), and a comprehensive literature review (MARSH & MIKHAYLOV, 2010) – and 17 (REIF & SCHMITT, 1980). The average of hypotheses per article was 4.38 considering the entire sample and 4.74 excluding articles without any hypotheses. However, as previously stated, not all articles explicitly defined their hypotheses. Considering only the articles that did so, the average number of hypotheses per article rises to 5.76. It is then noticed that the coding of the variable "Description of hypotheses" may have been biased downwards in articles that did not clearly define their hypotheses, whose average was 3.62, excluding the four articles without any hypotheses.

Only 15 articles provided details on the methodology adopted in the abstract or in the introduction. Regarding the studied elections, 24 articles studied only one election (44.44% of the total), while seven studied two elections. Another four studied three elections and seven studied four elections. Moreover, five articles studied at least five elections and

another six studied six elections. Only one article studied seven elections, the maximum number of elections possible for the temporal scope of the review, which was precisely Marsh and Mikhaylov (2010), the second major literature review on our research subject. Thus, the elections analysed in the final sample were those from 1979 until the first papers published about the 2009 elections. Since the number of elections available for analysis increases over time, the number of papers that analysed all available elections was of 22 (40.74% of the total). The most analysed election, whether in just one country or in all member countries, either individually or together with other election rounds; was the 1989 EP election, mentioned in 29 works. However, considering only the 24 works that analyse only one election, the election round most analysed individually was that of 2004, mentioned 8 times, followed by 1994, mentioned individually seven times. Chart 16 illustrates the number of elections studied in each work on a time scale. The dot size illustrates the amount of papers that analysed a given number of elections in that particular year. This size varies from one to four, indicating that, for example, in 2007, four articles analysed at least six elections, while another article analysed four elections and, finally, another one analysed two elections. Similarly, in 2004, two articles analysed five elections, another two analysed three elections and another one analysed only one election.

CHART 16 – Number of elections analysed per article per year

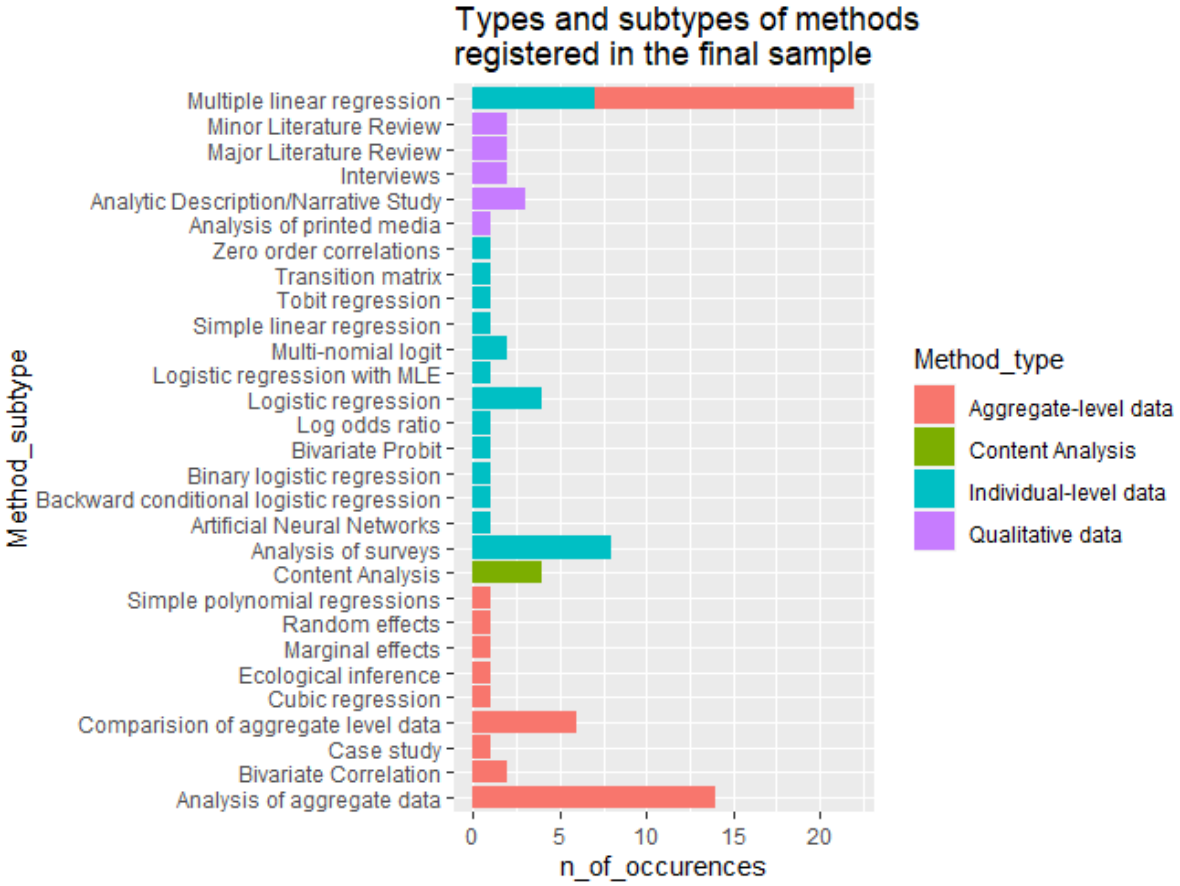


Regarding the number of countries analysed in each work, 11 articles analysed only one country (20.37% of the total), which demonstrates the preference of researchers to

analyse the European elections at the European level, or, at least, comparatively. Among the eleven works that analysed only one country, the ones that stood out were the Netherlands and the United Kingdom, with three articles each, followed by Ireland with two articles. The number of case-years in the final sample ranged between 1 and 118 (MARSH & MIKHAYLOV, 2010). In turn, the first articles to analyse a restricted number of countries on subregional criterion only appeared after the 2004 elections, more specifically in 2006, with Koepke and Ringe (2006). A possible explanation for this is that only nine countries were part of the then European Community (EC) when the first EP elections took place. This number rose gradually until 2004, when eight countries entered the Union at once. Furthermore, until the 2004 expansion, all the countries that made up the Union were considered "Western countries". The 2004 expansion brought with it a new group of countries to the integration process, to which it was possible to make comparisons at the sub-regional level, a more consistent criterion than the year of entry into the Union.

Among the 54 works in the sample, 41 adopted quantitative methods (75.92%), while 9 adopted qualitative methods (16.66%) and another three used content analysis (5.55%). More specifically, among the 88 cases – some articles had more than one submethod –, the most reported subtype of method was multiple linear regression with aggregate-level data, reported 15 times (17.04%), followed by descriptive analysis of aggregate-level data, reported 14 times (15.90%) and descriptive analysis of survey data, reported 8 times (9.09%). It was noticed that logistic regressions and their subtypes were only registered with articles that analysed individual-level data. This is explained by the type of data available or chosen for analysis. Articles that analyse turnout at the aggregate level, for example, typically use official election turnout records, a specific value positioned on a continuous scale ranging from 0% to 100%. Articles that analyse the turnout with data at the individual level usually use surveys, in which respondents answer whether they attended or intend to attend the polls, in a binary response model that favours the use of logistic regressions. Chart 17 illustrates these findings. Methods in blue are quantitative methods that were operationalized with aggregate-level data; while in orange are quantitative methods that were operationalized with individual-level data. Methods in purple are of the qualitative type. Finally, we chose to code articles that used content analysis in a separate category, considering that this method is comprehensive and can be operationalized both quantitatively and qualitatively. Articles that used content analysis typically studied parties' campaigns and media content such as news related to these campaigns.

CHART 17 – Types and subtypes of methods registered in the final sample



As for the level of analysis of the quantitative works, it is clear that, in fact, most of the articles catalogued in the final sample brought theory tests with individual-level data, whether restricted to this type of data, whether analysing both individual- and aggregate-level data. More specifically, among the 42 articles that used quantitative methods, 19 of them analysed only aggregate-level data (45.23%), while 14 analysed only individual-level data (33.33%) and another 9 used both types of data (21.42%). However, not all 23 articles with individual-level data brought robust statistical tests and neither their results necessarily gained prominence in the following years.

An important finding was that as early as the 1980s there were articles with tests that used micro-level data. In 1984, Rabier (1984) analysed the determinants of turnout of the two European elections that had taken place so far, using both aggregate-level and individual-level variables, with data from the Eurobarometer 19. However, the impact of his article on the research agenda on second-order elections can be considered low for three reasons. First, his tests took into account only one aspect of the theory: the turnout rates. Second, Rabier (1984) only unravelled the Eurobarometer 19 results, comparing the proportions of respondents who declared to have voted or not based on some parameters such as citizens'

cognitive mobilization and their attitudes towards European integration and towards the EP. Regression statistical tests were not performed. Finally, the fact that this article was written in French might have affected its reach, as it is the only article in the sample with zero citations by other articles in Scopus. Rabier (1984) also has no citations in Web of Science and Crossref.[8] Thus, as much as Rabier (1984) uses survey data in his analysis, his initial tests did not bring new developments to the research agenda, nor were they widely disseminated afterwards.

In addition to providing a detailed descriptive analysis of the Dutch political scenario prior to the elections, Eijk (1989) also used a survey carried out at the national level to explain the low turnout of the European elections in the Netherlands in 1989. His data corroborated hypotheses that were already well known in the research agenda on turnout in general, such as the fact that greater political interest is correlated with greater participation and that voters attached to a political party are more likely to turn out to the polls. Like Rabier (1984), Eijk's (1989) tests did not bring major new findings to the research agenda, as this article was rarely cited by its peers (there are only two citations in Scopus).

On the other hand, numerous tests of the second-order elections theory were carried out in the 1990s, either with aggregate- and individual-level data. Niedermayer (1990), for example, used data from Eurobarometers 30 and 31 to test whether the low turnout in EP elections was caused by voters' knowledge of EP's limited power and whether the difference in turnout between member countries was explained by EC approval level in each country. Niedermayer (1990) found that pro-integration attitudes have an impact on turnout, but it is considerably smaller compared to variables such as mandatory voting and other elections held in conjunction with European ones. Furthermore, the author also refuted the hypothesis that knowledge of the weak powers of the EP led to a smaller turnout, as only a small part of the interviewees was aware of the relative impotence of the EP in the EC decision-making system; and yet, those who knew this were more likely to vote compared to those who truly believed that the EP had strong powers similar to those of national parliaments.

Also in 1990, Schmitt used Eurobarometers 30, 31 and 31A to analyse whether larger and/or government parties actually had more difficulty obtaining votes in European elections than smaller parties. Schmitt (1990) found that, in terms of electoral mobilization, larger parties proportionately mobilized more loyal voters than small parties, but not as much as in national elections. Governing parties also performed slightly better than opposition

[8] However, as stated before, we found at least one citation of Rabier (1984) in Schmitt's and Mannheimer's (1991) work.

parties in mobilizing their voters. However, the mobilization of indecisive or indifferent voters seems to weigh more in European elections. According to Schmitt (1990), what most affects mobilization is mandatory voting and the occurrence of simultaneous elections, results similar to those found by Niedermayer (1990) and other authors. Furthermore, voters linked more strongly to a specific party made up only 10% of the sample on average, so the results obtained by Schmitt (1990) do not invalidate the theoretical model as a whole. In any case, these last two works had only 14 and 6 citations in Scopus, respectively, which indicates that they did not have a great impact on the research agenda either.

Schmitt also published another article in 1991, with Renato Mannheimer, analysing the EP elections at the micro level. The authors focused on the determinants of turnout, listing variables such as regular vote, party attachment, subjective political interest, EC approval, EP institutional prominence, salience of EC-related issues, in addition to specific variables such as age, gender, education, and whether the respondent shared post-materialist values. Schmitt and Mannheimer (1991) found strong evidence that turnout in EP elections is directly affected by party affiliation and habitual voting. Subjective political interest only indirectly affects electoral participation, as well as social variables. However, these results were valid only for the predisposition to vote. Based on the data from Eurobarometer 31A, carried out after the elections, only regular voting had a great impact on the effective turnout, while party attachment had an indirect impact. This would indicate that voters would overwhelmingly vote only out of voting habit or out of a sense of civic duty, at least in the case of European elections. However, as will be detailed later, these results were later refuted by Sinnott and Whelan (1992) and also by Eijk, Franklin and Marsh (1996). The relatively higher number of Schmitt and Mannheimer (1991) citations in Scopus – 60 citations – may also be a consequence of these controversial results, but this hypothesis was not formally tested in this research.

Despite using survey data to test their hypotheses, Eijk and Oppenhuis (1991) aimed to study the determinants of party competition, which means that their analysis is centred on parties' level and the degree to which they manage to mobilize their voters. The dependent variable in this case is the conversion of each party's electoral potential into votes, measured at the aggregate level, but operationalized with data at the individual level. Therefore, this work has been coded as an approach that uses both aggregate- and individual-level data, even though the former is the main level of analysis.

Galen Irwin (1995) also carried out tests at the micro level with data from Eurobarometer 41. The author aimed to better understand the low turnout of European

elections and tested hypotheses regarding the lack of disagreement between parties and within the electorate in relation to European issues. Irwin (1995) found that there were indeed controversial points in the electorate, especially regarding the future of integration – such as support for the single currency – that could have been mobilized by political parties. However, weak campaigns for European elections reinforced the less at stake effect of second-order elections. In the case of the 1994 elections, in addition to this effect, voters themselves were not encouraged to vote, as their choices were limited to national parties most of whom had similar positions on European issues, with the exception of Denmark and the United Kingdom. Despite the little difference between parties' positions, they insisted on putting up separate lists for voters, which made the system more confusing and compelled voters to vote based on national issues. Voters hesitant to further integration had few options in most countries, concentrated on new and/or far-right parties. Thus, Irwin (1995) states that voting in European elections was basically an act of support for integration. However, like Rabier (1984), Irwin (1995) only broke down the data from the Eurobarometer 41, not performing statistical or regression tests. Its number of citations – 39 – despite being higher than that of Niedermayer (1990) and Schmitt (1990) – is the median number of citations of the final sample, which does not indicate a strong impact on the research agenda either.

McLean, Heath and Taylor (1996) compared British local elections with European elections, aiming to find out if voters voted differently in each one of them. Their results showed that British voters considered some issues pertaining to their respective local arenas when voting in the 1996 local elections, and a greater number of voters said they were concerned about the results of such elections. In turn, European issues had little impact on turnout and vote choice in EP elections, indicating that some elections may have a greater second-order character than others. In any case, this work appears to have had little impact on the research agenda: despite not being listed in Scopus, it has 0 citations in Crossref.

The next work in the final sample to bring micro-level data analysis is Eisinga, Franses and Dijk (1998), who use artificial neural networks to determine how many weeks before elections the number of undecided voters starts to drop. From weekly-collected surveys, the authors found that the number of undecided voters begins to decline, on average, nine weeks before a first-order election and just four weeks before a second-order election. This period coincides with the duration of campaigns for each election, which are usually longer and more intense in first-order elections. This article, however, appears to have had little impact on the second-order election research agenda, as Crossref lists only 5 citations (this work is not listed in Scopus).

The following year, Heath *et al* (1999) studied how the less at stake dimension in second-order elections changes depending on the arena under analysis, in addition to understanding how voters choose their votes depending on the electoral arena. To do this, they analysed the 1994 European elections in Great Britain compared to local British elections held in the same year. The authors concluded that voters do believe there is "more at stake" in local elections than in European ones. Therefore, issues in the local political arena affect voting choice in local elections more than issues in the European political arena in relation to EP elections, reaffirming the results of McLean, Heath and Taylor (1996). The shift in voting between first- and second-order elections is dominated by national considerations in both arenas, although this occurs to a lesser degree at the local level. Second-order arena issues can also affect vote choice in that arena, but not on a large scale. Furthermore, Heath *et al* (1999) demonstrate that, in second-order elections, sincere voting and protest voting are not mutually exclusive at the individual level, confirming the main postulates of the theory and anticipating the argument of Schmitt *et al* (2020) that "it all happens at once".

Brug, Fennema and Tillie (2000), on the other hand, restrict their analysis to anti-immigration parties, seeking to delimit whether the votes they receive are based on ideological issues or if they are protest votes, since most anti-immigration parties at that time were not part of the European party-political mainstream. Based on data from the 1994 European Election Survey, the authors found that European elections have an impact on the national political arena and that voters are aware of this impact, which influences turnout and their vote choice. The authors also point out that, among the socio-structural characteristics they researched, only religious affiliation and social class influenced the preference for anti-immigration parties. Furthermore, attitudes towards the European Union had little impact on voters' preference for anti-immigration parties, indicating the opposite of what most of the academy had asserted so far. According to the authors, "this is particularly surprising because the leaders of anti-immigration parties generally show a strong dislike for the institutions of the European Union" (BRUG, FENNEMA & TILLIE, 2000, p. 96).

Although this is formally the most cited article with individual-level data in the final sample, with 324 citations and an average of 15.42 citations per year, it is believed that this is due to its analytical approach. Although it addresses important aspects for second-order election theory, such as the strength of anti-immigration parties in European elections, the focus of this article is not on the theory itself, but on studying anti-immigration parties. It is noteworthy that this theme gained greater momentum and relevance after successive migratory crises and subsequent terrorist attacks in the second decade of the 20th century, as

well as after the rise of such parties to the European political mainstream. Therefore, it is believed that its high number of citations is due to its analytical approach restricted to anti-immigration parties, and not to its findings related to the second-order elections theory itself. However, to test this hypothesis, it is necessary to check the list of more than 320 works that cite this article and check their research subject, which was not possible to be done in this review.

The next work to carry out tests with individual-level data was Siaroff (2001), who questions if the second-order elections theory better predicted the EP elections results than a model he called "expression of anti-integration sentiments", which is basically a hypothesis that voters Euroscepticism on the one hand and the lack of Eurosceptic parties on the other affects EP elections' turnout and results. Siaroff (2001) found mixed results, since both hypotheses were partially supported, as the models tested fit better for some but not for all countries. The second-order elections model fit best in Germany, Austria, the United Kingdom and Greece. On its turn, the model relating to the expression of Eurosceptic feelings by the electorate fit better in Luxembourg, Ireland, Denmark, France, the Netherlands and Portugal. Belgium and Spain fit both models while Sweden and Finland did not fit either. However, it is necessary to emphasize that Siaroff (2001) only tested the assumption regarding the loss of votes of the governing parties, without testing the other propositions of the theory. Siaroff (2001) also does not explain why so many countries fit the model of expressing Eurosceptic sentiments if only France and Denmark had electorally viable catch-all Eurosceptic parties. Finally, while testing his hypotheses with individual-level data, Siaroff (2001) only performed a simple regression test with the electoral cycle as a sole independent variable. As for citations, Siaroff (2001) only had five citations in Scopus until December 18th, 2021, which indicates a low impact in the research agenda.

