The Comparative Study of Electoral Systems (CSES)
Module 5 Macro & District Data Subcommittee report[1]

Rachel Gibson *(rachel.gibson@manchester.ac.uk)*
Stephen Quinlan *(stephen.quinlan@gesis.org)*
Orit Kedar *(orit.kedar@mail.huji.ac.il)*
Chi Huang *(chihuang@nccu.edu.tw)*

September 2016

Executive Summary

This report presents a review of the macro and district data components of the Comparative Study of Electoral Systems (CSES) project in preparation for Module 5. Based on this review we offer the following six key recommendations:

1. Collection of **15 net new macro variables**. Eight of these relate to the general objectives of the project. The remaining additions cater to the module theme on populism, views on 'in/out groups', and attitudes to elites and the public sector.

2. **Discontinuation of three variables** that capture government-spending levels as a percentage of GDP in certain policy areas.

3. **Retention of the district level data collection**, subject to available project resources. **Collection of two extra district level variables**. However, the Subcommittee advises against collecting additional contextual district level because of concerns over data availability, comparability, and the intensive resources that such an endeavor would involve. We suggest that the decision not to collect

---

[1] The report is not for publication and may not be cited or reproduced without the written permission of the CSES Secretariat (*cses@umich.edu*). All errors are those of the Committee.

contextual level district data and our recommendation to retain the district data for Module 5 should be reviewed in advance of Module 6.

4. **Addition of five questions to the collaborator macro report**.

5. **Addition of two new polity identifiers to allow data bridging** with other macro data. We also suggest that these new polity identifiers are included in previous CSES modules upon re-releases of the data.

6. **Production of a dedicated special issue or edited volume of papers that makes use of the macro/district data** as an effective way to better promote these components of CSES.

1. Mission of the report and work of the committee

The objective of this report is to review and assess the macro and district data components of the Comparative Study of Electoral Systems (CSES) project. One of the key parts of the CSES project is the inclusion of macro and district level variables. These are obtained from two main sources:

1) A 'macro report', where each national collaborator provides contextual information about the specific election and macro/institutional features of their country.

2) Aggregate level data[2] compiled by the CSES secretariat and drawn from an array of publicly available sources such as the World Bank, the Inter Telecommunications Union, and the OECD.

To date there have been two reports commissioned by previous CSES Planning Committees to review the macro data component, the most recent of these focused on the data collected for Module 4 in 2010. Neither it, nor the first report reviewed the district data component. The terms and conditions of this Subcommittee review were as follows:

1. To advise on the addition of new macro variables relevant to the module 5 theme and new macro variables compatible with the general goals of the CSES project.

2. To recommend discontinuation and improvements to the existing suite of macro variables.

3. To review the current macro report and clarify /improve the instructions and reduce collaborator burden where possible.

---

[2] The CSES Secretariat classifies these data as aggregate level macro data.

4. To review the use of district level data, what new data (if any) might be collected, and the consistency of conventions CSES currently applies for this data collection across countries.

The Subcommittee met on numerous occasions between February 2015 and October 2015. We presented a first draft of the report to the CSES Planning Committee (PC) in March 2015 in Taipei. Arising from this meeting, two further charges were added to the Committee's remit, namely:

5. To identify new ways of enabling users to bridge CSES data with other macro data sources.
6. To identify new ways for promoting the macro and district data in the project.

The report proceeds as follows: Section 2 examines in detail the addition of new macro variables and deletion of existing variables. Section 3 provides a comprehensive assessment of the CSES district data. Section 4 reviews the CSES macro report and makes several suggestions as to how collaborator report can be improved and how collaborator burden can be lessened. Section 5 proposes two recommendations that will allow bridging of CSES data and other data sources. Section 6 puts forward several measures as to how the CSES might better promote the macro and district data components of the study. Section 7 concludes with a summary of our recommendations.

2. Macro Data

In identifying the variables for discontinuation or addition, the following considerations guided our recommendations:

1. Our starting point was the need to include variables that would be compatible with both the CSES project in general (i.e. to understand the influences on vote choice and turnout cross-nationally) and the Module 5 theme. To achieve the latter aim, we relied on the interim report of the Module 5 Subcommittee of June 2015, which developed the theme of populism and individual attitudes to political elites and 'in/out' groups in society.

2. We also sought to ensure that collaborator burden in completing the macro report is minimized where possible (i.e.: variables that can be accessed from existing reputable sources should be the responsibility of the Secretariat).

3. We were mindful of the need to ensure our recommendations were in line with the resources available to the project, especially with regard to any extension of data collection at the district level.

4. We were cautious about recommending the inclusion of indices because of the complications associated with these measures. On the one hand, they are an efficient way of summarizing multidimensional data. On the other hand, they themselves are a black box. For example, the judgment calls that went into the calculation of indexes are often lacking in clarity with the components considered for inclusion, exclusion, and the reasons behind it often unknown. Furthermore, details about missing data, and how the indicators included in the indices correlate with one another, are often missing. Therefore, we decided to recommend indices for inclusion only when we were confident about the appropriate data sources, where the methodology was made clear, and where there was a consensus within the community about their use (for e.g.: Consumer Price Index, Effective Number of parties).

5. We were sensitive about recommending inclusion of measures based on expert judgments. While expert surveys are an undoubtedly a potential excellent resource, they also have limits. For example, the identity and training of experts is often unknown, as is variation among the answers different experts provide. Thus we recommend only including expert surveys where their benefit is large (i.e.: in trying to measure corruption) but also where sufficient information about the experts and their answers are available.

2.1 Selection strategy

We began by reviewing existing core comparative datasets with macro variables likely to be relevant to the emergent theme of Module 5 and the CSES project. These included the *European Social Survey*[3], *World Values Survey*[4], the *Quality of Governance* (QoG)

---

[3] *http://www.europeansocialsurvey.org/data/multilevel/*

[4] *http://www.worldvaluessurvey.org/wvs.jsp/*

dataset[5], the *Varieties of Democracy* project[6], the ACE Electoral Knowledge Network[7], and the World Bank[8]. To enable us to sort through the lengthy list of variables identified for potential inclusion, we devised a traffic light decision-making system to guide our recommendations. Specifically:

- Those variables that were seen as a high priority for inclusion were coded green. These were considered highly relevant to CSES/Module 5 and also readily available and accessible. Tables 1 and 2 detail the variables in this category.
- Variables that merited further consideration for inclusion were coded amber. These are variables that are moderately relevant to CSES/Module 5 but face some problems of availability across time and space and/or of cross-national comparability. We detail these variables in Appendices A3 and A4 of this report.
- Finally, those variables that were considered for inclusion but ultimately rejected were coded as red. These presented significant problems of availability and/or relevance to the core theme. Appendices A1 and A2 of the report details the variables falling into this category.

2.2 Recommended additional macro variables: Module 5 theme and general

Table 1 provides a list of the 15 new macro variables we recommend for inclusion in Module 5 that are specifically relevant to the module theme. Table 2 contains new macro variables we recommend for inclusion in relation to the general CSES project. The list in table 1 includes several items from the *Quality of Government (QoG)* expert survey dataset (10-15). Despite some reservations about using expert survey data (see footnote 2) we considered the QoG dataset to meet the required standards of quality and transparency required for inclusion in CSES. Specifically, demographic data are provided on individual respondents and reliability checks on their responses. Furthermore, as we discovered almost all measures of corruption involve some element of subjective assessment (e.g. the Corruptions Perception Index that is already collected by CSES).

---

5   *http://qog.pol.gu.se/data/*

6   *https://v-dem.net/en/*

7   *http://aceproject.org/*

8   *http://databank.worldbank.org/data/home.aspx/*

Table 1 Theme specific macro variables being recommended for inclusion in CSES M5

	VARIABLE	POTENTIAL SOURCES
1	Net migration rates	World Bank;
2	Population by citizenship (foreigners)	UN;
3	Linguistic Fractionalization[9]	Alesina *et al.* 2003.
4	Religious Fractionalization[9]	Alesina *et al.* 2003.
5	Ethnic Fractionalization[9]	Alesina *et al.* 2003.
6	Polity Fragmentation Index[10]	Polity IV; QoG.
7	Gini coefficient of equalized disposable income	OECD; World Bank.
8	Direct Democracy: Referendum Mandatory/Optional	ACE Electoral Know. Network; Centre for Democracy CH.
9	Direct Democracy: Referendums by citizen initiative	ACE Electoral Know. Network; Centre for Democracy CH.
10	Control of Corruption Index**[11]**	QoG.
11	QOG Expert Judgment of Public Sector: Firms provide kickbacks to public servants	QoG.
12	QOG Expert Judgment of Public Sector: Public sector employees and how they treat society	QoG.
13	QOG Expert Judgment of Public Sector: Treat cases impartially	QoG.
14	QOG Expert Judgment of Public Sector: Strive to implement policies to help citizens	QoG.
15	QOG Expert Judgment of Public Sector: Strive to follow rules	QoG.

Please note: The order in of the variables listed are arbitrary. The intention is not to reflect the preference of one variable over another.

---

[9] We recognize one particular set of difficulties with these data, namely that they are based on measures of fractionalization from many years ago and thus may not capture the current situation in several countries. In spite of this sizeable drawback, we are unaware of any other measures of fractionalization cross-nationally and as scholars in the field still use these measures, we have recommended their inclusion.

[10] This variable codes the operational existence of a separate polity, or polities, comprising substantial territory and population within the recognized borders of the state and over which the coded polity exercises no effective authority.

[11] The Control of corruption index captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests. The particular aspect of corruption measured by the various sources differs somewhat, ranging from the frequency of additional payments to get things done, to the effects of corruption on the business environment, to measuring grand corruption in the political arena or in the tendency of elite forms to engage in state capture. The WGI are composite governance indicators based on 32 underlying data sources and come available with standard errors.

Table 2 General macro variables being recommended for definite inclusion in CSES M5

	VARIABLE	POTENTIAL SOURCES
1	Percentage of women in parliament (%)	IPU; World Bank;
2	Do parties receive direct/indirect public funding?	ACE Electoral Know. Network
3	Voting operations: Early/advance voting?	Macro report
4	Voting operations: Vote by mail / postal?	Macro report
5	Voting operations: Vote online/Internet?	Macro report
6	Voter registration: Compulsory or not?	ACE Electoral Know. Network
7	Country subject to IMF conditionality at election?	IMF (MONA); Macro report
8	Unemployment rates by age 15-24 years in %	World Bank

Please note: The order in which the variables are listed is arbitrary and is not intended to reflect the preference of one variable over another.

2.3 Recommended macro variables for discontinuation

Table 3 details the variables that we are proposing to discontinue for Module 5. These are variables D5089, D5091 and D5092 in Module 4, all of which measure government expenditure in various areas of policy. Each measure has three sub-components (t, t-1, t-2). We argue that these variables were of particular relevant to Module 4 but are of little relevance to Module 5. Furthermore, these data are often not available for the required time point (t) on publication of CSES data.

Table 3 Macro variables being recommended for discontinuation in CSES M5

	VARIABLE	CSES M4 CODE
1	General Govt. Expend. (%GDP) – T, T-1, T-2.	D5089
2	Health Expenditure (% GDP) – T, T-1, T-2.	D5091
3	Military Expenditure (% GDP) – T, T-1, T-2.	D5092

3. District Data

This section of the report reviews the rationale and provision of CSES district data.[12] This is the first time such a review has been undertaken and thus our task was more exploratory in nature than was the case in other sections of the report. The district data

---

[12] This section of the report refers to the data component in the CSES denoted by the variables A/B/C/D4000-4005. We use the descriptions 'constituency' and 'riding' interchangeably with district.

have been collected for CSES since the start of the project, although the rationale for their inclusion never has been formally stated.

Our review addressed the following three broad issues:

1. Should the CSES continue to collect district level data (i.e.: what is its value to the project vis-a-vis costs?)

2. What (if any) new district variables might be collected – this covered both electoral and non-electoral / contextual district level information.[13]

3. Examine the consistency of the CSES district data in order to provide greater clarity to the CSES Secretariat about the precise district data that is collected for each country/election.

3.1 Should the CSES continue to collect district data?

While district data are widely available for most countries in the CSES, we identified the following key problems/issues with continuation of the data:

- Logistically, finding the appropriate data is challenging. Oftentimes the data do not always come in a format readily compatible for merging with the CSES dataset. This results in the CSES Secretariat devoting significant resources to this task.

- Use of CSES district data is limited. Research conducted by the secretariat on use of CSES data in published work estimated is no more than 5% of those studies used district data.[14] This raises a question about the return on investment of secretariat time in collecting such data.

- In addition, the classification of a district within and across countries is not always a simple task.

Countering these points, we identified several key reasons to retain these data:

- Discontinuation would break a time series link with previous CSES modules.

---

[13] As we did with the macro data of the project and in accordance with CSES conventions, we consider two types of district data. The first is what we already collect – that is data related to election results and what are classified as systematic district data. The second, which is aggregate level measures of districts, corresponds to contextual information about the district – for example information from censuses about the demographics of the district, or contextual information such as unemployment data etc.

