

Introducing COMEPELDA: Comprehensive European Parliament electoral data covering rules, parties and candidates

European Union Politics

2022, Vol. 23(2) 351–371

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14651165211053439

journals.sagepub.com/home/eup**Thomas Däubler** 

School of Politics and International Relations, University College
Dublin, Dublin, Ireland

Mihail Chiru

Department of Politics and International Relations, University of
Oxford, Oxford, United Kingdom; Median Research Centre,
Bucharest, Romania

Silje SL Hermansen 

iCourts – The Danish National Research Foundation’s Centre of
Excellence for International Courts, University of Copenhagen,
Copenhagen, Denmark

Abstract

We introduce a new collection of data that consolidates information on European Parliament elections into one comprehensive source. It provides information on formal electoral rules as well as national-level and district-level election results for parties and individual politicians (including full candidate lists). The use of existing and new key variables makes it easy to link the data across the different units of observation (country, party, candidate, member of parliament) and join them with external information. Currently, the data cover four elections (1999–2014). Among other aspects, the collection should facilitate research on the European Parliament’s allegedly weak electoral connection. In this article, we outline the main features of the datasets, describe patterns

Corresponding author:

Thomas Däubler, School of Politics and International Relations, University College Dublin, Belfield, Dublin 4, Ireland,

Email: thomas.daubler@ucd.ie

of intra-party competition and preference voting and conduct exploratory analyses of individual-level changes in list positions.

Keywords

European Parliament, electoral systems, election results, candidate selection, political careers

Introduction

The European Parliament (EP) operates in an intricate political setup regarding both its relations with other European Union (EU) institutions and its intra-chamber dynamics. Its members (MEPs) come from up to 28 member states and represent different national parties. At the same time, most national parties are affiliated to European party groups. Moreover, while MEPs deal with European matters, their recruitment occurs in elections that are formally and practically independent national-level contests (De Vreese and van der Brug, 2016; Reif and Schmitt, 1980; Van der Brug et al., 2022).

Empirical research into the behavior of MEPs has expanded considerably in the past decades, matching the empowerment of the legislature (for comprehensive reviews, see Hix and Høyland, 2014; Ripoll Servent, 2018; Yordanova, 2011). Most of these studies assume that the behavior of MEPs lacks an electoral connection in the sense that it is not motivated by re-(s)election concerns and is without electoral consequences. This assumption remains the standard view among EP scholars (Hix and Høyland, 2013; Judge and Earnshaw, 2008). However, systematic empirical studies of such electoral causes and consequences are scarce and somewhat limited in scope.¹

This remains a limitation of the literature on this topic because the large variety of electoral systems used in EP elections (Däubler and Hix, 2018; Farrell and Scully, 2007) may result in considerable heterogeneity in representational behavior (e.g. Bowler and Farrell, 2011; Hix, 2004; Hix and Hagemann, 2009) and electoral accountability (Sorace, 2021; Wilson et al., 2016). For example, several larger member states use subnational districts, which have become more prevalent since 2004 with their introduction in France (Navarro, 2016) and the EU accession of Poland. In addition, the use of preferential voting is now more widespread. No fewer than 10 of the 13 countries that have joined the EU since 2004 use flexible or open rather than closed lists.

An obstacle to empirical research into electoral incentives and effects is the lack of a comprehensive data source. While plenty of excellent data on EP elections exist, there is no consolidated and easy-to-use collection covering electoral rules, national- and district-level results for parties and electoral information for individual candidates and MEPs. What is already available from scattered sources includes classifications of electoral systems (Farrell and Scully, 2007) and party lists (Däubler and Hix, 2018), national-level election results for parties from the ParlGov database (Döring and Manow, 2020),

information on MEPs' formal parliamentary roles (Høyland et al., 2009) and several replication datasets (e.g. Daniel, 2015; Meserve et al., 2015). The various components of the European Election Studies are of limited use in this context.

This article introduces a new 'one-stop shop' for such data.² The information comes in the form of three datasets: an aggregate dataset with country-level information (e.g. electoral rules) and party-level results (at the national and district levels), a dataset for candidates (including list positions and preference votes where applicable) and a dataset for MEPs (with mandate types, district affiliation and convenient links to the other two datasets). The data collection covers four European elections (1999, 2004, 2009 and 2014).

This new data collection can help address a number of core research themes on the EP. To highlight their potential, in this article we use these data to describe recent EP elections with regard to intra-party competition, the take-up of preference voting and the extent to which MEPs are elected on the basis of personal votes. An exploratory analysis of the link between parliamentary activity (questions, reports and speeches) and list rank change for first-term MEPs reveals a complex pattern.

Electoral incentives, research themes and the added value of a new data collection

This new collection of data will enable researchers to explore several research themes on electoral incentives and the intra-party dimension in the EP. The first of these concerns voters' choice of candidates in systems that allow for it. How do voters make such choices and on what kind of information do they base their decision (Christensen et al., 2020; Gattermann and De Vreese, 2017, 2022)? Not only is this an interesting question about voting behavior, but it also has implications for MEPs' electoral incentives. At the same time, survey data rarely provide information on candidate choice, since it is too expensive to collect in this context. Many interesting questions can thus be answered using the candidate-level preference voting results we provide. Our data allow to examine preference vote take-up, candidate performance and the list-level characteristics of preference votes such as their concentration (Dodeigne and Pilet, 2021).

A broad second field comprises the link between electoral incentives and the behavior of politicians. Existing work has, for example, analyzed the impact of electoral systems on candidates' campaign strategies (Hix and Hagemann, 2009; Lorenzo Rodríguez and Garmendia Madariaga, 2016; Obholzer and Daniel, 2016). For MEPs, the focus has been on representational roles (Brack and Costa, 2019; Farrell and Scully, 2010;), on the type of parliamentary activities they prioritize (Høyland et al., 2019; Sorace, 2018), on parliamentary voting (Däubler and Hix, 2018; Hix, 2004) and on their interactions with interest groups (Ibenskas and Bunea, 2021).³

A third theme examines the consequences of parliamentary behavior on the electoral performance of parties and candidates. For example, analyzing data from six EP elections, Hobolt and Høyland (2011) show that parties that nominate highly experienced national politicians as leading candidates tend to be rewarded by voters. Their focus on the top candidate from each list reflects the data availability constraints that have

characterized extant research at the intersection between the EP legislative and electoral arenas. While there is evidence for MEP productivity-based retrospective voting being a feature of the 2019 EP elections (Sorace, 2021), no study has tested longitudinally whether the vote shares of party lists vary with the different quality of MEP incumbents or whether citizens punish legislators for their voting record on controversial issues. Our datasets offer the possibility to easily match information on MEPs, candidates and election results, thereby providing a starting point for answering such research questions.