Three years later, Kousser (2004) analysed the motives of voters switching parties between first-order national elections and second-order European elections. More specifically, the author tests four hypotheses, the first relating to the electoral cycle, the second relating to the retrospective voting by economic criteria, the third relating to the protest voting and the fourth relating to sincere voting. Of these, only the second and third hypotheses were fully supported. The results obtained for the first and fourth hypotheses were mixed. According to Kousser (2004), the effect of the electoral cycle exists, but it is not enough to explain vote switching alone. This variable must be complemented with economic performance. On its turn, concerning sincere voting, although voters generally have a greater incentive to vote for smaller parties in second-order elections because there is less at stake, they do so to a greater

degree when electoral regulation favours such parties. This indicates that vote switching, in addition to being guided by considerations about economic performance, is also guided by strategic calculations and not only by sincere preferences. Voters do not want to waste their votes, even in second-order elections. Therefore, voting for smaller parties is more encouraged in the face of favourable political contexts and regulations. This does not mean that the sincere voting hypothesis should be discarded, but contextualized. As for citations, Kousser (2004) had 43 until December 18th, 2021, a relatively higher number than most articles released that far with individual-level data tests.

In the same year, Freire (2004) studied elections in Portugal based on second-order elections theory. The results showed that there was a majoritarian drive both in first- and second-order elections in Portugal and that electors use second-order elections as a way of expressing satisfaction or discontent with national government. This article, however, had only 26 citations as of December 18th, 2021, in addition to being a case study, which may indicate a lower impact of its findings on the research agenda as a whole.

As previously reported, Carruba and Timpone (2005) released one of the first high-impact articles to bring theory tests with data at an individual level, reaching a total of 71 citations as of December 18th, 2021. In this article, the authors ask the following question: Why else – apart from strategic voting – might individuals vote for different parties in different elections? The authors tested the theory with different variables and models, but only got mixed results. The hypotheses regarding opposition voting and balancing between government and opposition parties in different political levels were only partially corroborated in one specific model, and the hypotheses related to sincere voting and different preferences at different political levels were only corroborated in the model which tested vote switching between green and non-green parties.

In the same year, Schmitt (2005) released one of the most cited articles in the entire final sample, with 231 citations. This article tests the theory with new data related to the 2004 election, considered by the author to be a very important election, as it was the first after the implementation of the Euro and after the entry of eight new members, two from Southern Europe and another six from Post-Communist Europe. With tests using both aggregate- and individual-level data, Schmitt (2005) corroborated the main hypotheses of the theory: lower turnout and worse performance of government parties were registered again in 2004. Schmitt (2005) also corroborated that the lower turnout rates do not mean an European legitimacy crisis, since it is not related to a greater Euroscepticism. However, two other important hypotheses, one related to the electoral cycle and the other related to the worst performance of

smaller parties, were not supported in the newly integrated post-communist countries. This indicates that the theory of second-order elections did not satisfactorily explain European elections results in these countries, at least for that first round of elections. As stated earlier, this article has had a major impact on the research agenda by providing a comprehensive test of the theory and approaching it from a comparative perspective.

The next work to bring tests with individual-level data was only released in 2007, by Weber (2007). He focused his analysis on protest voting, seeking to unravel why it occurs to a greater degree in the middle of the electoral cycle and what are the incentives for voters to exchange their votes between one election and another due to discontent or anger with the party in which they voted before. For this, Weber (2007) places the perceptual agreement between voters and parties' policies when it comes to left-right ideology and integration issues as his dependent variable, and tests his propositions with data from the 1999 and 2004 European Election Studies surveys (ESS). Weber (2007) found out that although campaigns polarize public opinion, they are stronger when elections are of the national kind. In fact, parties avoid mobilizing voters in the middle of the electoral term, and above all, in European elections; and for that, they end up taking vague and ambiguous positions on issues related to integration. In addition, this text adopts a very critical tone of the way second-order elections theory was developed, stating that one of its main assumptions, that government parties lose more votes in the middle of the electoral cycle due to low popular approval, adopts an almost tautological logic. The author does not intend to disqualify the theory entirely, but rather to explain why government parties do not perform well in European elections in a methodologically and theoretically consistent manner. He also writes that:

> Once parties are assigned an active role, their strategic interest in maximizing support towards election day in their domestic arenas leads to a campaign cycle. As a result, high defection rates in EP elections held during the midterm can be explained by a mobilization deficit. More specifically, parties fail both in communicating their policy positions and in persuading voters to adopt them. This has far-reaching consequences because electoral mobilization can be shown to work only if parties present distinct programmatic alternatives to the public. Because prevailing retrospective evaluations of policy outcomes restrict the capacity of national government parties to follow this logic during the midterm, their cyclical losses ultimately reflect the intensity and effectiveness of domestic campaigning (WEBER, 2007, p. 531).

In any case, Weber (2007) had 46 citations in 14 years of publication, which gives an average of 3.28 citations per year, a value close to but still below the average for the sample, which, excluding the seminal article, was 3.83.

Also in 2007, Freire and Teperoglou (2007) tested the theory in some of the young southern democracies of the EU: Portugal, Spain and Greece. Their results showed that, in accordance with previous works, EP elections are used by voters to express content or discontent with national government, and this is especially true when governments are re-elected for a second or a third term. Data shows that governments do not benefit from a honeymoon period on re-elections, which explains some government losses on EP elections at the beginning of the electoral cycle. The propositions that large parties perform worse and small parties perform better on second-order election were only corroborated in Greece and Portugal. Unfortunately, this article is not indexed in Scopus, but has only five citations in Crossref, which indicates a low impact on the research agenda.

The following year, Magni-Berton (2008) analysed the effects of the electoral cycle on the 2004 European elections. He found out that the electoral cycle heavily influences the electoral results on the EP elections. Moreover, his results also showed that the governing parties tend to lose EP elections due to a greater differential abstention and that voters who place more emphasis on the role of the EU in their daily lives tend to vote more. However, this article had only five citations, which demonstrates not only its low impact on the research agenda but also the lower number of citations for articles published in languages other than English, since it was written in French.

The next article to study European elections as second-order ones with micro-level data was Marsh (2009). In this article, the author compares the performance of Campbell's (1960) Surge and Decline theory, more applied to US midterm elections, as well as the “referendum” theory with the second-order elections theory to see which one explains the EP elections better. However, he makes a confusing division of the hypotheses, distinguishing them by theories, despite the fact that the three theories studied have very similar propositions. Thus, although the fully corroborated hypotheses were not qualified by Marsh (2009) as specific hypotheses of the second-order elections theory, they are clearly also part of it; such as the hypothesis that governing parties lose, classified by Marsh (2009) as a specific hypothesis of the Surge and Decline theory. Still, the author found that second-order concerns do not affect voters decision to switch, but they do affect voters decision to abstain: if you support the integration, you are less likely to abstain. Alongside with Schmitt (2005) findings, this was only valid in old member states. Either way, Marsh (2009) does not test the postulate that smaller parties gain more votes, arguing that this will be tested in future work. Thus, despite some important tests having been carried out, the methodological design and the confusing and arbitrary distinction between the postulates of the theories studied are some of

the shortcomings of this work. Finally, Marsh (2009) only had 9 citations until December 18th, 2021.

Finally, there are two more papers published in 2009 with individual-level data that had a high impact on the research agenda. Clark and Rohrschinder (2009) divide voters into four types, seeking to define which variables impact the decision of these voters to vote against or in favour of the government or to abstain in European elections. They found evidence that voters' evaluations of the national government and the national economy influence not only defection but also levels of support for governmental parties in EP elections. But they also found out that the parties' performance at the EU level influences the extent to which voters defect or stay with the same party in EP elections. Contrasting with Schmitt (2005), Koepke and Ringe (2006), and Hix and Marsh (2007), these hypotheses were also valid for the countries of Eastern Europe, when tested with data from the EES 2004. Clark and Rohrshcinder (2009) had 67 citations on Scopus, an average of 5,58 citations per year, which is above the average of the final sample.

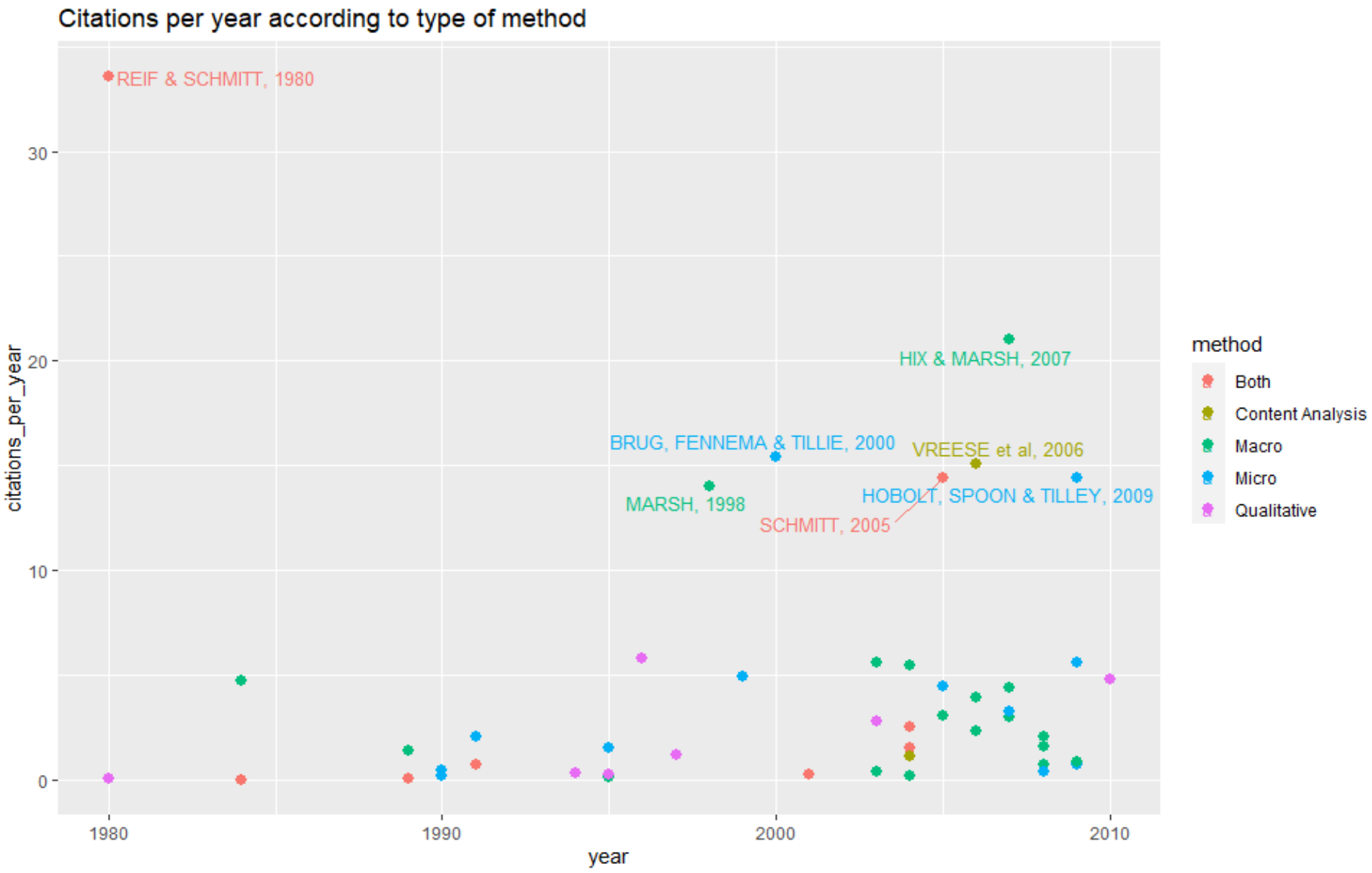
On their turn, disregarding the article by Brug, Fennema and Tillie (2000) by the motives already mentioned above, Hobolt, Spoon and Tilley (2009) wrote the most cited article with tests restricted to individual-level data of the final sample, with 173 citations until December 18th, 2021 and an average of 14.41 citations per year. They corroborated important propositions of the theory, such as the fact that defection or abstention from a party a voter chose previously is positively correlated with the distance between the voter's and the party's attitudes on European integration. In other words, when voters are less supportive of European integration than their preferred party, they are more likely to defect or abstain. Hobolt, Spoon and Tilley (2009) also found out that campaigns with a more Eurosceptic tone do not affect abstention rates but increase the likelihood of defection from governing parties. These are important findings, as previous works questioned to whom the votes lost by the governing parties were going to (did these voters abstain, or did they switch their votes?).

It can be seen then, from this brief exposition of the literature dedicated to analysing the theory of second-order elections with individual-level data, that a considerable part of these articles has a low number of citations, which denotes a low impact on the literature, at least until the middle of the first decade of the 20th century. In fact, as highlighted above, some of the first articles dealing with this type of data in the 1980s and 1990s did not bring robust and consistent statistical tests. Another possible explanation is that the emergence of online repositories may have facilitated access to survey data; as well as the emergence and improvement of softwares dedicated to analysing such data may have made

this process easier and cheaper. Indeed, chart 18 shows the distribution of citations per year of articles indexed in Scopus by type of data or method adopted. The visual illustration clearly demonstrates that articles that analysed the European elections with individual-level data existed, but they were not cited very often by other authors. Each colour represents a type of method. Articles that exclusively analyse aggregate-level data are marked in green, while articles that exclusively analyse individual-level data are marked in blue. Articles in red analyse both types of data. Articles in purple are those that have opted for qualitative methods. Finally, articles that opted for content analysis are shown in yellow. It is noticed that the articles by Brug, Fennema and Tillie (2000) and by Hobolt, Spoon and Tilley (2009) stand out. However, as highlighted previously, the article by Brug, Fennema and Tillie (2000) can be considered as an outlier, due to its analytical approach.

The table below describes the mean and the median of citations per year for each category. We chose to exclude the seminal article by Reif and Schmitt (1980) from the calculations, as this article is an outlier. We also calculated the mean and median for the group of articles that exclusively analysed individual-level data excluding Brug, Fennema and Tillie (2000), for the reasons previously mentioned. The results demonstrate that studies that used aggregate- and individual-level data obtained a similar mean and median, considering Brug, Fennema and Tillie (2000) in the calculations. Excluding this article, the mean and median for the group of articles that use only individual-level data fell by 22.40% and 22.77%, respectively. The mean and median of articles using data at both levels was surprisingly low; however, a good part of the articles in this group were published in the 1980s and 1990s, a period in which articles on the subject tend to have a lower number of citations. Finally, the mean and median of articles that opted for content analysis was very high, as only two of the four articles that adopted this method were indexed in Scopus, which invalidates the results obtained for this group.

CHART 18 – Citations per year according to type of method



Type of method	Nº of articles registered	Mean nº of citations per year	Median nº of citations per year
Aggregate-level data	18	4.161174	2.666667
Individual-level data	12	4.453298	2.67619
Individual-level data excluding Brug, Fennema and Tillie (2000)	11	3.455546	2.066667
Both levels	7	2.786975	0.7
Qualitative	8	2.041191	1.166667
Content analysis	2	8.092157	8.092157

Nevertheless, it is noteworthy that the sample consists of only 47 articles – 36 excluding Brug, Fennema and Tillie (2000), qualitative articles and those that used content analysis – which gives us a relatively low number of cases for so many categories. It is important, therefore, that further works seek to analyse a greater number of articles to confirm the results obtained in this review.

Finally, another indirect finding of this review was that the years 2007 to 2009 appear to be a turning point in the research agenda, in which more works sought to analyse the theory with individual-level data in a more consistent way. This becomes clearer when taking into account working papers presented at events in these two years in particular. In the second list sample, there were three working papers presented in 2007, two in 2008 and another three in 2009 that tested the theory with individual-level data. This finding may indicate that the research agenda was being renewed in this period and starting to be interested in analysing this type of data. However, to confirm these hypotheses, it is necessary that future works study the academic production on the subject from the year 2010 and afterwards, verifying whether the research agenda really started to focus more frequently and consistently on the analysis of individual-level data and whether this change persisted in subsequent years.

The following chapter will bring the data and results related to the variables of the substantive dimension. A brief discussion of the articles that mentioned the issue of the ecological fallacy also follows. Finally, this chapter will also contain a summary of the two reviews published between 1980 and 2010.

CHAPTER 04: Substantive Dimension – Analysis and Results

This chapter aims to present the variables related to the substantive dimension of the systematic review, as well as their respective results. These variables are related to aspects such as the dependent and independent variables studied in each article, if the articles mentioned the term "ecological fallacy" or not, what were the hypotheses and how many were corroborated, etc. In this chapter, the number of hypotheses supported in each article will also be compared to verify the existence of publication bias, as suggested by Cooper (2016) as a systematic review's fifth step.

Regarding the dependent variables studied by the articles in the final sample, it was noticed that they varied a lot according to the analytical approach adopted by the authors. Up to three dependent variables were catalogued per article, although the vast majority only analysed one (37 out of 54 articles, 68.51% of the final sample). For purposes of comparison and counting, we opted for a second cataloguing, in which we grouped the different dependent variables into thematic groups. For example, about 17 articles dealt with vote switching between elections; however, some focused only on government parties, others on small parties, others analysed all parties. Even so, all of these articles had vote switching as their dependent variable, and were regrouped into a single thematic group. Similarly, previous reviews or works that analysed the second-order character of these elections were grouped together as having the theory itself as the dependent variable.

Among the 78 dependent variables found, 19 of them were related to turnout in general (about 24.35% of the sample). Another 17 were categorized as related to vote switching between elections, as mentioned above (21.79%). Seven cases of dependent variables related to the theory were registered, as well as another five dealt with electoral results in general. The other categories had three or fewer cases. 20 categories were catalogued, and all of them are described in the table below with the respective number of occurrences in the final sample.