[14] This estimate comes from the CSES Secretariat Bibliography audit project.

- Cessation would also limit the potential for multilevel analyses for those countries where data were available. In particular, since in vast majority of democratic countries the electoral system is districted (the Netherlands and Israel are two of several exceptions), omitting district-level data will preclude analyses of topics such as strategic voting, electoral competition, party-systems at the district level, which is becoming particularly pertinent with the evolvement of political geography in recent years.
- Low usage of the district data may be because of lower user community awareness of their availability. If more variables were collected at the district level than currently is the case, it might increase the likelihood of this data being used more and that district level measures, merged with individual level observations, as the CSES offers, allows for innovative analyses, which no other cross-national dataset provides.

Discussion of these issues at the Taipei and Seattle PC meetings in 2015 indicated that retention was the better option, despite the problems noted. The consensus was that efforts should be directed toward helping improve their quality and consistency and increase user awareness of them. That said there was also an understanding that this commitment to continuation should be subject to ongoing review. Given the labor-intensive nature of district data collection and the possible increasing overlap with projects such as the Constituency Level Elections Archive (CLEA) there is a need to monitor the resources devoted by the Secretariat to this task. We thus recommend that in preparation for Module 6 a new Subcommittee of the project revisit the merits of including district data in the CSES. However, our recommendation is to retain the district data for Module 5 as the arguments for discontinuation were not sufficient and potential exists for this data to be utilized more going forward. The remainder of this section of our report deals with a) means of improving the consistency and clarity of the CSES district data across modules and b) potential ways of improving its saliency to the user community.

3.2 Clarity and consistency of CSES district data

While the CSES has a definition of what constitutes a district, namely: *The data should indicate the respondent's primary electoral district…The goal of this variable, wherever*

possible, is to identify electoral constituencies, the definition of which is "the smallest unit for which there is to the national parliament." i.e. the point at which votes are translated into seats,[15] the decision about the level at which district data should be collected is far from straightforward and there are a number of anomalies or conceptual/ operational inconsistencies that we have identified below:

- Most of the district data CSES collects refers to the lower house parliamentary elections. However, there are exceptions to this that are detailed in the CSES codebook and usually taken in consultation with national collaborators. One such exception is Japan where the district data refers to the upper house as it is on this election that Japan runs the CSES.

- Problems can arise when simultaneous elections occur in one country at the same time. The question then becomes about which election should the district data refer to? Take the US for example, which has up to three elections occurring simultaneously with two potentially overlapping differentiations of district data depending on whether to relate the district data to the congressional elections or the Presidential election. To date the CSES has collected district data about the Presidential elections (1996, 2000, 2004, 2008, and 2012) and thus the district data is based on state level data as it is this that decides victory in the US Electoral College. A similar reasoning has also led the CSES to collect district data referring to Presidential elections in Peru.

- There is an issue whether countries that operate a so-called national district (for e.g.: Israel, the Netherlands, Serbia) should be classified as countries who operate districts or not. The current CSES definition might suggest yes and thus the CSES convention has been to include the national level results in these countries as district level data. However, it is questionable whether this data is equivalent to the type of constituency data one might have had in mind if we consider countries like Australia, Britain, Germany, or Ireland.

- In addition to problems of defining a district, there is also the question of whether all district data should be collected or whether data collected corresponds to the micro data sample available? Currently the CSES collects district data for constituencies

---

[15] This definition is borrowed from the CLEA project, most probably to allow for the potential merging of the two data sources at some point. For more see *http://www.electiondataarchive.org/*

only where we have respondents. Most of the time most, all districts are sampled and this is not a problem. However, there are countries where this is not the case and these are always noted in the codebook. For example, in CSES Module 4, only 135 of the 435 US congressional districts were sampled, equating to just 32% of all congressional districts. This may limit the possibility of analyses using district level variables.

- A secondary problem to that of micro correspondence is that the district-level results may not add up to national level results reported (potentially an issue in countries where the country acts as a district - the Netherlands for example). One potential means of addressing this would be to include all district level data regardless of whether it is linked to an individual observation or not in the CSES data. However, such a strategy would result in additional rows in the dataset that would have missing values for all other variables and thus would be devoid of any use for individual level analysis. Thus, we rule out this suggestion as a potential solution.

Clearly there are important conceptual and operational questions raised by the CSES district data that need to be addressed. To help the user understand the current CSES Appendix B details the current CSES practices concerning district data. The table shows the level of aggregation for each country, the election to which the district data refers to, as well as the availability of each variable for each country.

In order to improve the clarity and consistency of the CSES district data going forward, the Subcommittee recommends the following:

1. The CSES district data should be collected for the lower house election unless a compelling reason exists to the contrary.[16] The unit of analysis should be consistent across all modules and should not change from election to election.

---

[16] We can think of two instances where this is legitimate. First, when the CSES survey is administered during an election that is not the lower house one (for example Japan would fall into this category as stated above). In this circumstance, we recommend including Japanese district data with an explicit election study note in the codebook referring to the upper house distinction, thus enabling practitioners to decide whether to include this data in their analysis or not.

The second instance would be in countries like the United States or France where Presidential elections are the "first-order" election and district data for these elections might link to more individual observations in the data and could be more easily available. We would recommend including district data relating to the "first order election" in these countries but again with the caveat that the CSES

2. CSES should specify explicitly the level at which the district data refers to for each country. To avoid ambiguity, the codebook should specify the following:

   - Type of election which electoral district variables refer to.
   - Level at which electoral district variables refer to.
   - Total number of electoral districts in country (e.g.: Britain 650) and the total number for which CSES has data for.

3. In mixed systems (such as Germany and New Zealand), we recommend retaining the current CSES practice that district data should refer to the constituency vote (as opposed to the list-PR vote) as we argue it is more useful for comparison.

4. District data should be collected for countries that operate the electoral district at the national level (for e.g.: Israel, Netherlands, Serbia). However, we suggest that a distinction is made between districts operating at the subnational level (e.g.: Australia, Britain, Germany, Ireland etc…) to take account of the different level at which the data is collected at. This could be achieved by the splitting of district variables that operate at the national from district level variables at subnational level. This would have the advantage of a) drawing user's attention to the difference in level and b) allowing user's greater flexibility in deciding to include national or subnational level district data in their analysis. The CSES coding scheme also would allow easy merging of these variables together should analysts not wish to make the national or subnational level distinction.

5. The CSES should provide a URL link/source to users where comprehensive district data can be accessed for the said country. This would allow users more power in dealing with the problem that district-level results may not add up to national level results reported (potentially an issue in countries where the country acts as a district). Our solution also preserves the integrity of the CSES micro-macro design.

3.3 Requiring collaborators to provide a source to the district data

We briefly explored the idea that national collaborators collate the district data themselves and provide it to the CSES Secretariat when depositing their data. We rejected this because of the excessive load it would put on collaborators. It was also felt

---

codebook makes explicit to users the differences in district data by country, enabling users to make the choice relevant to their analysis.

that such a strategy might create further definitional problems outlined in the previous section. However, feedback from the CSES Secretariat suggested that resources could be conserved with some assistance from the national collaborator as to the appropriate sources of district level data. To address this issue, we are proposing that collaborators provide a URL link/source to a country's district data in the Macro Report.

3.4 Additional system level measures of electoral districts

The CSES district level data currently measures five different things, namely:

- Number of seats in the district.
- Number of candidates in the district.
- Number of party lists in the district.
- Percentage vote in the district for Parties A-I.
- Turnout in the district.

We recommend the collection of two extra variables, which we assume, would be readily available in the existing corpuses of data used for collation of the existing district data. They are:

- *New variable 1:* Seats won by each party in each district.
- *New variable 2:* Size of the electorate (or population) in the district.

We feel that these extra variables will provide users with more information about the districts and might result in use of the data increasing. We also feel that the collection of this additional data would not put any additional load on the CSES Secretariat.

3.5 Additional Contextual Variables to be collected at District Level

At the PC Meeting in Taipei in March 2015, there was strong support for exploring the possibility of adding aggregate level measures of district context to the CSES to increase the utility and appeal of the district data. This would include basic demographic data (e.g. age, income, gender, population density) as well as economic measures by district (e.g. unemployment rates). To assess the viability of adding these contextual variables in the dataset, the subcommittee took several steps.

First, we chose a selection of countries to establish whether we could access data easily at the district level. This selection took account of all relevant regions covered by the CSES to avoid any selection bias. We contacted the national collaborators in each

country chosen and individuals who had country/regional specific knowledge were also relied upon to complement coverage where possible. Based on the responses received, we classified our sample of countries into four different tiers. The tiers correspond to the difficulty of obtaining contextual variables for electoral districts (1 being the easiest and 4 showing that no such data was accessible). Fuller details of the cases included are given in Appendix B2. Table 4 reports the results for 17 CSES contributor nations, which cover a wide range of geographic regions.

Table 4 Classification of countries according to ease of collecting
contextual variables for electoral district

Tier	Countries
Tier 1	Australia, Britain, Germany, Iceland, Switzerland, Serbia, Netherlands, Israel
Tier 2	Austria. France, Poland, Slovenia
Tier 3	Greece, South Africa, Kenya
Tier 4	Taiwan, Japan,

Tier 1 countries are those countries where contextual level district variables are on the surface readily available, thus making it a straightforward task to merge with district data and include them in the CSES in theory. *Tier 2 countries* are those countries where census and/or other major socioeconomic and demographic data are available but would require some work to disaggregate to the district level. The task is possible but might involve the Secretariat devoting more resources to it than it currently does. *Tier 3 countries* are those countries where census and/or other major socioeconomic demographic data are available but the work required to ensure they mapped appropriately to the electoral district would be considerable and probably beyond the bounds of the CSES Secretariat. *Tier 4 countries* are those where it is not possible to obtain contextual variables corresponding to electoral district because either the data are unavailable or the existing data is irreconcilable to the electoral district level.

Having established that there was a possibility for contextual variables to be extracted and merged with district level data for the CSES for Tier 1 countries we examined how easy collection would be. With the help of the Secretariat, an exercise was undertaken

on four of the Tier 1 countries, namely Australia, Britain, France, and Germany to establish:

- Whether aggregate level data corresponding to electoral districts was available?
- If yes, what format does this data come in?
- An estimate of the resources that the CSES Secretariat might need to devote to finding this data and incorporating into the CSES dataset.

We selected two variables likely to be in high demand: unemployment rates and size of population. Full details of the results of the exercise are given in Appendix B3. To summarize our findings, we did identify several challenges for data collection among this set of 'ideal' cases. Most obvious of these were a) language - the data was not in English and thus would require translation b) format - the data was in a format not immediately compatible for merging with the CSES. While these challenges could clearly be overcome, to do so would obviously require significant additional resources from the Secretariat. A further substantive challenge identified by this exercise was varying time lags between measurement of the contextual variables at district level and the election itself. Unemployment statistics for example may only cover a period of several years prior to the election, and thus be irrelevant to understanding individual level voting behavior. Based on our findings from this review, our recommendation is not to include any aggregate level measures of districts for Module 5. In doing so, we acknowledge that district data sources are likely to expand over time and that the issue of inclusion of this data should be revisited in advance of Module 6.

4. Collaborator Macro report

A key charge for this Subcommittee has been to review the existing CSES Macro report, with a particular emphasis on reducing collaborator burden wherever possible. Collaborators submit the report to the CSES Secretariat along with the data deposit. It provides system level information on the election and country in question and has the purpose of providing a coherent link between national level specialists and the compilation of system level and election specific data to that country. The report contains questions on cabinet composition in the country both before and after the election, information on electoral rules and electoral alliances in the election as well as electoral

results. The answers to the questions posed in the report form the basis of a portion of the macro data included for country in the CSES dataset.

Arising from our review of the existing report, the Committee advises the following:

1. The majority of existing questions in the macro report are kept.

2. The following questions be discontinued in the macro report:

   - *Question 4 a c d e*[17]

   This question asks collaborators to provide detailed information about the specific election result. This information is easily obtained by the CSES secretariat from a variety of public sources. We recommend keeping question 4b, however, which asks collaborators to provide an official source of the election results complete with a URL link.

   - *Question 8*[16]

   This question asks collaborators to provide the party name and the party leader. This information is also widely available from a variety of public sources.

3. A clarification is applied to *Question 9(d)*[16], that asks collaborators on which date the election was originally scheduled to be held. Currently, there is some ambiguity about whether the question refers to the legal date on which the election was scheduled to take place (i.e.: the date set out in the election writ or by law or whether it means an intention to hold the election on a specific date which is not specified by law (for example an intended date stated by a politician).[18]

   For the purposes of clarity and comparability, we propose that the question be amended to the following:

   On what date was the election originally scheduled to be held legally?