The COMPELDA collection

Starting point, aims and approach

A major obstacle to research on the electoral connection in the context of the EP appears to be the lack of a consolidated and comprehensive source of both electoral data on different levels or actors (i.e. MEPs, parties and candidates) and information on the institutional context of EP elections in the various countries. Data do exist on various aspects of EP elections. Institutions are described, for example, by Däubler and Hix (2018), Farrell and Scully (2007) and Stöver and Wüst (2006). The ParlGov database (Döring and Manow, 2020) provides aggregate election results at the national level, but does not include information on the level of subnational districts. Data on the electoral performance of politicians are often available on the web, but the sources are scattered. Hence, rather than collecting genuinely new data, our main aim was to assemble existing data from different sources and integrate them into a coherent collection.

Guiding our efforts was a focus on electoral data. Rather than taking a more encompassing approach toward electoral incentives (e.g. trying to collect biographical data on individual politicians or including information about what happens inside the EP such as party switching), we concentrated exclusively on the electoral context and a rather narrow set of variables. This helped us both keep the scope manageable and build a data collection characterized by very high coverage and quality as a result of thorough checking. Further, because our collection has a somewhat narrow focus on the electoral arena, it is important that scholars can easily link it to existing data and extend it in the future. Therefore, we tried to set up the data in a way that makes it simple to combine our datasets with others. We will describe this in more detail below.

In line with the idea of keeping the task focused, we also delineated the case coverage early in the project. First, we decided to study EP elections since 1999. We presumed that locating information for contests before that year could be challenging, since online availability is limited. The 2019 election occurred while the project was already ongoing, and we opted to finish the work on the earlier cases rather than take on another task, which can be addressed in future research. Second, motivated by an interest in how previous and upcoming elections affect politics in the EP, we chose to include those parties that: a) win at least one seat in the current election, b) won at least one seat in the previous election or c) present as candidate at least one MEP who served during the outgoing term of

the EP. The first of these criteria is simple to implement, the second slightly harder (for instance in case of party renaming) and the third turned out to be occasionally challenging (since some outgoing MEPs may run for parties that garner few votes).

Our data collection starts from three main sources. First, we imported national-level data on votes and seats from the ParlGov database (Döring and Manow, 2020), which also helped delineate the cases to cover following the rules outlined above.⁴ Second, we employed the data from Høyland et al. (2009), which allowed us to identify MEPs and the basic information on their parliamentary mandates. Third, large parts of the information on electoral systems was retrieved from Däubler and Hix (2018), who cover the elections between 2004 and 2014. The remaining data, especially party-level results for member states with subnational districts and politician-level information on list positions and preference votes, were harder to assemble. We partly reused the data provided by Däubler and Hix (2018) and Hermansen (2018), but much information had to be collected from online (and occasionally offline) sources. Whenever possible, we relied on information available from the EP itself (e.g. European Parliament, 1999, 2004) or on the official election results provided by election authorities, home affairs ministries and national statistical offices.

Some readers might think that compiling these data requires only a few days of webscraping and some data wrangling. Yet the necessary steps are numerous, and due to various problems, time intensive. To start with, some of the data were difficult to find. For example, more detailed information such as candidate-level results is not always available on English-language websites. The 28 countries use a broad range of languages, some of which are written in non-Latin alphabets (Bulgarian and Greek). A particular challenge in collating the data was correctly matching parties and politicians across data sources given the variations in formatting, naming and spelling due to the inconsistent handling of party alliances, language-specific characters, and additional forenames. Other difficulties arose from various institutional or behavioral idiosyncrasies. For example, candidates running as independents or the fact that Italian candidates may be listed (and are often elected) in more than one district caused additional complexity.

Introducing the three datasets

Our data collection consists of three datasets, which refer to the different entities of interest: member states and parties (aggregate data), candidates and MEPs. Table 1 provides an overview of the content of the three datasets

The *aggregate-level dataset* comprises approximately 1800 observations. The unit of analysis is the party as it contested the election (i.e. including alliances and any team of candidates running under the same label) in an EP election in a district. In other words, the unit of analysis is an electoral party, because all the variables (e.g. votes, seats) are well defined at this level in a given election. Complications arise for ID variables or when comparing units over time, since alliances, which are frequent in certain countries (e.g. Croatia, Portugal and Spain), are often short-lived. Similar problems are caused by party name changes and mergers. In general, we use the same party ID and party name as

Table 1. Structure and content of the datasets.

	Aggregate-level data (AD)	Candidate-level data (CD)	MEP-level data (MD)
<i>Basic unit</i>	Combination of party, election year and electoral district	Combination of person, district and list type (main or successor list)	Combination of MEP, EP session and mandate episode
<i>Coverage</i>	Parties (and independent candidates) that win any seats in the current election, have done so in the previous election, or run any MEPs serving in the outgoing EP session	All candidates running for the parties included in the AD	All MEPs
<i>Types of information provided</i>	<ul style="list-style-type: none"> • Votes • Number of candidates • Seats • List-level features of preference voting • Electoral system features 	<ul style="list-style-type: none"> • List rank • Preference votes • Electoral success and order of election 	<ul style="list-style-type: none"> • Episode information (start, end, mandate type) • Indicators for running and being elected in the previous and subsequent elections
<i>Identifiers for internal linkage</i>	<ul style="list-style-type: none"> • ID for entry in this dataset (AD) 	<ul style="list-style-type: none"> • ID for the corresponding entry in AD • IDs for the corresponding entries in MD (for previous EPs and subsequent EPs, with separate variants that are non-missing only if the candidacy was linked to an EP mandate) 	Previous election: <ul style="list-style-type: none"> • ID of the focal entry in AD • ID of the focal candidacy in CD Subsequent election: <ul style="list-style-type: none"> • ID of the focal entry in AD (if running again) • ID of the focal candidacy in CD (if running again)
<i>Identifiers for external linkage</i>	<ul style="list-style-type: none"> • Party ID from ParlGov • Election ID from ParlGov 	<ul style="list-style-type: none"> • MEP ID as assigned by the EP (currently only for members of the two EPs before and after the respective election) 	<ul style="list-style-type: none"> • MEP ID as assigned by the EP • Wikidata ID • For previous and subsequent elections: Party ID from ParlGov

ParlGov, which need not be the same as the label actually used in the respective election. As explained in the codebook, our data include an indicator variable for party alliances and additional IDs that allow users to look up component parties in ParlGov. In addition,

the dataset provides information on party lists that run separately but form apparante-ments for inter-party seat allocation.

The substantive data encompass information on the number of votes, seats and candidates, complemented by list-level summaries of preference voting. For several variables, we present both country-level and district-level versions. If a country does not use sub-national districts, the entries in the latter are identical to those in the former. This dataset also holds information on the electoral systems used in the various countries with regard to both the between-party dimension (e.g. the electoral formula used or whether multi-tier allocation takes place) and the within-party dimension (e.g. ballot structure or whether preference voting is obligatory). The included ParlGov IDs for party, alliance and election can be used for external linkage in combination with the Party Facts database (Döring and Regel, 2019).

