

The Comparative Study of Electoral Systems (CSES)

Module 6

3M (micro, meta, macro) Data Subcommittee report

Masahiro Yamada (*myamada@kwansei.ac.jp*)

Joshua Tucker (*jat7@nyu.edu*)

Catherine De Vries (*catherine.devries@unibocconi.it*)

December 15, 2021

Executive Summary

This report presents a review of the macro, micro, and meta data components of the Comparative Study of Electoral Systems (CSES) project in preparation for Module 6 (M6).

Based on this review we offer the following eight key recommendations:

1. Macro and district data

- (1) Collection of three new macro variables that relate to theme of M6: quality of democracy, gender.
- (2) Discontinuation of seven variables related to the theme of Module 5 (M5): country subject to IMF conditionality at election, net migration rate, population by citizenship, linguistic fractionalization index, religious fractionalization index, ethnic fractionalization index, and polity fragmentation index.
- (3) District-level data collection should be continued.

2. Micro data

- (4) Suggest to other committees to consider introducing variables on people's attitude towards climate change.
- (5) Suggest to other committees to consider introducing variables to measure a household's wealth.

3. Meta data

- (6) Ask the Secretariat to hire an expert to undertake an assessment of the weight data and a report for the user community.
- (7) Do not add a simple ethnicity variable with interviewers being identified as whether they are a member of the majority ethnic group, a minority ethnic group, or if they consider themselves of mixed ethnic descent.
- (8) Ask the Secretariat to revisit how survey mode is captured and recorded in the data.

1. Mission of the report and work of the committee

The objective of this report is to review and assess three components of Comparative Study of Electoral Systems (CSES) project. The components consist of the macro data including district data, micro data composed of answers from survey respondents, and meta data also known as ‘administrative data’ in cses.org.

Our charge from Elizabeth J. Zechmeister, the chair of the Module 6 planning committee, as the subcommittee is to review micro-level questions (e.g., demographic questions not already covered in Module 6 or core items), meta-data fields (e.g., administrative variables such as interviewer characteristics and date of interview), and macro-level content (e.g., district data, electoral system data) and suggest possible revisions, deletions, or additions for Module 6. Therefore, this report addresses these three parts: macro with district data, micro data, and meta data, in order.

2. Macro and District-Level Variables

2.1. Selection Strategy

We started to consider selection of variables based on the “Module 5 Macro and District Data Subcommittee report” (Gibson et al. 2016)¹ and “The CSES Bibliography project” (Quinlan and CSES Secretariat 2016)². Moreover, discussion in the planning committee and further communication among the member of the 3M subcommittee, the chair of the planning committee, the chair of the theme subcommittee contributed to our selection.

2.2. Recommended additional macro variables

Table 1 is a list of three new variables we recommend for inclusion in Module 6 that are relevant to the module theme. Among of them, the first two variables, “share of women in incumbent government” and “share of independent ministers in incumbent government”, require each country team to fill the form in macro report in the M6. The third variable, “share of women in parliament”, is provided at the site of Inter-Parliamentary Union.³

VARIABLE	POTENTIAL SOURCES
Share of women in incumbent government	Macro report
Share of independent ministers in incumbent government	Macro report
Share of women in parliament	IPU Parline

¹ https://cses.org/wp-content/uploads/2019/03/CSES5_MacroDistrictSubcommittee_FinalReport.pdf

² https://cses.org/wp-content/uploads/2019/04/CSES_2016Philadelphia_Bibliography.pdf

³ <https://data.ipu.org/content/parline-global-data-national-parliaments>.

Moreover, the recent global pandemic of the COVID-19 urges us to include some related variables. The CSES Secretariat is actively considering the inclusion of administrative and macro variables related to the pandemic that may be useful to users of CSES Module 5. Our subcommittee proposes that these variables from CSES Module 5 which continue to be reliable and useful also be added for CSES Module 6.

2.3. Recommended macro variables for discontinuation

Table 2 shows the list of variables that we recommend discontinuing for M6. These seven variables (E5010, E5105, E5016, E5107, E5108, E5010 in M5) were introduced as theme specific variables in M5, but we regard them as weak relevance to M6.⁴

VARIABLE	CSES M5 CODE
COUNTRY SUBJECT TO IMF CONDITIONALITY AT ELECTION	E5101
NET MIGRATION RATE - 2000-2005, 2005-2010, 2010-2015	E5105
POPULATION BY CITIZENSHIP: PERCENTAGE OF POPULATION WHO ARE CITIZENS	E5106
LINGUISTIC FRACTIONALIZATION INDEX	E5107
RELIGIOUS FRACTIONALIZATION INDEX	E5108
ETHNIC FRACTIONALIZATION INDEX	E5109
POLITY FRAGMENTATION INDEX	E5110

In the discussion in the M6 planning committee, we argue the possibilities to drop other macro variables. However, we could not make consensus which variables should be dropped on M6 or not. Therefore, our conclusion here is to exclude the seven variables in Table 2.

2.4. District Data

The M5 has district-level variables numbered as E4001 to E4007. In the discussion of the planning committee, use of the Constituency Level Elections Archive (CLEA)⁵ was proposed. The 3M subcommittee checked the CLEA dataset, and received a memorandum from David Howell (CSES) and Julia Lippman (CLEA), which is appended in this report as Appendix IV.