4. Notwithstanding the macro variables measuring the official voter turnout rates cross-nationally, we noted that the CSES macro component lacks several relevant variables that might be of relevance to the study of voter turnout comparatively. Therefore, we suggest the following additions to the macro report to address this shortcoming:

---

[17] See Appendix C for question wording.

[18] Such a circumstance occurred with respect to the 2013 Australian Federal Election. On 30 January 2013, the then Prime Minister Julia Gillard had originally scheduled the election for 14 September 2013. However, when Gillard was replaced as Prime Minister by Kevin Rudd as a consequence of a change in the leadership of the Labor Party, Rudd changed the election date and the poll was held on 7 September 2013.

a. *New question 1:* Early/advance voting

A phenomenon becoming more common with voting is that electors in many countries are able to cast their ballot in some form prior to Election Day. This is mostly a consequence of attempts to incentivize electors to vote by removing impediments that might make it unfeasible for electors to cast a ballot on Election Day. In line with this development, the CSES has observed some election studies sampling frames which have included respondents who have cast their ballot before polling day. As there is no consistent cross-national public database that tracks whether countries have some form of early or advance voting, we propose the addition of the following question to the macro report:

Can electors (voters) cast their ballot before Election Day polls open?

[ ] Yes

[ ] No

[ ] Other[19]

Please specify other: _____

b. *New questions 2 and 3:* Mail and Internet voting

As more and more electors now have the ability to cast a ballot, so too has the means of casting a ballot become more varied. Cross-nationally, a multitude of options exists as to how electors might cast their ballot ranging from voting via mail to voting online. It might be plausible to expect that the probability of voting might be influenced by the means and/or the multitude of means available to an elector. As there is little consistent cross-national public information that tracks the different options open to electors to vote and thus we propose the addition of the following question to the macro report:

Can electors (voters) cast their ballot via mail?

[ ] Yes

[ ] No

[ ] Other

---

[19] The "other" category might apply to a country like the United States where voting operations might differ by state. We recommend the answer 'other' be explained to the user community by means of an election study note in the CSES codebook.

Please specify other: _____

Can electors (voters) cast their ballot online?

  [ ] Yes

  [ ] No

  [ ] Partially

Please specify other: _____

5. We recommend asking national collaborators to provide a source complete with URL for the district data for the said country (new question 4). We suggest asking for this information as an addition to Question 4 of the current macro report when a source is being sought for the election results.

6. With the theme of Module 5 in part focusing on populism, there is merit in asking national collaborators to provide their expert ratings of parties contesting the election on a populism scale. Our suggested question provides a definition of populism that is line with the one advocated by the Module Theme Subcommittee and with the literature. We suggest the addition of the following question:

<u>New question 5:</u> Expert rating of populist parties

As Module 5 focuses in part on populism, please indicate the degree to which each of the parties (in the expert judgment of the CSES Collaborator) can be characterized as a <u>populist party</u>? Please use the same parties used in the CSES Module 5 respondent questionnaire and label them the same way (A-I). A definition of populism is below.

DEFINITION: Populism can be defined as a thin-centered ideology that pits a virtuous and homogeneous people against a set of elites and dangerous 'others' who are depicted as depriving "the people" of their rights, values, prosperity, identity, and voice. The emphasis on anti-elite/ anti-establishment rhetoric and the contrast between the "pure people" and the "corrupt elite" are thus indications of the degree to which a party is populist. Populist parties can be found across the left-right ideological spectrum. On a scale of 0 to 10, where 0 is "not at all populist" and 10 is "very populist", where would you place each of the parties in your country?

Party Name	Not at all populist								Very populist		
	0	1	2	3	4	5	6	7	8	9	10
A.											
B.											
C.											
D.											
E.											
F.											
G.											
H.											
I.											

After taking into account feedback from the CSES Plenary meeting in Philadelphia in September 2016, we also recommend that in answering the above question, collaborators be asked the following two questions:

- *How many country experts/national collaborators have assisted in the classification of parties on the populist scale?*
- *We encourage collaborators to provide any further specific or relevant information (if applicable) regarding the above classifications, which might help users of the data understand them to a greater extent.*

We suggest that the answers to these questions be included in the CSES codebook and will provide users with more information to enable them make their own judgment on the validity of the expert judgments. We also suggest that the same two questions be asked of collaborators with respect to their classification of parties on the left-right scale and the alternative scale (where applicable).

5. Data Bridging

At the PC Meeting in Taipei in March 2015, some members of the PC advocated that the CSES should explore how the CSES data could be more directly linked with other macro

datasets. This would increase the appeal of CSES and allow scholars the option to undertake new analysis beyond the scope of the CSES data. It would also relieve the CSES of a commitment to becoming a universal macro data source, which we have already established resources do not permit.

One means of addressing this gap would be to have polity identifiers that were compatible with other macro data sources. The CSES currently has two primary country identifiers. The first is the election study identifier coupled with the numeric polity code (in Module 4 this is variable D1003). This eight-digit code is constructed from three different components. The first three digits are the polity code assigned by the United Nations Statistics Division to each country. The fourth digit distinguishes between multiple studies conducted within a single country, for the same election. The final four digits correspond to the election year. For example, the Australia 2013 election study is coded: "03602013" – "036" being the UN identifier for Australia, "0" accounting for the fact that there were not multiple studies conducted within Australia, and "2013" being the year in which the election was held. The second identifier is the Alphabetical Polity Code and Election Year (in Module 4 this is variable D1004). Two components make up this variable. The first three characters are the alphabetical version of the country codes assigned by the United Nations Statistics Division. The fourth digit distinguishes between multiple studies conducted within a single country, for the same election. The final four digits correspond to the election year. So again taking the Australian 2013 example, it is coded as AUS_2013.

The CSES Secretariat explored five prominent macro datasets to find common country identifiers. The details of this exercise are in Appendix D. The exercise concluded that there is scope for the CSES to add extra polity level identifiers that would aid users in merging the CSES data with other macro data sources. We thus recommend the creation of two new country level identification variables to stand alongside the existing polity identifiers (in Module 4 D1003 and D1004):

- *New Country identifier 3*[20]*:* Country name verbatim in English (E.g.: AUSTRALIA)
- *New Country identifier 4:* Country 3-letter ISO identifier (E.g.: AUS).

---

[20] We refer to these as country identifiers 3 and 4 because there are two country identifiers already included in the CSES and we are recommending preservation of these items.

We also advocate that these new country identifiers be included in previous modules of the CSES when the Secretariat re-releases them.

6. Promotion of Macro/District data components

While the CSES is known for its micro-macro research design, our review of the literature suggested that its macro and district data components could be more widely used and promoted. Accordingly, we suggest are several ways in which use of the macro and district data components could be maximized. They include:

1. Promotion of macro and district data in e-mails sent to the user community list.
2. Greater promotion of macro and district data on the CSES website. One means of doing this might be to have specific sections on the website that are more noticeable and accessible to the user community.
3. A special one-day workshop (e.g.: at APSA or MPSA) on the macro and district data components of the project.
4. Special journal issue targeting papers that make extensive use of the macro/district data. Members of this Subcommittee are willing to take the lead in this endeavor.

We also considered whether the macro and district data components of the data should be collated into a separate dataset and offered to the user community separate independent of the individual level data. This is a practice of the ESS for instance. While the benefits for promoting CSES macro data are evident, we recommended against this on several number of grounds. First, the current approach maximizes the ease of use of macro data by integrating into the micro file. Second, collection and collation of macro/district data is not the sole goal of the project. Rather our goal is to ensure that macro/district can be merged with individual level observations to allow for cross level analysis. Our concern is that a new focus on the creation of additional datasets might distract the Secretariat and the PC from the project's core objectives. Third, such a separation risks the CSES becoming a repository for macro level data and the value of the individual level data diminishing. Finally, given that the district data currently collected is based on the presence of individual observations for those districts (i.e. not all districts will be included), producing a separate macro/district dataset from current files would risk being incomplete.

7. Summary

This report has reviewed the macro and district data components of the CSES for the Module 5 component of the project, scheduled to begin in autumn 2016. We compiled our report between February 2015 and September 2015 with assistance of the CSES Secretariat. Our key objectives were to conduct a thorough review and viability test of the macro and district data currently collected. Identify new macro and district data that would enhance both the CSES generally and support Module 5 in particular. Secondary aims were to rationalize the macro report to reduce collaborator burden where possible and identify new ways promoting the use of macro and district data. Our key recommendations are:

- Adding fifteen new macro variables focusing on the module theme and the general objectives of CSES.

- Discontinuation of variables related to government spending levels in a range of policy areas as a percentage of GDP.

- Retention of the current district data variables and addition of two new variables.

- Not adding new measures of context for electoral districts because of data availability, comparability, and resources.

- Retention of most questions on the collaborator macro report, minor clarifications of some and the addition of several new questions. We recommend dropping the requirement for collaborators to provide election results verbatim in the macro report.

- To enable data bridging, the addition of two new polity identifiers specifying the country name verbatim in English and ISO 3-digit code for Country.

- To promote the macro and district data components of the project, a special issue or edited volume of papers that makes use of the macro/district data specifically.

The Comparative Study of Electoral Systems (CSES) Module 5 Macro & District Data Subcommittee, *9 September 2016.*

Appendix A1

No	Genre	System or Aggregate	Variable	Description	Possible Source I	Possible Source II	A... ava...
1	Indices	A	Bertelsmann Index: Status Index	*The Status Index ranks the countries according to their quality of democracy and market economy as of January 31 in a given year. It is based in the average score for the dimensions of political (Democracy Status) and economic (Market Economy Status) transformation.*			I
2	Political	S	Gallagher disproportionality index	*Also known as least squares index and is used to measure the disproportionality of an electoral outcome*			I
3	Political	S	Government fractionalization	*The probability that two deputies picked at random from among the government parties will be of different parties.*			I
4	Political	S	Opposition fractionalization	*The probability that two deputies picked at random from among the opposition parties will be of different parties.*			I
5	Media	S	TV debates	*Are televised debates between candidates or party representatives normally conducted?*	ACE Electoral Knowledge Network		
6	Media	S	Blackout period for release of opinion polls	*Is there a period where opinion polls being published before election day are limited?*	ACE Electoral Knowledge Network		
7	Indices	A	Gender Inequality Index (GII)	*The Gender Inequality Index (GII) was introduced in 2010 and reflects women's disadvantage in three dimensions: reproductive health, empowerment and the labour market for as many countries as data of reasonable quality allow. The index shows the loss in human development due to inequality between female and male achievements in these dimensions. It ranges from 0, which indicates that women and men fare equally, to 1, which indicates that women fare as poorly as possible in all measured dimensions. The health dimension is measured by two indicators: maternal mortality ratio and the adolescent fertility rate. The empowerment dimension is also measured by two indicators: the share of parliamentary seats held by each sex and by secondary and higher education attainment levels. The labour dimension is measured by women's participation in the work force. The Gender Inequality Index is designed to reveal the extent to which national achievements in these aspects of human development are eroded by gender inequality, and to provide empirical foundations for policy analysis and advocacy efforts.*	UN Development Programme		I
8	Indices	A	Bertelsmann Index: Democracy Status	*Democracy Status is measured in terms of five criteria: Stateness, Political participation, Rule of law, Stability of democratic institutions and Political and social integration.*			I
9	Indices	A	Bertelsmann Index: Market Economy Status	*Market Economy Status is measured in terms of seven criteria: Level of socioeconomic development, Organization of the market and competition, Currency and price stability, Private property, Welfare regime, Economic performance and Sustainability*			I

10	Indices	A	Bertelsmann Index: Management Index	*The Management Index ranks the countries according to their leadership's political management performance. It is formed by calculating the average of scores given for the management criteria (Management Performance), which is then offset against the assigned level of difficulty.*	I
11	Indices	A	Bertelsmann Index: Management Performance	*Management Performance is based in five critera: Level of difficculty, Steering capability, Resource efficiency, Consensus-building and International cooperation.*	I
12	Indices	A	Democracy Barometer: Individual liberties	*Individual liberties primarily secure the inviolability of the private sphere. This requires the right to physical integrity and the right to free conduct of life.*	I
13	Indices	A	Democracy Barometer: Rule of law	*Rule of law designates the independence, the primacy, and the absolute warrant of and by the law. This requires equality before the law and quality of the legal system.*	I
14	Indices	A	Democracy Barometer: Public Sphere	*The communication about politics and moral norms takes place in the public sphere. A vital civil society and a vivid public sphere requires freedom of association and freedom of opinion.*	I
15	Indices	A	Democracy Barometer: Competition	*Vertical control of the government is established via free, regular, and competitive elections. This requires vulnerability (uncertainty of the electoral outcome) and contestability (electoral competitors have to meet in order to be allowed to enter the race).*	I
16	Indices	A	Democracy Barometer: Mutual constraints	*The horizontal and institutional dimension of control of the government is encompassed by mutual constraints of constitutional powers.*	I
17	Indices	A	Democracy Barometer: Governmental Capabilities	*Governmental capability entails the availability of resources and conditions for efficient implementation of policy.*	I
18	Indices	A	Democracy Barometer: Transparency	*Transparency requires no secrecy on the part of political representatives as well as provisions for a transparent political process.*	I
19	Indices	A	Democracy Barometer: Participation	*Citizens must have equal participation rights. All persons who are affected by a political decision should have the right to participate in shaping that decision.*	I
20	Indices	A	Democracy Barometer: Representation	*All citizens' preferences are adequately represented in the political decision-making process*	I
21	Indices	A	Worldwide Governance Indicators: Voice and accountability	*Capturing perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.*	I
22	Indices	A	Worldwide Governance Indicators: Control of Corruption	*Capturing perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.*	I
23	Indices	A	Worldwide Governance Indicators: Government effectiveness	*Capturing perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.*	I
24	Indices	A	Worldwide Governance Indicators: Rule of law	*Capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.*	I