The *candidate-level dataset* provides information on all the candidates running for those parties included in the aggregate-level dataset. The file covers more than 29,000 observations. The main substantive variables are pre-electoral list rank (for systems in which parties make that decision), preference votes (where applicable), indicators for electoral success and order of election. Within the collection, the dataset can easily be linked to the other files using the internal IDs for the corresponding entries in the aggregate-level and MEP-level data (here with different variants referring to the previous and subsequent terms and – in the case of multiple candidacies – indicating whether the candidacy was linked to an actual mandate).

The *MEP-level dataset* includes more than 3700 observations. The units refer to a specific ‘mandate episode’ of an MEP in an EP session, which is defined by a certain path to becoming an MEP. Table 2 shows the frequency of the mandate types across sessions. Although there is only one episode per MEP and EP in most instances, some politicians may have two or three episodes within a session, especially in the context of EU enlargement. For example, some Bulgarian and Romanian legislators had up to three statuses in EP 6: as observers until January 1 2007, then as appointed MEPs until elections were organized in each country and, finally, as elected MEPs afterward.

Table 2. Mandate types of episodes (row %, absolute number in parentheses).

<i>Parliamentary term</i>	<i>Initially elected</i>	<i>Replacement from the beginning of the session</i>	<i>Replacement during the session</i>	<i>Appointed (EU enlargement)</i>	<i>Appointed (Lisbon Treaty change)</i>	<i>Observer</i>
EP 5	60%	1%	8%	16%	NA	15%
1999–2004	(620)	(6)	(86)	(163)		(154)
EP 6	74%	3%	12%	6%	NA	5%
2004–2009	(750)	(35)	(122)	(57)		(47)
EP 7	84%	3%	11%	0%	2%	1%
2009–2014	(724)	(24)	(92)	(0)	(18)	(5)
EP 8	86%	2%	13%	0%	NA	0%
2014–2019	(737)	(14)	(109)	(0)		(0)

The MEP-level dataset holds basic episode information (start, end and episode type) and indicators for whether the MEP was a candidate and was elected in the previous and subsequent elections.⁵ It also provides information on the *electoral* party (ParlGov Party ID) and district with which the MEP was affiliated in the previous and subsequent elections (where applicable). For MEPs with several candidacies in the same election, these variables refer to the ‘focal’ candidacy. For the previous election, it refers to the district the MEP eventually represented and the list type – main or successor list – on which the mandate was based.⁶ For the subsequent election, it refers to the district represented in the outgoing EP and the main list. Adding further information from the other datasets of the collection is easy using the simple numeric ID variables included. To facilitate joining information from external sources, The data also come with the MEP’s Wikidata ID to enable scholars to join additional information from Wikidata, the database hosted by the Wikimedia Foundation. This could be done by using the R package WikidataR (Keyes et al., 2021). It is also possible to link to the Comparative Legislator Database (Göbel and Munzert, 2021), which at the time of writing covers 10 national-level parliaments (but not the EP).

Illustration of data features and example analyses

We first discuss the electoral systems used in the 2014 EP election as visualized in Figure 1. The classification follows that of Däubler and Hix (2018). In our data, we distinguish 29 units for this purpose, the 28 member states and Northern Ireland (which has a different electoral system and party system than Great Britain). In line with the EU Treaty, elections need to be conducted using some form of proportional representation system (subsuming STV, which is used in Ireland, Northern Ireland and Malta). Of the list-based systems, seven employed closed lists (intra-party seat allocation based on the pre-electoral list rank), ten had open lists (intra-party seat allocation based on preference votes exclusively), and nine flexible ones (candidates can ‘jump the queue’ given by the pre-electoral ranking if they win a certain number of personal votes). In eight of the 29 cases, a candidate vote is obligatory. Some form of subnational district is found in six countries (France, Germany, Great Britain, Ireland, Italy and Poland).

Intra-party competition

Intra-party competition at the ballot, induced by the existence of preferential voting, generates the strongest electoral incentive for personal vote seeking, because candidates need to distinguish themselves from co-partisans that benefit equally from the same party reputation (André et al., 2014; Carey and Shugart, 1995; Hix and Hagemann, 2009). The standard approach to testing the argument follows the ‘conditional effect of district magnitude’ hypothesis introduced by Carey and Shugart (1995). Intra-party competition is assumed to increase with district magnitude (M) in systems allowing for a candidate vote, whereas under closed lists personal vote-seeking incentives should decrease in larger districts, since candidates become more anonymous.

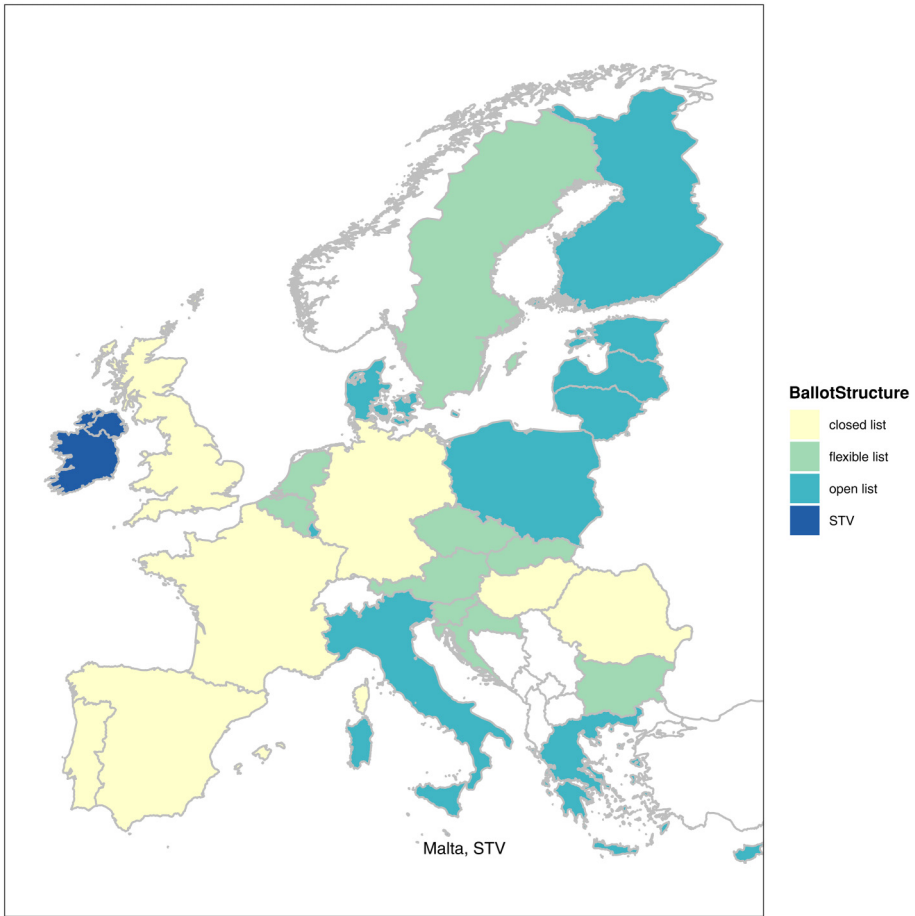


Figure 1. Ballot structure in EP elections, 2014.

