CODEBOOK:
Voters' Perceptions of Party Brands

David Fortunato*
Thiago N. Silva†
Laron K. Williams‡

December 21, 2022

Abstract

This is a codebook for the dataset generated for the project "Voters' Perceptions of Party Brands." This dataset was generated based on the compilation of several surveys (e.g., European Election Studies and Comparative Study of Electoral Systems). The data frame is a nxp matrix where n is the number of survey respondents and p is the number of parties in their respective surveys. Therefore, each row of data refers to a respondent-party within a given survey, and for each row we have information on voters' socioeconomic and demographic data, voters' self placement on a left-right political spectrum, and voters' perceptions of political parties, as described in the sections of this codebook. The dataset comprises 4,955,312 observations, across 39 countries from 1956 to 2019.

Keywords: Party Ideological Position, Electoral Behavior, Voter Perception, Survey, Party Competition.

*Associate Professor, School of Global Policy and Strategy, University of California, San Diego. Email: dfortunato@ucsd.edu.

†Assistant Professor, School of Politics and International Relations, The Australian National University. Email: thiago.silva@anu.edu.au.

‡Associate Professor, Department of Political Science, University of Missouri. Email: williamslaro@missouri.edu.

Variables

- **country**

  [String variable] Name of the country in English. See list of countries below.

Table 1: List of Countries and Frequencies

Country	Observations	Percent	Cumulative
australia	107,046	2.16	2.16
austria	51,113	1.03	3.19
belgium	220,035	4.44	7.63
canada	118,777	2.4	10.03
croatia	24,716	0.5	10.53
cyprus	12,509	0.25	10.78
czech republic	73,655	1.49	12.27
denmark	394,595	7.96	20.23
estonia	38,931	0.79	21.02
finland	91,074	1.84	22.85
france	167,945	3.39	26.24
germany	208,686	4.21	30.45
greece	108,652	2.19	32.65
hungary	55,414	1.12	33.76
iceland	85,569	1.73	35.49
ireland	117,742	2.38	37.87
israel	28,850	0.58	38.45
italy	317,317	6.4	44.85
japan	34,163	0.69	45.54
latvia	49,668	1	46.55
lithuania	42,726	0.86	47.41
luxembourg	53,841	1.09	48.49
malta	8,147	0.16	48.66
montenegro	16,227	0.33	48.99
netherlands	342,380	6.91	55.9
new zealand	160,136	3.23	59.13
northern ireland	25,719	0.52	59.65
norway	203,468	4.11	63.75
poland	93,662	1.89	65.64
portugal	127,090	2.56	68.21
romania	78,659	1.59	69.79
serbia	10,976	0.22	70.02
slovakia	53,216	1.07	71.09
slovenia	67,040	1.35	72.44
spain	670,755	13.54	85.98
sweden	385,480	7.78	93.76
switzerland	115,184	2.32	96.08
uk	175,593	3.54	99.63
united states	18,556	0.37	100
Total:	**4,955,312**	**100**	

- **oecdmember**

  [Binary variable] Country's OECD membership status:

  0 = Not a OECD member; 1 = OECD member.

OECD Membership	Freq.	Percent	Cum.
0. no oecd member	2,067,461	41.72	41.72
1. oecd member	2,440,314	49.25	90.97
. (missing value)	447,537	9.03	100.00
Total	4,955,312	100.00	

- **eumember**

  [Continuous variable] Country's European Union (EU) membership status:

  0 = Not a EU member; 1 = EU member; 2 = EU applicant.

EU Membership	Freq.	Percent	Cum.
0. no eu membership	2,917,523	58.88	58.88
1. eu membership	1,471,541	29.70	88.57
2. eu application	118,711	2.40	90.97
. (missing value)	447,537	9.03	100.00
Total	4,955,312	100.00	

- **survey**

  [String variable] Name of the survey project. See list below.

