

**CONSTITUENCY-LEVEL  
ELECTIONS ARCHIVE (CLEA):  
European Union Parliamentary Elections**

**Codebook**

**Version: November 15, 2023 (20231115)**

## TABLE OF CONTENTS

INTRODUCTION.....	3
Project Description .....	3
Funding Support .....	3
Bibliographic Citation .....	3
Credits.....	4
Data Procedures.....	4
EUROPEAN UNION PARLIAMENT DESCRIPTION .....	5
DOCUMENTATION OF VARIABLES .....	6
Variable List .....	6
Variable Descriptions .....	7

# INTRODUCTION

## Project Description

The central aim of the Constituency-Level Elections Archive (CLEA) project is to produce a repository of detailed results – i.e., votes received by each candidate/party, total votes cast, number of eligible voters – at the constituency level. In this dataset, we focus our attention on elections to the European Union parliament. Our motivation is to preserve and consolidate these valuable data in one comprehensive reference resource that is publicly available at no cost. This public good is expected to be of use to a range of audiences for purposes of research, education, and policy-making and evaluation.

## Funding Support

We gratefully acknowledge the American National Science Foundation ([www.nsf.gov](http://www.nsf.gov)) and the Taiwan Foundation for Democracy for their funding of the broader CLEA project. We thank the Institute of Political Science at the University of St. Gallen and the Department of Political Science at the University of Zurich for research assistance.

## Bibliographic Citation

APA (7th edition)

Kollman, K., Hicken, A., Caramani, D., Backer, D., & Lublin, D. (2020). *Constituency-Level Elections Archive: European Union Parliamentary Elections*. [Data set and codebook]. Ann Arbor, MI: Center for Political Studies, University of Michigan [producer and distributor]. Retrieved from <http://www.electiondataarchive.org>.

MLA (8th edition)

Kollman, Ken, Allen Hicken, Daniele Caramani, David Backer, and David Lublin. *Constituency-Level Elections Archive: European Union Parliamentary Elections*. Ann Arbor, MI: Center for Political Studies, University of Michigan [producer and distributor], 2020. Web. 16 December 2020. <<http://www.electiondataarchive.org>>

Chicago (17th edition)

Kollman, Ken, Allen Hicken, Daniele Caramani, David Backer, and David Lublin. 2020. *Constituency-Level Elections Archive: European Union Parliamentary Elections*. Produced and distributed by Ann Arbor, MI: Center for Political Studies, University of Michigan, <http://www.electiondataarchive.org>.

These materials are based on work supported by the National Science Foundation (under grants SES-0617101, SES-1021915, and SES-1528262). Any opinions, findings and conclusions or recommendations expressed in these materials are those of the author(s) and do not necessarily reflect the views of the funding organizations.

## Credits

The co-directors of the CLEA project are Ken Kollman of the University of Michigan, Allen Hicken of the University of Michigan, Daniele Caramani of the University of Zurich, David Backer of the University of Maryland, and David Lublin of American University. The project manager is Julia Lippman of the Center for Political Studies, University of Michigan.

Data were contributed by Daniele Caramani.

Research assistance was provided by Anna Torell, Katie Malloure, Magali Serra Duran, Morgan Vanderleest, Nerea Gándara Guerra, Nicole Zhong, Pablo Cañete Perez, Pedro Martin, Rebeca Gonzale Antuña, Sarah Kent, and Tarek Jaziri Arjona,

## Data Procedures

In order to construct the EU dataset, we have collected constituency-level election results from a variety of data sources. Once collected, the election results were checked for errors both internally and through comparison against other sources. Any identified errors were corrected. The datasets were then transformed into a common format and merged into a single data set for analysis. Although great care has been taken to prepare the data and codebook as accurately as possible prior to release, we would appreciate receiving your comments, feedback and notifications of any remaining errors by email to: [thea-project@umich.edu](mailto:thea-project@umich.edu)

## EUROPEAN UNION PARLIAMENT DESCRIPTION

The European Economic Community (EEC), established in 1958, included six initial countries: Belgium, Germany, France, Italy, Luxembourg, and the Netherlands. As part of the EEC, the European Parliament (EP) seated representatives from the respective national parliaments. The first EP had 142 members, weighted by size of country. What was eventually called the European Community (EC) agreed in 1979 on a new system whereby Members of the European Parliament (MEPs) were to be democratically elected. The Maastricht Treaty in 1992 established the European Union (EU) as we know it today, and the EP has continued to evolve, growing in size and reallocating seats to accommodate new member countries. The EU now has 27 member countries, and the EP has 707 MEPs and a president.