Table 1 in the memorandum shows that the CLEA lacks four variables (number of candidates, number of party lists, seats in district, and number of seats per part) in the CSES dataset. Although users could calculate the four variables using other variables in the CLEA dataset, this would require “some background knowledge about the countries’

⁴ We appreciate with Bojan Todosijevic’s wonderful suggestion at the plenary session on October 28, 2021

⁵ <http://www.electiondataarchive.org/>

electoral systems to know when this is appropriate.”⁶

Such background knowledge about electoral districts in each country is brought to the project by each national collaborator. In this sense, our subcommittee proposes that collecting district data by each national collaborator is desirable.

Additionally, the CLEA does not necessarily gather upper house election data as the Table 3 in the memorandum shows. This implies that dependence of the CLEA may generate works for supplementing district-level data for the upper house election in the CSES. This is another reason that the 3M subcommittee propose district data collection by national collaborators.

3. Micro data

Module 5 includes a wide range of micro-levels questions that broadly include two sets of questions:

A. Core questions that tap into political attitudes, electoral behavior and preferences as well as political knowledge and efficacy of respondents. The “core” refer to questions from the main module (questions labelled beginning with “Q” – i.e. question Q1, Q2, Q3, etc.) which should be considered to be repeated for every module of the CSES.

B. Demographic questions that tap into demographic characteristics of respondents. The demographic questions are those labelled beginning with “D” (i.e. question D1, D2, D3...).

The Module 5 Core & Demographic Subcommittee report⁷ provided an evaluation of both sets of questions. Their recommendations included:

A. Core questions:

- 1) Add question about political interest: “How interested would you say you are in politics?”
- 2) Add additional internal efficacy question: “Please tell me whether you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree with each of the following statements: You feel you understand the most important political issues of this country.”
- 3) Drop one external efficacy question: “Please tell me whether you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree with each of the following statements: Who is in power can make a difference.” *This recommendation was reversed at the plenary.*
- 4) Add a media exposure question: “I WOULD LIKE TO ASK YOU how closely you follow political news, whether on TV, radio, newspapers or internet. Please tell me if you

⁶ Julia’s reply on September 24, 2021 to questions via e-mail.

⁷ https://cses.org/wp-content/uploads/2019/03/CSES5_CoreDemographicsSubcommittee_FinalReport.pdf

follow this news very closely, fairly closely, not too closely, or not at all closely?”

5) Add retrospective government evaluation question: “Now thinking about the performance of the [government in [CAPITAL]/president] in general, how good or bad a job do you think the [government/president in [CAPITAL]] has done over the past [NUMBER OF YEARS SINCE LAST GOVERNMENT TOOK OFFICE, BEFORE THE CURRENT ELECTION] years?”

B. Demographic questions:

- 1) Drop the question of level of religiosity (D23).
- 2) Drop organizational membership questions, with the exception of trade union membership.
- 3) Drop the set of questions (employment status, type, occupation, SES) about respondents’ spouses.
- 4) Add a simple question of whether respondents’ parents are born outside of country, as a follow up to the questions of respondents’ country of birth to allow for the identification of 1st and 2nd generation immigrants.

The recommendations of Module 5 Core & Demographic Subcommittee report were largely based on an analysis of the actual use of core and demographic variables in publications based on previous rounds of the CSES. When it comes to core questions, according to the presentation by Stephen Quinlan at the CSES Plenary Session in Philadelphia in August 2016, left/right ideological positioning, party ID, political efficacy and vote choice were the mostly used questions by researchers (Quinlan 2016, slide 16). When it comes to demographic questions, education, age, gender, household income, group membership, and urban/rural were the most frequently used questions (Quinlan 2016, slide 15). Since 2016, 48 new articles were published using the CSES data and these virtually all include the micro data (for a list of the articles see the appendix). Given the frequent use of the micro data, consistency in questions is an important criterion. When reviewing the core results of the articles, it is interesting to note that most of the articles include left/right ideological positioning, party ID, political efficacy, vote choice or satisfaction with democracy as their dependent variables (as highlighted in Quinlan 2016), but the battery of elite questions is also used quite frequently as indicators of populist sentiment. In addition, when it comes to demographics there has been considerable interest in socio-economic characteristics of individuals in the crisis/post-crisis period.

When it comes to the core questions, the questionnaire already covers a lot of ground and questions could be retained in order to guarantee a time series. Yet, one thing that seems blatantly missing from the core questionnaire is any question relating to people’s attitudes towards climate change. Given the importance of this issue globally, this seems like an omission. Perhaps considering to add a question tapping into this could be worthwhile. The American National Election Study for example includes the following question that could be used:

“Some people think the federal government needs to regulate business to protect the environment. They think that efforts to protect the environment will also create jobs. Let us say this is point 1 on a 1-7 scale. Others think that the federal government should not regulate business to protect the environment. They think this regulation will not do much

to help the environment and will cost us jobs. Let us say this is point 7 on a 1-7 scale. And of course, some other people have opinions somewhere in between, at points 2,3,4,5, or 6.”