Note: Map created using Lahti et al. (2017). © EuroGeographics for the administrative boundaries.

Using *M* as an indicator of intra-party competition suffers from a number of problems. First, the measure also taps into other concepts, especially the size of the district electorate. The various mechanisms linked to the correlates of *M* prove hard to disentangle (André and Depauw, 2014). Second, as proposed by Crisp et al. (2007), there is no guarantee that *M* captures co-partisan crowdedness. In the original contribution, Carey and Shugart (1995: 431) discuss that the ratio calculated as the number of a party’s candidates over *M* is actually the ‘key determinant’ of the need for intra-party distinction; nevertheless, they settle on *M* as the ‘fixed and identifiable’ part, arguing that the number of a party’s candidates ‘tends to rise with *M*’. Crisp et al. (2007) take up the ratio idea again and make the valid point that the notion of crowdedness would require an appropriate denominator equaling the expected number of seats a party is going to win.

In our data, there is a high correlation between M and the number of candidates parties run (Pearson's $r = 0.80$, Spearman's rank correlation $\rho = 0.81$). Figure 2 plots the ratio of party candidates and party seats against M (the axes use a \log_{10} scale). The value of this ratio may be the best prediction of intra-party competition that MEPs have, at least at the beginning of the term (their party's electoral prospects may change over time, of course). As can be seen, the relationship between the two variables is rather complex. While the smoothing line points to a linear relation in the middle range of M (approximately between 4 and 15), the overall relationship between M and intra-party crowdedness has a sigmoid shape (Spearman's rank correlation $\rho = 0.34$). Researchers must thus be careful when using M as a measure of intra-party competition in EP elections (and beyond). Indeed, the ratio of candidates and seats might not be the ideal solution; one problem of this approach is that typically few candidates are serious contenders, so the indicator tends to overestimate intra-party competition, and possibly not equally in all instances (André et al., 2014: 236).

There is another thorny issue with M in the context of EP elections. Three of the member states with subnational districts (Germany, Italy and Poland) operate a two-tier seat allocation system under which seats are first assigned between parties at the national level and then within parties to the districts. The German case (whose observations are excluded from Figure 2) is even trickier, since typically only the Christian Democrats run 16 state-level lists (CSU in Bavaria, CDU in the other states; lists are connected for inter-party allocation), whereas the other parties employ one national list. Hence, in Italy, Poland and Germany, how to code M remains debatable. Overall, there is scope for new theoretical and empirical suggestions about measuring intra-party

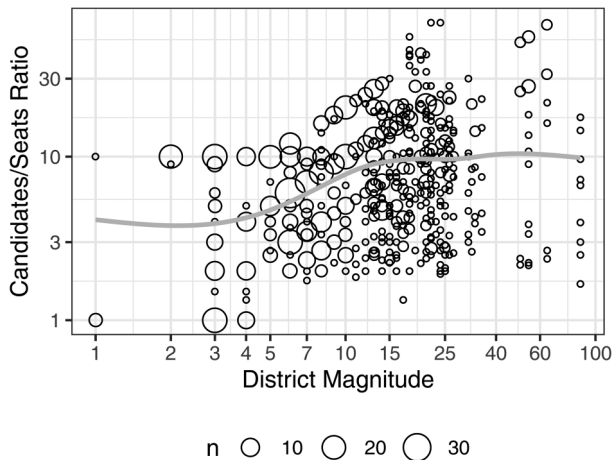


Figure 2. The non-linear relationship between district magnitude and intra-party competition. Note: Based on aggregate-level data for EP elections 1999 to 2014. M is coded at the district level for France, Great Britain and Ireland and for the two-tier allocation cases for Italy and Poland. German cases are excluded (see the text). The smoothing line is based on a generalized additive model, with a 95% confidence interval.

competition (for a discussion of vote dispersion measures see Dodeigne and Pilet, 2021), some of which can easily be implemented with our data collection.

Preference voting

As pointed out above, candidate voting by citizens is of interest both in itself and in terms of its consequences for MEPs’ incentive structures. A first question relates to the extent to which voters use the preference vote option if they can choose between casting a party vote or supporting a candidate. Figure 3 displays the share of party voters casting preferences at the party level in those member states in which preference voting is optional and that grant citizens one candidate vote.⁷ Each circle represents the value of a national party in an EP election; the mean across the parties in each election is shown as a solid larger point. Member states are arranged in descending order by the mean across all party-level

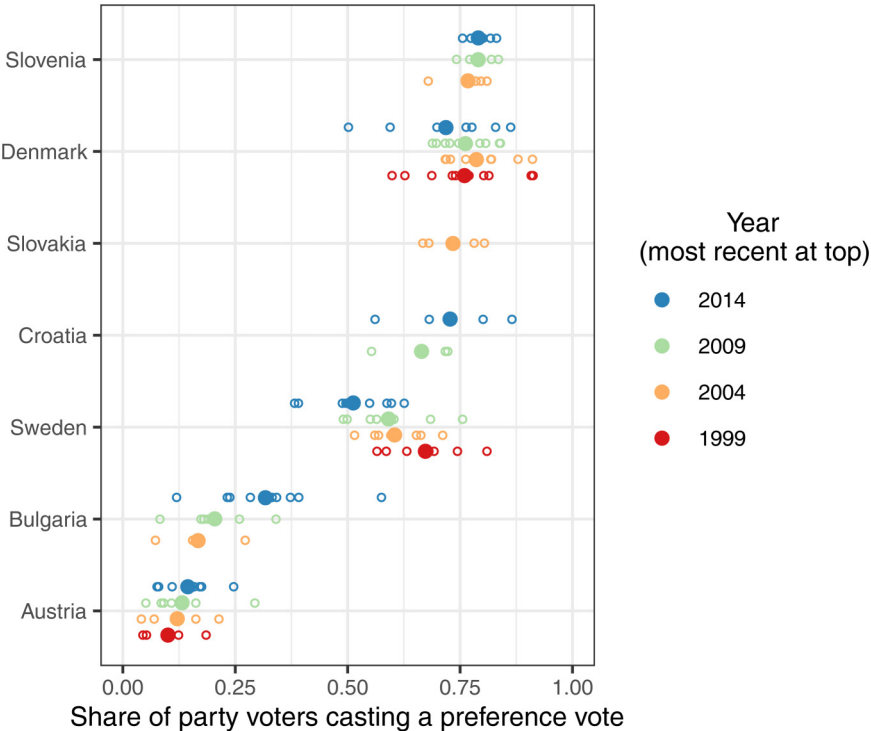


Figure 3. Preference vote use across countries, across parties and over time.
Note: Shown are member states with one optional candidate vote (Slovakia increased the number of admissible preference votes to two for the 2009 election). Election years (most recent at the top within each country) refer to the previous main EP election year (Bulgaria voted in 2007 rather than 2004, Croatia in 2013 rather than 2009).

data points. Within each country, the most recent election is shown at the top and the earliest EP election at the bottom.