  - **cses**: The Comparative Study of Electoral Systems (1996–2019);
  - **ees**: The European Election Studies (1989–2019);
  - **es**: European Studies (1965–2019).
  - **ev**: The European Voter (1956–1998).

survey	Freq.	Percent	Cum.
cses	1,505,287	30.38	30.38
ees	1,531,993	30.92	61.29
es	1,236,129	24.95	86.24
ev	681,903	13.76	100.00
Total	4,955,312	100.00	

- **adm_year**

  [Continuous variable] Year the survey was conducted (min = 1956; max = 2020).

- **adm_month**

  [Continuous variable] Month of the year the survey was conducted.

- **election_year**

  [Continuous variable] The year the election covered by the survey was held (min = 1965; max = 2018).

- **election_month**

  [Continuous variable] The month the election covered by the survey was held (min = 1; max = 12).

- **cmp**

  [Discrete variable] The party identification code according to the Comparative Manifestos Project (CMP). It consists of five or six digits: the first two resemble the CMP country code, the third digit is the CMP party family code (see below), and the last digits are running numbers.

- **party_name**

  [String variable] Name of the party in English .

- **party_abbrev**

  [String variable] Abbreviation (acronym) of party's name.

- **party_fam**

  [String variable] Party family membership. Tentative grouping of political parties and alliances into the following party families, according to the CMP. The variable is constant over time and does not account for possible changes of party family. The variable is assigned when a party is included in the data set for the first time and is generally

not changed afterwards:

10	ECO	Ecological parties
20	LEF	Socialist or other left parties
30	SOC	Social democratic parties
40	LIB	Liberal parties
50	CHR	Christian democratic parties
60	CON	Conservative parties
70	NAT	Nationalist parties
80	AGR	Agrarian parties
90	ETH	Ethnic and reginal parties
95	SIP	Special issue parties
98	DIV	Electoral alliances of diverse origin without dominant party

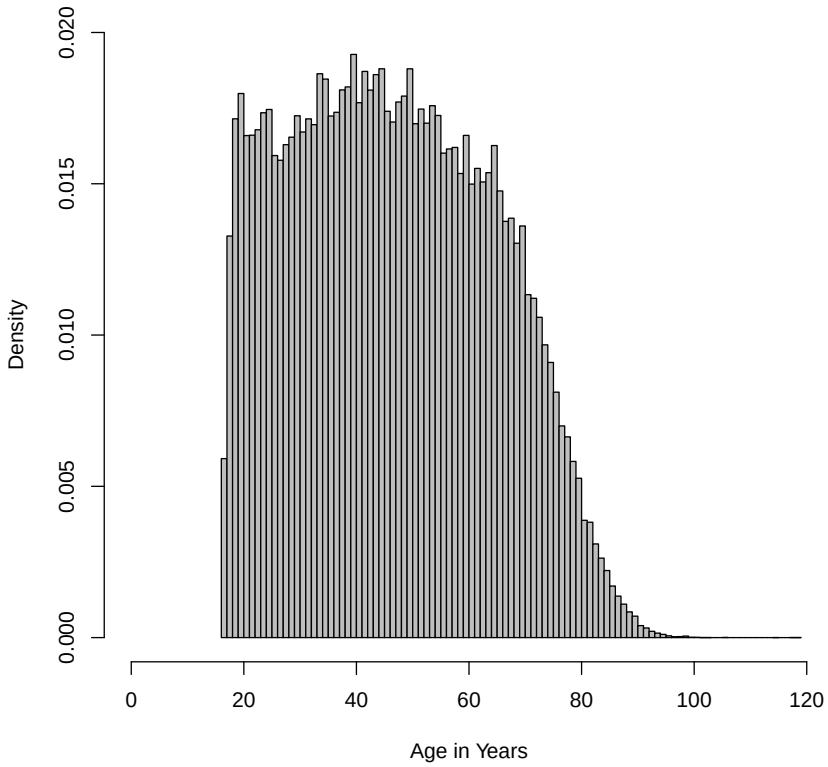
- **id**

  [Discrete variable] Unique respondent's identification by survey.