Today, elections to the EP are held across the EU every five years. If a nation joins the EU mid-term (between EP elections), such as when Croatia joined in 2013, a special election is held to elect their parliamentary representatives. Countries have flexibility to choose their electoral methods, but they must be based on proportional representation (PR).

A few countries in the initial election to the EP used methods similar to their own national parliaments. The U.K, for instance, after joining used single member districts for the first EP elections. Thereafter while it was a member, though, the U.K. used PR like the other countries.

All countries except four elect representatives from a single national constituency. The four exceptions (Belgium, Ireland, Italy and Poland) have subnational constituencies that use PR rules. The countries vary in their minimum thresholds for election—for instance, some countries require 5% for a party to be represented in the EP, while others require 4%, 3%, or 1.8%. Details on these rules can be found here:

<https://www.europarl.europa.eu/factsheets/en/sheet/21/the-european-parliament-electoral-procedures>

The allocation of MEPs to each member country continues to be based on population. Regardless of population size, however, each is guaranteed at least 6 seats, and no country under the current treaties can have more than 96. The president holds approval authority over the EU budget and acts as the parliament's representative on the global stage.

For more information about the EU, visit [https://european-union.europa.eu/index\\_en](https://european-union.europa.eu/index_en)

## DOCUMENTATION OF VARIABLES

### Variable List

RELEASE	Dataset Release
ID	Election Identifier
RG	Region
CTR_N	Country Name
CTR	Country Code
YR	Election Year
MN	Election Month
SUB	Sub-National Geographical Unit
CST_N	Constituency Name
CST	Constituency Code
MAG	District Magnitude
PTY_N	Party Name
PTY	Party Code
CAN	Candidate Name
PEV1	Number Of Eligible Voters (First Round)
VOT1	Votes Cast (First Round)
VV1	Valid Votes (First Round)
IVV1	Invalid Votes (First Round)
TO1	Turnout (First Round)
CV1	Candidate Votes (First Round)
CVS1	Candidate Vote Share (First Round)
PV1	Party Votes (First Round)
PVS1	Party Vote Share (First Round)
SEAT	Seats Won

## Variable Descriptions

---

### **RELEASE**

### *Dataset release*

This variable indicates in which release the constituency-level data were added to the archive.

1. November 15, 2023

---

### **ID**      *Election Identifier*

This variable uniquely identifies each election in the archive.

---

### **RG**      *Region*

The following two regions are used in the CLEA: European Union Parliamentary Elections dataset:

Western Europe  
Eastern Europe

---

### **CTR\_N**    *Country Name*

Names of countries. The following countries are included in CLEA: European Union Parliamentary Elections dataset:

Austria	Germany	Poland
Belgium	Greece	Portugal
Bulgaria	Hungary	Romania
Croatia	Ireland	Slovakia
Cyprus	Italy	Slovenia
Czech Republic	Latvia	Spain
Denmark	Lithuania	Sweden
Estonia	Luxembourg	United Kingdom
Finland	Malta	
France	Netherlands	

---

### **CTR**      *Country Code*

Country codes developed by the UN: <https://unstats.un.org/unsd/methodology/m49/>

040 Austria	380 Italy
056 Belgium	428 Latvia
100 Bulgaria	440 Lithuania
191 Croatia	442 Luxembourg
196 Cyprus	470 Malta
203 Czech Republic	528 Netherlands
208 Denmark	616 Poland
233 Estonia	620 Portugal
246 Finland	642 Romania
250 France	703 Slovakia
276 Germany	705 Slovenia
300 Greece	724 Spain
348 Hungary	752 Sweden
372 Ireland	

---

***YR***      ***Election Year***

Year of election.

---

***MN***      ***Election Month***

Month of election

- 01. January
- 02. February
- 03. March
- 04. April
- 05. May
- 06. June
- 07. July
- 08. August
- 09. September
- 10. October
- 11. November
- 12. December

---

***SUB***      ***Sub-National Geographical Unit***

Sub-national geographical or administrative unit that is greater than constituency. This variable is given if the original data collected contain such information.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

---

***CST\_N***    ***Constituency Name***

Name of geographical area that a particular elected representative or group of elected representatives represents.