Dependent variables' categorization	Nº of occurrences
Turnout	19
Vote switching	17
Theory	7
Electoral results in general	5
Campaigns	3
Specific cases analysed as a whole (i.e., 2004 Elections in Romania and Bulgaria – STEFANOVA, 2008)	3
News related to the EP elections	3
Probability of voting or abstaining	3
General political mobilization	2
Voters' interests	2
Voters' party preferences	2
Past votes (survey data)	2
Government popularity / support	2
Satisfaction with EU / national democracy	2
Party competition	1
Election polls	1
Electoral results + electoral cycle	1
Parties' manifestos	1
Parties' and voters' agreement on European issues	1
EU and national institutions political power	1

Regarding independent variables, these are even more diverse, as for each type of dependent variable it is possible to list numerous independent variables. This made it impossible to categorically catalogue independent variables, making it difficult to systematize so much information. What is possible to determine is the average amount of independent variables per article. The article with the most catalogued independent variables was Clark and Rohrschinder (2009), with 24 of them; followed by the seminal article by Reif and Schmitt (1980), with 17 independent variables. Some articles were coded with zero independent variables (e,g, case reports, some literature reviews, etc.), since their research design dispenses the relationship between dependent and independent variables. The mean for the entire sample was 5.65 independent variables per article while the median was 5.

Excluding the nine articles without independent variables, the mean rises to 6.82, while the median rises to 6.

On the other hand, some articles distinguished their variables between independent and control. In this small group of 12 articles, the control variables were coded as a separate variable on the coding sheet. Thus, for a more reliable assessment, the number of independent variables was added to the number of control variables of these 12 articles. Including the control variables, the sample mean rose to 6.59 and the median rose to 6. Excluding again the nine articles without any variable, the mean rose to 7.91 and the median rose to 7.

Regarding the issue of ecological fallacy, only five articles addressed the topic. The first of these was the seminal article, which demonstrates that Reif and Schmitt (1980) had been concerned with this issue since 1980. In their article, they claim that inferring from aggregate-level data was necessary given the numerous problems of survey methods. According to these authors, surveys carried out before two weeks of the elections tend to overvalue support for the opposition, as some voters would make their decisions just before the elections. Furthermore, polls carried out on election days provide only indirect information about absentee voters – a crucial factor in analysing second-order elections. Post-electoral surveys, on the other hand, tend to overvalue voting for the winning parties and candidates. Finally, the authors point out that all surveys would have problems in correctly capturing the relationship between attitudes and the behaviour that actually occurred, but they do not detail this problematic aspect on the article.

The second work to mention the issue of ecological fallacy was Sinnott and Whelan (1992). In this article, the authors analyse turnout for European elections, but only for the city of Dublin in 1984 and 1989. By reducing the scope of the article to the number of voters in Dublin, it was possible to carry out tests at the individual level of analysis but without compromising the results with the inherent flaws of the survey methods, since the authors were working with the real turnout of the elections, at least for the Dublin city level. However, the authors operationalize their model at the level of the city's electoral sections, reaching an *n* of 447 sections, considering the two elections together. This made it counterproductive to code Sinnott and Whelan (1992) as an article testing the theory with individual-level data. Still, the article deals with the issue of the ecological fallacy in relative depth. More specifically, the authors criticize the research design adopted by Schmitt and Mannheimer (1991) to measure the determinants of turnout in the 1989 European elections. For Sinnott and Whelan (1992), Schmitt and Mannheimer (1991) used survey data that largely

conflicted with official data on participation at the aggregate level, in addition to operationalizing the variable "habitual voting" in a methodologically questionable way. Using individual-level data to make inferences about aggregate-level data may have led to the result obtained by Schmitt and Mannheimer (1991) that voters would vote largely out of habit, something disputed by Sinnott and Whelan (1992). When reviewing the literature developed so far, Eijk, Franklin and Marsh (1996) also discarded the "habitual voting" hypothesis.

In turn, the third work to mention the issue of the ecological fallacy was Matilla (2003). He quickly addresses it, stating that the use of aggregate-level data to make inferences at the individual level can be problematic. But he does not delve into this issue, citing only the work of King (1997). The author also briefly discusses the methodological division existing in the area of studies on European elections and claims to be in favour of methods that merge different levels of analysis (although he restricts his analysis to the aggregate level, according to him, for practical reasons).

The next work to comment on this topic was Fauvelle-Aymar and Stegmaier (2008). The authors recognize the difficulty in finding survey data indicating the location of the respondent, using this as a justification for restricting their analysis to aggregate-level data; however, not at the country level, but at national sub-regions level. About the issue of ecological fallacy, the authors write the following:

> Nevertheless, since our analysis is at the aggregate level, even if the use of regional data offers the possibility to better assess the influence of the economic situation on turnout, the aggregated nature of the data does not allow us to make any inference about individual-level participation in order to avoid the ecological fallacy. In fact, our analysis is what is called ''an ecological analysis'' where the aim is to examine the relationship between the characteristics of the voters' territories and the electoral results. How does the environment influence voters' behavior? This is a very important question that has been largely neglected with the dominant use of individual survey data, where there is no indication of localization (FAUVELLE-AYMAR & STEGMAIER, 2008, p. 666).

Finally, the last article to mention this issue in our final sample was Clark and Rohrschinder (2009). In this article, the authors argue that aggregate-level data only show that party size matters for European election results: larger parties usually lose votes, and smaller parties usually do better. However, there is no way to assess why this happens only with aggregate-level data in a consistent and reliable way. The theory's interpretation is that voters aim to punish the governing parties, or else they end up voting for the parties they like best, since they don't have to worry about which party is more likely to form a future government. However, it would be entirely possible, according to Clark and Rohrschinder (2009), for

voters to opt for different parties because the European arena is different. In other words, voters would vote with European issues in mind, which would minimize the second-order character of EP elections. They could vote for green parties, for example, because they understand that environmental issues are better discussed and dealt with at the community level. Thus, for the authors, it was possible that the excessive focus on analysing the elections with aggregate-level data would lead to incorrect inferences. However, the results obtained by the authors with data at the individual level actually corroborated the assumptions of the theory, which further consolidated the theory almost 30 years after the publication of the seminal article.

4.1. Cooper's (2016) Fourth Stage - Hypotheses Bias Test

As stated in the previous chapters, one of the steps in the systematic review process is the analysis and synthesis of collected studies, in the case of a meta-analysis. It is noteworthy here that the synthesis is more than a mere exposition of the results already obtained in the literature, which was done here in the last three chapters. Synthesis involves, for example, in the case of experimental works with similar research designs, summing up the sample of all the experiments and comparing whether the proportion of findings was the same for each of the works. This type of synthesis is easier to be carried out in works in the Medical and Health Sciences field, for example, in which experiments with drugs or some type of treatment are carried out by different authors with different sample sizes and cuts. The objective of this step in a meta-analytic review is precisely to compare whether such works obtained similar results, and if not, to seek answers to these differences.

Figueiredo Filho *et al* (2014) also suggest other forms of synthesis, which mostly involve observing the significance of the results obtained in relation to the null hypothesis to determine the direction of these results, as well as their graphic distribution. Other more complex and sophisticated means involve the aggregation of regression coefficients; however, this form of synthesis requires that different studies have measured their dependent, independent and control variables in a similar way, which makes this type of synthesis difficult to make.

As this review deals with a very comprehensive topic, in which the articles not only have different dependent variables but also different ways of measuring these variables, it becomes impractical to make a quantitative synthesis of the results obtained in the literature.

Furthermore, not all articles included in the review adopt quantitative methods, and even those that do end up using different types of data and analysis methods.

Although this review is not meta-analytic precisely because of its research object, it is still possible to adapt some type of evaluation to this research design. In this case, instead of testing the proportion of findings in the expected direction, it was decided to analyse the proportion of hypotheses corroborated or not for the total of hypotheses tested in the literature until 2009, the time limit of this review. One of the reasons for this is to check for the existence of publication bias. Publication bias occurs when the result of a particular research influences the decision of editors or reviewers to publish it, usually leaning towards statistical significance or, in other words, in favour of positive results. This phenomenon was already extensively studied, especially in the Health Sciences literature (SONG *et al*, 2010; DWAN *et al*, 2013; SONG, HOOPER & LOKE, 2013). Publication bias can also lead to direct or indirect manipulation of data or redesign of the research until its findings reach the expected significance, compromising the advancement of scientific knowledge (PEARCE & DERRICK, 2019).

However, one of the problems in carrying out this stage of the research was that not all authors explicitly define their studied hypotheses. As highlighted in chapter three; no matter how careful the coding of this variable has been, the interpretation of hypotheses based on the text itself and its analysis remains a summarily subjective criterion that can impact the reliability of this step's results. The finding described in chapter three, that articles that clearly define which and how many their hypotheses are, tend to have, on average, one more hypothesis than articles that do not reinforce these suspicions.

Thus, this step was divided into two parts. The first analysis of the proportion of supported hypotheses considers the total number of articles in the final sample, excluding some case reports and some reviews, which do not have any hypothesis. Thus, the total number of works included in this first part was 50. The second part analyses only the 26 articles that explicitly defined how many and what their hypotheses were.

Regarding the 50 articles in the sample with hypotheses tests in general, 232 tested hypotheses were coded, of which 140 were fully corroborated, approximately 60.34% of the total. Another 48 hypotheses were partially supported (20.68%) while 44 of them were refuted (18.96%). Among the articles, 14 of them (28%) fully corroborated all their hypotheses, while another six (12%) did not fully corroborate any of their hypotheses. Considering these numbers, it is not possible to affirm that there is publication bias.

However, as more than half of the articles did not clearly define their studied hypotheses, it is necessary to compare these results with the numbers and proportions obtained considering only the 26 articles that delimited their hypotheses. In this subset, 145 hypotheses were catalogued, of which 88 were fully corroborated (60.68%), 32 were partially corroborated (22.06%) and 25 were refuted (17.24%). The proportion of articles in which all variables were fully supported was also like that of the full sample: seven out of 26 articles, approximately 26.92%. The same occurred with the proportion of articles that did not fully corroborate any hypothesis: three out of 26 articles, approximately 11.53%.

Thus, the tests carried out indicate the absence of publication bias in the research agenda regarding second-order elections theory and EP elections until 2009.

4.2. What Was Already Known: a Summary of the two Previous Literature Reviews

Our search found only two major literature reviews about second-order elections and EP elections, neither of which were of the systematic kind. The first review was published by Eijk, Franklin and Marsh (1996), while the second one was published by Marsh and Mikhaylov (2010).

As for the first review, Eijk, Franklin, and Marsh (1996) first briefly explain the assumptions of the second-order elections theory, highlighting the influence that first-order issues have in the second-order arena. Although the four main propositions of the theory (lower turnout, larger parties perform worse, governing parties perform worse, newer and more radical parties perform better) had already been supported by several studies in each of the election rounds, the authors demonstrate that much of this variation had not yet been explained. Furthermore, many of the variables of interest to the theory are so interconnected that it had not yet been possible to distinguish between their effects individually. For example, knowing that government parties also tend to be larger and that newer parties also tend to be smaller and more radical, which of these factors ends up influencing vote choice more: party size, government vs. opposition status, party age, or party radicalism? Another issue that was not clear at the time was the impact of the media and campaigns on elections. It was known that parties spent less on second-order campaigns, and that media coverage was not as intense either; but large-scale studies of the effects of these variables were still a gap in the research agenda. Eijk, Franklin, and Marsh (1996) also draw attention to other issues not addressed by Reif and Schmitt (1980) in their seminal article: why voters would seek to impact national politics in an European election (e.g., not supporting parties' government officials)? What is

the impact of this behaviour on national policy (if any)? Does this behaviour change if a national election is held shortly after the European elections?

In this sense, Eijk, Franklin and Marsh (1996) add to their review tests of hypotheses related to the theory. Therefore, despite the article being a literature review, it was also catalogued as having hypotheses and quantitative methods of analysis. The authors state that, in addition to reviewing what had been done so far, another objective of the article was to take a deeper look at turnout and party choice in European elections and what were the consequences of these elections both for the national arena and for the European arena. The authors aimed to analyse how much second-order elections theory explains these variations, but also how much European elections themselves can help to explain issues and behaviours related to national elections. Thus, this article already assumes that it is possible for European issues to exert a relative impact on campaigns and voters' choices, relativizing the second-order character of these elections.

As for turnout, one of the main research objects of the theory, Eijk, Franklin and Marsh (1996) state that previous works usually attributed the low turnout of European elections to its lower salience among voters and its lower impact on European Community/European Union policy-making. They then raise several factors that impact the turnout and help explain its variation over time and between member countries, such as the novelty factor of the first elections and compulsory voting. They cite Blumler and Fox (1982), who argued that the culture politicization and, especially, positive views on Europe positively impacted turnout. Other works, such as Eijk and Oppenhuis (1990) and Eijk and Schmitt (1991), showed that being more favourable to integration not only foresees voting in European elections, but also in national elections; which indicates that, in fact, the main predictor of voting in these elections would be a personal interest in politics, and not necessarily being pro-integration. Schmitt and Mannheimer (1991), in turn, argue that age, interest in politics and, above all, attachment to political parties contribute to predict turnout, while the salience of European issues did not have a general effect. These findings confirmed the theory's main postulates, as in elections with low salience, voters would vote out of force of habit or social norms, or else because of partisan attachment and personal interest in politics. In elections with high salience, other factors would come into play. However, Eijk, Franklin and Marsh (1996) state that Schmitt and Mannheimer (1991) do not consider that these factors were also present in Greece and Ireland in 1989, countries that held national elections together with European ones, indicating that the factors they listed would also predominate in first-order elections. Eijk, Franklin and Marsh (1996) conclude that attempts

to understand turnout in European elections based on what was already known about turnout in national elections proved insufficient.

They then propose the opposite path: analysing turnout in national elections based on what was already known about turnout in European elections. In this sense, they cite Franklin *et al* (1996), who stated that turnout is highly dependent on the context in which elections are held in each country. In other words, in addition to compulsory voting, factors such as the proportionality of the electoral system and the day on which elections are held have a significant impact on turnout. In the case of European elections, another strong impact factor is the holding of local or national elections in conjunction with EP elections. Marsh and Franklin (1996) also reported that turnout for European elections increases when national elections are held shortly thereafter. Finally, Franklin (1991) and Franklin *et al* (1992) also point out that the decline in the strength of party cleavages was at the time the main responsible for the decline in turnout, both nationally and in Europe, as political parties were having more difficulty in mobilizing voters by appealing only for the loyalty of their supporters.

Finally, Eijk, Franklin and Marsh (1996) state that the tendency of turnout to be greater in European elections held shortly before national elections is not only due to the greater salience of the national arena for voters, as stated by Patterson and Caldeira (1993) and Mitchell and Wlezien (1995). Citing Marsh and Franklin (1996), Eijk, Franklin and Marsh (1996) argue that this salience is also heightened by the holding of the European elections themselves, as they give the electorate the chance to respond on parties' performance before the national elections. But to state this clearly, according to Eijk, Franklin and Marsh (1996), it was necessary to look at the voters' party choice.

Regarding this last topic, the authors state that explanations about party choice at the European level are even less satisfactory compared to the already existing explanations about turnout. Even when carried out together or closely, party choice in European elections seems to follow a different path from that in national elections. However, this difference is not based on party supply, as it is basically the same in both types of elections and transnational party groups have a marginal impact on party choice at EP elections.

Another well-known fact about European elections is that the results tend to disproportionately benefit smaller parties, and that larger parties and those that make up the government tend to perform worse. The authors ask themselves, however, what determines the degree of these gains and losses among these different sets of parties or even between parties within the same set. They mention that previous works such as Eriksson (1988), Reif

(1984, 1985) and the seminal article by Reif and Schmitt (1980) itself suggest a "punishment effect", which would explain the low performance of government parties. Reif (1984, 1985) argues that this effect only occurs in some circumstances, more specifically towards the middle of the electoral cycle, when government popularity is at its lowest. However, Eijk, Franklin and Marsh (1996) draw attention to the fact that this punishment effect hypothesis contradicts another hypothesis of the theory: that voters would vote "with their hearts", choosing the parties closest to their ideologies without taking into account government formation processes or whether their vote will be "wasted". Even in second-order elections held shortly after national ones, the tendency is for the results of the first-order election to be supported in the next election, rather than for voters to vote "with their hearts" after they have already voted tactically in the previous election.

Eijk, Franklin, and Marsh (1996) then criticize the tests done by Reif (1984) on this issue, due to the low number of cases and the way he coded the variable "decrease in government support", stating that his results are only suggestive. When comparing them with the results obtained by Marsh and Franklin (1996) years later, the authors see more clearly the deficiencies of Reif's (1984). Marsh and Franklin's (1996) model, with a bigger sample and defining "decrease in government support" as the difference between the votes obtained by the governing parties in national and European elections, showed only partial support for the electoral cycle hypothesis, which would explain only 18% of the variation in the difference between the votes for governing parties in each type of election. Despite this, neither model accurately answers whether voters actually vote differently in both elections.

A possible explanation for this question is in the work of Oppenhuis *et al* (1994, 1996), who found that the electoral cycle mattered but not as a proxy for government popularity, as described by Reif (1984). More specifically, when European elections are held together or shortly after national elections, they appear to be "throw-away" elections, as political parties' position markers in the view of the electorate have already been established in national elections. In this sense, any distinction between the electoral results of these elections will make no difference, as the focus is on the results of the national elections. Apparently, voters know this and end up voting more sincerely in the European elections, which benefits smaller parties. On the other hand, when European elections are held at a more considerable temporal distance from national elections (especially in a scenario of preparation for the following national elections), European elections gain greater importance, as they are seen – by the media, by the parties and by the voters – as a marker of parties' strength in the electorate's view. Although there is still "less at stake" compared to first-order elections,

voters have a greater incentive to vote not "with their heart", but to protest against the government. Although Oppenhuis *et al's* (1994, 1996) model predicts the "protest voting" phenomenon, it does so in a different way from Reif's (1985), since in the latter "protest voting" would occur in the middle of the electoral cycle, while Oppenhuis *et al's* (1994, 1996) model opens more time space for this type of voting. Thus, the results obtained by Oppenhuis *et al* (1995, 1996) do not necessarily disprove Reif's and Schmitt's (1980) second-order elections theory. They actually refine it, as they specify in which situations the "sincere voting" and the "protest voting" take place.