This figure shows that the uptake of the candidate vote option varies substantially across countries. Broadly speaking, in some countries, around three out of four voters regularly choose a candidate (Slovenia, Denmark, Slovakia and Croatia). In Bulgaria and Austria, on the contrary, typically only one out of four voters does so. This may be because these two countries had flexible lists with hard-to-reach thresholds before 2014 (Däubler and Hix, 2018). Both citizens and candidates may care less about personal votes if these are unlikely to change the list order. In Bulgaria, we find the expected increase in take-up following the reduction of the preference vote threshold for candidates. Sweden occupies a central position in the graph. Interestingly, the Scandinavian country also features the clearest differences over time. Perhaps surprisingly, the take-up of the candidate vote has decreased considerably over time in Sweden. Overall, the graph illustrates that a larger part of the variation is found across countries rather than between parties or over time.

Another question is whether preference votes ‘make a difference’. While providing a detailed answer to this question is beyond the scope of this paper, we show some results based on a simple intuitive approach. For countries that use lists pre-ranked by parties and allow for preference voting, we can identify MEPs who were elected based on the actual rules in place, but who would not have been so under closed lists. In other words, given the total number of seats won in the given EP election, the selectors would have preferred other candidates to take them. Figure 4 illustrates how often that happens, distinguishing for each country-election the share of MEP candidates who owe their election to their personal vote (in color) and those who would also have been elected under a closed list rule (in light gray). It could be argued that party selectors at the time of nominations cannot clearly predict the number of seats the party will win. Using a stricter definition, we therefore further distinguish candidates (with darker colors) whose list position was at least three ranks lower than the number of seats won (e.g. list position five or lower if the party won two seats).⁸ Member states are sorted according to the share of candidates elected out of order (pooled across elections) and within country by election year (most recent at the top). The circles indicate the average size of the national delegation.

Candidates can become MEPs even if they were ranked at a list place insufficient for election according to the selectorate’s ranking. This phenomenon is most frequent in Italy. In the four elections covered, the relative frequency of Italian MEPs ‘jumping the queue’ was 40–47% of initially elected candidates. Austria and Bulgaria are again at the bottom end of the graph; this is not surprising since it is hard to reach the preference vote threshold in flexible list systems if few voters are inclined to use the candidate vote option, as discussed above. In most countries, out-of-order election occurs regularly, although infrequently. However, caution is warranted when interpreting the number of such candidates. Low figures may not imply that parties always ‘get their way’; often selectors’ and voters’ preferences may simply be aligned, or parties even anticipate who will be the most popular candidates in the electorate. Again, these patterns point to interesting research possibilities from the perspectives of both voting and parliamentary behavior.

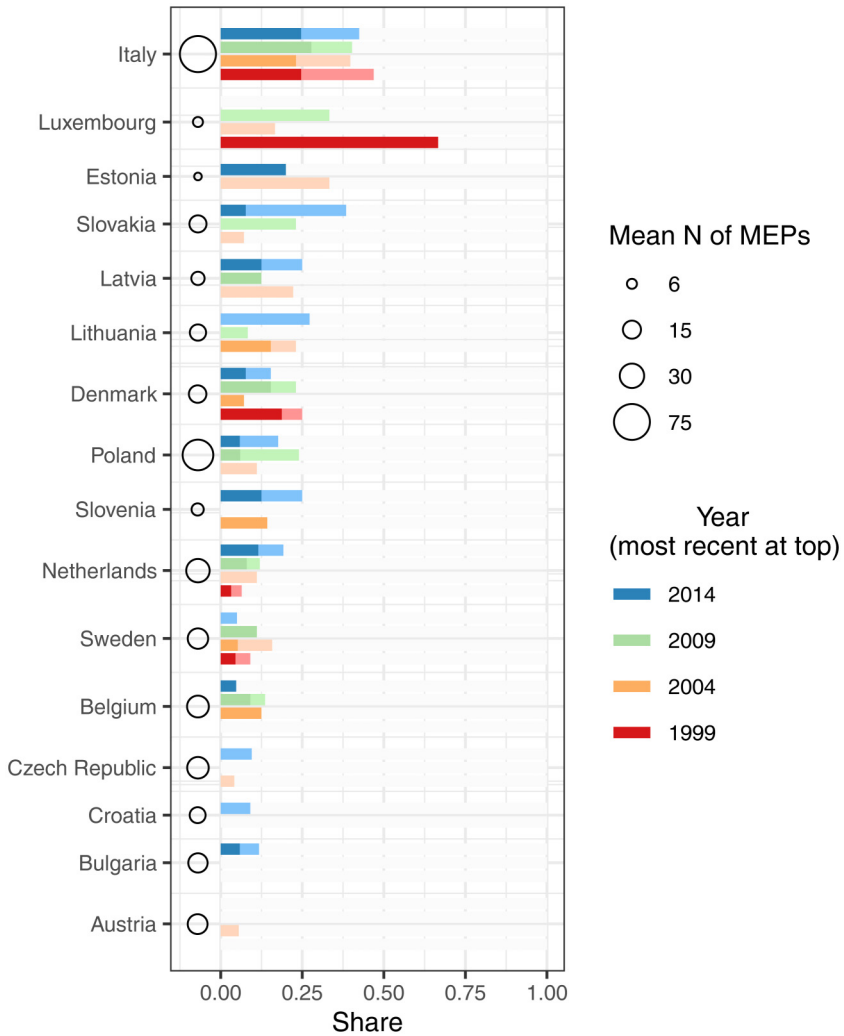


Figure 4. MEPs owing their election to preference votes.

Note: Data refer to MEPs initially elected (i.e., not considering replacement MEPs). Election years (most recent at the top within each country) refer to the previous main EP election year (Bulgaria voted in 2007 rather than 2004, Croatia in 2013 rather than 2009).

Parliamentary activity and re-selection

From a principal–agent perspective (Hix, 2004; Hug, 2016; Klüver and Spoon, 2015; Mühlböck, 2012), the national party selectorate is one of the key principals of MEPs, since it decides on renomination and therefore crucially affects re-election prospects. If an electoral connection is at work, the basic expectation would be that party selectors

Table 3. Regression analyses of the change in list rank.