When it comes to the demographic questions, the questionnaire includes the core sets of demographics used by scholars of electoral behaviour, participation and political attitudes. Given that about a third of articles use household income as a key demographic variable, it may be worthwhile considering additional aspects that tap into a household’s wealth, such as assets or savings for example. Household income might not be a sufficient enough indicator to capture a household’s socio-economic status. Wealth inequality has increased in many countries across the globe and across generations. As a result, questions could be included to capture a household’s wealth more directly by capturing assets, debt or savings. One way to make space for this is to take out spouse related questions.

4. Meta data⁸

Overview of Administrative Variables

The CSES’s administrative variables can be divided up into four broad categories: “Identifier”, “Weights”, “Election Logistics”, and “Study Logistics” variables.

Identifiers

Identifier variables describe the dataset and data so that researchers can identify the dataset or subset of the dataset that is pertinent to their research questions. Identifier variables for example include country codes to identify which observations are from a given country, election year codes to identify observations in a given time period, or respondent id numbers to identify respondents within an election study.

Some identifier variables are redundant. For example, there are currently five different variables that make it possible to identify which country a given study was fielded in: Polity CSES Code, Polity UN Code, Polity Name, Election Study (Numeric Polity), and Election Study (Alphabetic Polity). Election Study (Numeric Polity) and Election Study (Alphabetic Polity) are constructed by combining country code and election year identifiers.

Since it is easy and fast to construct these variables, removing them from the study would not improve efficiency particularly. Furthermore, their inclusion makes it easy for researchers to work with multiple waves of the study or to combine the CSES data with other data sources. For these reasons, it should not be a priority to remove these from the study.

⁸ We are very grateful to Tine Paulsen (<https://www.tinepaulsen.com/>) for assistance in preparing this section of the report.

Weights

The CSES includes a variety of weight variables divided up into sample, demographic, and political weights. Sample weights are included so researchers can correct for unequal selection probabilities, while demographic weights ensure that researchers can weigh the data, so the sample's socio-demographic characteristics more closely resemble the population's. Finally, political weights provide the researcher with a tool to correct for the sample's reported voting behavior not fitting with the official vote count.

There are three different types: Original Weights, Polity Weights, and Dataset Weights. The Original Sample Weights are the weights reported by the individual election studies. Polity Weights are standardized so that the original weights have a mean of 1 within a given country. Dataset Weights are calculated so that researchers can use all observations in the CSES regardless of country and can correct for not all sample sizes being the same across election studies.

The Weights category also includes certain variables that are used to construct the weights. "Factor: Sample Size Adjustment", for example, reports the ratio of the average sample size to the given election study's sample size.

The Weights variables are fairly technical. Even though researchers could theoretically construct all the weight variables from the set of Original Weights, this might be quite demanding and result in calculation errors. Therefore, removing them could impact the quality of research being done using the CSES. This negative effect would especially hit research using cross-country comparisons since the Dataset Weights are the hardest to calculate.

Election Logistics

This is the smallest category of administrative variables. The variables describe the type of election taking place and the timing of the election.

The timing variables are divided up in "day", "month", and "year" variables, which is fairly inefficient. However, this division avoids possible standardization mistakes that could happen when multiple election study teams differ in their election timing descriptions.

Study Logistics

The last category of administrative variables describes how the specific election study was administered. This set of variables allows a researcher to see whether the study was conducted by telephone, how long an interview lasted, and whether the study was conducted during or after an electoral period.

These variables are important for researchers who use the CSES for cross-country comparisons. If a researcher wants to compare two different countries' responses to the same question, it could for example be important that both of the country's samples were asked the particular question over the phone.

Comparing Administrative Variables to Other Studies

The CSES's number of identifier variables is above average for the studies surveyed. This seems to be because the CSES wants to make it easy to use the included studies for both national analysis and cross-national analysis. Other cross-national studies, with the exception of the World Values Survey and LAPOP's AmericasBarometer, do for example not include a unique "study id" variable that combines the year of study and the country the study took place in. Instead, the logic seems to be that researchers can construct these types of variables on their own. This obviously increases the likelihood of a researcher making a mistake and not analyzing the exact observations they are interested in. The CSES does not include a respondent's region of residence or electoral district of residence under its list of administrative variables. However, these smaller geographic units are included as ordinary variables in the CSES, which means that researchers have a chance to divide respondents up in smaller geographical units for analysis.

The CSES's number and diversity of weight variables is large compared to all other studies surveyed. However, the weight variables make it possible to use the national studies in their original form and also use them for cross-country analyses. Part of the reason the CSES's weight variables are so numerous is that some of the other studies surveyed are not cross-national. Therefore, the set of weighting variables that are constructed to make sure that individual observations in the CSES can be studied across countries are not present in these other studies. However, even compared to the cross-national studies, the CSES's number of weight variables is large. Most cross-national studies do not include the original country weights and the intermediate variables that transform them to cross-national weights. Therefore, it is possible to use the CSES dataset to extract single studies in their original form or to use a multitude of studies for cross-national analysis.