Then, Eijk, Franklin and Marsh (1996) state that most works until then ignored party size as a factor influencing party choice, although this relationship is relatively clear, as larger parties will certainly have a greater impact on public policies, for example. In this sense, the authors highlight their findings in another work (EIJK & FRANKLIN, 1996) that party size was one of the most important determinants of vote choice – even more important than party loyalties and ideology. Institutional differences between countries, such as the number of effective parties or the type of electoral system, had very little effect on vote choice, contrary to what was previously known about turnout.

Although the second-order election model suggests that national elections are more likely to affect European elections than the other way around, Eijk, Franklin, and Marsh (1996) argue that the two affect each other mutually. In this regard, they cite numerous works reporting the varied consequences that European elections had on the national political arenas of member countries of the European Community or the European Union. One of the most notorious consequences is that the European elections played a key role in the rise of the Front National in France and the Green Parties in Germany. Oppenhuis (1996) draws attention to the encouragement of certain electoral pacts, while Cayrol and Ysmal (1996) emphasize the encouragement of party divisions due to European elections. Castillo (1996) highlights that the Spanish government even anticipated elections in view of its popularity in European elections. In turn, Marsh and Franklin (1996) demonstrate that when government parties do much worse or much better than expected in European elections, the same occurs in national elections, confirming that EP elections can be seen as markers of the performance of these parties.

However, Eijk, Franklin and Marsh (1996) point out that European elections failed to exert substantial impacts on the European political arena. Such elections do not fulfil the normative expectations of common democratic processes. They were expected to increase citizens' knowledge of European institutions, in particular the EP itself, and stimulate debate

on European issues, bringing legitimacy both to the institutions and to the community agenda. This unfortunately did not happen: the turnout had been in decline since 1979, the general interest in integration did not grow and the campaigns barely discussed European issues. The authors agreed that many of these expectations were unrealistic as they were based on what took place in national elections. Although some authors claim that the EP elections are flawed because voters unhappy with integration have little or no opportunity to express this discontent through their vote, as most of the main parties are, to a greater or lesser degree, favourable to integration; Eijk, Franklin and Marsh (1996) affirmed that there was a great degree of correspondence between the parties' and their voters' views on integration; which gives a certain degree of legitimacy and representation to the EP. These last results were found in Eijk and Franklin (1991) and Schmitt (1994). However, Eijk, Franklin and Marsh (1996) claim that while European elections remain second-order ones, voters' preferences will also remain irrelevant to decision-making in the European arena, which is a serious problem in European institutional design.

Finally, Eijk, Franklin and Marsh (1996) take a very critical stance on the dichotomization of the second-order elections theory by Reif and Schmitt (1980). They agree that the theory has generated important propositions about the main patterns in the results of European elections, but argue that it is not possible to say that European elections, like other types of elections, are unquestionably second-order; and that it is always national elections that affect second-order elections, not the other way around. In this sense, the authors propose an evolution of the theory, which instead of understanding European elections invariably as second-order national elections, now understands them from different political contexts, focusing on how much there is "less at stake" and on how much each election was more or less second-order.

The second review, by Marsh and Mikhaylov (2010), adopts a less critical tone, aiming to demonstrate how the theory of second-order elections became the dominant approach in the analysis of EP elections 30 years after the realization of the first elections. The authors emphasize that the theory has not only been consolidated in the area of European elections, but has also expanded to other topics such as referendums and regional elections. Initially, Marsh and Mikhaylov (2010) highlight two important findings in the post-1996 literature: some elections are more second-order than others (HEATH *et al*, 1999; RALLINGS and THRASHER, 2005; SKIRINIS and TEPEROGLOU, 2008); and the theory did not fit so well in post-communist countries, at least for the year 2004 (SCHMITT, 2005).

Regarding the performance of certain groups of parties in European elections, post-1996 work corroborated that green parties do not necessarily always perform better in EP elections compared to national elections (MARSH, 1998; HIX and MARSH, 2007). However, tests with data at the individual level showed that voters who valued environmental issues and who perceived the EP as an important institution tended to shift their vote to green parties in European elections (CARRUBA & TIMPONE, 2005). The hypothesis concerning the better performance of extremist parties had mixed results, with some papers finding no evidence of this (MARSH, 1998) and others finding that Eurosceptic parties typically perform slightly better in European elections than in national elections (MANOW & DORING, 2008). By analysing survey data, Hobolt, Spoon and Tilley (2009) refined this argument, corroborating that voters who are more Eurosceptic than the parties in power in their respective country tend to vote for more Eurosceptic parties in European elections.

As for the turnout, subsequent works showed that it can be affected by issues such as: the fact that the country is a net beneficiary of EU investments and/or hosts an important European institution (JESUIT, 2003; MATTILA, 2003; STUDLAR, FLICKINGER & BENNETT, 2003); regional unemployment rates (FAUVELLE-AYMAR & STEGMAIER, 2008); Euroscepticism in post-communist countries (SCHMITT, 2005); among others. Nevertheless, the literature found only mixed evidence regarding retrospective voting by economic criteria in European elections. Kousser (2004) found evidence that a higher unemployment rate is correlated with a better performance of opposition parties in EP elections; however, economic growth and inflation had no statistical significance. On the other hand, Marsh (2003, 2005, 2007) did not find evidence of retrospective voting by economic criteria. In a more sophisticated research design, Tilley, Garry and Bold (2008) found that only the most politically sophisticated voters in specific contexts such as single-party countries with frequent government shifts vote with clear economic criteria in mind.

On the other hand, Marsh and Mikhaylov (2010) emphasize that the theory also had its criticisms and rebuttals. More specifically, Blondel, Sinnott and Svensson (1997) question whether there is actually less at stake in European elections, stating that the difference in opinion among voters regarding the degree of importance of who wins national and European elections does not vary much. However, Marsh and Mikhaylov (2010) demonstrate that later studies corroborated this hypothesis, such as Orford *et al* (2009), who found that the distance between polling centres has a more negative impact on turnout in local and European elections than in national elections.

Another central criticism of the theory is the argument that European issues matter, whether for parties or voters; since the theory predicts that national issues will be dominant. According to Marsh and Mikhaylov (2010), this issue had already been addressed by studies with data both at the aggregate level (HIX & MARSH, 2007; ARNOLD & PENNINGS, 2009) and at the individual level (TILLEY & WLEZIAN, 2008; MATHER, 2001; NIELSEN, 2001; WORRE, 1996; YSMAL & CAYROL, 1996). Typically, such works shows that European issues matter, but only to a few voters and to a lesser degree, particularly in the case of tests with individual-level data. Other works that analyse this type of data demonstrate that voters know little about parties' positions on European issues, either because most larger parties are in favour of integration in general or because since they focus on national issues during the campaign, they do not end up differentiating themselves on European issues (HOBOLT, SPOON & TILLEY, 2009; CLARK & ROHRSCHNEIDER, 2008; BRUG, EIJK & FRANKLIN, 2007, BRUG *et al*, 2007).

Sanz (2008) states that ideology impacts more on vote choice in European than in other types of second-order elections. Ferrara and Weishaupt (2004) argue that party cohesion around European issues also affects their electoral results. Finally, Weber (2007) and Tóka (2007) corroborate that, as parties invest less in the campaign in European elections, voters become less interested, which decreases the turnout.

Marsh and Mikhaylov (2010) conclude their review by stating that the research agenda needed to move forward in two directions. Firstly, it was necessary to better study the mechanisms that fuelled the second-order effects in these elections, to further analyse how voters choose their votes. Secondly, it was necessary to provide more quality data, especially panel data that would facilitate analyses regarding vote change and turnout patterns in these elections.

After this summary of the results reported in previous reviews, the conclusion of this dissertation follows, in which the results described in the previous chapters will be summarized. The limits of this review will also be discussed, highlighting our time frame choice – between 1980 and 2009.

CONCLUSION: Comparing Reviews and Limits of our Design

As stated earlier, one of the objectives of this research is to publicize the benefits and importance of systematic reviews for the advancement of scientific knowledge. And one way to do this is to compare the results obtained here with what has already been accumulated by previous reviews.

Firstly, it is expected that the amount of literature cited will be greater in literary reviews due to the lack of specific and standardized excluding criteria while searching and selecting the texts. Furthermore, in our specific case, this review chose to exclude books and book chapters for reasons already specified before, which at once considerably reduces the relevant literature available for analysis.

It was noticed that among the 52 works cited by Eijk, Franklin and Marsh (1996), only 7 were cited in this review (approximately 13.46%), while at least 18 articles that were published in this period entered our final sample. In other words, the systematic review method found at least 11 other potentially relevant articles not addressed by the first comprehensive literature review on the subject. The fact that most of these early works on the subject are old articles that could still be found in full in digitized format further demonstrates the importance of systematized collection methods. The existence of online repositories today also facilitated such searches. Only one article cited by the first review was not found (EIJK & OPPENHUIS, 1990). Another five works cited by Eijk, Franklin and Marsh (1996) were excluded by the second list criteria adopted by this review, while 36 works were excluded for being books or book chapters, articles that do not mention the term “second-order” or have been published before 1980. Finally, three working papers mentioned by the authors of the first review were not found, as expected due to their age.

The second review by Marsh and Mikhaylov (2010) mentioned a much larger amount of work: 145 in total. Among them, 32 entered our final sample (approximately 59.25% of our final sample). Another 19 articles were excluded by the second-list criteria adopted by this review, while 66 articles were mostly excluded because they were books or book chapters or articles that did not mention the term “second-order”. Other 27 works cited by Marsh and Mikhaylov (2010) were published either before 1980 or in 2010, and therefore were also excluded from this review, although some of them were cited in the introduction, such as Tufte (1975) and Campbell (1960). Only one article was not found in our searches (JESUIT, 2003); however, by the criteria adopted here, it would not have entered the final sample.

This small comparison illustrates that systematic reviews can find texts that have not been much covered in the literature and not cited by previous literary reviews. This is important for the advancement of scientific knowledge as literary reviews usually only consider the author's prior knowledge of the published literature, which usually favours more commented and impactful works, or even the author's personal tastes. Taking this logic to an extreme, newer authors could have difficulties in inserting themselves in the research agendas, as well as new hypotheses and analyses could not receive due attention in the literature, which would be focused on rewriting and reanalysing old concepts.

On the other hand, systematic reviews can also be limited by the adoption of strict exclusion criteria. This was precisely the case of this review, which, in addition to having adopted numerous exclusion criteria, especially those relating to books and book chapters, it was also analytically designed to explore the issue of the level of analysis of published articles in the field of second-order elections and EP elections. It was noticed that important analyses and findings were left out because they were published in books.

However, precisely because of the analytical approach and the search tools adopted, it was demonstrated here that tests with individual-level data were already carried out in the 1990s. In addition, the greater amount of information collected in this systematic review allowed us to discover important findings. Collecting and analysing the articles' number of citations, something that had not yet been done in the literature, allowed us to discover that some of the first tests with individual-level data had no great repercussions in the literature in general. Analysing the proportion of corroborated hypotheses allowed us to verify that there was no publication bias in this research agenda, at least until 2009. Collecting data on universities and on the authors' gender allowed us to verify the centrality of German, Dutch, North-American and British academic production on this topic, as well as pointing out the large gender disparity that characterized – and probably still characterizes – this research agenda. Collecting and exposing data in chronological order allowed us to show how the theory has been refined over the years and how hypotheses such as the "habitual voting" lost strength over time while others such as "some elections are more second-order than others" and "European issues matter, but at a minor level" have been gaining prominence and relevance.

Even some limitations in our research design allowed us to reveal interesting findings about the linguistic bias that exists in adopting the impact factor of journals as a criterion to measure the quality of publications. Non-English-language journals traditionally have lower impact factors, and adopting this feature as an exclusion criterion may not be an

optimal choice in multi-lingual reviews. Furthermore, journals from outside the main scientific circuits – notably, those from countries on the periphery of global scientific production – may also have a lower impact factor, which should be taken into account in the research design of reviews that include journals from countries with different degrees of development. In turn, although this work is restricted to EP elections, our broad search criteria allowed us to find works that apply the theory in other countries and in other types of elections, as reported in the introduction.

However, one of the most defining cuts of this review is the time frame, which was restricted from 1980 to 2009. One of the reasons that this cut was so precise is the huge amount of published works that mention the second-order elections theory from 2010 onwards. Chart 19 illustrates the amount of works found in Google Scholar per year when searching for the terms “second-order elections” and “European Parliament” together. These numbers refer to December 18th, 2021.

As can be seen from the graph, the number of works found on Google Scholar on this topic enters an accelerated pace of growth from the turn of the 20th-21st century until the year 2015, with a peak of 319 entries found only in this year alone, 285 of them with the search terms in English. Even though such numbers are skewed upwards, since the search with terms in English also finds works in other languages, as described in the first chapter; such numbers still are impressive. The low amount of works found in the three remaining languages for the following years also makes us believe that the percentage of non-English works found in the search with terms in English is not that significant. Another interesting pattern highlighted by the image is that the number of works on the subject increases in the years in which there are elections and in the years after the elections, and then usually stabilizes or decreases and starts to grow again when a new election approaches; showing that interest in the subject of second-order elections and European elections naturally increases when elections approach and, above all, when a new parliamentary cycle begins.

Therefore, one of the reasons this review was restricted to articles published between 1980 and 2009 was the immense amount of texts found in the following years, which made it impossible for more years to enter the review, since the collection and analysis of all texts was done manually. Until 2009, Google Scholar searches found 984 works (counting duplications), while from 1980 to 2020 this number increases to 3579 (also counting duplications). This demonstrates that the amount of texts found increased by 363% in just 11 years, making it impossible to check, collect, filter and analyse so much data in the time allotted for completing this research.

CHART 19 – Number of entries found on Google Scholar by year

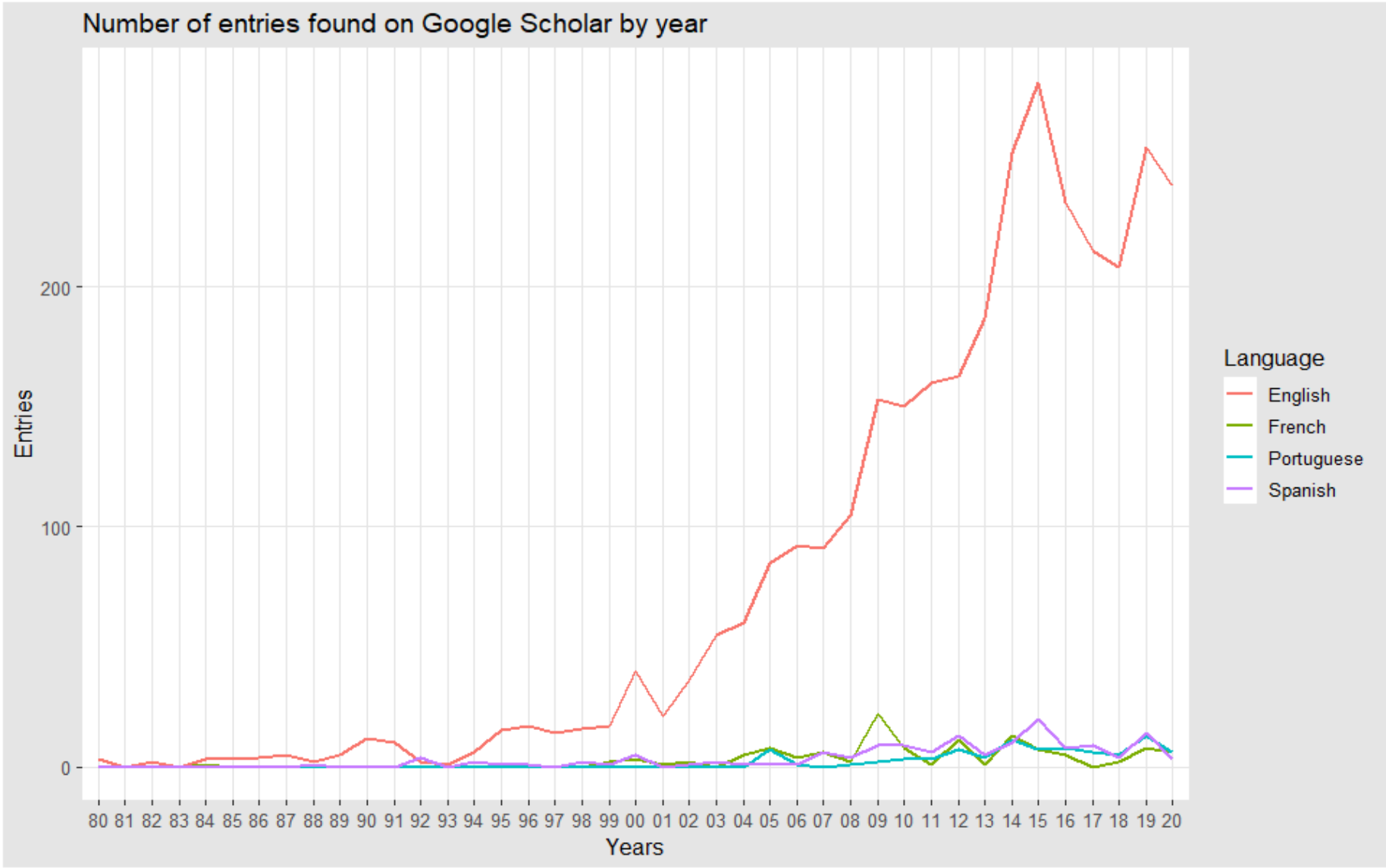
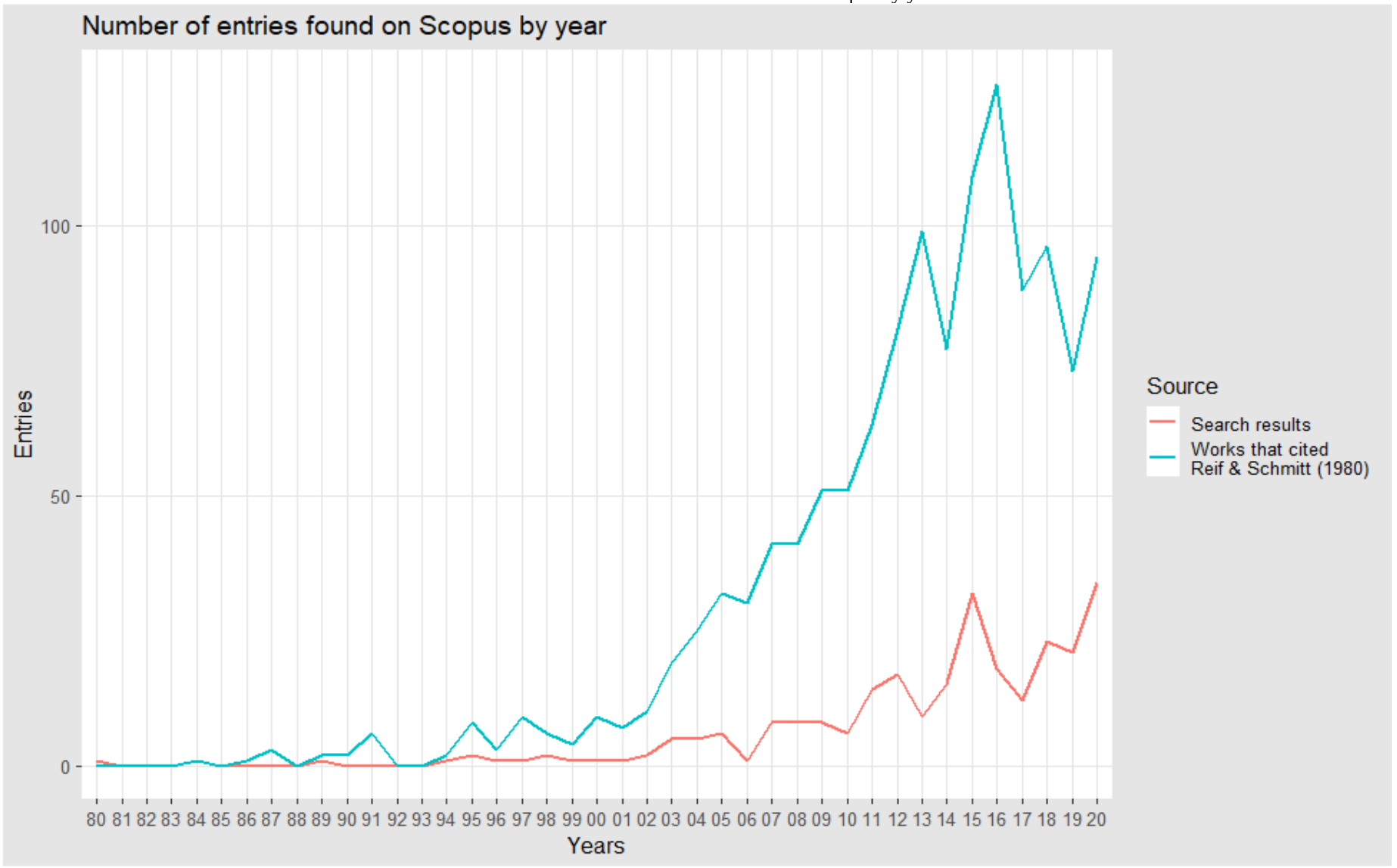