	Binary logit		Multinomial logit	
	Demotion vs No change/ Promotion		Demotion vs Promotion	No change vs Promotion
Questions				
2 nd tercile	-0.15 (0.36)		-0.04 (0.44)	-0.77* (0.31)
3 rd tercile	-0.66 (0.37)		-0.58 (0.51)	-0.40 (0.33)
Reports				
2 nd tercile	-0.07 (0.40)		-0.04 (0.52)	-0.19 (0.32)
3 rd tercile	-0.33 (0.35)		-0.52 (0.49)	-0.15 (0.32)
Speeches				
2 nd tercile	-0.11 (0.34)		-0.55 (0.41)	-0.52 (0.49)
3 rd tercile	0.70* (0.33)		0.52 (0.46)	-0.28 (0.45)
Leadership 0/1	-0.01 (0.38)		-0.17 (0.47)	-0.03 (0.45)
Age	0.01 (0.02)		0.02 (0.02)	0.02 (0.02)
Female	-0.17 (0.27)		-0.22 (0.36)	0.04 (0.33)
Replacement MEP	2.82* (1.22)		3.19** (1.13)	1.84 (1.71)
Log candidates/seats	-0.44 (0.29)		-0.31 (0.35)	0.49 (0.31)
First list rank 0/1	0.00 (0.96)			
Normalized list rank	-4.02** (1.54)		-4.93** (1.55)	-3.35 (2.03)
First in election order 0/1	-0.13 (0.96)		-0.55 (1.30)	-1.64 (1.25)
Normalized election order	-0.63 (0.64)		-0.75 (0.82)	-1.27* (0.64)
Reselection year				
2009	0.22 (0.40)		0.17 (0.47)	0.27 (0.41)
2014	-0.17		-0.22	0.23

(continued)

Table 3. Continued.

	Binary logit	Multinomial logit	
	Demotion vs No change/ Promotion	Demotion vs Promotion	No change vs Promotion
	(0.39)	(0.65)	(0.54)
Constant	-0.02	0.33	-0.38
	(1.04)	(1.58)	(1.38)
N	475		353

Note: Standard errors clustered for party in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

reward those MEPs who work hard and sanction those who do not. Previous empirical evidence for such a relationship is mixed, though.⁹

COMPELDA allows us to examine this relationship from a variety of perspectives. As an example, we present the results from a simple model analyzing the change in list position for outgoing MEPs (who were running and renominated). The analysis focuses on first-time MEPs, since their initial list position cannot have been affected by prior EP work in a narrow sense. Hence, selectors may update their assessment of these politicians based on their performance during their first term in the EP. As dependent variables, we consider a binary indicator for demotion (lower list place than the previous one) and, for MEPs who did not start off from the first list place already, a ternary indicator for demotion, no change or promotion. To improve comparability, the analyzed set of MEPs is restricted to those who rerun for the same party (i.e. those whose party ID remains constant) in the same district and whose party did not change between running on its own and as part of an alliance over time. To avoid further assumptions about functional form, MEPs are grouped based on the terciles of the distributions of the number of reports, questions and speeches.¹⁰ As control variables, we include age, sex, the co-partisan crowdedness measure discussed above (in logged form) and indicators for EP leadership position and replacement MEP status. To capture the variation in starting points and marginality, we consider the ex-ante list rank and ex-ante rank of election, each with indicators for being first and with a relative measure dividing the rank by the total number of candidates and seats (André et al., 2015).

Table 3 summarizes the results that, as in previous work, are somewhat mixed. In Model 1, for the binary demotion variable, five of the six activity-related coefficients have the expected negative sign, suggesting that MEPs are less likely to be demoted if their activity does not fall in the bottom third, but none of these coefficients is statistically significant. Surprisingly, MEPs who give many speeches face a higher risk of being demoted ($p < 0.05$). For the analysis of the trichotomous measure (demotion, no change and promotion), Figure 5 displays the predicted probabilities based on Model 2. We can see that MEPs asking at least an intermediate number of questions are more likely to be promoted. The differences for reports are fairly muted, while non-linear patterns are clearly visible for speeches. MEPs who speak moderately often are most likely

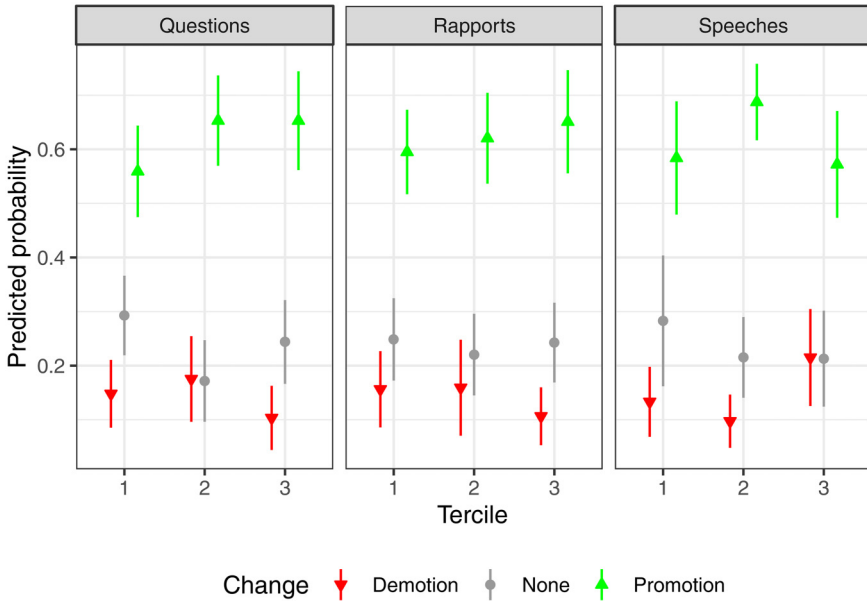


Figure 5. Mean predicted probability for types of list rank changes.

Note: Depicts the mean predicted probabilities (and 95% confidence intervals) across the sample, with other covariates as observed.

to be promoted. Again, those speaking most frequently are most likely to receive a *worse* list position. It is an interesting question whether this is an endogeneity problem in the sense that we do not adequately capture the renomination-related incentives behind the need to speak or if these MEPs often say things disliked by their national selectorate.

This analysis is merely exploratory, and further research should examine more thoroughly, both in theoretical and in empirical terms, how accountability patterns vary with such factors as ballot structure, citizens’ preference-voting behavior, and party type. Our data either provide the relevant measures to do so or make it easy to add such information from other sources.

Conclusion

The growing salience of the EP in EU decision-making and in the domestic politics of its member states are paralleled by an expanding scholarship focusing on its position and role compared with other EU institutions as well as on its members’ behavior and relationships with their parties and voters. A wide range of this literature implicitly or explicitly adheres to assumptions and hypotheses related to the electoral environment in which MEPs and national delegations operate, which could not be tested previously due to the absence of data.