National electoral studies unsurprisingly tend to have less information about the elections they are studying than the CSES. The set of election logistics variables makes it clear that the CSES is trying to help researchers who compare certain features of political systems to do so without having to do much outside research. This ensures that an inexperienced researcher does not accidentally include studies that focus on a type of election the researcher is not interested in.

The study logistics set of variables are designed so researchers can identify national differences in the way of conducting interviews that could lead to systematically different respondent responses. A lot of other cross-national studies are missing this component. An alternative strategy would be to make sure that all studies included in the CSES conduct their studies in the same way. However, this would mean that some country studies would be forced to change their study logistics. This is not a viable strategy since

such a change would make it impossible to ensure consistency in the responses to questions over time. Finally, some studies that use face-to-face interviews, the ANES and LAPOP's AmericasBarometer, include a battery of interviewer demographics variables. These are included out of a concern for an interviewer's ethnicity or gender affecting the respondent's answers to certain questions. The CSES includes the interviewer's gender but does not include other demographic variables.

Table 3. CSES Administrative Variables in Comparison to Other Studies

	CSES	ANES	BES	WVS	ESS	LAPOP	AfroBarometer	EBRD Life in Transitions
	Identifier Variables							
Dataset ID	Yes	No	No	No	Yes	Yes	No	No
Dataset Version	Yes	Yes	No	No	Yes	Yes	No	No
Respondent ID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Study ID	Yes	No	No	Yes	No	Yes	No	No
Country ID	Yes	No	No	Yes	Yes	Yes	Yes	Yes
Country subset ID	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Subsample ID	Yes	No	No	No	No	No	No	No
Year ID	Yes	No	No	Yes	No	Yes	No	Yes
	Weights							
Study Sample W	Yes	No	No	No	No	Yes	No	Yes
Study Demographic W	Yes	No	No	No	No	No	No	Yes
Study Political W	Yes	No	No	No	No	No	No	No
Mean of S Sample W	Yes	No	No	No	No	No	No	No
Mean of S Demographic W	Yes	No	No	No	No	No	No	No
Mean of S Political W	Yes	No	No	No	No	No	No	No
Standardized S Sample W	Yes	No	No	No	No	No	No	No
Standardized S Demographic W	Yes	No	No	No	No	No	No	No
Standardized S Political W	Yes	No	No	No	No	No	No	No
Sample Size adjustment	Yes	No	No	Yes	No	Yes	No	Yes
Dataset Sample W	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Dataset Demographic W	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Dataset Political W	Yes	Yes	Yes	No	No	No	No	No

	Election Logistics							
Election Type	Green	Green	Red	Red	Red	Red	Red	Red
Election Timing	Green	Red	Red	Red	Red	Red	Red	Red
	Study Logistics							
Study Timing	Green	Red	Red	Red	Red	Red	Red	Red
Study Context	Green	Red	Red	Red	Red	Red	Red	Red
Interview Mode	Green	Green	Red	Red	Red	Red	Red	Red
Self-Selection into Interview Mode	Green	Red	Red	Red	Red	Red	Red	Red
Duration of Interview	Green	Green	Red	Red	Green	Red	Green	Red
Interviewer ID	Green	Red	Red	Red	Red	Red	Green	Red
Interviewer Gender	Green	Green	Red	Red	Red	Red	Green	Red
Interviewer Demographics (other)	Red	Green	Red	Red	Red	Red	Green	Red
Interview Timing	Green	Green	Green	Green	Green	Green	Green	Green
Fieldwork Timing	Green	Red	Red	Red	Red	Red	Red	Red
Fieldwork Duration	Green	Red	Red	Red	Red	Red	Red	Red
Language of Interview	Green	Green	Red	Green	Red	Green	Green	Red

Green = included; Red = not included; Data collected by Tine Paulsen and Joshua Tucker.

Recommendations

Our survey of other election studies suggests that -- far from lacking administrative data found on other studies -- CSES is actually at the forefront of providing researchers with valuable administrative data. We therefore have only three recommendations:

- 1) If there is a belief that the administrative data requirements are proving too onerous for national election teams, it is certainly the case that CSES provides many, many more potential weights than different surveys. That being said, we find good reason for these weights, and think that by and large they provide useful tools for researchers that increase the value of the data. Nevertheless, it is possible that some of these weights are rarely used. Therefore, if there was to be a CSES wide survey of existing literature to check for usage of different variables or types of data, it might be worth including in that review a measure of how often different weights are applied, with an eye towards possibly not including certain weights if they are never used. To be clear, this is a second best

option to continuing to include the weights, and one which we think should only be pursued if there is a mandate to *reduce* administrative data.

Because there are a number of weight measures, and these have not been assessed in a comprehensive fashion within the project, the 3M subcommittee therefore recommends that the Secretariat pay for a methodological expert to review the weights and write a report for the Planning Committee that lays out how and when the different weights should be used. The subcommittee will then write a blog post announcing the report to the user community and summarizing some of the key recommendations. Thus unless the methodological expert recommends cutting any of the weights, this is not a proposal to change any of the provided weights, but rather simply to provide more guidance to the user community as to the contexts in which the different weights should be applied.