CHART 20 – Number of entries found on Scopus by year



Although there is a greater amount of uninteresting work for this review in Google Scholar searches, as demonstrated in Chapter 1, the same pattern described above also occurred in the other two search engines used in this review. Chart 20 above illustrates this pattern. The blue line is the list of articles and books indexed in Scopus that cite the seminal article by Reif and Schmitt (1980), and the orange line shows the result of the search with the terms "'second-order' AND 'second-order elections' OR 'European Parliament' OR European elections'" also in Scopus.

Thus, despite the amount of data collected and the analyses carried out in this review, it is worth noting that it is also temporally restricted and that it cannot be assumed that the patterns found for the first 30 years of academic research on the subject necessarily continued in the following 12 years. It is possible that, in this new and immense group of works, some publication bias has emerged, favouring works with a greater number of hypotheses whose results were statistically significant. Likewise, it is possible that the gender disparity has not reduced as much as expected in the course of this time. Furthermore, it would also be interesting to analyse whether, in these years of intense production on the subject, new authors and new universities began to stand out in the literature. It is possible that, with EU's expansion to the East, new authors from Eastern Europe have inserted themselves in this research agenda, dynamizing and diversifying it.

Finally, it is known that new and better panel data, as well as greater academic attention to tests with data at the individual level, encouraged the publication of more works in this topic in the following 12 years. Although the trend apparently was to confirm and refine the original postulates, the politicization of European issues, the advance of Eurosceptic parties and the raise in turnout rates in 2019 brought new challenges to the theory.

Such aspects must be considered in future research and, above all, in future reviews. It is believed that this review was an important first step towards the systematic accumulation of knowledge already produced on the theory of second-order elections and EP elections. Finally, it is expected that future works cover not only the remaining literature, but also opt for systematized review designs, overcoming the limitations of this work and including other languages and research subtopics.

BIBLIOGRAPHIC REFERENCES

ARNOLD, Christine; PENNINGS, Paul. Party Positions on the European Constitution During the 2004 European Parliament Elections. *Journal of Contemporary European Research*, v. 5, n. 1, p. 24–42, 2009.

BLONDEL, Jean; SINNOTT, Richard; SVENSSON, Palle. Representation and voter participation. *European Journal of Political Research*, v. 32, n. 2, p. 243–272, 1997. Available at: http://dx.doi.org/10.1023/A:1006832419724. Accessed in: January 3rd, 2022.

BLONDEL, Jean; SINNOTT, Richard; SVENSSON, Palle. *People and Parliament in the European Union: Participation, Democracy, and Legitimacy*. Oxford: University of Oxford Press, 1998.

BLUMLER, Jay G. FOX, Anthony D. *The European Voter: Popular Responses to the First European Community Elections*. London: Policy Studies Institute, 1982.

BRUG, Wouter van der; EIJK, Cees van der; FRANKLIN, Mark N. EU support and party choice. In: BRUG, Wouter van der; EIJK, Cees van der (Eds.). *European Elections & Domestic Politics: Lessons from the Past and Scenarios for the Future*. Notre Dame: University of Notre Dame Press, p. 168–189, 2007.

BRUG, Wouter van der; EIJK, Cees van der; SCHMITT, Hermann; FRANKLIN, Mark N. *et al*. European Elections, Domestic Politics and European Integration. In: BRUG, Wouter van der; EIJK, Cees van der (Eds.). *European Elections & Domestic Politics: Lessons from the Past and Scenarios for the Future*. Notre Dame: University of Notre Dame Press, p. 226-262, 2007.

BRUG, Wouter van der; FENNEMA, Meindert; TILLIE, Jean. Anti-immigrant parties in Europe: Ideological or protest vote? *European Journal of Political Research*, v. 37, n. 1, p. 77-102, 2000. Available at: https://link.springer.com/article/10.1023/A:1007013503658. Accessed in: January 3rd, 2022.

CAMPBELL, Angus. Surge and decline: a study of electoral change. *Public Opinion Quarterly*, v. 24, n. 3, p. 397–418th, 1960. Available at: https://www.jstor.org/stable/2746724. Accessed in: December 3rd, 2021.

CAYROL, Roland; YSMAL, Colette. France: the midwife comes to call. In: EIJK, Cees van der; FRANKLIN, Mark N. (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: Michigan University Press, p. 111-133, 1996.

CARRERAS, Miguel. Presidential Institutions and Electoral Participation in Concurrent Elections in Latin America. *Political Studies*, v. 66, n. 3, p. 541-559, 2017. Available at: https://journals.sagepub.com/doi/pdf/10.1177/0032321717723502. Accessed in: Dec 3rd, 2021.

CARRUBBA, C.; TIMPONE, R. J. Explaining vote switching across first- and second-order elections: Evidence from Europe. *Comparative Political Studies*, v. 38, n. 3, p. 260–281, 2005. Available at: https://journals.sagepub.com/doi/10.1177/0010414004272693. Accessed in: Dec 3rd, 2021.

CASTILLO, Pilar del. Spain: a dress rehearsal. In: EIJK, Cees van der; FRANKLIN, Mark N. (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*, Ann Arbor: Michigan University Press, p. 254-273, 1996.

CLARK, Nick; HUSLEY, John. The salience of EU issues: Explaining political behavior in European Parliament elections. In: 10th Biennial Meeting of the European Union Studies Association, Montreal, Canada, May 17-19, 2007. *Working paper…* Montreal: *[s.n.]*, 2007, p. 1-23. Available at: http://aei.pitt.edu/7786/1/clark-n-02h.pdf. Accessed in: January, 3rd, 2022.

CLARK, Nick; ROHRSCHNEIDER, Rohr. Second-order elections versus first-order thinking: How voters perceive the representation process in a multi-layered system of governance. Journal of European Integration, v. 31, n. 5, p. 645-664, 2009. Available at: https://doi.org/10.1080/07036330903145906. Accessed in: January 3rd, 2022.

COOPER, Harris. *Research synthesis and meta-analysis*: a step-by-step approach (applied social research methods). 5. ed. Thousand Oaks: Sage Publications [Kindle format], 2016.

COSTA, O; BRACK, N. The Role of the European Parliament in Europe's Integration and Parliamentarization Process. In: COSTA, O; STAVRIDIS, S; DRI, C. *The parliamentary dimension(s) of regionalization and globalization*. London: Palgrave, p. 45-69, 2013. Available at:https://www.researchgate.net/publication/304817687_The_Role_of_the_European_Parliament_in_Europe's_Integration_and_Parliamentarization_Process. Accessed in: Dec 03rd, 2021.

CUTLER, Fred. One voter, two first-order elections? *Electoral Studies*, v. 27, n. 3, p. 492-504, 2008. Available at: https://doi.org/10.1016/j.electstud.2008.01.002. Accessed in: Dec 3rd, 2021.

DACOMBE, Rod. Systematic Reviews in Political Science: What Can the Approach Contribute to Political Research? *Political Studies Review*, v. 16, n. 2, p. 148-157, 2018. Available at: https://journals.sagepub.com/doi/pdf/10.1177/1478929916680641. Accessed in: Dec 3rd, 2021.

DELGADO, Danielton Jorge Jesus. Eleições de Segunda Ordem em Cabo Verde, 1991-2016. Master Thesis (Master of Science) – Political Science and Public Policy Department, Lisbon University Institute. Lisbon, p. 51, 2017. Available at: https://repositorio.iscte-iul.pt/bitstream/10071/15074/4/master_danielton_jesus_delgado.pdf. Accessed in: Dec 3rd, 2021.

DENYER, D.; TRANFIELD, D. Producing a systematic review. In: BUCHANAN, D. A.; BRYMAN, A. (Eds.). *The Sage handbook of organizational research methods*. Thousand Oaks: Sage Publications Ltd., p. 671–689, 2009.

DINKEL, Reiner Hans. Der Zusammenhang zwischen Bundes- und Landtagswahlergebnissen, in Wahlsoziologie heute: Analysen aus Anlass der Bundestagswahl 1976 [The connection between federal and state election results in electoral sociology today: Analyses on the occasion of the 1976 federal election]. *Politische Vierteljahresschrift*, v. 18th, p. 348-359, 1977.

DWAN, K.; GAMBLE, C.; WILLIAMSON, P. R.; KIRKHAM, J.J. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias — An Updated Review. *PLoS ONE*, v. 8, n. 7, 2013. Available at: https://doi.org/10.1371/journal.pone.0066844. Accessed in: Jan. 3rd, 2022.

EHIN, Piret; TALVING, Liisa. Still second-order? European elections in the era of populism, extremism, and Euroscepticism. *Politics*, v. 41, n. 4, p. 467-485, 2021. Available at: https://doi.org/10.1177/0263395720986026. Accessed in: Dec 3rd, 2021.

EIJK, Cees van der; BRUG, Wouter van der (eds.). *Voting in European Parliament elections: Lessons from the past and scenarios for the future*. Notre Dame: University of Notre Dame Press, 2007.

EIJK, Cees van der; FRANKLIN, Mark N. European Community politics and electoral representation: evidence from the 1989 European elections study. *European Journal of Political Research*, v. 19, n. 1, p. 105-128, 1991.

EIJK, Cees van der; FRANKLIN, Mark N. *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: University of Michigan Press, 1996.

EIJK, Cees van der; FRANKLIN, Mark N.; MARSH, Michael. What Voters Teach Us About European Elections: What Europe-Wide Elections Teach Us About Voters, *Electoral Studies*, v. 15, n. 2, p. 149-166, 1996. Available at: https://www.sciencedirect.com/science/article/abs/pii/0261379496000091?via%3Dihub. Accessed in: Dec 03rd, 2021.

EIJK, Cees van der; OPPENHUIS, Erik V. European parties' performance in electoral competition. *European Journal of Political Research*, v. 19, n. 1, p. 55-80, 1991. Available at: https://ejpr.onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6765.1991.tb01177.x. Accessed in: January 3rd, 2022.

EIJK, Cees van der; OPPENHUIS, Erik V. Turnout and second-order effects in the European elections of June 1989 - evidence from The Netherlands. *Acta Politica*, v. 25, n. 1, p. 67-94, 1990.

EIJK, Cees van der; SCHIMITT, Hermann. The role of the Eurobarometer in the study of European elections and the development of comparative electoral research. In: REIF, Karlheinz and INGLEHART, Ronald (Eds.). *Eurobarometer: The Dynamics of European Public Opinion*. London: Palgrave Macmillan, p. 257-274, 1991.

EIJK, Cees Van Der. The Netherlands. *Electoral Studies*, v. 8, n. 3, p. 305-312, 1989. Available at: https://doi.org/10.1016/0261-3794(89)90012-7. Accessed in: January 3rd, 2022.

EISINGA, Rob; FRANSES, Philip Hans; DIJK, Dick van. Timing of Vote Decision in First and Second Order Dutch Elections 1978–1995: Evidence from Artificial Neural Networks. *Political Analysis*, v. 7, p. 117-142, 1998. Available at: https://doi.org/10.1093/pan/7.1.117. Accessed in: January 3rd, 2022.

ELGIE, Robert; FAUVELLE-AYMAR, Christine. Turnout Under Semi-presidentialism: First- and Second-Order Elections to National-Level Institutions. *Comparative Political*

Studies, v. 45, n. 12, p. 1598-1623, 2012. Available at: https://journals.sagepub.com/doi/pdf/10.1177/0010414012463903. Accessed in: Dec 3rd, 2021.

ERIKSON, Robert. The puzzle of midterm loss. *Journal of Politics*, v. 50, n. 4, p. 1011- 1029, 1988.

ERKAN, Ozgur. Spain's Referendum on the European Constitutional Treaty: A Quantitative Analysis Within the Conceptual Framework of First and Second Order Elections. *LSE Europe in Question*, n. 25, p. 1-41, 2010. Available at: http://eprints.lse.ac.uk/53298/1/__Libfile_repository_Content_European%20Institute_LEQS%20Discussion%20Papera_LEQSPaper25.pdf. Accessed in: Dec 3rd, 2021.

ERVIK, B. Second-Order Arguments for Second-Order Elections? Measuring "Elections Stakes" at a Multilevel Context: The Case of Norway. *Journal of Elections, Public Opinion and Parties*, v. 22, n.1, p. 27-50, 2012. Available at: https://www.tandfonline.com/doi/abs/10.1080/17457289.2011.642241. Accessed in: Dec 3rd, 2021.

FAUVELLE-AYMAR, Christine; LEWIS-BECK, Michael S. Second-Order Elections and Economic Voting: The French Regional Example. *Rivista Italiana di Scienza Politica*, v. 41, n. 3, 2011.

FAUVELLE-AYMAR, Christine; STEGMAIER, Mary. Economic and political effects on European Parliamentary electoral turnout in post-communist Europe. *Electoral Studies*, v. 27, n. 4, p. 661-672, 2008. Available at: https://doi.org/10.1016/j.electstud.2008.05.008. Accessed in: January 3rd, 2022.

FERRARA, Federico; WEISHAUPT, J. Timo. Get your act together: Party performance in European Parliament elections. *European Union Politics*, v. 5, n. 3, p. 283–306, 2004. Available at: http://dx.doi.org/10.1177/1465116504045154. Accessed in: January 3rd, 2022.

FIGUEIREDO FILHO, D. B. *et al*. O que é, para que serve e como se faz uma meta-análise? *Teoria e Política*, v. 23, n. 2, p. 205-228, 2014. Available at: http://doi.editoracubo.com.br/10.4322/tp.2014.018. Accessed in: Dec 3rd, 2021.

FORTES, Braulio Gómez; MAGALHÃES, Pedro. As eleições presidenciais em sistemas semi-presidenciais: participação eleitoral e punição dos governos. *Análise Social*, v. 11, n. 177, p. 891-922, 2005. Available at: http://analisesocial.ics.ul.pt/documentos/1218711910H1xMA5lp6Sb55FV1.pdf. Accessed in: Dec 3rd, 2021.

FRANKIN, Mark N. Getting out the vote: social structure and the mobilization of partisanship in the 1989 European elections. *European Journal of Political Research*, v. 19, n. 1, p. 129-148, 1991.

FRANKLIN, Mark N.; MACKIE, Thomas T.; VALEN, Henry *et al*. *Electoral Change: Responses to Evolving Social and Attitudinal Structures in Western Nations*. Cambridge: Cambridge University Press, 1992.

FRANKLIN, Mark N.; EIJK, Cees van der; OPPENHUIS, Erik. Turnout decline and the motivational basis of British participation in the European elections of 1994. In: RALLINGS, Colin *et al* (Eds.). *British Elections and Parties Yearbook 1995*. London: Frank Cass, 1996.

FREIRE, A. Second-Order Elections and Electoral Cycles in Democratic Portugal. *South European Society and Politics*, v. 9, n. 3, p. 54-79, 2004. Available at: https://www.tandfonline.com/doi/abs/10.1080/13608740420002718 61?journalCode=fses20. Accessed in: Dec 3rd, 2021.

FREIRE, André; TEPEROGLOU, Eftichia. European Elections and National Politics: Lessons from the "New" Southern European Democracies. *Journal of Elections, Public Opinion and Parties*, v. 17, n. 1, p. 101-122, 2007. Available at: https://doi.org/10.1080/13689880601148709. Accessed in: Januray 3rd, 2022.