The purpose of this paper is to introduce a new comprehensive data collection on EP elections. This collection complements existing datasets and provides information useful

for a broad range of subfields, encompassing voting behavior, political representation, and legislative politics. As an in-built feature, the datasets can easily be linked both within the collection and to external sources, and they are amenable to future extensions. We are planning to add information from the 2019 election once we receive feedback from end users on the 1999–2014 data.

We also sketch a research agenda on how electoral incentives shape behavior in the EP and its corollary scenario: voters rewarding legislative efforts. The data should enable researchers to examine the supply side of European elections from a longitudinal perspective. This would help re-evaluate the often-mentioned claim of a missing electoral connection. We hope that this data collection will facilitate a large number of comparative analyses and raise the understanding of the functioning of the EP and EU politics more generally.

Acknowledgements

The data collection endeavor described in this paper would not have been possible without the diligent research assistance by Ertan Bat. We are grateful to him. We also would like to thank all the people who helped with country-specific questions and other queries. These include Patrick Dumont, Kostas Gemenis, George Hajipavli, Bjørn Høyland, Raimondas Ibenskas, Tom Louwerse, Julien Navarro and Tevfik Murat Yildirim. We apologize to anybody we have forgotten to mention.

Author contributions

The order of the authors reflects the contributions made to the project.

Declaration of conflicting interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Mihail Chiru's work was supported by a grant of the Romanian Ministry of Education and Research, CNCS - UEFISCDI, project number PN-III-P4-ID-PCE-2021-233, within PNCDI III.

ORCID iDs

Thomas Däubler  <https://orcid.org/0000-0001-5106-4918>

Silje SL Hermansen  <https://orcid.org/0000-0002-1919-4365>

Supplemental material

Supplemental material for this article is available online.

Notes

1. For example, Frech (2016) and Lo (2013) focus on just one and Navarro (2010) on three member states. Hermansen (2018) studies countries with one type of ballot structure, closed

- lists. Other studies have broader country coverage but consider just one election (Meserve et al., 2015; Wilson et al., 2016).
2. The data are available at <https://doi.org/10.7910/DVN/GNRMTO>.
 3. Some related literature branches cover MEPs' career trajectories (Dodeigne, 2014; Whitaker, 2014) and the determinants of EP committee membership and roles (Chiru, 2020; Daniel and Thierse, 2018; Treib and Schlipphak, 2019).
 4. This part was also assisted by the data on MEPs' decisions to rerun available from Daniel (2015).
 5. The rerunning rates are similar across terms (MEPs from EP 5: 61.8%; from EP 6: 60.2%; from EP 7: 64.5%), and so are the re-election rates among those who stand again (MEPs from EP 5: 63.6%; from EP 6: 66.6%; from EP 7: 64.8%).
 6. Some countries such as Belgium recruit successor MEPs from a separate list if candidates elected from the main list do not take up or at a later point resign their mandate.
 7. The sum of candidate votes divided by the total party votes is not directly comparable across systems that vary with regard to the number of admissible preference votes.
 8. Approximately 90% of the lists in the relevant set actually won four or fewer seats (the larger countries of Italy and Poland use subnational districts).
 9. Hermansen (2018) finds that MEPs who authored more reports are more likely to obtain safe list positions (in closed list systems). Drawing on data from the 7th and 8th terms of the EP, De Connick (2021) concludes that MEPs elected from party-centered systems have a higher chance of being renominated. According to Wilson et al. (2016), reports and parliamentary attendance contribute to renomination per se but not to subsequent re-election. Navarro (2010) reports that the number of reports and parliamentary questions predicts overall re-election in Belgium, France and Luxembourg, but not consistently so in the 2004 and 2009 elections. Parliamentary work also has no consistent positive association with the quality of the list position MEPs obtain in Germany (Frech, 2016).
 10. The parliamentary activity data come from Daniel (2015), Høyland et al. (2019), Hurka and Kaeding (2012), and Hurka et al. (2015). We initially normalize the counts by mandate duration, for robustness considering only MEPs whose mandate spans at least a quarter of the full term.

References

- André A and Depauw S (2014) District magnitude and the personal vote. *Electoral Studies* 35: 102–114.
- André A, Depauw S and Shugart MS (2014) *The effect of electoral institutions on legislative behavior*. In: Martin S, Saalfeld T and Strøm K (eds) *The Oxford Handbook of Legislative Studies*. Oxford: Oxford University Press, 231–249.
- André A, Depauw S and Martin S (2015) Electoral systems and legislators' constituency effort: The mediating effect of electoral vulnerability. *Comparative Political Studies* 48(4): 464–496.
- Bowler S and Farrell DM (2011) Electoral institutions and campaigning in comparative perspective: Electioneering in European Parliament elections. *European Journal of Political Research* 50(5): 668–688.
- Brack N and Costa O (2019) Parliamentary questions and representation of territorial interests in the EP. In: Costa O (eds) *The European Parliament in Times of EU Crisis*. Basingstoke: Palgrave Macmillan, 225–254.
- Carey JM and Shugart MS (1995) Incentives to cultivate a personal vote. A rank-ordering of electoral formulas. *Electoral Studies* 14(4): 417–439.

- Chiru M (2020) Loyal soldiers or seasoned leaders? The selection of committee chairs in the European Parliament. *Journal of European Public Policy* 27(4): 612–629.
- Christensen HS, La Rosa MS and Grönlund K (2020) How candidate characteristics affect favorability in European Parliament elections: Evidence from a conjoint experiment in Finland. *European Union Politics* 21(3): 519–540.
- Crisp BF, Jensen KM and Shomer Y (2007) Magnitude and vote seeking. *Electoral Studies* 26(4): 727–734.
- Daniel W (2015) *Career Behaviour and the European Parliament: All Roads Lead Through Brussels?* Oxford: Oxford University Press.
- Daniel W and Thierse S (2018) Individual determinants for the selection of group coordinators in the European Parliament. *Journal of Common Market Studies* 56(4): 939–954.
- Däubler T and Hix S (2018) Ballot structure list flexibility and policy representation. *Journal of European Public Policy* 25(12): 1798–1816.
- De Connick I, 2021. *Rewarding Parliamentary Work? The effect of MEP performance on re(s)election*. PhD dissertation. KU Leuven.
- De Vreese C and van der Brug W (2016) Conclusion. In: van der Brug W and De Vreese C (eds) *(Un)Intended Consequences of European Parliamentary Elections*. Oxford: Oxford University Press, 277–288.
- Dodeigne J (2014) (Re-)assessing career patterns in multi-level systems: Insights from Wallonia in Belgium. *Regional & Federal Studies* 24(2): 151–171.
- Dodeigne J and Pilet J-B (2021) Centralized or decentralized personalization? Measuring intra-party competition in open and flexible list PR systems. *Party Politics* 27(2): 234–245.
- Döring H and Manow P (2020) Parliaments and governments database (ParlGov): Information on parties, elections and cabinets in modern democracies. Development version. Available at parl-gov.org (accessed 17 September 2021).
- Döring H and Regel S (2019) Party facts: A database of political parties worldwide. *Party Politics* 25(2): 97–109.
- European Parliament (1999) European Parliament elections 1999. Names of candidates. Available at: <https://www.europarl.europa.eu/election/listcand/cceu01.htm> (accessed 17 September 2021).
- European Parliament (2004) European elections 10–13 June. Sitemap with links to candidate lists. Available at: <https://www.europarl.europa.eu/elections2004/ep-election/sites/en/sitemap/> (accessed 17 September 2021).
- Farrell DM and Scully R (2007) *Representing Europe's citizens? Electoral Institutions and the Failure of Parliamentary Representation*. Oxford: Oxford University Press.
- Farrell DM and Scully R (2010) The European Parliament: One parliament, several modes of political representation on the ground? *Journal of European Public Policy* 17(1): 36–54.
- Frech E (2016) Re-electing MEPs: The factors determining re-election probabilities. *European Union Politics* 17(1): 69–90.
- Gattermann K and De Vreese CH (2017) The role of candidate evaluations in the 2014 European Parliament elections: Towards the personalization of voting behaviour? *European Union Politics* 18(3): 447–468.
- Gattermann K and de Vreese CH (2022) Understanding leader evaluations in European Parliament elections. *European Union Politics*. 23(1). DOI:10.1177/14651165211036263
- Göbel S and Munzert S (2021) The comparative legislators database. *British Journal of Political Science*. Epub ahead of print 23 February 2021. DOI:10.1017/S0007123420000897
- Hermansen SSL (2018) Striving for influence: The effect of performance on candidate selection. *Party Politics* 24(4): 38–396.