- **age**

  [Continuous variable] Age of the respondent in years (Min = 16; max = 119. Missing values: 190,983 observations).

Figure 1: Age of the Respondents in Years



- **female**

  [Binary variable] Gender of the respondent, where 0=male and 1=female.

female	Freq.	Percent	Cum.
0. male	2,389,317	48.22	48.22
1. female	2,532,509	51.11	99.32
. (missing value)	33,486	0.68	100.00
Total	4,955,312	100.00	

- **education**

  [Ordinal variable] The level of the education of the respondent. 1 = Primary Education; 2 = Secondary Education, and; 3 = Post-Secondary Education.

  Different scales across surveys:

  > The EES variable "education" in the raw data is measured as: 1 = Up to 14 years; 2 = 15 years; 3 = 16 years; 4 = 17 years; 5 = 18 years; 6 = 19 years; 7 = 20 years; 8 = 21 years; 9= 22 years or older; 10. = Still studying. We, then, rescaled the original scale based on the following ranges: *education = 1* (primary) if ees_education == 1 (up to 14 years); *education = 2* (secondary) if ees_education>=15 & ees_education<=18; *education = 3* (post-secondary) if ees_education>=19. *Note:* We have some observations from age 15 to 88 when education == 10 ["still studying"]. Then, we distributed the value of 10 (still studying) as follows: When age < 18 = 1 (primary); when age is between (inclusive) 18 and 22 = 2 (secondary); when agen > 22 = 3 (post-secondary).
  >
  > In the EES 2009, the variable education was based in the question "what age you stopped full-time education?" (variable q100 in the raw data). There are some values in which q100 is greater than the respondent's age. We kept the values as they are, because we can not know if a mistake is the respondent's age or the respondent's education answer. We just accepted the self reports at face value.

education	Freq.	Percent	Cum.
1. Primary Education	1,552,363	31.33	31.33
1. Primary Education	2,090,134	42.18	73.51
3. Post-Secondary Education	1,128,392	22.77	96.28
. (missing value)	184,423	3.72	100.00
Total	4,955,312	100.00	

- **partisan**

  [Binary variable] Variable identifying whether respondent identifies as partisan of any party, being 0 = no and 1 = yes.

  Different questions across surveys:

  > *European Election Studies (EES):* Do you feel yourself to be very close to this party, fairly close, or merely a sympathiser?
  >
  > *The Comparative Study of Electoral Systems (CSES):* Are you close to any political party? 0 = No; 1 = Yes.
  >
  > *The European Voter (EV):* Partisanship to any political party. Member of a party? 0 = No; 1 = Yes.

partisan	Freq.	Percent	Cum.
0. no identification	1,845,130	37.24	37.24
1. identify with a party	2,290,164	46.22	83.45
. (missing value).	820,018	16.55	100.00
Total	4,955,312	100.00	

- **pid**

  [Discrete variable] Party the respondent identifies with (cmp code used to identify the party).

  When party information is not provided, the following codes were used for pid:

  > 66666 = No CMP code
  >
  > 77777 = Other / Independent
  >
  > 88888 = None of the candidates or parties
  >
  > 99995 = N/A, Not Applicable
  >
  > 99996 = Did not vote
  >
  > 99997 = Not eligible/too young to vote
  >
  > 99998 = Blank-Spoilt ballot paper
  >
  > 99999 = Prefer not to say/Refuse/Don't know
  >
  > . = missing values.

Different questions for pid across surveys:

> *European Election Studies (EES):* Do you consider yourself to be close to any particular party? If so, which party do you feel close to?
>
> *The Comparative Study of Electoral Systems (CSES):* Which party do you feel closer to? From party 1 to party 65. [See Appendix A of CSES dataset for party labels.]
>
> *The European Voter (EV): Party identification.* To which party do you feel close? [See Appendix A of EV dataset for party codes in each country.]

- **voteint**

[Discrete variable] Vote choice (vote intention) of the respondent (cmp code used to identify the party).