GIEBLER, H.; WAGNER, A. Contrasting First- and Second-Order Electoral Behaviour: Determinants of Individual Party Choice in European and German Federal Elections. *German Politics*, v. 24, n. 1, p. 46-66, 2015. Available at: https://www.tandfonline.com/doi/abs/10.1080/09644008.2014.949684. Accessed in: Dec 3rd, 2021.

GLASS, G. V. Primary, secondary and meta-analysis of research. *Educational Researcher*, v. 5, p. 3-8, 1976.

GLENCROSS, Andrew; TRECHSEL, Alexander. First or Second Order Referendums? Understanding the Votes on the EU Constitutional Treaty in Four EU Member States. *West European Politics*, v. 34, n. 4, p. 755-772, 2011. Available at: https://www.tandfonline.com/doi/pdf/10.1080/01402382.2011.572390. Accessed in: Dec, 3rd, 2011.

HEATH, Anthony; MCLEAN, Iain; TAYLOR, Bridget; CURTICE, John. Between first and second order: a comparison of voting behaviour in European and local elections in Britain. *European Journal of Political Research*, v. 35, n. 3, p. 389-414, 1999. Available at: https://ejpr.onlinelibrary.wiley.com/doi/abs/10.1111/1475-6765.00454. Accessed in: Dec. 3rd, 2021.

HIX, Simon; MARSH, Michael. Punishment or Protest? Understanding European Parliament Elections. *The Journal of Politics*, v. 69, n. 2, p. 495-510, 2007. Available at: https://www.journals.uchicago.edu/doi/10.1111/j.1468-2508.2007.00546.x. Accessed in: January 3rd, 2022.

HIX, Simon.; MARSH, Michael. Second-order effects plus pan-European political swings: An analysis of European Parliament elections across time. *Electoral Studies*, v. 30, n. 1, p. 4-15, 2011.

HOBOLT, Sara Binzer, SPOON, Jae-Jae; TILLEY, James. A Vote Against Europe? Explaining Defection at the 1999 and 2004 European Parliament Elections. *British Journal of Political Science*, v. 39, n. 1, p. 93-115, 2009. Available at: https://doi.org/10.1017/S0007123408000422. Accessed in: January 3rd, 2022.

HOBOLT, Sara Binzer; WITTROCK, Jill. The second-order election model revisited: An experimental test of vote choices in European Parliament elections. *Electoral Studies*, v. 30, n. 1, p. 29-40, 2011. Available at: https://doi.org/10.1016/j.electstud.2010.09.020. Accessed in: Dec 3rd, 2021.

IDROVO, A. J. Three Criteria for Ecological Fallacy. *Environmental Health Perspectives*, v. 119, n. 8, p. 332, 2011. Available at: https://ehp.niehs.nih.gov/doi/full/10.1289/ehp.1103768. Accessed in: Dec 3rd, 2021.

IMAI, Ryosuke; HINO, Airo. Second-ordering the Upper House Election in Japan. *Japanese Journal of Electoral Studies*, v. 27, n. 2, p. 5-19, 2012. Available at: https://www.jstage.jst.go.jp/article/jaes/27/2/27_5/_article/-char/en. Accessed in: Dec 3rd, 2021.

IMBEAU, L. M.; PÉTRY, F. & LAMARI, M. Left-right party ideology and government policies: A meta-analysis. *European Journal of Political Research*, v, 40, n. 1, 2001.

IRWIN, Galen. Second-order or third-rate?: Issues in the campaign for the elections for the European Parliament 1994. *Electoral Studies*, v. 14, n. 2, p. 183-199, 1995. Available at: https://www.sciencedirect.com/science/article/abs/pii/0261379495968437?via%3Dihub. Accessed in: January 3rd, 2022.

JADOT, Anne. (Ne pas) être un électeur européen: une analyse multiniveaux des déterminants individuels et contextuels de l'abstention en 1999. *Revue Internationale de Politique Comparée*, v. 9, n. 1, 2002. Available at: https://www.cairn.info/revue-internationale-de-politique-comparee-2002-1-page-31.htm?contenu=article. Accessed in: Dec 23rd, 2021.

JESUIT, David. The Regional Dynamics of European Electoral Politics Participation in National and European Contests in the 1990s. *European Union Politics*, v. 4, n. 2, p. 139–164, 2003. Available at: https://journals.sagepub.com/doi/10.1177/1465116503004002001. Accessed in: January 3rd, 2022.

KING, Gary. *A Solution to the Ecological Inference Problem. Reconstructing Individual Behaviour from Aggregate Data*. Princeton: Princeton University Press, 1997.

KING, Gary. Replication, Replication. *Political Science and Politics*, v. 28, p. 444-452, 1995. Available at: https://j.mp/2oSOXJL. Accessed in: Dec 3rd, 2021.

KING, G.; ROSEN, O.; TANNER, M. A. *Ecological Inference: New Methodological Strategies*. Nova York: Cambridge University Press, 2004.

KOEPKE, Jason R.; RINGE, Nils. The Second-order Election Model in an Enlarged Europe. *European Union Politics*, v. 7, n. 3, p. 321-346, 2006. Available at: https://doi.org/10.1177/1465116506066259. Accessed in: January 3rd, 2022.

KOUSSER, Thad. Retrospective voting and strategic behavior in European Parliament elections. *Electoral Studies*, v. 23, n. 1, p. 1-21, 2004. Available at: https://doi.org/10.1016/S0261-3794(02)00030-6. Accessed in: January 3rd, 2022.

KRITZINGER, Sylvia; PLESCIA, Carolina; RAUBE, Kolja; WILHELM, James; WOUTERS, Jan (eds.). *Assessing the 2019 European Parliament Elections*. London: Routledge, 2020.

LAŞAN, N. Testing the Second-Order Elections Model on the 2019 European Elections in Romania. Studia Universitatis Babes-Bolyai. Studia Europaea, v. 65, n. 2, p. 359-374, 2020. Available at: http://www.studia.ubbcluj.ro/download/pdf/1344.pdf. Accessed in: Dec 3rd, 2021.

LINEK, L.; LYONS, P. What can ecological inference tell us about the Second-Order-Election-Thesis in the Czech Republic and Slovakia? In: MARSH, M.; MIKHAYLOV, S.; SCHMITT, H. (eds.). *European Elections after Eastern Enlargement: Preliminary Results from the European Election Study 2004*. Mannheim: CONNEX, p. 327-369, 2007. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.378.5922&rep=rep1&type=pdf#page=335. Accessed in: Dec 3rd, 2021.

LIU, B. EI Extended Model and the Fear of Ecological Fallacy. *Sociological Methods & Research*, v. 36, n. 1, p. 3-25, 2007. Available at: https://doi.org/10.1177/0049124106295797. Accessed in: Dec 3rd, 2021.

LODGE, Juliet. *The 2004 Elections to the European Parliament*. London: Palgrave Macmillan, 2005.

MAGNI-BERTON, Raùl. Pourquoi les partis gouvernementaux perdent-ils les élections intermédiaires ? Enquête Eurobaromètre 2004 et élections européennes. *Revue Française de Science Politique*, v. 58, n. 4, p. 643-656, 2008. Available at: https://www.cairn.info/revue-francaise-de-science-politique-2008-4-page-643.htm. Accessed in: January 3rd, 2022.

MAIER, Michaela. *Campaigning in Europe - Campaigning for Europe, Political parties, campaigns, mass media and the European Parliament elections in 2004*. London: LIT Publisher, 2006.

MANOW, Philip and DÖRING, Holger. Electoral and Mechanical Causes of Divided Government in the European Union. *Comparative Political Studies*, v. 41, p. 1349–1370, 2008. Available at: https://journals.sagepub.com/doi/10.1177/0010414007304674. Accessed in: January 3rd, 2022.

MANOW, Philip. National Vote Intention and European Voting Behavior, 1979-2004: Second Order Election Effects, Election Timing, Government Approval and the Europeanization of European Elections. In: MPIfG Discussion Paper, No. 05/11, Max Planck Institute for the Study of Societies, Cologne, 2005, *Discussion paper…* Cologne: MPIfG, 2005, p. 1-31. Available at: https://www.econstor.eu/bitstream/10419/19925/1/dp05-11.pdf. Accessed in: January 3rd, 2022.

MARIANO, Karina L. Pasquariello; LUCIANO, Bruno Theodoro; BRESSAN, Regiane Nitsch. *Integração com Democracia: os Desafios para os Parlamentos Regionais*. Rio de Janeiro: Konrad Adenauer Stiftung, 2016.

MARIEN, Sofie; DASSONNEVILLE, Ruth; HOOGHE, Marc. How Second Order Are Local Elections? Voting Motives and Party Preferences in Belgian Municipal Elections. *Local Government Studies*, v. 41, n. 6, p. 898-916, 2015. Available at: https://doi.org/10.1080/03003930.2015.1048230. Accessed in: Dec, 3rd, 2021.

MARSH, Michael. European Parliament elections and losses by governing parties. In: BRUG, Wouter van der; EIJK, Cees van der. *European Elections & Domestic Politics: Lessons from the Past and Scenarios for the Future*. Notre Dame: University of Notre Dame Press, p. 51–73, 2007.

MARSH, Michael; FRANKLIN, Mark N. Understanding European elections, 1979-1994. In: EIJK, Cees van der; FRANLIN, Mark N. *et al* (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: Michigan University Press, pp. 23-45, 1996.

MARSH, Michael; MIKHAYLOV, Slava. European Parliament elections and EU governance. *Living Reviews on European Governance*, v. 5, 2010. Available at: http://europeangovernance-livingreviews.org/Articles/lreg-2010-4. Accessed in: January 3rd, 2022.

MARSH, Michael; MIKHAYLOV, Slava; SCHMITT, Hermann (eds.). *European Elections after Eastern Enlargement: Preliminary Results from the European Election Study 2004*. Mannheim: CONNEX, 2007. Available at: http://www.mzes.uni-mannheim.de/projekte/typo3/site/fileadmin/BookSeries/Volume_One/European_Elections_gesamt.pdf#page=401. Accessed in: Dec 7th, 2021.

MARSH, Michael. Testing the Second-Order Election Model after Four European Elections. *British Journal of Political Science*, v. 28, p. 591–607, 1998. Available at: https://doi.org/10.1017/S000712349800026X. Accessed in: Dec 3rd, 2021.

MARSH, Michael. Theories of less important elections: explanations of electoral behaviour in European parliament elections. *Revue de la Maison Française d'Oxford*, v. 1, n. 1, 2003.

MARSH, Michael. The Results of the 2004 European Parliament Elections and the Second-order Model. In: NIEDERMAYER, Oskar; SCHMITT, Hermann (Eds.). *Europawahl 2004*. Wiesbaden: Verlag für Sozialwissenschaften, p. 142–158, 2005.

MARSH, Michael. Vote Switching in European Parliament Elections: Evidence from June 2004. *Journal of European Integration*, v. 31, n. 5, p. 627-644, 2009. Available at: https://doi.org/10.1080/07036330903145898. Accessed in: January 3rd, 2022.

MATHER, Janet. The United Kingdom. In: LODGE, Juliet (Ed.). *The 1999 Elections to the European Parliament*. London: Palgrave Macmillan, p. 214–231, 2001.

MATILLA, Mikko. Why bother? Determinants of turnout in the European elections. *Electoral Studies*, v. 22, n. 3, p. 449-468, 2003. Available at: https://doi.org/10.1016/S0261-3794(01)00054-3. Accessed in: January 3rd, 2022.

MCLEAN, Iain; HEATH, Anthony; TAYLOR, Bridget. Were the 1994 Euro- and local elections in Britain really second-order? Evidence from the British election panel study.

British Elections and Parties Yearbook, v. 6, n. 1, p. 1-20, 1996. Available at: https://www.tandfonline.com/doi/abs/10.1080/13689889608412969?journalCode=fbep18. Accessed in: January 3rd, 2022.

MITCHELL, Glen; WLEZIEN, Christopher. The impact of legal constraints on voter registration, turnout, and the composition of the American electorate. *Political Behavior*, v. 17, n. 2, p. 179-202, 1995.

MÜLLER, Jochen. German regional elections: Patterns of second-order voting. *Regional & Federal Studies*, v. 28, n. 3, p.301-324, 2018. Available at: https://doi.org/10.1080/13597566.2017.1417853. Accessed in: Dec 3rd, 2021.

NACHMIAS, David; ROSENTHAL, Maoz; ZUBIDA, Hani. Inverted First- and Second-Order Elections: The Israeli Case. *Israel Studies Review*, v. 27, n. 1, p. 31-54, 2012.

NEIVA, Pedro. Revisitando o calcanhar de Aquiles metodológico das ciências sociais no Brasil. *Sociologia, Problemas e Práticas*, n. 79, 2015. Available at: http://journals.openedition.org/spp/2232. Accessed in: Dec 3rd, 2021.

NIEDERMAYER, Oskar. Turnout in the European elections. *Electoral Studies*, v. 9, n. 1, p. 45-50, 1990. Available at: https://doi.org/10.1016/0261-3794(90)90041-6. Accessed in: January 3rd, 2022.

NIELSEN, Hans Jorgen. Denmark. In: LODGE, Juliet (Ed.). *The 1999 Elections to the European Parliament*. London: Palgrave Macmillan, p. 60–72, 2001.

NIELSEN, Julie Hassing; FRANKLIN, Mark N. (eds.). *The Eurosceptic 2014 European Parliament Elections: Second Order or Second Rate?* London: Palgrave Macmillan, 2016.

NORRIS, Pippa. Second-order elections revisited. *European Journal of Political Research*, v. 31, p. 109-114, 1997. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6765.1997.tb00768.x. Accessed in: Dec. 3rd, 2021.

OPPENHUIS, Erik; EIJK, Cees van der; FRANKLIN, Mark. Consulting the oracle: consequences of treating European elections as "markers" of domestic political developments. In: Dutch Political Science Association Annual Meeting, Amsterdam, Netherlands, 1994. *Working paper…* Amsterdam: *[s.n.]*, 1944.

OPPENHUIS, Erik; EIJK, Cees van der; FRANKLIN, Mark N. The party context: outcomes. In: EIJK, Cees van der; FRANKLIN, Mark N. *et al* (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: Michigan University press, p. 303-329, 1996.

OPPENHUIS, Erik. The Netherlands: small party evolutions. In: EIJK, Cees van der; FRANKLIN, Mark N. (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: University of Michigan Press, p. 210-228, 1996.

OPPENHUIS, Erik. *Voting behavior in Europe: a comparative analysis of electoral participation and party choice*. Amsterdam: Het Spinhuis, 1995.

ORFORD, Scott; RALLINGS, Colin, THARSHER, Michael; BORISYUK, Galina. Electoral salience and the costs of voting at national, sub-national and supra-national elections in the UK: a case study of Brent, UK. *Transactions of the Institute of British Geographers*, v. 34, n. 2, p. 195–214, 2009. Available at: http://dx.doi.org/10.1111/j.1475-5661.2008.00335.x. Accessed in: January 3rd, 2022.

PATTERSON, Samuel; CALDEIRA, Gregory. Getting out the vote: participation in gubernatorial elections. *American Politics Science Review*, v. 78, n. 1, p. 675-689, 1993.

PEARCE, Jack; DERRICK, Ben. Preliminary testing: The devil of statistics?. *Reinvention: An International Journal of Undergraduate Research*, v. 12, n. 2, 2019. Available at: https://doi.org/10.31273/reinvention.v12i2.339. Accessed in: Jan. 3rd, 2022.

PERRINEAU, Pascal; GRUNBER, Gérard; YSMAL, Colette (eds.). *Europe at the Polls: The European Elections of 1999*. London: Palgrave Macmillan, 2002.

PETTICREW, Mark; ROBERTS, Helen. *Systematic reviews in the social sciences: a practical guide*. Oxford: Blackwell Publihsing, 2006.

RABIER, Jacques-René. Les paradoxes d'un scrutin: Les citoyens de dix pays élisent le parlement Européen. *Journal of European Integration*, v. 8, n. 1, p. 5-32, 1984. Available at: https://www.tandfonline.com/doi/abs/10.1080/07036338408428859. Accessed in: January 3rd, 2022.

RALLINGS, Colin; THRASHER, Michael. Not All 'Second-Order' Contests are the Same: Turnout and Party Choice at the Concurrent 2004 Local and European Parliament Elections in England. *British Journal of Politics and International Relations*, v. 7, p. 584-597, 2005. Available at: https://journals.sagepub.com/doi/pdf/10.1111/j.1467-856x.2005.00207.x. Accessed in: Dec 3rd 2021.

REIF, Karlheinz. *European elections 1979/81 and 1984: Conclusions and perspectives from empirical research (European electoral studies)*. Berlin: Quorum Books, 1984a.

REIF, Karlheinz. European elections as member state second-order elections revisited. *European Journal of Political Research*, v. 31, p. 115-124, 1997. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6765.1997.tb00768.x. Accessed in: Dec 3rd, 2021.

REIF, Karlheinz. National electoral cycles and European elections 1979 and 1984. *Electoral Studies*, v. 3, n. 3, p. 244–255, 1984b.

REIF, Karlheinz.; SCHMITT, Hermann. Nine Second-Order National Elections - A Conceptual Framework for the Analysis of European Election Results. *European Journal of Political Research*, v. 8, n. 1, p. 3–44, 1980. Available at: http://doi:10.1111/j.1475-6765.1980.tb00737.x. Accessed in: Dec 3rd, 2021.

REIF, Karlheinz. Ten second-order elections. In: REIF, Karlheinz (Ed.). *Ten European Elections*. Aldershot: Gower, p. 1-36, 1985.

ROBINSON, W. S. Ecological correlations and the behaviour of individuals. *American Sociological Review*, v. 15, n. 3, p. 351-357, 1950. Available at https://www.jstor.org/stable/2087176?origin=crossref&seq=1. Accessed in: Dec 3rd, 2021.

ROCHA, Enivaldo C. da; PARANHOS, Ranulfo; FIGUEIREDO FILHO, Dalson B.; CARMO, Erinaldo F. de. A Importância da Replicabilidade na Ciência Política: O Caso do SIGOBR. *Política Hoje*, v. 22, n. 2, p. 213-229, 2013. Available at: https://periodicos.ufpe.br/revistas/politicahoje/article/view/3770. Accessed in: Dec 3rd, 2021.

ROSCOE, D. D.; JENKINS, S. A Meta-Analysis of Campaign Contributions' Impact on Roll Call Voting. *Social Science Quarterly*, v. 86, n. 1, 2005.