- Hix S (2004) Electoral institutions and legislative behavior: Explaining voting defection in the European Parliament. *World Politics* 56(2): 194–223.
- Hix S and Hagemann S (2009) Could changing the electoral rules fix European Parliament elections? *Politique européenne* 28(2): 37–52.
- Hix S and Høyland B (2013) Empowerment of the European Parliament. *Annual Review of Political Science* 16: 171–189.
- Hix S and Høyland B (2014) Political behavior in the European Parliament. In: Martin S, Saalfeld T and Strøm K (eds) *The Oxford Handbook of Legislative Studies*. Oxford: Oxford University Press, 592–605.
- Hobolt SB and Høyland B (2011) Selection and sanctioning in European Parliamentary elections. *British Journal of Political Science* 41(3): 477–498.
- Høyland B, Hobolt SB and Hix S (2019) Career ambitions and legislative participation: The moderating effect of electoral institutions. *British Journal of Political Science* 49(2): 491–512.
- Høyland B, Sircar I and Hix S (2009) An automated database of the European Parliament. *European Union Politics* 10(1): 143–152.
- Hug S (2016) Party pressure in the European Parliament. *European Union Politics* 17(2): 201–218.
- Hurka S and Kaeding M (2012) Report allocation in the European Parliament after Eastern enlargement. *Journal of European Public Policy* 19(4): 512–529.
- Hurka S, Kaeding M and Obholzer L (2015) Learning on the job? EU enlargement and the assignment of (shadow) rapporteurships in the European Parliament. *JCMS: Journal of Common Market Studies* 53(6): 1230–1247.
- Ibenskas R and Bunea A (2021) Legislators organizations and ties: Understanding interest group recognition in the European Parliament. *European Journal of Political Research* 60(3): 560–582.
- Judge D and Earnshaw D (2008) *The European Parliament*. Basingstoke: Palgrave Macmillan.
- Keyes O, Signorelli S, Graul C, et al. (2021) WikidataR: API client library for ‘Wikidata’. R package version 2.3.1. available at: <https://CRAN.R-project.org/package=WikidataR> (accessed 17 September 2021).
- Klüver H and Spoon J-J (2015) Bringing salience back in: Explaining voting defection in the European Parliament. *Party Politics* 21(4): 553–564.
- Lahti L, Biecek P, Kainu M, et al. (2017) Retrieval and analysis of Eurostat open data with the eurostat package. *R Journal* 9(1): 385–392.
- Lo J (2013) An electoral connection in European Parliament voting. *Legislative Studies Quarterly* 38(4): 439–460.
- Lorenzo Rodríguez J and Garmendia Madariaga A (2016) Going public against institutional constraints? Analyzing the online presence intensity of 2014 European Parliament election candidates. *European Union Politics* 17(2): 303–323.
- Meserve SA, Pemstein D and Bernhard WT (2015) Brussels Bound: Policy experience and candidate selection in European elections. *Comparative Political Studies* 48(11): 1421–1453.
- Mühlböck M (2012) National versus European: Party control over members of the European Parliament. *West European Politics* 35(3): 607–631.
- Navarro J (2010) Le travail parlementaire, un investissement payant? Les élections comme évaluation rétrospective du bilan des députés sortants. *Revue internationale de politique comparée* 17(4): 141–160.
- Navarro J (2016) France. In: Viola DM (eds) *Routledge Handbook of European Elections*. London: Routledge, 51–75.
- Obholzer L and Daniel WT (2016) An online electoral connection? How electoral systems condition representatives’ social media use. *European Union Politics* 17(3): 387–407.

- Reif K and Schmitt H (1980) Nine second-order national elections—a conceptual framework for the analysis of European election results. *European Journal of Political Research* 8(1): 3–44.
- Ripoll Servent A (2018) *The European Parliament*. Basingstoke: Palgrave.
- Sorace M (2018) Legislative participation in the EU: An analysis of questions, speeches, motions and declarations in the 7th European Parliament. *European Union Politics* 19(2): 299–320.
- Sorace M (2021) Productivity-based retrospective voting: Legislative productivity and voting in the 2019 European Parliament elections. *Politics*. Epub ahead of print 3 March 2021. DOI:10.1177/2F0263395721991403
- Stöver P and Wüst A (2006) Electoral systems. In: Deloyé Y and Bruter M (eds) *Encyclopedia of European Elections*. Basingstoke: Palgrave, 109–114.
- Treib O and Schlipphak B (2019) Who gets committee leadership positions in the European Parliament? Evidence from the 2014 selection process. *European Union Politics* 20(2): 219–238.
- Van der Brug W, Gattermann K and De Vreese CH (2022) Introduction: Electoral responses to the increased contestation over European integration. The European elections of 2019 and beyond”. *European Union Politics* 23(1).
- Whitaker R (2014) Tenure turnover and careers in the European Parliament: MEPs as policy-makers. *Journal of European Public Policy* 21(10): 1509–1527.
- Wilson SL, Ringe N and van Thomme J (2016) Policy leadership and re-election in the European Parliament. *Journal of European Public Policy* 23(8): 1158–1179.
- Yordanova N (2011) The European Parliament: In need of a theory. *European Union Politics* 12(4): 597–617.