25	Indices	A	Worldwide Governance Indicators: Regulatory quality	*Capturing perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.*			I
26	Indices	A	Worldwide Governance Indicators: Political stability, no violence	*Capturing perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism.*			I
27	Indices	A	Gender Inequality Index	*All scores are reported on a scale of 0 to 1, with 1 representing maximum gender equality. The study measures the extent to which women have achieved full equality with men in ve critical areas: Economic participation; Economic opportunity; Political empowerment; Educational Attainment; Health and well-being*	World Economic Forum	QoG	
28	Indices	A	Political Stability Index	*Political Stability combines several indicators which measure perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism. The WGI are composite governance indicators based on 32 underlying data sources. Available with standard errors*	World Bank World Governance Indicators	QoG	
29	Indices	A	Rule of Law Index	*Rule of Law includes several indicators which measure the extent to which agents have con dence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts. Together, these indicators measure the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions and the extent to which property rights are protected.*	World Bank World Governance Indicators		

Appendix A2

APPENDIX A2 THEME SPECIFIC VARIABLES – TRAFFIC LIGHT RED

No.	Genre	System or Aggregate	Variable	Description	Possible Source I	Possible Source II
1	Electoral Integrity	A	Electoral laws were unfair to smaller parties	*Strongly Disagree, 1; Disagree, 2; Neither agree nor disagree, 3; Agree, 4; Strongly Agree, 5.*	Electoral Integrity Project	
2	Electoral Integrity	A	Elections laws restricted citizens' rights	*Same as #1*	EIP	
3	Electoral Integrity	A	Elections were well managed	*Same as #1*	EIP	
4	Electoral Integrity	A	Information about voting procedures was widely available	*Same as #1*	EIP	
5	Electoral Integrity	A	Boundaries discriminated against some parties	*Same as #1*	EIP	
6	Electoral Integrity	A	Boundaries favored incumbents	*Same as #1*	EIP	
7	Electoral Integrity	A	Boundaries were impartial	*Same as #1*	EIP	
8	Electoral Integrity	A	Some opposition candidates were prevented from running	*Same as #1*	EIP	
9	Electoral Integrity	A	Women had equal opportunities to run for offic	*Same as #1*	EIP	
10	Electoral Integrity	A	Some parties/candidates were restricted from holding campaign rallies	*Same as #1*	EIP	
11	Electoral Integrity	A	Parties/candidates had equitable access to public political subsidies	*Same as #1*	EIP	
12	Electoral Integrity	A	Parties/candidates had equitable access to political donations	*Same as #1*	EIP	
13	Electoral Integrity	A	Parties/candidates publish transparent financial accounts	*Same as #1*	EIP	
14	Electoral Integrity	A	Some voters were threatened with violence at the polls	*Same as #1*	EIP	
15	Electoral Integrity	A	Some fraudulent votes were cast	*Same as #1*	EIP	
16	Electoral Integrity	A	National citizens living abroad could vote	*Same as #1*	EIP	
17	Electoral Integrity	A	The election authorities were impartial	*Same as #1*	EIP	
18	Electoral Integrity	A	The authorities distributed information to citizens	*Same as #1*	EIP	
19	Electoral Integrity	A	The authorities allowed public scrutiny of their performance	*Same as #1*	EIP	

20	Electoral Integrity	S	Electoral Disputes Agency	*Who is responsible at the first level for electoral complaints?*	ACE Electoral Knowledge Network	Macro report
21	Corruption	A	QoG Expert Jud of Public Sector: Officials recruited within	Senior public officials are recruited from within the ranks of the public sector	QoG	
22	Corruption	A	QoG Expert Jud of Public Sector: Public sectir employees and personal contacts	When granting licenses to start up private firms, public sector employees favor applicants with which they have strong personal contacts	QoG	
23	Corruption	A	QoG Expert Jud of Public Sector: Salaries 1	Senior officials have salaries that are comparable with the salaries of private sector managers with roughly similar training and responsibilities	QoG	
24	Corruption	A	QoG Expert Jud of Public Sector: Salaries 2	The salaries of public sector employees are linked to appraisals of their performance	QoG	
25	Corruption	A	QoG Expert Jud of Public Sector: Reprimand	When found guilty of misconduct, public sector employees are reprimanded by proper bureaucratic mechanisms	QoG	
26	Corruption	A	QoG Expert Jud of Public Sector 10yrs: Reprimand Public sector recrutiment: skills and merits of applicants deciding factor	When recruiting public sector employees, the skills and merits of the applicants decide who gets the job	QoG	
27	Corruption	A	QoG Expert Jud of Public Sector 10yrs: When recruiting public sector employees, formal examination system	Public sector employees are hired via a formal examination system	QoG	
28	Corruption	A	QoG Expert Jud of Public Sector 10yrs: Salaries 1	Senior officials have salaries that are comparable with the salaries of private sector managers with roughly similar training and responsibilities	QoG	
29	Corruption	A	QoG Expert Jud of Public Sector 10yrs: political leadership matters in hiring/firing	The top political leadership hires and fires senior public officials	QoG	
30	Corruption	A	QoG Expert Jud of Public Sector 10yrs: Officials recruited within	Senior public officials are recruited from within the ranks of the public sector	QoG	
31	Corruption	A	QoG Expert Jud of Public Sector 10yrs: When recruiting public sector employees, one stays in the system for rest of life	Once one is recruited as a public sector employee, one stays a public sector employee for the rest of one's career	QoG	

32	Corruption	A	QoG Expert Jud of Public Sector 10yrs: Public sector employees and how they treat society	When deciding how to implement policies in individual cases, public sector employees treat some groups in society unfairly	QoG
33	Corruption	A	QoG Expert Jud of Public Sector: $1000 reach the poor	The needy poor	QoG
34	Corruption	A	QoG Expert Jud of Public Sector: $1000 reach middlemen/consultants	Middlemen/consultants	QoG
35	Corruption	A	QoG Expert Jud of Public Sector: $1000 reach others	Others	QoG
36	Corruption	A		The terms of employment for public sector employees are regulated by special laws that do not apply to private sector employees	QoG
37	Corruption	A		The provision of public services is subject to competition from private sector companies, NGOs or other public agencies	QoG
38	Corruption	A		Women are proportionally represented among public sector employees	QoG
39	*Indices*	A	Political Stability Index	*Political Stability combines several indicators which measure perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism. The WGI are composite governance indicators based on 32 underlying data sources. Available with standard errors*	World Bank QoG World Governance Indicators

Appendix A3

APPENDIX A3 GENERAL VARIABLES – TRAFFIC LIGHT ORANGE

	VARIABLE	POTENTIAL SOURCES
1	Voting operations: Overseas pop. able to vote	ACE Electoral Know. Network Macro report;
2	World Press Freedom Index	Reporters Without Borders
3	Criteria for allocating broadcast time to parties	ACE Electoral Know. Network
4	World Bank income category	World Bank
5	GDP per capita on purchasing power parity	World Development Indicators
6	Duration of Education - Compulsory	UNESCO
7	Duration of Primary/Secondary Education	UNESCO

Please note: The order in which the variables are listed is arbitrary and is not intended to reflect the preference of one variable over another.

Appendix A4

APPENDIX A4 THEME SPECIFIC VARIABLES – TRAFFIC LIGHT ORANGE

VARIABLE	POTENTIAL SOURCES
1 Refugee Population	World Bank; QoG.
2 EIP Expert Jud: PEIIndexp Perceptions of Electoral Integrity Index (individual level)	Electoral Integrity Project.
3 EIP Expert Jud: Electoral laws favoured the governing party or parties	Electoral Integrity Project.
4 Electoral Management Body: Nat v regional	ACE Electoral Know. Network;
5 Electoral Management Body: Independent or not	ACE Electoral Know. Network;
6 QOG Expert Jud of Public Sector: Public sector recruitment - skills and merits of applicants deciding factor	QoG.
7 QOG Expert Jud of Public Sector: Public sector recruitment: political connections the factor	QoG.
8 QOG Expert Jud of Public Sector: When recruiting public sector employees, formal examination system	QoG.
9 QOG Expert Jud of Public Sector: political leadership matters in hiring/firing	QoG.
10 QOG Expert Jud of Public Sector: When recruiting public sector employees, one stays in the system for rest of life	QoG.
11 QOG Expert Jud of Public Sector: Strive to be efficient	QoG.
12 QOG Expert Jud of Public Sector: Strive to implement policies by politicians	QoG.
13 QOG Expert Jud of Public Sector: $1000 reach public employees own pocket	QoG.
14 QOG Expert Jud of Public Sector: $1000 superior public employee	QoG.
15 KOF Globalization Index[1]	Swiss Federal Instit. Of Tech.

Please note: The order in which the variables are listed is arbitrary and is not intended to reflect the preference of one variable over another.

---

[1] Measures globalization on three main dimensions: economic, political, and social. An overall index of globalization also possible. For more see: *http://globalization.kof.ethz.ch/*

Appendix B1

Table B1 CSES District Data Conventions: Election type, district unit of analysis and availability of district data.

No.	COUNTRY (Elec Year)	ELEC.	DISTRICT UNIT OF ANALYSIS	DATA AVAILABILITY				
				4001	4002	4003	4004	4005
1	Australia (2013)	Parl. LH	Single Member Electoral Division (150)	X	X	O	X	X
2	Austria (2013)	Parl. LH	Multi-Mem Constit.(9) into regional Constit. (43)	X	O	X	X	X
3	Brazil (2010)	Pres.	Multi-Mem Constit.(27)	X	O	X	X	X
4	Canada (2008)	Parl. LH	Single Member Constit. (308)	X	X	O	X	O
5	Chile (2009)	Parl. LH	Multi-Mem Constit.(60)	X	O	O	X	O
6	Croatia (2007)	Parl. LH	Multi-Mem Constit.(12)	X	O	X	X	X
7	Czech Rep. (2010)	Parl. LH	Multi-Mem Constit.(14)	X	O	X	X	X
8	Denmark (2007)	Parl. LH	Multi-Mem Constit.(10)	X	X	X	X	X
9	Estonia (2011)	Parl. LH	Multi-Mem Constit.(12)	X	X	X	X	X
10	Finland (2011)	Parl. LH	Multi-Mem Constit.(14) & Single-Mem Constit.(1)	X	X	X	X	X
11	France (2012)	Parl. LH	Single-Mem Constit.(577)	O	O	O	O	O
12	Germany (2013)	Parl. LH	Single Member Constit. (299)	X	X	X	X	X
13	Great Britain (2005)	Parl. LH	Single Member Constit. (650)	X	X	X	X	X
14	Greece (2012)	Parl. LH	Multi-Mem Constit.(48) & Single-Mem Constit.(8)	X	X	X	X	X
15	Iceland (2013)	Parl. LH	Multi-Mem Constit.(6)	X	O	X	X	X
16	Ireland (2011)	Parl. LH	Multi-Mem Constit.(43)	X	X	O	X	X
17	Israel (2006)	Parl. LH	National	X	X	X	X	X
18	Japan (2013)	Parl. UH	Multi-Mem Constit.(47)	X	O	O	X	X
19	Latvia (2010)	Parl. LH	Multi-Mem Constit.(5)	X	X	X	X	O
20	Mexico (2012)	Parl. LH	Single-Mem Constit.(300)	X	O	O	X	X
21	Montenegro (2012)	Parl. LH	National	X	O	X	X	X
22	Netherlands (2010)	Parl. LH	National	X	X	X	X	X
23	New Zealand (2011)	Parl. LH	Single-Member Constit. (70)	X	X	X	X	X
24	Norway (2009)	Parl. LH	Multi-Mem Constit.(19)	X	O	X	X	X
25	Peru (2011)	Pres.	Multi-Mem Constit.(25)	O	O	X	X	X
26	Philippines (2010)	Parl. LH	Single-Mem Constit.(233)	O	O	O	O	O
27	Poland (2011)	Parl. LH	Multi-Mem Constit.(41)	X	X	X	X	X
28	Portugal (2009)	Parl. LH	Multi-Mem Constit.(22)	X	X	X	X	X
29	Serbia (2012)	Parl. LH	National	X	O	X	X	X

No.	COUNTRY (Elec Year)	ELEC.	DISTRICT UNIT OF ANALYSIS	DATA AVAILABILITY				
				4001	4002	4003	4004	4005
30	Slovenia (2008)	Parl. LH	Electoral Units (8)	X	X	X	X	X
31	South Africa (2009)	Parl. LH	Multi-Mem Constit.(9)	X	X	X	X	X
32	South Korea (2008)	Parl. LH	Single-Mem Constit.(246)	X	O	X	O	O
33	Spain (2008)	Parl. LH	Multi-Mem Constit.(50) & 2 Single-Member.	X	X	X	X	X
34	Sweden (2006)	Parl. LH	Multi-Mem Constit.(29)	X	O	O	X	X
35	Switzerland (2011)	Parl. LH	Multi or Single-Mem Constit.(26)	X	X	X	X	O
36	Taiwan (2012)	Parl. LH	Multi-Mem Constit.(75)	O	O	O	O	O
37	Thailand (2011)	Parl. LH	Single-Mem Constit.(375) and National Constit.	O	O	O	O	O
38	Turkey (2011)	Parl. LH	Multi-Mem Constit.(79)	X	O	X	X	X
39	USA (2012)	Pres.	States (50+DC)	X	X	O	X	X
40	Uruguay (2009)	Parl. LH	Multi-Mem Constit.(19)	O	O	O	X	X

Please note: X = Data available. O = Data unavailable; 4001-4005 refers to variables C4001-C4005 or D4001-D4005 in the CSES district data component. Number after definition of district unit of analysis is the number of districts in total for the said country. It does not reflect the number for which we have data.