SAMPAIO, Daniel; FILHO, Dalson B. F.. Como o dinheiro influencia as eleições municipais no Brasil: uma revisão sistemática. *Revista Brasileira de Informações Bibliográficas*, v. 1, n. 88, p. 1-25, 2019. Available at: https://www.researchgate.net/publication/331261455_Como_o_dinheiro_influencia_as_eleicoes_municipais_no_Brasil_uma_revisao_sistematica_1. Accessed in: Dec 3rd, 2021.

SANZ, Alberto. Split-Ticket Voting in Multi-Level Electoral Competition: European, National and Regional Concurrent Elections in Spain. EIJK, Cees van der; SCHMITT, Hermann (Eds.). *Multi-level electoral systems of the European Union vol. 4 of The CONNEX Report Series*. Mannheim: CONNEX, p. 101–136, 2008.

SCHÄFER, C. Indifferent and Eurosceptic: The motivations of EU-only abstainers in the 2019 European Parliament election. *Political Studies Association*, 2021. Available at: https://journals.sagepub.com/doi/pdf/10.1177/0263395720981359. Accessed in: Dec 3rd, 2021.

SCHMITT, Hermann; EIJK, Cees van der. There is Not Much Euro-hostile Non-voting in European Parliament Elections. In: ECPR Joint Sessions of Workshops, Turin, Italy, March 22-27, 2002. *Working paper…* Turin: *[s.n.]*, 2002, p. 1-28.

SCHMITT, Hermann; MANNHEIMER, Renato. About voting and non-voting in the European elections of June 1989. *European Journal of Political Research*, v. 19, n. 1, p. 31-54, 1991. Available at: https://ejpr.onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6765.1991.tb01176.x. Accessed in: January 3rd, 2022.

SCHMITT, Hermann. Party attachment and party choice in the European elections of June 1989: A cross-national comparative analysis of the post-electoral surveys of the European Voters Study 1989. *International Journal of Public Opinion Research*, v. 2, n. 2, p. 169-184, 1990. Available at: https://doi.org/10.1093/ijpor/2.2.169. Accessed in: January 3rd, 2022.

SCHMITT, Hermann. National party systems and the politics of the European Union: first results from the 1994 European Election Study. In: XVIth World Congress of the International Political Science Association, Berlin, August 1994. *Working paper…* Berlin: *[s.n..]*, 1994.

SCHMITT, Hermann; SANZ, Alberto; BRAUN, Daniela. Motive individuellen Wahlverhaltens in Nebenwahlen: Eine theoretische Rekonstruktion und empirische Überprüfung [Motives of individual voting behaviour in second-order elections: A theoretical

reconstruction and empirical test]. IN: FALTER, J. W.; GABRIEL, O.; WESSELS, B. (Eds.). Wahlen und Wähler. Analysen aus Anlass der Bundestagswahl 2005 [Elections and voters: Analyses at the occasion of the Federal Election of 2005]. Wiesbaden: Verlag für Sozialwissenschaften, p. 585–605, 2009.

SCHMITT, Hermann; SANZ, Alberto; BRAUN, Daniela. The Micro-Foundations of Second-order Election Theory: Theoretical Reconstructions and Empirical Tests Based Upon the European Election Study 2004. In: Citizen Politics Workshop, University of Manchester, December 9-11, 2008. *Article…* Manchester: *[s.n.]*, 2008, p. 1-26. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.566.1686. Accessed in: Dec 3rd, 2021.

SCHMITT, Hermann; TEPEROGLOU, E. The 2014 European Parliament elections in Southern Europe: Second-order or critical elections? *South European Society and Politics*, v. 20, n. 3, p. 287–309, 2015.

SCHMITT, Hermann; TEPEROGLOU, E. The study of less important elections. IN: ARZHEIMER, K. EVANS, J. LEWIS-BECK, M. (Eds.). *The Sage handbook of electoral behaviour*. London: Sage, p. 56–79, 2017.

SCHMITT, Hermann; SANZ, A.; BRAUN, D.; TEPEROGLOU, E. It All Happens at Once: Understanding Electoral Behaviour in Second-Order Elections. *Politics and Governance*, v. 8, n. 1, p. 6-18th, 2020. Available at: https://www.cogitatiopress.com/politicsandgovernance/article/view/2513/2513. Accessed in: Dec 3rd, 2021.

SCHMITT, Hermann. The European Parliament Elections of June 2004: Still Second-Order? *West European Politics*, v. 28, n. 3, p. 650–679, 2005. Available at: https://doi.org/10.1080/01402380500085962. Accessed in: December 3rd, 2021.

SELVIN, H. C. Durkheim's suicide and problems of empirical research. *American Journal of Sociology*, v. 63, n. 6, p. 607-619, 1958. Available at: https://www.journals.uchicago.edu/doi/10.1086/222356. Accessed in: Dec 3rd, 2021.

SIAROFF, Alan. Elections to the European parliament: Testing alternative models of what they indicate in the member nations. *Journal of European Integration*, v. 23, n. 3, p. 237-255, 2001. Available at: https://www.tandfonline.com/doi/abs/10.1080/07036330108429104. Accessed in: January 3rd, 2022.

SINNOTT, Richard; WHELAN, Brendan J. Turnout in Second Order Elections: The Case of EP Elections in Dublin 1984 and 1989. *The Economic and Social Review*, v. 23, n. 2, p. 147-166, 1992. Available at: http://www.tara.tcd.ie/bitstream/handle/2262/64459/23?sequence=1. Accessed in: January 3rd, 2002.

SKRINIS, S.; TEPEROGLOU, E. Studying and Comparing Second-Order Elections. Examples from Greece, Portugal and Spain. IN: VAN DER EIJK, C.; SCHMITT, H. *The Multilevel Electoral System of the EU*. Mannheim: Mannheim Centre for European Social Research, CONNEX Report Series Nr 04, p. 169-189, 2008. Available at: http://www.mzes.uni-

mannheim.de/projekte/typo3/site/fileadmin/BookSeries/Volume_Four/CONNEX%20Series%20Book%204.pdf#page=169. Accessed in: Dec 3rd, 2021.

SMITH, Julie. The 1994 European elections: Twelve into one won't go. *West European Politics*, v. 18th, n. 3, p. 199-217, 1995. Available at: https://doi.org/10.1080/01402389508425098. Accessed in: January 3rd, 2022.

SOARES, Gláucio Ary Dillon. O Calcanhar Metodológico da Ciência Política no Brasil. *Sociologia, Problemas e Práticas*, n. 48, p. 27-52, 2005. Available at: https://sociologiapp.iscte-iul.pt/pdfs/48/512.pdf. Accessed in: Dec 3rd, 2021.

SONG, F.; HOOPER, L.; LOKE, Y. Publication bias: what is it? How do we measure it? How do we avoid it?. *Open Access Journal of Clinical Trials*, v. 5, p. 71-81, 2013. Available at: https://doi.org/10.2147/OAJCT.S34419. Accessed in: Jan 3rd, 2022.

SONG, F.; PAREKH, S.; HOOPER, L.; LOKE, Y. K.; RYDER, J.; SUTTON, A. J.; HING, C.; KWOK, C. S.; PANG, C.; HARVEY, I. Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment*, v. 14, n. 8, 2010. Available at: https://doi.org/10.3310%2Fhta14080. Accessed in: Jan. 3rd, 2022.

STEFANOVA, Boyka. The 2007 European elections in Bulgaria and Romania. *Electoral Studies*, v. 27, n. 3, p. 566-571, 2008. Available at: https://doi.org/10.1016/j.electstud.2008.02.009. Accessed in: Januray 3rd, 2022.

STIMSON, J. A. Public support for American presidents: a cyclical model. *Public Opinion Quarterly*, v. 40, n. 1, p. 1–21, 1976.

STUDLAR, Donley; FLICKINGER, Richard S.; BENNETT, Stephen. Turnout in European Parliament elections: Towards a European-centred model. *Journal of Elections, Public Opinion and Parties*, v. 13, n. 1, p. 195–209, 2003. Available at: https://www.tandfonline.com/doi/abs/10.1080/13689880308413094. Accessed in: January 3rd, 2022.

SUBRAMANIAN, S. V.; JONES, K.; KADDOUR, A.; KRIEGER, N. Revisiting Robinson: The perils of individualistic and ecologic fallacy. *International Journal of Epidemiology*, v. 38, n. 2, p. 342–360, 2009. Available at: https://doi.org/10.1093/ije/dyn359. Accessed in: Dec 3rd, 2021.

THORLAKSON, Lori. Explaining the Federal-Provincial Turnout Gap in the Canadian Provinces. *Canadian Political Science Review*, v. 9, n. 1, p. 164-176, 2015.

TILLEY, James; GARRY, John; TESSA, Bold. Perceptions and reality: Economic voting at the 2004 European Parliament elections. *European Journal of Political Research*, v. 47, p. 665–686, 2008. Available at: http://dx.doi.org/10.1111/j.1475-6765.2008.00780.x. Accessed in: Januray 3rd, 2022.

TILLEY, James; WLEZIEN, Christopher. Does Political Information Matter? An Experimental Test Relating to Party Positions on Europe. *Political Studies*, v. 56, p. 192–214, 2008. Available at: http://dx.doi.org/10.1111/j.1467-9248.2007.00698.x. Accessed in: January 3rd, 2022.

TOKA, Gabor. Information Effects on Vote Choices in European Elections. In: MARSH, Michael; MIKHAYLOV, Slava; SCHMITT, Hermann (Eds.). *European Elections after Eastern Enlargement: Preliminary Results from the European Election Study 2004 vol. 1 of The CONNEX Report Series*. Mannheim: CONNEX, p. 141–180, 2007.

TUFTE, Edward R. Determinants of outcomes of midterm congressional elections. *American Political Science Review*, v. 69, n. 3, p. 812–826, 1975. Available at: https://doi.org/10.2307/1958391. Accessed in: Dec 3rd, 2021.

VIOLA, Donatella M (eds.). *Routledge Handbook of European Elections*. London: Routledge, 2016.

VREESE, Claes H. de; BANDUCCI, Susan A.; SEMETKO, Holli A.; BOOMGAARDEN, Hajo G. The News Coverage of the 2004 European Parliamentary Election Campaign in 25 Countries. *European Union Politics*, v. 7, n. 4, p. 477–504, 2006. Available at: https://doi.org/10.1177/1465116506069440. Accessed in: January 3rd, 2022.

WEBER, Till. Campaign Effects and Second-Order Cycles: A Top-Down Approach to European Parliament Elections. *European Union Politics*, v. 8, n. 4, p. 509-536, 2007. Available at: https://doi.org/10.1177/1465116507082812. Accessed in: January 3rd, 2022.

WEBER, Till. Exit, Voice, and Cyclicality: A Micro-Logic of Voting Behaviour in European Parliament Elections. In: 2009 EUSA conference, Los Angeles, USA. *Working paper…* Los Angeles: *[s.n.]*, 2009, p. 1-37. Available at: http://aei.pitt.edu/33160/1/weber._till.pdf. Accessed in: January, 3rd, 2022.

WORRE, Toben, Denmark: Second-order Containment. In: EIJK, Cees van der; FRANKLIN, Mark N. (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: University of Michigan Press, p. 97–114, 1996.

YSMAL, Colette; CAYROL, Roland. France: The Midwife Comes to Call. In: EIJK, Cees van der; FRANKLIN, Mark N. (Eds.). *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: University of Michigan Press, p. 115–136, 1996.

APPENDIX A - Table of the Formal Dimension Variables

VARIABLE	DESCRIPTION	MEASUREMENT
Title	Indicates the title of the work	Nominal
Source	Name of the journal, congress or institution where the work was presented/published	Nominal
Type	Classification of the work as an article published in a journal, working paper, monograph, master thesis or dissertation.	Nominal
Type code	Categorization of the "Type" variable	Categorical (1=Journal article, 2 = working paper, 3 = monograph, 4 = master thesis, 5 = dissertation)
Authors	Name of the author(s)	Nominal
Number of authors	Number of authors of the work	Discrete
First author	Name of the first author	Nominal
Second author	Name of the second author	Nominal
Third author	Name of the third author	Nominal
Fourth author	Name of the fourth author	Nominal
Fifth author	Name of the fifth author	Nominal
Year	Year which the work was published or presented	Nominal
Reference	Author(s) surname(s) and year of publication	Nominal
Female	Indicates the number of female authors in the work	Discrete
Male	Indicates the number of male authors in the work	Discrete
Number of pages	Number of pages of the work	Discrete
URL	link to the source of the work	Nominal
Age	Time from publication/presentation to 2021 (when data were collected)	Discrete (in years)
First author affiliation	First author's home university	Nominal
City of first author's university	City where the first author's home university is located	Nominal
Country of first author's university	Country where the first author's home university is located	Nominal
Second author affiliation	Second author's home university	Nominal
City of second author's university	City where the second author's home university is located	Nominal
Country of second author's university	Country where the second author's home university is located	Nominal
Third author affiliation	Third author's home university	Nominal
City of third author's university	City where the third author's home university is located	Nominal
Country of third author's	Country where the third	Nominal

university	author's home university is located	
Fourth author affiliation	Fourth author's home university	Nominal
City of fourth author's university	City where the fourth author's home university is located	Nominal
Country of foirth author's university	Country where the fourth author's home university is located	Nominal
Fifth author affiliation	Fifth author's home university	Nominal
City of fifth author's university	City where the fifth author's home university is located	Nominal
Country of fifth author's university	Country where the fifth author's home university is located	Nominal
Country of the first author	Indicates the first author’s country of origin	Nominal
Country of the second author	Indicates the second author’s country of origin	Nominal
Country of the third author	Indicates the third author’s country of origin	Nominal
Country of the fourth author	Indicates the fourth author’s country of origin	Nominal
Country of the fifth author	Indicates the fifth author’s country of origin	Nominal
Country of publication	Country where the works were published or presented	Nominal
Methodology Section	Indicates whether the work has a section dedicated only to the applied methods	Categorical (1 = yes, 0.5 = in the appendix, 0 = no)
Citations on Scopus	In the case of works present in the Scopus database, indicate its number of citations	Discrete
Language	Indactes the language of the work	Nominal
Seminal article	Indicates whether the work cites the seminal article in the references	*Dummy* (1=yes, 0=no)
Database	Indicates whether the work was found in Scopus or Google Scholar	Categorical

APPENDIX B - Table of the Methodological Dimension Variables

VARIABLE	DESCRIPTION	MEASUREMENT
Abstract	If the work has an abstract	*Dummy* (1 = yes, 0 = no)
Research question	Indicates whether the research question is clear in the abstract or in the first section of the work	Categorical (1 = yes, question is clear, 0.5 = research question in another following section, 0 = survey question is not explicit in any part of the text)
Description of the research question	Describe the question asked by the author(s)	Nominal
Hypothesis	Indicates whether the hypotheses are clear in the abstract or in the first section of the work	Categorical (1 = yes, hypotheses are clear, 0.5 = hypotheses mentioned in the following sections, 0 = hypotheses were not made explicit in the text)
Hypothesis description	Describes the work's hypothesis	Nominal
Methodology	Information about the methods used in the abstract or in the first section of the work	*Dummy* (1 = yes, 0 = no)
Year of the elections	Indicates the year(s) of the election(s) analyzed	Nominal
Year 1979	Indicates whether the work analyzes the 1979 election	*Dummy* (1 = yes, 0 = no)
Year 1984	Indicates whether the work analyzes the 1984 election	*Dummy* (1 = yes, 0 = no)
Year 1989	Indicates whether the work analyzes the 1989 election	*Dummy* (1 = yes, 0 = no)
Year 1994	Indicates whether the work analyzes the 1994 election	*Dummy* (1 = yes, 0 = no)
Year 1999	Indicates whether the work analyzes the 1999 election	*Dummy* (1 = yes, 0 = no)
Year 2004	Indicates whether the work analyzes the 2004 election	*Dummy* (1 = yes, 0 = no)
Year 2009	Indicates whether the work analyzes the 2009 election	*Dummy* (1 = yes, 0 = no)
Number of elections	Informs the number of analyzed elections	Discrete
Number of countries	Informs the number of countries analyzed	Discrete
Number of country-years	Informs the number of countries-years analyzed in total	Discrete
Number of main country-years	Informs the number of country-years of the main analysis	Discrete
Regional analysis	Informs if the work analyzed only countries from the same region	*Dummy* (1 = yes, 0 = no)
Analysis level	Specifies the level of the data analysed	Categorical (Micro, macro, both, qualitative, contente analysis)
Sample size	Indicates whether the work is a	Categorical (case study, small

	single case study, small n, medium n, large n or very large n	n, medium n, large n, very large n)
Sample size code	Codes the sample sizes categories	Categorical (0 = literature review, 1 = case study, 2 = small n, 3 = medium n, 4 = large n, 5 = very large n)
Number of methods	Informs the number of methods used in the work	Discrete
Method 01	Informs the single or the most adopted method in the work as a whole	Nominal
Method 02	Informs the second method adopted in the work	Nominal
Method 03	Informs the third method adopted in the work	Nominal

APPENDIX C - Table of the Substantive Dimension Variables

VARIABLES	DESCRIPTIONS	MEASUREMENT
DV1	First dependent variable	Nominal
DV1 measurement	How the first dependent variable is measured	Nominal
DV2	Second dependent variable	Nominal
DV2 measurement	How is the second dependent variable measured, if any	Nominal
DV3	Third dependent variable	Nominal
DV3 measurement	How is the third dependent variable measured, if any	Nominal
Data source	Description of where the authors got their data	Nominal
IVs	Independent variables	Nominal
Number of independent variables	Indicates the total number of independent variables	Discrete
Control variables	Control variables, if any	*Dummy* (1 = yes, 0 = no)
Number of control variables	Indicates the total number of control variables	Discrete
IVs and CVs	Sum of the number of independent variables with the number of control variables	Discrete
Sample (aggregate data)	Sample size (aggregate level)	Discrete
Sample (individual data)	Sample size (ndividual level)	Discrete
Ecological fallacy	Informs whether the work mentions the terms “ecological fallacy” throughout the text	*Dummy* (1 = sim, 0 = não)
Supported hypotheses	Description of the hypotheses supported to a large extent by the tests performed	Nominal
Mixed results	Description of hypotheses with mixed results	Nominal
Refuted hypotheses	Description of hypotheses not supported by the tests performed	Nominal
Untested Assumptions	Description of hypotheses that have not been extensively tested	Nominal
Total number of hypotheses	Total number of hypotheses mentioned in the study	Discrete
Number of tested hypotheses	Number of hypotheses tested in the study	Discrete
Number of supported hypotheses	Number of hypotheses widely supported by the tests	Discrete
Proportion between tested and supported hypotheses	Result of dividing the number of supported hypotheses by the total tested hypotheses	Continous (between 0 and 1)
Fraction of the number of supported hypotheses over the number of tested hypotheses	Fraction representation of the previous variable	Nominal