When party information is not provided, the following codes were used for *voteint*:

> 66666 = No CMP code.
>
> 77777 = Other / Independent party.
>
> 88888 = None of the candidates or parties.
>
> 99995 = N/A, Not Applicable.
>
> 99996 = Did not vote.
>
> 99997 = Not eligible/too young to vote.
>
> 99998 = Blank-Spoilt ballot paper.
>
> 99999 = Prefer not to say/Refuse/Don't know.
>
> . = missing values in the survey.

Different questions for voteint *across surveys:*

> *European Election Studies (EES):* And if there was a general election tomorrow, which party would you vote for? [See EES Appendix "Parties" for codes]
>
> *The Comparative Study of Electoral Systems (CSES) [Not included for CSES Module 4]:* Four

different variables for vote intention:

- Current election: Vote choice - Lower house 1 (Party list). From party 1 to party 51. See Appendix A of CSES dataset for party and leader codes.
- Current election: Vote choice - Lower house 2 (District candidate). From party 1 to 21. See Appendix A of CSES dataset for party and leader codes.

The European Voter (EV): Vote choice (vote intention). [See Appendix A of EV dataset for party codes in each country.]

- **lastvote**

[Discrete variable] Vote choice of the respondent in last election (cmp code used to identify the party).

When party information is not provided, the following codes were used for *lastvote*:

66666 = No CMP code.

77777 = Other / Independent party.

88888 = None of the candidates or parties.

99995 = N/A, Not Applicable.

99996 = Did not vote.

99997 = Not eligible/too young to vote.

99998 = Blank-Spoilt ballot paper.

99999 = Prefer not to say/Refuse/Don't know.

. = missing values in the survey.

Different questions for lastvote across surveys:

> *European Election Studies (EES):* Which party did you vote for at the [General Election] of [Year of Last General Election]? [See EES Appendix "Parties" for codes].
> *The Comparative Study of Electoral Systems (CSES) [Not included for CSES Module 4]:* Four different variables for last vote:
>
> - Previous election: Vote choice - Lower house 1 (Party list). From party 1 to party 46. See Appendix A of CSES dataset for party and leader codes.
> - Previous election: Vote choice - Lower house 2 (District candidate). From party 1 to 16; and parties 20; 27; 66; 68; 69; 70. See Appendix A of CSES dataset for party and leader codes.
>
> *The European Voter (EV):* Vote choice in last election. [See Appendix A of EV dataset for party codes in each country.].

- **income**

[Ordinal variable] Income of the respondent. Standardized to make it comparable across surveys: 1 = Low; 2 = Middle; 3 = High.

Different scales across surveys:

> *European Election Studies (EES):* Variable based on respondent's social class self-placement If you were asked to chose one of these five names for your social class, which would you say you belong to: 1 = the working class, 2 = the lower middle class, 3 = the middle class, 4 = the upper middle class or, 5 = the upper class? Income = 1 if respondent's is in 1 (working class); = 2 if respondent's is between (inclusive) 2 (the lower middle class) and 4 (the upper middle class); = 3 if respondent is in 5 (the upper class).
> *The Comparative Study of Electoral Systems (CSES):* Variable based on respondent's household

income (in groups). 1. lowest household income group; 2. second lowest household income group; 3. middle household income group; 4. second highest household income group; 5. highest household income group. Income = 1 if respondent's is in 1 (lowest household income group); = 2 if respondent's is between (inclusive) 2 (second lowest household income group) and 4 (second highest household income group); = 3 if respondent is in 5 (highest household income group).

The European Voter (EV): Based on household income groups. 1. low income; 2. fairly low income; 3. neither low nor high income; 4. fairly high income; 5. high income. Income = 1 if respondent's is in 1 (low income); = 2 if respondent's is between (inclusive) 2 (fairly low income) and 4 (fairly high income); = 3 if respondent is in 5 (high income).