---

***CST***        ***Constituency Code***

A unique numeric code assigned to each constituency in each election in a country. In general, all constituencies in a country are sorted alphabetically, according to their names, and then assigned a constituency code. This code assignment is repeated in each election in the country. Thus, the same code may or may not belong to the same constituency across elections, depending upon whether redistricting occurs between elections or if constituency names change.

---

***MAG***        ***District Magnitude***

Number of seats allocated in a given constituency.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e. a single candidate contested the election)
- 994. Suspended Election

---

***PTY\_N***    ***Party Name***

Name of a party or electoral alliance. If possible, the official name in the original language is used. If this name is not available, the transliterated or English-translated party name is used. For more information, refer to *Appendix II: Party Codes*.

In some cases where the original data sources we collected have small parties grouped under an “*Others*” category, “others” is used. In several countries, special kinds of party groupings are used in reported election results. For instance, categories such as “*miscellaneous right-wing*” and “*regionalists and separatists*” are used in France. For those special categories, their names are used for this variable and unique numeric codes are assigned to each such category (see “PTY” below for more information about these categories and also refer to *Appendix II: Party Codes*).

Independent candidates are handled in two different ways when election results are reported. For much of the data we have collected, all independent candidates are grouped under a single category. In such cases, “*Independents*” is used. However, when each and every individual independent candidate is identified and his/her votes received are reported separately in the election returns, “*Independent*” is used instead. Different numeric codes are assigned with these different methods. See “PTY” below for more information about numeric codes and Appendix for the labels for independent candidates.

In a few countries, there are independent candidates who are affiliated with a party, but cannot officially stand under its label. As a result, they may be labeled in a manner that reflects both their independent status and their party affiliation (e.g., “*Independent Labour*” or “*Independent Greens*”). In the CLEA dataset, we treat these candidates as independents in the coding, but keep their label under the *PTY\_N*.

For a full list of political parties, see *Appendix II: Party Codes*.

---

**PTY**      **Party Code**

A unique numeric code is assigned to each party that runs a candidate in any given election. In general, political parties in a country are sorted alphabetically according to *PTY\_N* and then assigned party codes. Parties have been matched across the lower chamber and upper chamber election results. Therefore, a party in a given country that runs in both chambers receives the same party code. To differentiate political parties and the aforementioned special and residual categories (see *PTY\_N* above), the CLEA dataset uses the following coding scheme:

0001-3999. Political parties

3996. None of these candidates (in some countries, voters have the option to express disapproval for all the candidates on the ballot)

3997. Elected

3998. No against for uncontested

3999. Unknown

4000. “Others” (i.e., more than two small parties are grouped)

4001-. Special kinds of ‘others’ (see Appendix II for more information)

4998. Write-in

4999. Blank/Scattering

5001-5999. Electoral coalitions or alliances between political parties

6000. “Independents” (i.e., more than two independent candidates are grouped)

6001-. Independent 1, Independent 2, and so on (i.e., a single unaffiliated candidate), including special kinds of ‘independents’.

For a full list of political parties and their codes in each country, see *Appendix II: Party Codes*.

Each party and electoral alliance is assigned a unique numeric code that remains consistent across elections. If a party changes its name, merges with other parties or splits into separate parties, a new numeric code is given to the party that emerges as a result of such changes.

Party codes for 'other' and 'independent' are assigned to parties or unaffiliated candidates in each election. This code assignment is repeated in each election in the country. Thus, the same code does not belong to the same minor party or independent candidate across elections.

---

**CAN**     ***Candidate Name***

Name of a candidate, if available.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

---

**PEVI**     ***Number of Eligible Voters (First Round)***

The number of eligible voters in a given constituency. If there is a runoff election, this indicates the number of eligible voters in the first-round election.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

---

**VOT1**     ***Votes Cast (First Round)***

The total number of votes cast for all candidates in a given constituency. If there is a runoff election, this indicates the number of total votes cast in the first-round election.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

---

**VVI**     ***Valid Votes (First Round)***

The total number of valid votes in a constituency. If there is a runoff election, this indicates the total number of valid votes in the first-round election.