Appendix B2

APPENDIX B2

Australia (Tier 1)

The Australian Bureau of Statistics provides census data broken down to constituency level. The Australian Data Archive makes files available for federal electorates at https://www.ada.edu.au/historical/abs-census

Sources: Steven McEachern, Deputy Director, Janet McDougall, Senior Data Archivist Australian Data Archive (ADA)

UK (Tier 1)

The BES team have helpfully set up a file with linked census data http://www.britishelectionstudy.com/data-object/2015-bes-constituency-results-with-census-and-candidate-data/ The full list of data available in Table A.1

Source: BES Team

Germany (Tier 1)

German Federal Returning Officer ("Bundeswahlleiter") provides contextual data at the level of the electoral districts. The data for the two most recent elections can be downloaded from http://www.bundeswahlleiter.de/de/bundestagswahlen/BTW_BUND_09/strukturdaten/index.html

http://www.bundeswahlleiter.de/de/bundestagswahlen/BTW_BUND_13/strukturdaten/index.html Data for earlier elections is available upon request. The information is extensive and includes provided includes: - number of municipalities – area - population (gender, age categories, German population) - population increase/decrease – education - stock of motor vehicles – dwellings - mining and manufacturing - statistics of business notification - trade tax revenue - insolvency procedures - employees subject to social insurance (incl. some information on type of employment)

- unemyployment rates

Source: CSES Collaborator Anne Schäfer, Rudiger Schmitt-Beck

Switzerland (Tier 1)

Constituencies are the cantons and the cantons are the main sub-national level of government. Data related to the election and the socio-economic context are widely available. Most of the data is available online by the Federal Statistical Office (www.bfs.admin.ch).

Source: CSES collaborator George Lutz, Swiss NES

Iceland (Tier 1)

Socio-economic data are available from Statistics Iceland (www.statice.is) webpage (see left hand side bar with key statistics). These data are available for the local level (municipalities) and on the level of postcodes but would need to be aggregated to electoral district. On the Statice site all information is in English, however, some local assistance would be needed to provide information about which postal and municipalities belong to each electoral district. Specifically a code can be applied that converts postal codes and municipalities names into electoral districts. When you have that it will be easy for cses staff to work with the statice data. Data can be downloaded as machine readable files.
Source: CSES collaborator Eva Heiða Önnudóttir <eho@hi.is. ICENES.

Austria (Tier 2)

In Austria, the lowest level of aggregation that contextual data can be collected for is political district (n=95) which is an administrative unit and not an electoral tier. Demographic and economic data are published by *Statistik Austria* at the political district level (e.g. age, ethnicity, income). Employment records are calculated according to specifical 'employment districts' i.e. different geographic boundaries. These can be aggregated to the level of regional electoral district (n=39) but would require some significant additional work and local knowledge. Costs would likely be incurred. We would also like to point out that the data that would be reported are likely to differ in dates/time when they were collected, e.g. some of the variables may be collected before (or after) the election. They are not necessarily tied to the election year. Data on the state-level (Bundesländer, n=9) is more readily available and these units coincide with the second electoral tier in Austrian national elections.
Source CSES collaborator Silvia Kritzinger. Kathrin Thomas Austrian National Election Study (AUTNES)

France (Tier 1)

The INSEE provides census data for each "circonscription" or constituency on their website
See http://www.insee.fr/fr/themes/detail.asp?reg_id=0&ref_id=circo_leg-2012 The information can be found by clicking on "département" (a département is more or less equivalent to a county and it's divided in circonscriptions) then on each "circonscription". There is a further division then into a list of cantons - cantons are

very small constituencies only used for some kind of local elections. Tables are provided that contain population data such as age, gender, occupation, citizenship (French/foreigners). The files are not downloadable and are written in English but you have tables that can be printed.

Serbia (Tier 1*)

In Serbia, the situation is specific since the entire country is a single electoral district (as in, for instance, the Netherlands). Hence, national-level contextual data are equivalent to district level data. Historically, Serbia probably would fit under Tier 3 in that it has operated several different types of electoral systems that have included sub-national districts (Majoritarian, PR list) but these districts have not matched directly to the administrative units for which socio-economic data are reported. Some limited demographic data are published for 'local communities' and aggregation is possible but would take time to accomplish. More extensive socioeconomic statistical data are available at the level of municipalities than on the level of local communities. Generally several municipalities would be within a single electoral district. The Serbian Statistical office publishes the data. Some of the data are available in English. http://webrzs.stat.gov.rs/WebSite/ The site has a searchable database – assistance would be needed locally to track down relevant local information if needed.
Source: CSES collaborator Bojan Todosijevic <bojan.todosijevic@gmail.com>

Poland (Tier 2)

Population data are available at the sub-national level from the Central Statistical Office of Poland (GUS). There is a searchable database available in English that provides regional/local data http://stat.gov.pl/bdlen/app/strona.html?p_name=indeks. These data collected at three administrative levels - gminas/poviats/voivodships – which do not map onto electoral districts. Thus collecting useful data (for the electoral studies that focus on constituency/distict level) requires some computing, merging, aggregation, disaggregation etc. However, it is seen as feasible.
Source: CSES collaborator Mikołaj Cześnik mczesnik@swps.edu.pl, Polish National Election Study

Slovenia (Tier 2/3)

No specific sources provided. Referred to publication "STAT'O'BOOK" available at: www.stat.si/eng/pub.asp / http://www.stat.si/doc/pub/Statobook.pdf produced by the Statistical office of the Republic of Slovenia (SURS). e-mail: info.stat@gov.

Local assistance likely to be needed to produce matching contextual data for constituencies.

Source: Karolina Kušević, svetovalka direktorja

Taiwan (Tier 3/4)

Economic indicators (e.g. per capita income or unemployment rates, etc.) are mainly based on administrative units such as township or counties. Electoral Districts for the national parliament. (Legislative Yuan) typically cut across these administrative units making it very difficult to transform social and economic data into congressional districts. Among the 73 SMDs nationwide only 20 of the least populated counties have district boundaries that overlap with their boundaries. At the other end of the spectrum is the highly populated Taipei City as an example (see Figure 1, Appendix A.2). Taipei is divided into 8 single-member districts for the 2012 Legislative Yuan (Three out of the 12 "chu" or administrative units are cut across by the SMD boundaries.

Japan (Tier 3/4)

As in Taiwan economic indicators are mainly based on administrative units such as township or counties. Electoral Districts for the Lower House in Japan also often cut across these administrative units making it difficult to transform social and economic data into congressional districts. Take Tokyo in Japan as an example (see Figure 2, Appendix A.2). The capital city is divided into 25 single-member districts (SMD, numbers printed in red color) for the 2014 House election. Many economic indicators, such as employment rate, refers to the whole city and therefore cannot be further broken down. A few indicators, such as labor participation rate, can be found at the "chu" level (a sub-city administrative unit in the east area) or village level. SMD boundaries often cut across "chu" in populated areas and make it difficult to aggregate them into SMD level.

Greece (Tier 3/4)

No data are readily available at the district level. The Hellenic Statistical Authority -ELSTAT (English version: http://www.statistics.gr/portal/page/portal/ESYE) provides some data at the level of the 13 Greek Prefectures ("perifereies"). A request can be raised for these data (see here: http://www.statistics.gr/pls/apex/f?p=106:1030:3325892154848172::NO:::) but this would need to be done through local sources.

Source: CSES collaborator Eftichia Teperoglou efteperoglou@gmail.com

South Africa (Tier 2/3)

For South Africa there are no real districts. 200 MPs are elected to represent the entire country from a closed party list, and the other 200 are selected from 9 closed provincial lists, with the number proportional to provincial population. For SA, the closest thing would be data on the provinces which is available from StatSA, the census bureau or the Electoral Commission. StatsSA makes machine readable data available on its webpage, but the IEC makes data available to province and much lower levels.

Source Robert Mattes

Kenya (Tier 3?)

In Kenya, the only other African country in CSES, matters are complicated by the fact that they recently revised all their parliamentary constituency boundaries before the most recent CSES survey. http://www.knbs.or.ke/index.php   . Details of municipal boundaries are available from the Kenyan Electoral Commission website http://www.iebc.or.ke/index.php/2015-01-15-11-10-24/downloads/category/bou-ndaries Source: Robert Mattes Founder and Senior Advisor on the Afrobarometer CSES collaborator for South Africa and Kenya.

African region more generally (Tier 3)

Beyond SA and Kenya the impression from local experts is that political data (and boundaries) is often not matched well to census data in Africa as a result of gerrymandering.   Overall census frames would in most cases only have the administrative districts, and not constituencies, so there would be some significant work required in converting any statistics into constituency level indicators

Source: Robert Mattes, Carolyn Logan Deputy Director of the Afrobarometer

Table B2 British Election Study 2015 Constituency Data File Variables

Name	Label	Census table
pano	Press Association Constituency ID	
ONSConstID	ONS Constituency ID	
Constituency Name	Constituency Name	
Country	Country	
Region	Region	
Winner15	2015 Winning party	
Con15	2015 Conservative vote share	
Lab15	2015 Labour vote share	
LD15	2015 Liberal Democrat vote share	
SNP15	2015 Scottish National Party (SNP) vote share	
PC15	2015 Plaid Cymru vote share	
UKIP15	2015 United Kingdom Independence Party (UKIP) vote share	
Green15	2015 Green Party vote share	
Other15	2015 Combined other vote share	
Majority15	2015 Majority	
Turnout15	2015 Turnout	
ConVote15	2015 Conservative number of votes	
LabVote15	2015 Labour number of votes	
LDVote15	2015 Liberal Democrat number of votes	
SNPVote15	2015 Scottish National Party (SNP) number of votes	
PCVote15	2015 Plaid Cymru number of votes	
UKIPVote15	2015 United Kingdom Independence Party (UKIP) number of votes	
GreenVote15	2015 Green Party number of votes	
BNPVote15	2015 British National Party (BNP) number of votes	
TotalVote15	2015 Total number of votes cast	
Electorate15	2015 Size of electorate	
SeatChange1015	2015 winning party gain from 2010 winning party	
Con1015	2010-15 Conservative vote share change	
Lab1015	2010-15 Labour vote share change	
LD1015	2010-15 Liberal Democrat vote share change	
SNP1015	2010-15 Scottish National Party (SNP) vote	

Name	Label	Census table
	share change	
PC1015	2010-15 Plaid Cymru vote share change	
UKIP1015	2010-15 United Kingdom Independence Party (UKIP) vote share change	
Green1015	2010-15 Green Party vote share change	
Winner10	2010 Winning party	
Con10	2010 Conservative vote share	
Lab10	2010 Labour vote share	
LD10	2010 Liberal Democrat vote share	
SNP10	2010 Scottish National Party (SNP) vote share	
PC10	2010 Plaid Cymru vote share	
UKIP10	2010 United Kingdom Independence Party (UKIP) vote share	
Green10	2010 Green Party vote share	
BNP10	2010 British National Party (BNP) vote share	
Majority10	2010 Majority	
Turn10	2010 Turnout	
Convote10	2010 Conservative number of votes	
Labvote10	2010 Labour number of votes	
LDvote10	2010 Liberal Democrat number of votes	
SNPvote10	2010 Scottish National Party (SNP) number of votes	
PCvote10	2010 Plaid Cymru number of votes	
UKIPvote10	2010 United Kingdom Independence Party (UKIP) number of votes	
Greenvote10	2010 Green Party number of votes	
BNPvote10	2010 British National Party (BNP) number of votes	
Elecorate10	2010 Size of electorate	
ConPPC	Conservative candidate name	
ConPPCsex	Conservative candidate sex	
ConPPCrace	Conservative candidate race	
LabPCC	Labour Candidate name	
LabPPCsex	Labour Candidate sex	
LabPPCrace	Labour Candidate race	
LDPCC	Liberal Democrat candidate name	