Electoral Studies (ES): Based on household income quartiles: 1 = lowest quartile, 2 = middle quartiles, 3 = highest quartile.

```
   income |      Freq.     Percent        Cum.
------------------+-----------------------------------
1. low            |  1,153,500       23.28       23.28
2. middle         |  2,031,406       40.99       64.27
3. high.          |    658,673       13.29       77.56
. (missing value) |  1,111,733       22.44      100.00
------------------+-----------------------------------
            Total |  4,955,312      100.00
```

- **medianincome**

  [Binary variable] Dummy variable identifying whether income is above or below the median income across respondents of the same survey, country, and year. 0 = below the median income; 1 = in the median income or above.

medianincome	Freq.	Percent	Cum.
0. below median	1,347,604	27.20	27.20
1. median or above	2,444,178	49.32	76.52
. (missing value)	1,163,530	23.48	100.00
Total	4,955,312	100.00	

- **econeval**

  [Ordinal variable] Retrospective economic evaluation. How does the respondent evaluate the country's economic condition in the day the survey was conducted in comparison to the country's economic condition 12 months ago. The scale for econeval was inverted to make the order from worse to better: 1 = Worse; 2 = Same; 3 = Better.

  Different questions across surveys:

  > *European Election Studies (EES):* What do you think about the economy? Compared to 12 months ago, do you think that the general economic situation in [country] is: 1. a lot better; 2. a little better; 3. stayed the same; 4. a little worse; 5. a lot worse.
  >
  > *The Comparative Study of Electoral Systems (CSES):* Economic evaluation. Economy improved 12 months? How do you evaluate the country's economic condition today in comparison to the country's economic condition 12 months ago? 1. very good; 2. good; 3. neither good nor bad; 4. bad; 5. very bad.
  >
  > *The European Voter (EV):* Economic evaluation. Retrospective socio-tropic evaluations. 1. much better; 2. a bit better; 3. same; 4. a bit worse; 5. much worse. [NOTE: The numbers assigned to respondents' answers don't have labels for Sweden. Thus, we need to double check the order of

> these numbers. No such variable for Denmark and Netherlands from EV survey.]

econeval	Freq.	Percent	Cum.
1. worse	1,166,449	23.54	23.54
2. same	1,009,864	20.38	43.92
3. better	594,502	12.00	55.92
. (missing value)	2,184,497	44.08	100.00
Total	4,955,312	100.00	

- **lrs**

  [Discrete data] Respondent's self-placement on a left-right ideology scale. From 0 (Left) to 10 (Right).

  When the information was not provided by the respondent, the following codes were used for *lrs*:

  > 666: No CMP code.
  >
  > 977: N/A, Not Applicable, Question Not Asked, etc.
  >
  > 987: Don't Place (or Consider) Myself as Left-Right (for Left-Right Self-Placement question); Refuses to Place Self; Refuse to Take Position.
  >
  > 988: Don't Know what Left-Right Is, Haven't Heard of Left-Right.
  >
  > 989: Haven't Heard of the Party (for Party Placement questions); Does not know party.
  >
  > 998: Refused, No Answer, No Response.
  >
  > 999: Don't Know, Don't Know Where to Place, No Opinion.
  >
  > . : Missing in the survey.

Scales for lrs per survey project:

> EES (1989-2004): From 1 (left) to 10 (right), integer.
>
> EES (2009): From 0 (left) to 10 (right), integer.
>
> EES (2014): From 0 (left) to 10 (right), integer.
>
> EES (2019): From 0 (left) to 10 (right), integer.
>
> ES (1965-2019): From 0 (left) to 10 (right), integer. With three exceptions:
>
> - ES France (1967): From 3 to 95, integer (e.g., 3, 5, 7, 8, 10, 15, 25, 27, and so on).
>
> - ES Italy (1968): From 2 to 99, integer (e.g., 2, 3, 4, 5, 6, 11, 15, 20, 21, 25, and so on).
>
> - ES Italy (1972): From 2 to 99, integer (e.g., 2, 4, 5, 10, 15, 20, 25, 30, 33, and so on).
>
> - ES UK (1983): From -10 to 10, integer.
>
> CSES I-IV (1996-2016): From 0 (left) to 10 (right), integer.
>
> CSES V (2015-2019): From 0 (left) to 10 (right), integer.
>
> EV (1956-1998): Different scales:
>
> - EV Sweden (1956-1998): From 0 (left) to 10 (right), integer.
>
> - EV Germany (1976-1998): From 1 (left) to 11 (right), integer.
>
> - EV Denmark (1971-1998): From 1 (left) to 10 (right), integer.
>
> - EV Norway (1965-1997): From 0 (left) to 10 (right), integer.
>
> - EV The Netherlands (1971-1972): From 1 (left) to 10 (right), rational numbers. Values: 1 (left), 2.5, 4, 5.5, 7, 8.5, 10 (right).
>
> - EV The Netherlands (1981-1998): From 1 (left) to 10 (right), integer.