When voters have multiple votes, the total number of valid votes in a constituency may be larger than the total votes cast or even the number of eligible voters.

NOTE: In some cases, this variable was calculated from the manually added sum of party votes (*PVI*) or candidate votes (*CVI*), if the original sources we collected do not contain information on valid votes but it is confirmed that no party or candidate is missing.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

---

***IVV1***     ***Invalid Votes (First Round)***

The total number of invalid and spoiled votes in a given constituency. If there is a runoff election, it indicates the total number of invalid and spoiled votes in the first-round election.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

NOTE: The treatment of blank votes depends on how an electoral commission reports them.

- If blank votes are reported separately *and* the electoral commission includes them as valid votes, then blank votes are included in CLEA as a separate row (*pty=4999*). CLEA *ivv1* is set equal to the invalid votes reported by the electoral commission. CLEA *vv1* includes blank vote totals.
- If blank votes are reported separately and the electoral commission *does not* include them as valid votes, then blank votes are not included separately in CLEA.
  - o If the electoral commission's invalid vote total includes blank votes, then CLEA *ivv1* is set equal to the electoral commission's invalid vote total.
  - o If the electoral commission's invalid vote total does not include blank votes, then CLEA *ivv1* is equal to sum of the invalid votes and blank votes reported by the electoral commission.

---

***TO1***     ***Turnout (First Round)***

The fraction of eligible voters who vote in a given constituency. If there is a runoff election, this indicates the turnout in the first round.

NOTE: This is not a turnout rate reported in an official election report. It is our own calculation from dividing *VOT1* by *PEV1*. In some instances, *TO1* is greater than 1 when the official results report more votes cast (*VOT1*) than eligible votes (*PEV1*).

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e., a single candidate contested the election)
- 994. Suspended Election

---

**CVI**      ***Candidate Votes (First Round)***

The total number of votes received by the candidate in a given constituency. If there is a runoff election, it indicates the total number of votes received by the candidate in the first-round election. This variable is most useful if there is more than one candidate running for election from the same party in a given constituency.

NOTE: In preferential vote systems, this variable refers to the number of first-preference votes a candidate received (see description for *Preferential Vote Counts* supplemental data file). In ordinal ballot systems, this variable refers to the number of votes a candidate's party list received, plus or minus the candidate-specific votes received (see description for *Ordinal Ballot Vote Counts* supplemental data file).

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e. a single candidate contested the election)
- 994. Suspended Election

---

**CVSI**      ***Candidate Vote Share (First Round)***

The fraction of votes received by the candidate in a given constituency. If there is a runoff election, it indicates the vote share of the candidates in the first-round election. This variable is most useful if there is more than one candidate running for election from the same party in a given constituency.

NOTE: In some instances, *CVSI* is greater than or equal to 1, depending on the officially reported candidate vote totals (*CVI*).

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e. a single candidate contested the election)
- 994. Suspended Election

---

**PV1**      ***Party Votes (First Round)***

The total number of votes received by the party in a given constituency. If there is a runoff election, it indicates the number of votes received by the party in the first-round election.

NOTE: If there are more than two candidates running for election from the same party in a given constituency, the sum of the votes received by all candidates from the same parties is used. Thus, party votes for the same party are repeated as many times in the column of Party Votes as there are candidates from that party.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e. a single candidate contested the election)
- 994. Suspended Election

---

***PVSI***     ***Party Vote Share (First Round)***

The fraction of votes received by a particular party.

In some cases, this variable was calculated by dividing *PVI* by *VVI* (or if *VVI* is not available, the sum of party votes), if the original data sources we collected do not report the party vote share (also see *VVI*) or if the original data sources have obvious errors in

NOTE: As with *PVI*, when there are more than two candidates from the same party in a given constituency, the party vote share of the same party is repeated as many times in the column of Party Vote Share as there are candidates from that party. Consequently, in such instances the sum of *PVSI* for a given constituency may exceed 1.

- 990. Missing Data (information not available / category not applicable)
- 992. Uncontested Election (i.e. a single candidate contested the election)
- 994. Suspended Election

---

***SEAT***     ***Seats Won***

Either the number of seats won by a party (under PR), or whether a party won or not (under SMP or MMP).

- 990. Missing Data (information not available / category not applicable)
  - 992. Uncontested Election (i.e. a single candidate contested the election)
  - 994. Suspended Election
-