- 2) The only category of administrative data that we could find for which CSES is lacking information provided by a (limited) number of other studies is in terms of interviewer demographic characteristics. In an era where there is greater attention to both identity politics and transparency surrounding data collection, more information regarding the identity of the interviewer might be useful. This is of course complicated by the lack of comparable identity characteristics across countries. However, we would like to suggest that perhaps a simple ethnicity variable could be included, with interviewers being identified as whether they are a member of the majority ethnic group, a minority ethnic group, or if they consider themselves of mixed ethnic descent. This seems like something that could be (a) simple to collect without being (b) overly intrusive and yet (c) provide value to researchers worried about interviewer effects on certain types of questions. The Board could also consider other types of interviewer characteristics (e.g., region of the country, primary language), but we suspect the most important would be ethnicity.

After further discussion with the planning committee, it was decided not to recommend an interviewer ethnicity variable due to the fact that interpretation would differ too extensively across different country context and might have a negative impact on ability to recruit those carrying out the surveys.

- 3) It was not clear to us from the documentation provided how exactly the interviewer mode variable was coded (is that from a list of categories? open ended questions?), but we would suggest that as survey research migrates online, if there are categories that are provided as part of this question, it might be worth revisiting them before Wave 6 to make sure they accurately reflect current modes of survey research.

The 3M committee therefore recommends that the CSES Secretariat to propose a revision to better capture the list of options for mode, including a "select all that apply" format for the question given the increased prevalence of mixed-mode studies.

We would also note that the question regarding the gender of the interviewer should probably reflect whatever updates we make in terms of thinking about how to question survey respondents regarding gender identity (i.e., do we want to add a non-binary category?).

Appendix II: Current Metadata (Administrative Data)

Name	Category	Website Description	Codebook Description (Shortened)
Dataset	Identifier	Dataset	This variable reports the CSES module applied in each election study.
Dataset Version	Identifier	Dataset version	The version number corresponds to the date of the dataset's release.
Digital Object Identifier	Identifier	Digital Object Identifier	This variable indicates the Digital Object Identifier (DOI) which is registered for the dataset. Each CSES dataset version has a unique, persistent DOI.
ID Variable - Election Study (Numeric Polity)	Identifier	Election Study Identifier: Numeric Polity Code and Election Year.	This eight digit variable uniquely identifies an election study within the CSES. The variable is constructed from two components 'CSES polity code' and 'election year'
ID Variable - Election Study (Alphabetic Polity)	Identifier	Election Study Identifier: Alphabetic Polity Code and Election Year.	This eight-character variable uniquely identifies an election study within the CSES. The variable is constructed from two components 'CSES polity code' and 'election year'.
ID Variable - Respondent	Identifier	Respondent Identifier.	This eighteen-character variable uniquely identifies a respondent within the CSES data file. The variable is constructed from three components: 'CSES polity code', 'election year', and 'respondent within election study'.
ID Component - Polity CSES Code	Identifier	Polity Identifier.	This four-character variable uniquely identifies a polity conducting an election study that is present in the CSES.
ID Component Polity - UN Code	Identifier	Polity Identifier UN Country Code.	This three-character variable uniquely identifies a polity conducting an election study that is present in the CSES.
ID Component - Polity - Name	Identifier	Polity Identifier Country Name.	This variable uniquely identifies a polity conducting an election study that is present in the CSES.
ID Component - Sample Component	Identifier	In some cases, analysts may wish to consider regions of countries or other sample components units of analysis, rather than the countries themselves. We	In some cases, analysts may wish to consider regions of countries or other sample components units of analysis, rather than the countries themselves. We use this variable to capture information about subsets of respondents that are meaningful but that are not

		use this variable to capture information about subsets of respondents that are meaningful but that are not captured by other variables. This may, for instance, refer to different sample components or respondents from different panel components. For all other cases, this variable is coded 001.	captured by other variables. This may, for instance, refer to different sample components or respondents from different panel components. For all other cases, this variable is coded 001.
ID Component - Election Year	Identifier	Election year.	Election year.
ID Component - Respondent within Election Study	Identifier	Respondent identifier.	This variable is ten characters in length. It is unique for each survey respondent within an election. While this variable uniquely identifies a respondent within an election study, it is not unique across the entire dataset.
Original Weight: Sample	Weights	Original Weight: Sample	These variables report the original weights provided with the respective deposited data files. Sample weights include those intended to correct for unequal selection probabilities resulting from "booster" samples, procedures for selection within the household, non-response, as well as other features of the sample design.
Original Weight: Demographic	Weights	Original Weight: Demographic	These variables report the original weights provided with the respective deposited data files. Demographic weights adjust sample distributions of socio-demographic characteristics to more closely resemble the characteristics of the population.
Original Weight: Political	Weights	Original Weight: Political	These variables report the original weights provided with the respective deposited data files. Political weights reconcile discrepancies in the reported electoral behavior of the survey respondents from the official vote counts.
Factor: Mean of Sample Weight	Weights	Factor: Mean of Sample Weight	These variables report the mean weight of each type, within each polity (election study). The resulting factors are then used to create the derivative "Polity Weights"
Factor: Mean of Demographic Weight	Weights	Factor: Mean of Demographic Weight	These variables report the mean weight of each type, within each polity (election study). The resulting factors are then used to create the derivative "Polity Weights"