ANNEX 01 - Rstudio script extract (charts 01 to 05)

The entire script, as well as the database and individual graphics images are available at the following link on Github: https://github.com/Victor-Santana-Santos/SOE_database. To obtain the personal access token to such documents, please send an email to victor.santanasantos@ufpe.br.

```
library(readxl)
library(dplyr)
library(ggplot2)
SOE_database <- read_excel("SOE_database.xlsx", sheet = "SAMPLE")
SOE_excluded <- read_excel("SOE_database.xlsx", sheet = "EXCLUDED WORKS")
SOE_google_scholar <- read_excel("SOE_database.xlsx", sheet = "GOOGLE SCHOLAR")

# PLOT 01

data_entries <- data.frame(
  Source=c("Google Scholar", "Scopus Search", "Scopus Citing List"),
  number_of_entries=c(984, 57, 314),
  Percentage=c(0.7262, 0.0421, 0.2317))

ggplot(data_entries, aes(x=Source, y=Percentage, fill = Source))+
  geom_bar(stat="identity")+
  geom_text(aes(label = number_of_entries), vjust = 1.5, colour = "black")+
  ggtitle("Proportion of entries found according to searching procedure")

# PLOT 02

data_exclusion_criteria <- data.frame(
  Exclusion_criteria=c("Promptly Excluded Entries", "Second List",
                       "Final Sample"),
  number_of_entries=c(746, 555, 54),
  Percentage=c(0.5505, 0.4096, 0.0398))

ggplot(data_exclusion_criteria, aes(x=Exclusion_criteria, y=Percentage,
                                    fill = Exclusion_criteria))+
  geom_bar(stat="identity")+
  geom_text(aes(label = number_of_entries), vjust = 1.5, colour = "black")+
  ggtitle("Proportion of entries found according to the exclusion criteria")

# PLOT 03

data_samples <- data.frame(
  Samples=c(rep("Total number of entries", 3), rep("Second List", 3),
```

```
          rep("Final Sample", 3)),
  Source=rep(c("Scopus Citing List", "Scopus Search", "Google Scholar"), 3),
  Entries=c(314, 57, 984, 186, 9, 360, 43, 3, 8))

ggplot(data_samples, aes(fill=Source, x=Samples, y=Entries)) +
  geom_bar(position="fill", stat="identity") +
  ggtitle("Proportion of works in each sample according to data source")

# PLOT 04

data_works_second_list <- data.frame(
  Works_catalogued=c("Articles", "Working Papers", "Docotral theses",
                     "Master theses", "Bachelor theses"),
  Entries=c(401, 117, 25, 9, 3))

ggplot(data_works_second_list, aes(x=Works_catalogued, y=Entries,
                                   fill = Works_catalogued))+
  geom_bar(stat="identity")+
  theme(legend.position="none")+
  ggtitle("Second list works according to type of work")

# PLOT 05

data_excluded_works_criteria <- data.frame(
  Works_coded_on_the_second_list=c("No SOEs theory,\ncites SA",
                                   "No SOEs theory,\ndoesn't cite SA",
                                   "No EP elections,\ncites SA",
                                   "No EP elections,\ndoesn't cite SA",
                                   "No SOEs theory, no EP\nelections, cites SA",
                                   "No SOEs theory, no EP\nelections, doesn't cite SA",
                                   "Monograph",
                                   "Master theses",
                                   "Doctoral theses",
                                   "Not on the 1st and 2nd\npercentiles of SJR",
                                   "Journal not\nregistred on SJR",
                                   "WPs that made up to the\nfinal sample as articles",
                                   "Translations of\nprevious works",
                                   "Working papers\nin general"),
  Works_registred=c(87, 47, 29, 4, 83, 52, 3, 9, 25, 63, 35, 5, 1, 112))

data_excluded_works_criteria %>%
  mutate(Works_coded_on_the_second_list  =  fct_reorder(Works_coded_on_the_second_list,
Works_registred)) %>%
  ggplot(aes(y=Works_coded_on_the_second_list, x=Works_registred,
             fill = Works_coded_on_the_second_list))+
  geom_bar(stat="identity")+
  theme(legend.position="none")+
  ggtitle("Works on second list according to exclusion criteria")

# PLOT 06
```

```
data_quality_criteria <- data.frame(
  Criteria=c(rep("Journals not on the\ntop half of SJR", 2),
             rep("Journals not indexed\nin SJR", 2)),
  Pass=rep(c("Would enter the final sample", "Would not enter the final sample"), 2),
  Entries=c(12, 51, 2, 32))

ggplot(data_quality_criteria, aes(fill=Pass, x=Criteria, y=Entries)) +
  geom_bar(position="fill", stat="identity")+
  ggtitle("Articles that would have made up the final sample \nif criteria 10 and 11 were not
adopted")

# PLOT 07

data_quality_language <- data.frame(
  Criteria=c(rep("2nd List", 4),
             rep("2nd List \nexcluding 10 and 11", 4),
             rep("Criteria 10 and 11", 4),
             rep("Criteria 10", 4),
             rep("Criteria 11", 4)),
  Language=rep(c("English", "French", "Portuguese", "Spanish"), 5),
  Articles=c(337, 30, 10, 25, 297, 4, 2, 1, 40, 26, 8, 24, 23, 18, 6,
             17, 17, 8, 2, 7))

ggplot(data_quality_language, aes(fill=Language, x=Criteria, y=Articles)) +
  geom_bar(position="fill", stat="identity") +
  ggtitle("Articles according to language and quality exclusion criteria")

# PLOT 08

data_quality_language_2 <- data.frame(
  Criteria=c(rep("Criteria 10\nWould enter\nthe final Sample", 4),
             rep("Criteria 10\nWould not\nenter the final sample", 4),
             rep("Criteria 11\nWould enter\nthe final sample", 4),
             rep("Criteria 11\nWould not\nenter the final sample", 4)),
  Language=rep(c("English", "French", "Portuguese", "Spanish"), 4),
  Articles=c(6, 3, 2, 1, 17, 15, 4, 16, 1, 0, 0, 1, 16, 8, 2, 6))

ggplot(data_quality_language_2, aes(fill=Language, x=Criteria, y=Articles)) +
  geom_bar(position="stack", stat="identity") +
  ggtitle("Second list articles excluded due to Criteria 10 and 11 \naccording to their
language")

# PLOT 09

data_journals_final <- data.frame(
  Journals=c("Electoral Studies",
             "Journal of European \nIntegration",
             "European Union \nPolitics",
             "Comparative Political \nStudies",
```

```
          "European Journal of \nPolitical Research",
          "West European \nPolitics",
          "Journal of Elections, \nPublic Opinion \nand Parties",
          "Other journals"),
  Articles=c(11, 8, 5, 4, 4, 4, 3, 15))

library(forcats)

data_journals_final %>%
  mutate(Journals = fct_relevel(Journals,
                        "Electoral Studies",
                        "Journal of European \nIntegration",
                        "European Union \nPolitics",
                        "Comparative Political \nStudies",
                        "European Journal of \nPolitical Research",
                        "West European \nPolitics",
                        "Journal of Elections, \nPublic Opinion \nand Parties",
                        "Other journals")) %>%
  ggplot(aes(y=Journals, x=Articles, fill = Journals))+
  geom_bar(stat="identity")+
  theme(legend.position="none")+
  geom_text(aes(label = Articles), hjust = 2, colour = "black")+
  ggtitle("Most recurrent journals in the final sample")

# PLOT 10

data_journals_country <- data.frame(
  Countries=c("United Kingdom",
    "United States",
    "Other countries"),
  Articles=c(44, 6, 4))

ggplot(data_journals_country, aes(x=Countries, y=Articles, fill=Countries))+
  geom_bar(stat="identity")+
  theme(legend.position = "none")+
  geom_text(aes(label = Articles), vjust = 1, colour = "black")+
  ggtitle("Journals registered in the final sample \naccording to the publisher's country")

# PLOT 11

data_gender_final <- data.frame(
  Years=c(rep("80", 2), rep("81", 2), rep("82", 2), rep("83", 2), rep("84", 2), rep("85", 2),
rep("86", 2), rep("87", 2), rep("88", 2), rep("89", 2), rep("90", 2), rep("91", 2), rep("92", 2),
rep("93", 2), rep("94", 2), rep("95", 2), rep("96", 2), rep("97", 2), rep("98", 2), rep("99", 2),
rep("00", 2), rep("01", 2), rep("02", 2), rep("03", 2), rep("04", 2), rep("05", 2), rep("06", 2),
rep("07", 2), rep("08", 2), rep("09", 2)),
  Gender=rep(c("Male", "Female"), 2),
  Mentions=c(4, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0,
          3, 0, 4, 0, 2, 0, 0, 0, 2, 0, 4, 1, 2, 1, 1, 1, 4, 0, 3, 1,
          3, 0, 1, 0, 0, 0, 2, 0, 7, 2, 5, 0, 5, 2, 10, 1, 3, 3, 5, 2))
```

```
data_gender_final %>%
  mutate(Years = fct_relevel(Years,
                        "80", "81", "82", "83", "84", "85", "86", "87", "88", "89", "90", "91", "92",
"93", "94", "95", "96", "97", "98", "99", "00", "01", "02", "03", "04", "05", "06", "07", "08",
"09")) %>%
ggplot(aes(fill=Gender, x=Years, y=Mentions)) +
  geom_bar(position="stack", stat="identity") +
  ggtitle("Authors per year by gender")

# PLOT 12

Universities %>% mutate(Country = fct_relevel(Country,"Netherlands", "Germany", "United
Kingdom", "United States","Ireland", "Portugal", "Italy", "France", "Belgium", "Canada",
"Finland", "Austria", "Greece", "Australia")) %>%
  ggplot(aes(x=Mentions, y=University, fill=Country))+
  geom_bar(stat="identity")+
  sort.by

# PLOT 13

data_method_section <- data.frame(
  Years=c(rep("80", 3), rep("81", 3), rep("82", 3), rep("83", 3), rep("84", 3), rep("85", 3),
rep("86", 3), rep("87", 3), rep("88", 3), rep("89", 3), rep("90", 3), rep("91", 3), rep("92", 3),
rep("93", 3), rep("94", 3), rep("95", 3), rep("96", 3), rep("97", 3), rep("98", 3), rep("99", 3),
rep("00", 3), rep("01", 3), rep("02", 3), rep("03", 3), rep("04", 3), rep("05", 3), rep("06", 3),
rep("07", 3), rep("08", 3), rep("09", 3)),
  Meth_section=rep(c("No meth. section",
                     "Meth. section \non appendix",
                     "With meth. section"), 3),
  Articles=c(1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 3, 0, 0,
1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 2, 0, 1, 2, 0, 0, 2, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0,
2, 2, 0, 3, 3, 0, 0, 0, 0, 3, 4, 0, 2, 3, 0, 1, 3, 1, 1))

data_method_section %>%
  mutate(Years = fct_relevel(Years, "80", "81", "82", "83", "84", "85", "86", "87", "88", "89",
"90", "91", "92", "93", "94", "95", "96", "97", "98", "99", "00", "01", "02", "03", "04", "05",
"06", "07", "08", "09"),
         Meth_section = fct_relevel(Meth_section,
                                    "With meth. section",
                                    "Meth. section \non appendix",
                                    "No meth. section")) %>%
  ggplot(aes(fill=Meth_section, x=Years, y=Articles)) +
  geom_bar(position="stack", stat="identity") +
  ggtitle("Works per year by methodological section")

# PLOT 14

install.packages("ggrepel")
library(ggrepel)
```

```
ggplot(Citations, aes(x=year, y=citations_scopus)) +
  geom_point(size = 1.5)+
  geom_text_repel(
    data=Citations %>% filter(citations_scopus>120),
    aes(label=reference),
    size = 3.5)+
  ggtitle("Articles citations on Scopus")

# PLOT 15

ggplot(Citations_per_year, aes(x=year, y=citations_per_year)) +
  geom_point(size = 2)+
  geom_text_repel(
    data=Citations_per_year %>% filter(citations_per_year>13),
    aes(label=reference),
    size = 3.5)+
  ggtitle("Citations per year")

# PLOT 16

ggplot(elections_per_work, aes(x=year, y=n_elections, size = n_of_articles)) +
  geom_point(alpha=0.7)+
  ggtitle("Number of elections analysed per article per year")

# PLOT 17

methods %>%
  mutate(Method_subtype = fct_reorder(Method_subtype, Method_type)) %>%
ggplot(aes(x=n_of_occurences, y=Method_subtype, fill=Method_type))+
  geom_bar(stat="identity")+
  ggtitle("Types and subtypes of methods\nregistered in the final sample")

# PLOT 18

ggplot(Citations_per_year, aes(x=year, y=citations_per_year, color = method)) +
  geom_point(size = 2.5)+
  geom_text_repel(
    data=Citations_per_year %>% filter(citations_per_year>13),
    aes(label=reference),
    size = 3.5)+
  ggtitle("Citations per year according to type of method")

# PLOT 19

data_google_scholar <- data.frame(
  Years=c(rep("80", 4), rep("81", 4), rep("82", 4), rep("83", 4), rep("84", 4), rep("85", 4),
rep("86", 4), rep("87", 4), rep("88", 4), rep("89", 4), rep("90", 4), rep("91", 4), rep("92", 4),
rep("93", 4), rep("94", 4), rep("95", 4), rep("96", 4), rep("97", 4), rep("98", 4), rep("99", 4),
rep("00", 4), rep("01", 4), rep("02", 4), rep("03", 4), rep("04", 4), rep("05", 4), rep("06", 4),
```

```
rep("07", 4), rep("08", 4), rep("09", 4), rep("10", 4), rep("11", 4), rep("12", 4), rep("13", 4), rep("14", 4), rep("15", 4), rep("16", 4), rep("17", 4), rep("18", 4), rep("19", 4), rep("20", 4)),
  Language=rep(c("Portuguese", "French", "Spanish", "English"), 41),
  Entries=c(0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 1, 0, 3, 0, 0, 0, 3, 0, 0, 0, 4, 0, 0, 0, 5, 0, 0, 1, 2, 0, 0, 0, 5, 0, 0, 0, 12, 0, 0, 0, 10, 0, 0, 4, 2, 0, 0, 0, 1, 0, 0, 2, 6, 0, 0, 1, 15, 0, 0, 1, 17, 0, 0, 0, 14, 0, 0, 2, 16, 0, 2, 1, 17, 0, 3, 5, 40, 0, 1, 0, 21, 0, 2, 1, 36, 0, 0, 2, 55, 0, 5, 1, 60, 7, 8, 1, 85, 1, 4, 1, 92, 0, 6, 6, 91, 1, 2, 4, 105, 2, 22, 9, 153, 3, 8, 9, 150, 3, 1, 6, 160, 7, 11, 13, 163, 4, 1, 5, 187, 11, 13, 10, 256, 7, 7, 20, 285, 8, 5, 8, 235, 6, 0, 9, 215, 5, 2, 4, 208, 13, 8, 14, 258, 6, 6, 3, 242))

install.packages("ggthemes")
library(ggthemes)

data_google_scholar %>%
  mutate(Years = fct_relevel(Years, "80", "81", "82", "83", "84", "85", "86", "87", "88", "89", "90", "91", "92", "93", "94", "95", "96", "97", "98", "99", "00", "01", "02", "03", "04", "05", "06", "07", "08", "09", "10", "11", "12", "13", "14", "15", "16", "17", "18", "19", "20")) %>%
ggplot(aes(x=Years, y=Entries, group=Language, color=Language)) +
  geom_line(size = 0.8) +
  theme_igray()+
  ggtitle("Number of entries found on Google Scholar by year")

# PLOT 20

data_scopus <- data.frame(
  Years=c(rep("80", 2), rep("81", 2), rep("82", 2), rep("83", 2), rep("84", 2), rep("85", 2), rep("86", 2), rep("87", 2), rep("88", 2), rep("89", 2), rep("90", 2), rep("91", 2), rep("92", 2), rep("93", 2), rep("94", 2), rep("95", 2), rep("96", 2), rep("97", 2), rep("98", 2), rep("99", 2), rep("00", 2), rep("01", 2), rep("02", 2), rep("03", 2), rep("04", 2), rep("05", 2), rep("06", 2), rep("07", 2), rep("08", 2), rep("09", 2), rep("10", 2), rep("11", 2), rep("12", 2), rep("13", 2), rep("14", 2), rep("15", 2), rep("16", 2), rep("17", 2), rep("18", 2), rep("19", 2), rep("20", 2)),
  Source=rep(c("Works that cited \nReif & Schmitt (1980)", "Search results"), 41),
  Entries=c(0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 3, 0, 0, 0, 2, 1, 2, 0, 6, 0, 0, 0, 0, 0, 2, 1, 8, 2, 3, 1, 9, 1, 6, 2, 4, 1, 9, 1, 7, 1, 10, 2, 19, 5, 25, 5, 32, 6, 30, 1, 41, 8, 41, 8, 51, 8, 51, 6, 63, 14, 81, 17, 99, 9, 77, 15, 109, 32, 126, 18, 88, 12, 96, 23, 73, 21, 94, 34))

data_scopus %>%
  mutate(Years = fct_relevel(Years,  "80", "81", "82", "83", "84", "85", "86", "87", "88", "89", "90", "91", "92", "93", "94", "95", "96", "97", "98", "99", "00", "01", "02", "03", "04", "05", "06", "07", "08", "09", "10", "11", "12", "13", "14", "15", "16", "17", "18", "19", "20")) %>%
  ggplot(aes(x=Years, y=Entries, group=Source, color=Source)) +
  geom_line(size = 0.8) +
  theme_igray()+
  ggtitle("Number of entries found on Scopus by year")
```