Name	Label	Census table
LDPPCsex	Liberal Democrat candidate sex	
LDPPCrace	Liberal Democrat candidate race	
UKIPPPC	UKIP candidate name	
UKIPPPCsex	UKIP candidate sex	
UKIPPPPCrace	UKIP candidate race	
SNPPPC	SNP candidate name	
SNPPPCsex	SNP candidate sex	
SNPPPCrace	SNP candidate race	
PCPPC	Plaid Cymru candidate name	
PCPPCsex	Plaid Cymru candidate sex	
PCPPCrace	Plaid Cymru candidate race	
GreenPPC	Green Party candidate name	
GreenPPCsex	Green Party candidate sex	
GreenPPCrace	Green Party candidate race	
c11Population	Census 2011: Usual resident population	KS101
c11PopulationDensity	Census 2011: Population density	KS101
c11Male	Census 2011: Male	KS101
c11Female	Census 2011: Female	KS101
c11Households	Census 2011: People living in households	KS101
c11Communal	Census 2011: People living in Communal establishments	KS101
c11Age0to4	Census 2011: Age 0 to 4	KS102
c11Age5to7	Census 2011: Age 5 to 7	KS102
c11Age8to9	Census 2011: Age 8 to 9	KS102
c11Age10to14	Census 2011: Age 10 to 14	KS102
c11Age15	Census 2011: Age 15	KS102
c11Age16to17	Census 2011: Age 16 to 17	KS102
c11Age18to19	Census 2011: Age 18 to 19	KS102
c11Age20to24	Census 2011: Age 20 to 24	KS102
c11Age25to29	Census 2011: Age 25 to 29	KS102
c11Age30to44	Census 2011: Age 30 to 44	KS102
c11Age45to59	Census 2011: Age 45 to 59	KS102
c11Age60to64	Census 2011: Age 60 to 64	KS102
c11Age65to74	Census 2011: Age 65 to 74	KS102
c11Age75to84	Census 2011: Age 75 to 84	KS102
c11Age85to89	Census 2011: Age 85 to 89	KS102

Name	Label	Census table
c11Age90plus	Census 2011: Age 90 plus	KS102
c11HouseOwned	Census 2011: Housing tenure - Owned	KS402
c11HouseOutright	Census 2011: Housing tenure - Owned outright	KS402
c11HouseMortgage	Census 2011: Housing tenure - Owned with a mortgage	KS402
c11HouseShared	Census 2011: Housing tenure - Shared ownership	KS402
c11HouseSocial	Census 2011: Housing tenure - Social rented	KS402
c11HouseSocialLA	Census 2011: Housing tenure - Social rented from council (Local Authority)	KS402
c11HouseSocialOther	Census 2011: Housing tenure - Other social rented	KS402
c11HousePrivate	Census 2011: Housing tenure - Private rented	KS402
c11HousePrivateLandlord	Census 2011: Housing tenure - Private rented from landlord or letting agency	KS402
c11HousePrivateOther	Census 2011: Housing tenure - Other private rented	KS402
c11HouseRentFree	Census 2011: Housing tenure - Living rent free	KS402
c11HouseholdOnePerson	Census 2011: Household - One person	KS105
c11HouseholdOnePerson65plus	Census 2011: Household - One person aged 65 and older	KS105
c11HouseholdOnePersonOther	Census 2011: Household - One person other	KS105
c11HouseholdOneFamily	Census 2011: Household - One family	KS105
c11HouseholdOneFamily65plus	Census 2011: Household - One family all 65 and older	KS105
c11HouseholdMarried	Census 2011: Household - Married	KS105
c11HouseholdMarriedNoChildren	Census 2011: Household - Married no children	KS105
c11HouseholdMarriedDependents	Census 2011: Household - Married with dependent children	KS105
c11HouseholdMarriedNondependents	Census 2011: Household - Married with no dependent children	KS105
c11HouseholdCohabit	Census 2011: Household - Cohabitating	KS105
c11HouseholdCohabitNoChildren	Census 2011: Household - Cohabitating no children	KS105
c11HouseholdCohabitDependents	Census 2011: Household - Cohabitating with dependent children	KS105

Name	Label	Census table
c11HouseholdCohabitNodependents	Census 2011: Household - Cohabitating with no dependent children	KS105
c11HouseholdLone	Census 2011: Household - Lone parent	KS105
c11HouseholdLoneDependents	Census 2011: Household - Lone parent with dependent children	KS105
c11HouseholdLoneNodependents	Census 2011: Household - Lone parent with no dependent children	KS105
c11HouseholdOther	Census 2011: Household - Other	KS105
c11HouseholdOtherDependents	Census 2011: Household - Other with dependent children	KS105
c11HouseholdAllStudents	Census 2011: Household - All full-time students	KS105
c11HouseholdAll65plus	Census 2011: Household - Other all 65 plus	KS105
c11HouseholdAnyOther	Census 2011: Household - Any other	KS105
c11CarsNone	Census 2011: No cars or vans in household	KS404
c11CarsOne	Census 2011: one car or van in household	KS404
c11CarsTwo	Census 2011: Two cars or vans in household	KS404
c11CarsThree	Census 2011: Three cars or vans in household	KS404
c11CarsFour	Census 2011: Four cars or vans in household	KS404
c11EthnicityWhite	Census 2011: Ethnicity - White	KS201
c11EthnicityMixed	Census 2011: Ethnicity - Mixed	KS201
c11EthnicityAsian	Census 2011: Ethnicity - Asian	KS201
c11EthnicityBlack	Census 2011: Ethnicity - Black	KS201
c11EthnicityOther	Census 2011: Ethnicity - Other	KS201
c11EthnicityWhiteBritish	Census 2011: Ethnicity detailed - White British/English/Scottish/Northern Irish	KS201
c11EthnicityWhiteIrish	Census 2011: Ethnicity detailed - White Irish	KS201
c11EthnicityWhiteTraveller	Census 2011: Ethnicity detailed - White Gypsy or Irish Traveller	KS201
c11EthnicityWhiteOther	Census 2011: Ethnicity detailed - White other	KS201
c11EthnicityMixedCaribbean	Census 2011: Ethnicity detailed - Mixed White and Black Caribbean	KS201
c11EthnicityMixedAfrican	Census 2011: Ethnicity detailed - Mixed White and Black African	KS201
c11EthnicityMixedAsian	Census 2011: Ethnicity detailed - Mixed White and Asian	KS201
c11EthnicityMixedOther	Census 2011: Ethnicity detailed - Mixed other	KS201

Name	Label	Census table
c11EthnicityIndian	Census 2011: Ethnicity detailed - Indian	KS201
c11EthnicityPakistani	Census 2011: Ethnicity detailed - Pakistani	KS201
c11EthnicityBangladeshi	Census 2011: Ethnicity detailed - Bangladeshi	KS201
c11EthnicityChinese	Census 2011: Ethnicity detailed - Chinese	KS201
c11EthnicityOtherAsian	Census 2011: Ethnicity detailed - Other Asian	KS201
c11EthnicityBlackAfrican	Census 2011: Ethnicity detailed - Black African	KS201
c11EthnicityBlackCaribbean	Census 2011: Ethnicity detailed - Black Caribbean	KS201
c11EthnicityBlackOther	Census 2011: Ethnicity detailed - Black Other	KS201
c11EthnicityArab	Census 2011: Ethnicity detailed - Arab	KS201
c11EthnicityAnyOther	Census 2011: Ethnicity detailed - Any other ethnic group	KS201
c11BornUK	Census 2011: Country of birth - UK	KS204
c11BornEngland	Census 2011: Country of birth - England	KS204
c11BornNI	Census 2011: Country of birth - Northern Ireland	KS204
c11BornScotland	Census 2011: Country of birth - Scotland	KS204
c11BornWales	Census 2011: Country of birth - Wales	KS204
c11BornUKNotSpecified	Census 2011: Country of birth - Northern Ireland	KS204
c11BornIreland	Census 2011: Country of birth - Ireland	KS204
c11BornOtherEU	Census 2011: Country of birth - Other EU (not Ireland)	KS204
c11BornOtherPre2004EU	Census 2011: Country of birth - Other pre 2004 EU (not Ireland)	KS204
c11BornPost2004EU	Census 2011: Country of birth - country joined EU 2004 or later	KS204
c11BornOther	Census 2011: Country of birth - Other	KS204
c11PassportNone	Census 2011: Passport held - None	KS205
c11PassportAny	Census 2011: Passport held - Any	KS205
c11PassportUK	Census 2011: Passport held - United Kingdom	KS205
c11PassportIreland	Census 2011: Passport held - Republic of Ireland	KS205
c11PassportEU	Census 2011: Passport held - Other Europe EU Countries	KS205
c11PassportEuropeNotEU	Census 2011: Passport held - Other Europe Non EU countries	KS205
c11PassportAfrica	Census 2011: Passport held - Africa	KS205

Name	Label	Census table
c11PassportAsia	Census 2011: Passport held - Middle East and Asia	KS205
c11PassportNorthAmerica	Census 2011: Passport held - North America and the Caribbean	KS205
c11PassportCentralAmerica	Census 2011: Passport held - Central America	KS205
c11PassportSouthAmerica	Census 2011: Passport held - South America	KS205
c11PassportOceania	Census 2011: Passport held - Antarctica and Oceania	KS205
c11EnglishAll	Census 2011: All 16+ have English (or Welsh in Wales) as main language	KS206
c11EnglishOne	Census 2011: At least one 16+ has English (or Welsh in Wales) as main language	KS206
c11EnglishChild	Census 2011: At least one 3-15 has English (or Welsh in Wales) as main language	KS206
c11EnglishNone	Census 2011: No one 16+ has English (or Welsh in Wales) as main language	KS206
c11Christian	Census 2011: Religion - Christian	KS209
c11Buddhist	Census 2011: Religion - Buddhist	KS209
c11Hindu	Census 2011: Religion - Hindu	KS209
c11Jewish	Census 2011: Religion - Jewish	KS209
c11Muslim	Census 2011: Religion - Muslim	KS209
c11Sikh	Census 2011: Religion - Sikh	KS209
c11ReligionOther	Census 2011: Religion - Other	KS209
c11NoReligion	Census 2011: Religion - No Religion	KS209
c11ReligionNotStated	Census 2011: Religion - Not Stated	KS209
c11NSSECHigherManager	Census 2011: NS-SeC Large employers, higher managerial and administrative	KS611
c11NSSECHigherProfessional	Census 2011: NS-SeC Higher professional occupations	KS611
c11NSSECLowerManager	Census 2011: NS-SeC Lower managerial, administrative and professional	KS611
c11NSSECIntermediate	Census 2011: NS-SeC Intermediate	KS611
c11NSSECSmallEmployer	Census 2011: NS-SeC Small employers and own account workers	KS611
c11NSSECLowerSupervisor	Census 2011: NS-SeC Lower supervisory and technical	KS611

Name	Label	Census table
c11NSSECSemiRoutine	Census 2011: NS-SeC Semi-routine occupations	KS611
c11NSSECRoutine	Census 2011: NS-SeC Routine	KS611
c11NSSECNeverWorked	Census 2011: NS-SeC Never worked	KS611
c11NSSECLongtermUnemploye d	Census 2011: NS-SeC Long-term unemployed	KS611
c11EconomicActive	Census 2011: Economically active - all	KS601
c11Employed	Census 2011: Economically active - Employed	KS601
c11EmployedPartTime	Census 2011: Economically active - Employee: Part-time	KS601
c11EmployedFullTime	Census 2011: Economically active - Employee: Full-time	KS601
c11SelfEmployed	Census 2011: Economically active - Self-employed	KS601
c11Unemployed	Census 2011: Economically active - Unemployed	KS601
c11EconomicallyActiveStudent	Census 2011: Economically active - Full-time student	KS601
c11EconomicInactive	Census 2011: Economically inactive - all	KS601
c11Retired	Census 2011: Economically inactive - Retired	KS601
c11EconomicallyInactiveStudent	Census 2011: Economically inactive - Student (including full-time students)	KS601
c11LookingAfterHome	Census 2011: Economically inactive - Looking after home or family	KS601
c11LongTermSick	Census 2011: Economically inactive - Long-term sick or disabled	KS601
c11EconomicInactiveOther	Census 2011: Economically inactive - Other	KS601
c11Unemployed16to24	Census 2011: Unemployed age 16 to 24	KS601
c11Unemployed50to74	Census 2011: Unemployed age 50 to 74	KS601
c11Neverworked	Census 2011: Never worked	KS601
c11LongTermUnemployed	Census 2011: Long-term unemployed	KS601
c11FulltimeStudent	Census 2011: Full-time students	KS611
c11IndustryAgriculture	Census 2011: Industry - Agriculture, forestry and fishing	KS605
c11IndustryMining	Census 2011: Industry - Mining and quarrying	KS605
c11IndustryManufacturing	Census 2011: Industry - Manufacturing	KS605