Factor: Mean of Political Weight	Weights	Factor: Mean of Political Weight	These variables report the mean weight of each type, within each polity (election study). The resulting factors are then used to create the derivative "Polity Weights"
Polity Weight: Sample	Weights	Polity Weight: Sample	These variables report standardized versions (with a mean 1 within the polity) of the original weights provided with the component election studies.
Polity Weight: Demographic	Weights	Polity Weight: Demographic	These variables report standardized versions (with a mean 1 within the polity) of the original weights provided with the component election studies.
Polity Weight: Political	Weights	Polity Weight: Political	These variables report standardized versions (with a mean 1 within the polity) of the original weights provided with the component election studies.
Factor: Sample Size Adjustment	Weights	Factor: Sample Size Adjustment	This variable reports the ratio of the average sample size to each election study sample.
Dataset Weight: Sample	Weights	Dataset Weight: Sample	These variables are intended for micro-level analyses involving the entire CSES sample. Using the sample size adjustment, the standardized weights are corrected such that each election study component contributes equally to the analysis, regardless of the original sample size.
Dataset Weight: Demographic	Weights	Dataset Weight: Demographic	These variables are intended for micro-level analyses involving the entire CSES sample. Using the sample size adjustment, the standardized weights are corrected such that each election study component contributes equally to the analysis, regardless of the original sample size.
Dataset Weight: Political	Weights	Dataset Weight: Political	These variables are intended for micro-level analyses involving the entire CSES sample. Using the sample size adjustment, the standardized weights are corrected such that each election study component contributes equally to the analysis, regardless of the original sample size.
Election Type	Election Logistics	Type of Election	Type of election. Eg. Parliamentary/Legislative or Head of Government
Date First	Election	Date [first round/second	Date [first round/second round] election began.

Round Election Began	Logistics	round] election began.	[Separate Variables for Month Day Year]
Date Second Round Election Began	Election Logistics	Date [first round/second round] election began.	Date [first round/second round] election began. [Separate Variables for Month Day Year]
Study Timing	Study logistics	Timing of study relative to election.	Timing of study relative to election. Eg. Post-Election Study, Between Rounds
Study Context	Study logistics	Study context in which CSES module was conducted.	Study context in which CSES module was conducted. Eg. CSES conducted as part of larger study
Mode of Interview - Study [First, Second, Third]	Study logistics	Mode of interview used in study.	Mode of interview used in study. Eg. Internet, Telephone. Mixed modes coded into [First, Second, Third]
Mode of interview - Respondent [First, Second, Third]	Study logistics	Mode of interview used by respondent.	Mode of interview used by respondent. Mixed modes coded into [First, Second, Third]
Self-Selection into Mode of Interview	Study logistics	Only applicable for mixed mode studies. This variable distinguishes between studies where respondents were assigned to the study mode and those studies where different characteristics of respondents led to de-facto self-selection into a survey mode.	Only applicable for mixed mode studies. This variable distinguishes between studies where respondents were assigned to the study mode and those studies where different characteristics of respondents led to de-facto self-selection into a survey mode.
Duration of Interview	Study logistics	Duration of interview.	Duration of interview in number of minutes
Interviewer within Election Study	Study logistics	Interviewer identification variable, within election study.	This variable uniquely identifies an interviewer within an election study. It is not unique across the entire dataset.
Interviewer Gender	Study logistics	Gender of interviewer.	Gender of interviewer.
Days Fieldwork Started Post Election	Study logistics	Number of days after the election fieldwork started.	Number of days after the election fieldwork started.
Duration of	Study logistics	Duration of fieldwork.	Duration of fieldwork in number of days.

Fieldwork			
Date Questionnaire Administered	Study logistics	Date questionnaire administered.	Date questionnaire administered. [Separate Variables for Month Day Year]
Days Interview Conducted Post First Round of Election	Study logistics	Number of days after the election interview conducted.	Number of days after the election interview conducted.
Days Interview Conducted Post First Second of Election	Study logistics	Number of days after the election interview conducted.	Number of days after the election interview conducted.
Language of Questionnaire Administration	Study logistics	Language of questionnaire administration.	Language of questionnaire administration. Eg. Afrikaans
Questionnaire Version	Study logistics	Version of the CSES Module 5 questionnaire that was fielded.	Version of the CSES Module 5 questionnaire that was fielded. Eg. Pilot Questionnaire