- **lrparty**

  [Discrete data] Respondent's Left-Right placement of the party. From 0 (Left) to 10 (Right).

  When the information was not provided by the respondent, the following codes were used for *lrparty*:

  > 666: No CMP code.
  >
  > 977: N/A, Not Applicable, Question Not Asked, etc.
  >
  > 987: Don't Place (or Consider) Myself as Left-Right (for Left-Right Self-Placement question); Refuses to Place Self; Refuse to Take Position.
  >
  > 988: Don't Know what Left-Right Is, Haven't Heard of Left-Right.
  >
  > 989: Haven't Heard of the Party (for Party Placement questions); Does not know party.
  >
  > 998: Refused, No Answer, No Response.
  >
  > 999: Don't Know, Don't Know Where to Place, No Opinion.
  >
  > . : Missing in the survey.

  Scales for lrparty per survey project:

  > EES (1989-2004): From 1 (left) to 10 (right), integer.
  >
  > EES (2009): From 0 (left) to 10 (right), integer.
  >
  > EES (2014): From 1 (left) to 11 (right), integer.
  >
  > EES (2019): From 0 (left) to 10 (right), rational numbers. Values: 0, .33, .5, 1, . . . , 9, 9.33, 9.5, 10.
  >
  > ES (1965-2019): From 0 (left) to 10 (right), integer. With three exceptions:
  >
  > - ES France (1967): From 1 to 90, integer (e.g., 1, 5, 8, 10, 11, 12, 14, 15, 17, 18, 20, and so on).
  >
  > - ES Italy (1968): From 2 to 99, integer (e.g., 2, 3, 4, 5, 6, 11, 15, 20, 21, 25, and so on).
  >
  > - ES Italy (1972): From 2 to 99, integer (e.g., 2, 4, 5, 10, 15, 20, 25, 30, 33, and so on).

> - ES UK (1983): From -10 to 10, integer.
>
> CSES I-IV (1996-2016): From 0 (left) to 10 (right), integer.
>
> CSES V (2015-2019): From 0 (left) to 10 (right), integer.
>
> EV (1956-1998): Different scales:
>
> - EV Sweden (1956-1998): From 0 (left) to 10 (right), integer.
>
> - EV Germany (1976-1998): From 1 (left) to 11 (right), integer.
>
> - EV Denmark (1971-1998): From 1 (left) to 10 (right), integer.
>
> - EV Norway (1965-1997): From 0 (left) to 10 (right), integer.
>
> - EV The Netherlands (1971-1972): From 1 (left) to 10 (right), rational numbers. Values: 1 (left), 2.5, 4, 5.5, 7, 8.5, 10 (right).
>
> - EV The Netherlands (1981-1998): From 1 (left) to 10 (right).

- **weight**

  [Continuous variable] Demographic weight per country.

  Original missing values for *weight* per *survey*:

  > CSES: 288,963 missing observations (19.19%).
  >
  > EES: 76,499 missing observations (5%).
  >
  > ES: 684,892 missing observations (55.4%).
  >
  > EV: 473,452 missing observations (69.44%).

  All missing observations were recoded as "1". Therefore, for elections/surveys without weight, all observations for the variable *weight* are equal to one.