Name	Label	Census table
c11IndustryElectricitySupply	Census 2011: Industry - Electricity, gas, steam and air conditioning supply	KS605
c11IndustryWaterSupply	Census 2011: Industry - Water supply	KS605
c11IndustryConstruction	Census 2011: Industry - Construction	KS605
c11IndustryWholesale	Census 2011: Industry - Wholesale and retail trade	KS605
c11IndustryTransport	Census 2011: Industry - Transport and storage	KS605
c11IndustryAccommadation	Census 2011: Industry - Accommodation and food service activities	KS605
c11IndustryCommunication	Census 2011: Industry - Information and Communication	KS605
c11IndustryFinance	Census 2011: Industry - Financial and insurance activities	KS605
c11IndustryRealEstate	Census 2011: Industry - Real estate activities	KS605
c11IndustryProfessional	Census 2011: Industry - Professional, scientific and technical activities	KS605
c11IndustryAdministrative	Census 2011: Industry - Administrative and support service activities	KS605
c11IndustryPublicAdministration	Census 2011: Industry - Public administration and defence	KS605
c11IndustryEducation	Census 2011: Industry - Education	KS605
c11IndustrySocialWork	Census 2011: Industry - Human Health and Social Work Activities	KS605
c11IndustryOther	Census 2011: Industry - Other	KS605
c11QualNone	Census 2011: Highest qualification - None	KS501
c11QualLevel1	Census 2011: Highest qualification - Level 1	KS501
c11QualLevel2	Census 2011: Highest qualification - Level 2	KS501
c11QualApprentice	Census 2011: Highest qualification - Apprenticeship	KS501
c11QualLevel3	Census 2011: Highest qualification - Level 3	KS501
c11QualLevel4	Census 2011: Highest qualification - Level 4 and above	KS501
c11QualOther	Census 2011: Highest qualification - Other	KS501
c11Degree	Census 2011: University Degree	QS502
c11HealthVeryGood	Census 2011: Very good health	KS301
c11HealthGood	Census 2011: Good health	KS301

Name	Label	Census table
c11HealthFair	Census 2011: Fair health	KS301
c11HealthBad	Census 2011: Bad health	KS301
c11HealthVeryBad	Census 2011: Very bad health	KS301
c11NoAdultsEmployed	Census 2011: No adults in employment in household	KS106
c11NoAdultsEmployedChildren	Census 2011: No adults in employment in household with dependent children	KS106
c11NoAdultsEmployedNoChildren	Census 2011: No adults in employment in household without dependent children	KS106
c11DeprivedNone	Census 2011: Household is not deprived on any dimension	QS119
c11Deprived1	Census 2011: Household deprived on one dimension	QS119
c11Deprived2	Census 2011: Household deprived on two dimensions	QS119
c11Deprived3	Census 2011: Household deprived on three dimensions	QS119
c11Deprived4	Census 2011: Household deprived on four dimensions	QS119

Figure B1 Map of Tokyo capital city
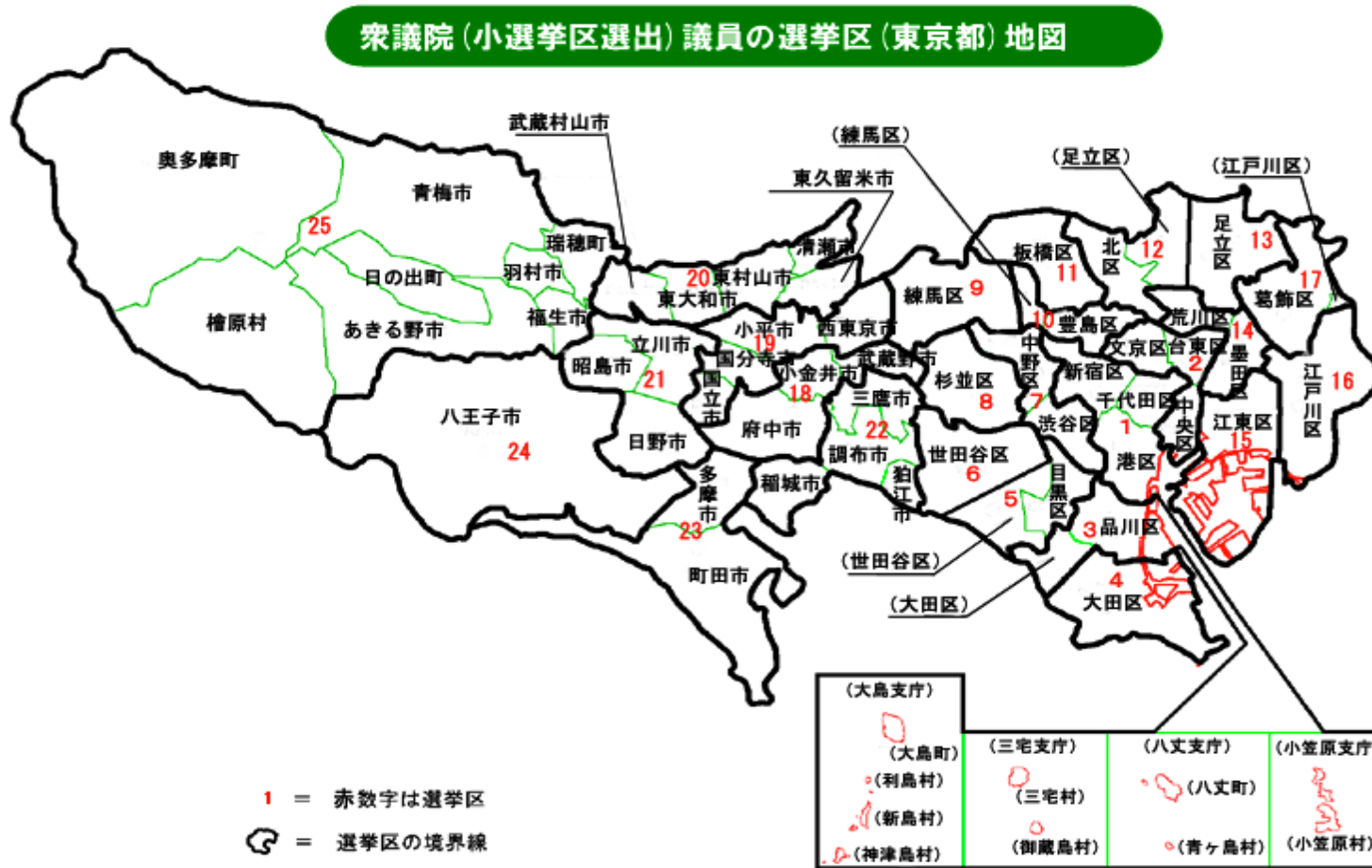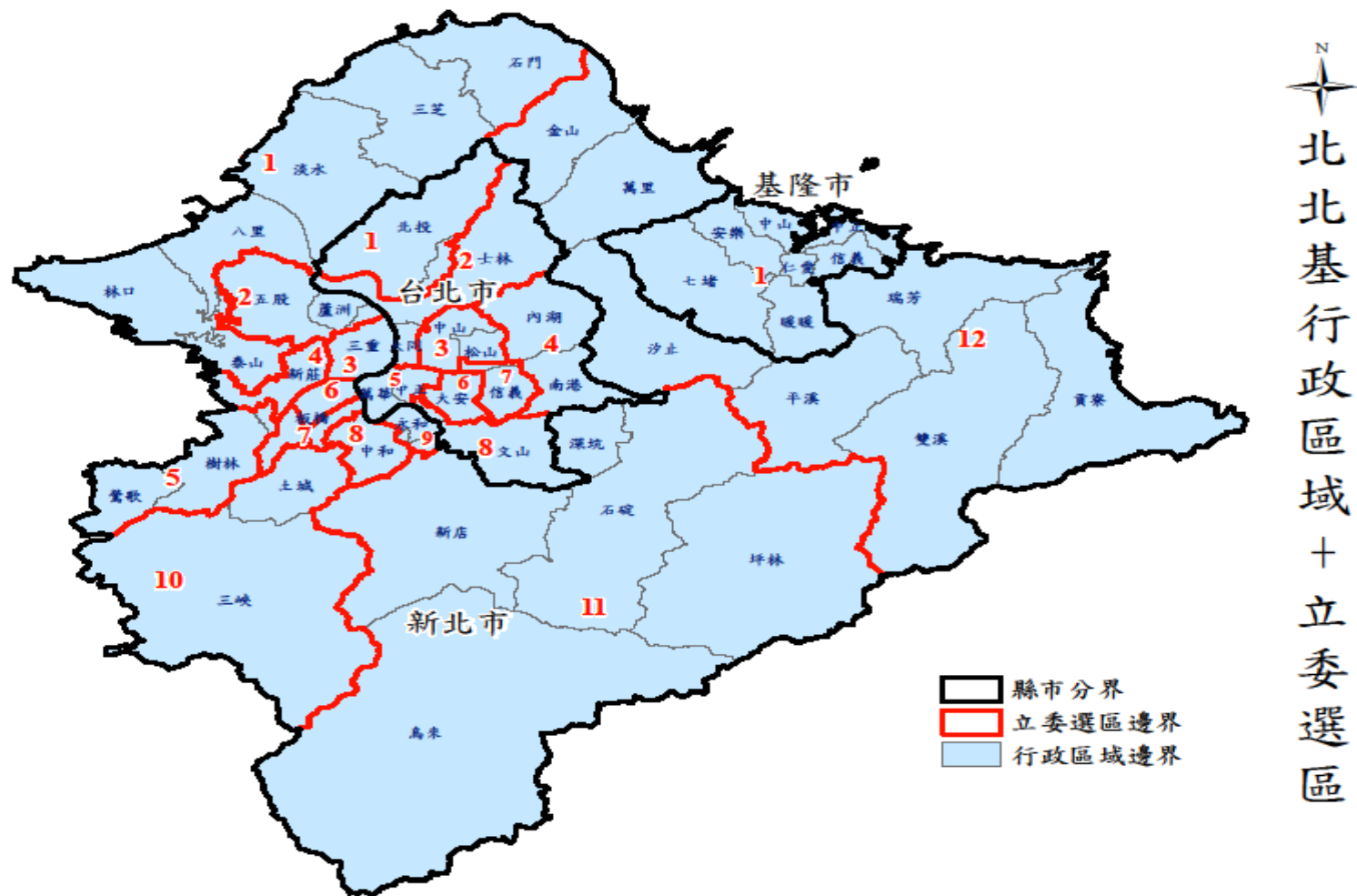

衆議院（小選挙区選出）議員の選挙区（東京都）地図

Figure B2 Map of northern Taiwan

Appendix B3

Access to aggregate level measures at the electoral district level

Memorandum prepared by the CSES Secretariat for the M5 PC Macro and District subcommittee

9 September 2015

Objective

This memorandum is for the CSES Module 5 PC macro and district variable sub-committee. The objective of this memo was to ascertain the following:

1. Is aggregate level data corresponding to electoral districts available?
2. If applicable, what format does this data come in?
3. An estimate of the resources that the CSES Secretariat might need to devote to finding the data and extracting it for incorporation into the CSES dataset.

The focus of inquiry was on four countries namely Australia, Britain, France, and Germany, who had been identified as "Tier 1" countries, where aggregate level measures corresponding to electoral divisions in each country might be obtained. Two potentially interesting variables: unemployment statistics for the district, and the district's population were chosen as examples.

Australia

Population data for each electoral district is available from the Australian Bureau of Statistics[1]. The data is available in MS excel format, and thus can be easily extracted by the CSES Secretariat.

However, finding data on *unemployment* was excessively time consuming. Having gone through the Australian Bureau of Statistics, we ascertained that only the censuses (conducted in 2011 and 2006) provide unemployment data by electoral district. As this was not ideal, we made an alternative search and found correspondence tables between the geographical/territorial units on which data for the census and the electoral divisions. While some of the geographical/territorial units on which data are available are substantially smaller than electoral divisions, aggregating these data using correspondence tables could be feasible and was done by the Parliament of Western Australia from 2010-2014, but data is only readily available for this state and there are numerous caveats attached to the methodology. A similar approach by the CSES Secretariat might prove difficult and resource intensive, and might require a large degree of nation specific knowledge.

Britain

---

[1] http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3218.02013-14?OpenDocument.

Population data and *unemployment data* by electoral district in Britain are easily accessible, the data having been located after approximately 30 minutes of searching.[2] Unemployment data by district is available for each year. The data is available in MS excel format, and thus can be easily extracted by the CSES Secretariat.