Appendix III: Articles using CSES Data After 2016 based on CSES website

1. Abou-Chadi, T. (2018). Electoral competition, political risks, and parties' responsiveness to voters' issue priorities. *Electoral Studies*, 55, 99-108. doi:10.1016/j.electstud.2018.08.012
2. Berz, J. (2019). Potent executives: the electoral strength of prime ministers in Central Eastern Europe. *East European Politics*. doi:10.1080/21599165.2019.1662400
3. Bischof, D. S., Roman. (2017). Simple politics for the people? Complexity in campaign messages and political knowledge. *European Journal of Political Research*. doi:10.1111/1475-6765.12235
4. Burlacu, D. E. (2018). Corruption and Ideological Voting. *British Journal of Political Science*. doi:10.1017/S0007123417000758
5. Carreras, M., & Castaneda-Angarita, N. (2019). Economic adversity and electoral participation of vulnerable socioeconomic groups. *Electoral Studies*, 57, 110-120. doi:10.1016/j.electstud.2018.11.004
6. Carroll, R., & Kubo, H. (2018). Explaining citizen perceptions of party ideological positions: The mediating role of political contexts. *Electoral Studies*, 51, 14-23. doi:10.1016/j.electstud.2017.11.001
7. Christmann, P. (2018). Economic performance, quality of democracy and satisfaction with democracy. *Electoral Studies*, 53, 79-89.
8. Christmann, P., & Torcal, M. (2018). The Effects of Government System Fractionalization on Satisfaction with Democracy. *Political Science Research and Methods*, 6(3), 593-611. doi:10.1017/psrm.2017.23
9. Claassen, C. (2019). Does Public Support Help Democracy Survive? *American Journal of Political Science*. doi:10.1111/ajps.12452
10. da Silva, F. F. (2018). Fostering turnout?: Assessing party leaders' capacity to mobilize voters. *Electoral Studies*, 56(December 2018), 61-79. doi:10.1016/j.electstud.2018.09.013
11. Daoust, J.-F., Blais, A., & Peloquin-Skulski, G. (2019). What do voters do when they prefer a leader from another party? *Party Politics*. doi:10.1177/1354068819845100
12. Dassonneville, R., & McAllister, I. (2018). Gender, Political Knowledge, and Descriptive Representation: The Impact of Long-Term Socialization. *American Journal of Political Science*, 62(2), 249-265. doi:10.1111/ajps.12353
13. Dassonneville, R., & McAllister, I. (2019). The party choice set and satisfaction with democracy. *West European Politics*, 43(1), 49-73. doi:10.1080/01402382.2019.1609286
14. Ferland, B. (2018). Ideological congruence over government mandates under majoritarian and proportional representation electoral systems. *West European Politics*, 41(2), 350-383. doi:10.1080/01402382.2017.1367589
15. Ferland, B. (2018). Party responsiveness to public opinion and party supporters – Revisiting the mechanisms and motivations. *Party Politics*. doi:10.1177/1354068818777889
16. Ferland, B. (2019). A Gender Gap in Party Congruence and

- Responsiveness? *Politics & Gender*. doi:10.1017/S1743923X18000818
17. Fortin-Rittberger, J. (2019). Political Knowledge: Assessing the Stability of Gender Gaps Cross-Nationally. *International Journal of Public Opinion Research*. doi:10.1093/ijpor/edz005
 18. Gomez, R., & Ramiro, L. (2019). Beyond the 2008 Great Recession – Economic factors and electoral support for the radical left in Europe. *Party Politics*, 25(3). doi:10.1177/1354068817718949
 19. Hellwig, T., & McAllister, I. (2019). Party Positions, Asset Ownership, and Economic Voting. *Political Studies*, 67(4). doi:10.1177/0032321718815781
 20. Herron, E. S., & Nishikawa, M. (2001). Contamination effects and the number of parties in mixed-superposition electoral systems. *Electoral Studies*, 20, 63-86. doi:10.1016/S0261-3794(00)00002-0.
 21. Herron, E.S., K. Nemoto, and M. Nishikawa. 2018. Reconciling Approaches in the Study of Mixed-Member Electoral Systems. In *The Oxford Handbook of Electoral Systems*, ed. E.S. Herron, R.J. Pekkanen, and M.S. Shugart, 445–471. Oxford: Oxford University Press.
 22. Houle, C. (2017). Inequality, ethnic diversity, and redistribution. *The Journal of Economic Inequality*, 15(1), 1-23. doi:10.1007/s10888-016-9340-8
 23. Jung, D.-J. (2017). Irrationalizing the rational choice model of voting: The moderating effects of partisanship on turnout decisions in Western and postcommunist democracies. *Electoral Studies*, 50, 26-38. doi:10.1016/j.electstud.2017.09.005
 24. Jung, D.-J. (2018). Does Partisanship Hurt Electoral Accountability? Individual-Level and Aggregate-Level Comparisons of Western and Postcommunist Democracies. *East European Politics and Societies and Cultures*. doi:10.1177/0888325417728772
 25. Lauka, A., McCoy, J., & Firat, R. B. (2018). Mass Partisan Polarization: Measuring a Relational Concept. *American Behavioral Scientist*, 1-20. doi:10.1177/0002764218759581
 26. Lupu, N. S., Lucía; Warner, Zach. (2017). A New Measure of Congruence: The Earth Mover's Distance. *Political Analysis*, 25(1), 95-113. doi:10.1017/pan.2017.2
 27. Maeda, K. 2008. Re-examining the contamination effect of Japan's mixed electoral system using the treatment-effects model. *Electoral Studies* 27, 723-731. doi:10.1016/j.electstud.2008.06.003.
 28. Magalhaes, P. C. A., John H. Gibson, Rachel K. (2018). New forms of mobilization, new people mobilized? Evidence from the Comparative Study of Electoral Systems. *Party Politics*. doi:10.1177/1354068818797367
 29. Marinova, D. M., & Anduiza, E. (2018). When Bad News is Good News: Information Acquisition in Times of Economic Crisis. *Political Behavior*. doi:10.1007/s11109-018-9503-3
 30. Mayer, S. J. S., Martin. (2017). Multiple Parteiidentifikation, politische Involvierung und Cross-Pressures. *Politische Vierteljahresschrift*, 58(3), 367-391. doi:10.5771/0032-3470-2017-3-367
 31. Moral, M. (2017). The Bipolar Voter: On the Effects of Actual and Perceived

- Party Polarization on Voter Turnout in European Multiparty Democracies. *Political Behavior*, 39(4), 935-965. doi:10.1007/s11109-016-9386-0
32. Nadeau, R., Daoust, J.-F., & Arel-Bundock, V. (2019). The market, the state and satisfaction with democracy. *West European Politics*, 43(1), 250-259. doi:10.1080/01402382.2019.1603655
 33. Nadeau, R., Lewis-Beck, M., & Foucault, M. (2019). Wealth and Voter Turnout: Investigating Twenty-Eight Democracies. *Polity*, 51(2). doi:10.1086/701961
 34. Park, J. Y. (2018). Punishing without rewards? A comprehensive examination of the asymmetry in economic voting. *Electoral Studies*, 57, 1-18. doi:10.1016/j.electstud.2018.10.003
 35. Riera, P. B., Damien. (2017). Ticket-splitting in mixed-member systems: on the importance of seat linkage between electoral tiers. *West European Politics*, 40(3), 584-597. doi:10.1080/01402382.2016.1271598
 36. Schmitt-Beck, R. (2019). Political Systems and Electoral Behavior: a Review of Internationally Comparative Multilevel Research. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 71, 343-373. doi:10.1007/s11577-019-00608-8
 37. Silva, B. C., Jungkunz, S., Helbling, M., & Littvay, L. (2019). An Empirical Comparison of Seven Populist Attitudes Scales. *Political Research Quarterly*. doi:10.1177/1065912919833176
 38. Singh, S. P., & Thornton, J. (2019). Elections Activate Partisanship across Countries. *American Political Science Review*, 113(1), 248-253. doi:10.1017/S0003055418000722
 39. Spoon, J.-J., & Kanthak, K. (2019). “He’s not my prime minister!”: negative party identification and satisfaction with democracy. *Journal of Elections, Public Opinion & Parties*. doi:10.1080/17457289.2019.1666271
 40. Spoon, J.-J. K., Heike. (2017). Does anybody notice? How policy positions of coalition parties are perceived by voters. *European Journal of Political Research*, 56(1), 115-132. doi:10.1111/1475-6765.12169
 41. Stiers, D. (2018). Beyond the distinction incumbent–opposition – Retrospective voting on the level of political parties. *Party Politics*. doi:10.1177/1354068817744201
 42. Stiers, D., & Dassonneville, R. (2019). Retrospective Voting and the Polarization of Available Alternatives. *Canadian Journal of Political Science*. doi:10.1017/S0008423919000556
 43. Sulmont, D. (2018). ¿Existe el voto programático en elecciones con un sistema de partidos políticos débil? Un análisis de las elecciones presidenciales peruanas de 2016. *Revista de ciencia política*, 38(3), 429-457. doi:10.4067/S0718-090X2018000300429
 44. Tiemann, G. (2019). The shape of utility functions and voter attitudes towards risk. *Electoral Studies*, 61. doi:10.1016/j.electstud.2019.102051
 45. Traber, D., Giger, N., & Hausermann, S. (2018). How economic crises affect political representation: declining party–voter congruence in times of constrained government. *West European Politics*, 41(5), 1100-1124.

- doi:10.1080/01402382.2017.1378984
46. VanDusky-Allen, J. (2017). Winners, losers, and protest behavior in parliamentary systems. *Electoral Studies*, 54(1), 30-38. doi:10.1016/j.sosci.2016.12.003
 47. Vlandas, T. (2018). Grey power and the Economy: Aging and Inflation Across Advanced Economies. *Comparative Political Studies*, 51(4), 514-552.
 48. Voogd, R., & Dassonneville, R. (2018). Are the Supporters of Populist Parties Loyal Voters? Dissatisfaction and Stable Voting for Populist Parties. *Government and Opposition*. doi:10.1017/gov.2018.24
 49. Voogd, R., & Dassonneville, R. (2018). Are the supporters of populist parties loyal voters? Dissatisfaction and stable voting for populist parties. *Government and Opposition*. doi:10.1017/gov.2018.24
 50. Ward, D. G., & Tavits, M. (2019). How partisan affect shapes citizens' perception of the political world. *Electoral Studies*, 60. doi:10.1016/j.electstud.2019.04.009
 51. Wuttke, A. (2019). New political parties through the voters' eyes. *West European Politics*, 43(1), 22-48. doi:10.1080/01402382.2019.1603940

Appendix IV

A PDF copy of the memorandum “Potential for using CLEA data within the CSES” by Julia Lippman and David Howell can be accessed at the following link (the link is to where the PDF file is stored on David Howell’s Dropbox):

https://www.dropbox.com/s/m8im4ths6ogk18a/Memo_CLEACSES_20211007.pdf