France

The French National Institute of Statistics and Economic Studies provide information pertaining to local districts only in French, which is likely to pose a particular difficulty to CSES Secretariat staff, given the language of the project is English.[3] However, having conducted more extensive investigations, *population data* for each district ("circonscriptions") were located elsewhere.[4] However, this information is relevant to only one year – 2008, the last time of a national census was conducted in France. Thus we conclude population data is not easily available for France.

Similarly, *unemployment data* by electoral district was also difficult to come by. France has data available on regions and departements but whether they correspond to electoral districts is not clear cut.

Germany

In the case of Germany, the data for both population and unemployment at the district level were easily obtained.[5] One drawback though was the website was only partly available in English. The data were available for each variable for each recent year. The data is available in MS excel format, and thus can be easily extracted by the CSES Secretariat. .

Summary

Aggregate level data at the district level for the two variables in question could be identified in Britain, and Germany. For Australia, while population data by district was obtainable, unemployment data by district is possible but not easily available and would require the CSES Secretariat or the collaborator to assist with ensuring correspondence with the electoral district. With respect to France, no data for either variable was easily obtainable.

Generally, deciphering this data was doable. Where available, it was not overly time consuming and for the most part data came in an easily extractable format. An exception to this was the unemployment data for Australia, which would probably need substantial construction from the CSES Secretariat if it were to correspond to electoral districts.

---

[2] http://www.ons.gov.uk/ons/guide-method/user-guidance/parliamentary-constituencies/data-catalogue-for-parliamentary-constituencies/index.html
[3] http://www.insee.fr/en/bases-de-donnees/default.asp?page=statistiques-locales.htm
[4] http://www.insee.fr/fr/themes/detail.asp?reg_id=0&ref_id=circo_leg-2012&page=donnees-detaillees/circo_leg/circo_leg-2012/tableau/circo_leg_1_1.htm (
[5] : www.regionalstatistik.de

A point of note for the Committee is that while cross-national sources such as the OECD and Eurostat do have district level aggregate measures available, for the most part these these data do not refer to districts, but rather to NUTS (Nomenclature of Territorial Units for Statistics, with NUTS 1 corresponding to German Länder, NUTS 2 corresponding to French regions, and NUTS 3 corresponding to French départments. This makes them unfit for our purposes.

Appendix C

Comparative Study of Electoral Systems
Module 4: Macro Report
September 10, 2012

Country:
Date of Election:

Prepared by:
Date of Preparation:

NOTES TO COLLABORATORS:

- The information provided in this report contributes to an important part of the CSES project. The information may be filled out by yourself, or by an expert or experts of your choice. Your efforts in providing these data are greatly appreciated! Any supplementary documents that you can provide (e.g., electoral legislation, party manifestos, electoral commission reports, media reports) are also appreciated, and may be made available on the CSES website.

- Answers should be as of the date of the election being studied.

- Where brackets [ ] appear, collaborators should answer by placing an "X" within the appropriate bracket or brackets. For example: [X]

- If more space is needed to answer any question, please lengthen the document as necessary.

Data Pertinent to the Election at which the Module was Administered

1a. Type of Election
- [ ] Parliamentary/Legislative
- [ ] Parliamentary/Legislative and Presidential
- [ ] Presidential
- [ ] Other; please specify: _____

1b. If the type of election in Question 1a included Parliamentary/Legislative, was the election for the Upper House, Lower House, or both?
- [ ] Upper House
- [ ] Lower House
- [ ] Both
- [ ] Other; please specify: _____

2a. What was the party of the president *prior* to the most recent election, regardless of whether the election was presidential?

2b. What was the party of the Prime Minister *prior* to the most recent election, regardless of whether the election was parliamentary?

2c. Report the number of cabinet ministers of each party or parties in cabinet, *prior* to the most recent election. (If one party holds all cabinet posts, simply write "all".) Ministers are considered those members of government who are members of the Cabinet and who have Cabinet voting rights. List also cabinet members that are 'independent'. If known, specify if the 'independents' are affiliated or close to certain parties.

Name of Political Party   Number of Cabinet Ministers

2d. What was the size of the cabinet *prior* to the election (total number of cabinet ministers detailed in 2c)? Please include only full ministers and the prime minister in the count. Ministers are considered those members of government who are members of the Cabinet and who have Cabinet voting rights.

3a. What was the party of the president *after* the most recent election, regardless of whether the election was presidential?

3b. What was the party of the Prime Minister *after* the most recent election, regardless of whether the election was parliamentary?

3c. Report the number of cabinet ministers of each party or parties in cabinet, *after* the most recent election. (If one party holds all cabinet posts, simply write "all"). Ministers are considered those members of government who are members of the Cabinet and who have Cabinet voting rights. List also cabinet members that are 'independent'. If known, specify if the 'independents' are affiliated or close to certain parties.

Name of Political Party   Number of Cabinet Ministers

3d. What was the size of the cabinet *after* the election (total number of cabinet ministers detailed in 3c)? Please include only full ministers and the prime minister in the count. Ministers are considered those members of government who are members of the Cabinet and who have Cabinet voting rights.

4a. How many political parties received votes in the election? In this answer, we want political parties, not merely alliances or coalitions of political parties. Please include all parties that received votes, but do not include independents. Where coalitions are present, please count all member parties separately – for instance, a coalition of three parties would count as three parties in your answer, not as one party. Please provide separate information for elections held contemporaneously (e.g., legislative and presidential voting), when voters cast separate ballots.

4b. Please provide a source of data and link to a website with official, detailed election results (votes and seats) for all parties participating in the election. If the data is not available electronically, please provide the information in paper format if possible.

4c. Please list all parties who received at least 1% of the vote nationally, and the applicable electoral results for each, in the following table. Please indicate the source (even if the same as in Question 4b), and add additional rows to the table as necessary. Please provide party names both in English and in the original language, if applicable.

Source:

Party Name	Number of Votes	% of Vote	Number of Seats	% of Seats
Total				

4d. What was the voter turnout in the election? Please also provide an official Internet address (preferably) or other official source where this information is available.

4e. Please provide the following six statistics for the country at the time of the election studied, so that we may calculate voter turnout in various ways. Some definitions, where provided, are based on those developed by International IDEA. Please also provide an official Internet address (preferably) or other official source where this information is available.

- Total Population:

  Definition: The total population includes all inhabitants, of all ages, both citizens and non-citizens (including illegal aliens).

- Total Number of Voting Age Citizens:

  Definition: This number is meant to approximate the eligible voting population.

- Total Vote:

  Definition: The total number of votes cast in the relevant election. Total vote includes valid and invalid votes, as well as blank votes in cases where these are separated from invalid votes.

- Total Number of Invalid and Blank Votes:

- Voting Age Population:

  Definition: Includes all persons above the legal voting age.

- Number of Registered Voters:

  Definition: The figure represents the number of names on the voters' register at the time that the registration process closes, as reported by the electoral management body.

5. Ideological family of political parties. For this question, please use the same parties that were used in the CSES Module 4 respondent questionnaire, and label them the same way (A-I).

Party Name	Ideological Family
A.	
B.	
C.	
D.	
E.	
F.	
G.	
H.	
I.	

Ideological Party Families: (These are suggestions only. If a party does not fit well into this classification scheme, please provide an alternative and some explanation.)

(A) Ecology Parties	(F) Left Liberal Parties	(K) Independents
(B) Communist Parties	(G) Liberal Parties	(L) Agrarian Parties
(C) Socialist Parties	(H) Right Liberal Parties	(M) Ethnic Parties
(D) Social Democratic Parties	(I) Christian Democratic Parties	(N) Regional Parties
(E) Conservative Parties	(J) National Parties	(O) Other Parties

6a. Ideological Positions of Parties:

Please indicate party positions on a left-right dimension (in the expert judgment of the CSES Collaborator). Please use the same parties that were used in the CSES Module 4 respondent questionnaire, and label them the same way (A-I).

Party Name	Left										Right
	0	1	2	3	4	5	6	7	8	9	10
A.											
B.											
C.											
D.											
E.											
F.											
G.											
H.											
I.											

6b. If you have asked respondents to rank political parties on an alternative dimension, other than the left-right dimension, please also provide your own rankings of the parties on this dimension. Please use the same parties that were used in the CSES Module 4 respondent questionnaire, and label them the same way (A-I).

Name of dimension:

    Label for left hand position:

    Label for right hand position:

Party Name	Left										Right
	0	1	2	3	4	5	6	7	8	9	10
A.											
B.											
C.											
D.											
E.											
F.											
G.											
H.											
I.											

7. In your view, what are the five most salient factors that affected the outcome of the election (e.g. major scandals; economic events; the presence of an independent actor; specific issues)? Rank them according to their salience (1 = most salient).

      1.

      2.

      3.

      4.

      5.

8. Party Leaders and Presidential Candidates:

In the table below, report the leader(s) of each party. Please use the same parties that were used in the CSES Module 4 respondent questionnaire, and label them the same way (A-I). If candidates were endorsed by more than one party, please indicate this.

Party Name	Name of Party Leader	Name of Presidential Candidate, if appropriate
A.		
B.		
C.		
D.		
E.		
F.		
G.		
H.		
I.		

9a. Fairness of the Election

How impartial was the body that administered the election law?
    [ ] Very impartial
    [ ] Mostly impartial
    [ ] Not very impartial
    [ ] Not impartial at all

9b. Was there a formal complaint against the national level results?
    [ ] Yes
    [ ] No

9c. Were there irregularities reported by international election observers?
    [ ] Yes
    [ ] No
    [ ] No international election observers

9d. On what date was the election originally scheduled to be held?

9e. On what date was the election actually held? If different from 9d, please explain why.

10a. Election Violence

To what extent was there violence and voter or candidate intimidation during the election campaign and the election day?
　　　[ ] No violence at all
　　　[ ] Sporadic violence on the part of the government
　　　[ ] Sporadic violence on the part of opposition groups
　　　[ ] Sporadic violence on all sides
　　　[ ] Significant violence on the part of the government
　　　[ ] Significant violence on the part of opposition groups
　　　[ ] Significant violence of all sides

10b. If there was violence, was it geographically concentrated or national?
　　　[ ] Geographically concentrated
　　　[ ] National

10c. Post-Election (and election-related) Violence

To what extent was there violence following the election?
　　　[ ] No violence at all
　　　[ ] Sporadic violence on the part of the government
　　　[ ] Sporadic violence on the part of opposition groups
　　　[ ] Sporadic violence on all sides
　　　[ ] Significant violence on the part of the government
　　　[ ] Significant violence on the part of opposition groups
　　　[ ] Significant violence of all sides

10d. Post-Election (and election-related) Protest

To what extent was there protest following the election?
　　　[ ] No protest at all
　　　[ ] Sporadic protest
　　　[ ] Significant protest

Questions about the Possibilities of Electoral Alliance

Definitions: A joint list refers to one on which candidates of different parties run together. Apparentement refers to a legal agreement between two or more lists to pool their votes for the purposes of an initial seat allocation, with seats initially allocated to the alliance then reallocated to the lists in the alliance.

11. Joint Lists/Candidates

There are multiple types of electoral alliances/coalitions, but we are explicitly interested in those involving joint lists or candidates - i.e. those where parties compete as a unit during the election.

Is this type of electoral coalition legally allowable?
　　　[ ] Yes
　　　[ ] No

Is this type of electoral coalition used in practice, even if not legally allowable?
　　　[ ] Yes
　　　[ ] No

If "Yes" was answered to either of the above questions, then please complete the following table for the election at which the Module was administered. Please mention only alliances that received at least 1% of the vote nationally. Add additional lines to the table as necessary.

Alliance Name	Participating Parties (please indicate dominant members with an "*")
Alliance 1:	
Alliance 2:	
Alliance 3:	
Alliance 4:	
Alliance 5:	

12. If joint lists are possible, are they subject to different regulations than single-party lists? For example, higher thresholds, different numbers of candidates that may appear on the list, etc. (please mark all applicable responses)
    [ ] Yes, joint party lists must satisfy higher thresholds
    [ ] Yes, joint party lists may present different numbers of candidates
    [ ] Yes, joint party lists are subject to other regulations that are different from the
        regulations governing independent parties; please specify: _____
    [ ] No, joint parties are governed by the same rules as the other parties
    [ ] Not applicable; no joint party lists are allowed

13a. Is there apparentement or linking of lists?
    [ ] Yes
    [ ] No

13b. If apparentement is possible, what lists can participate in such agreements:
    [ ] lists of the same party in the same constituency
    [ ] lists of the same party from different constituencies
    [ ] lists of different parties in the same constituency

14a. Can candidates run with the endorsement of more than one party?
    [ ] Yes
    [ ] No

14b. If candidates can run with the endorsement of more than one party, is this reflected on the ballot?
    [ ] No
    [ ] No party endorsements are indicated on the ballot paper
    [ ] Yes, candidate's name appears once, together with the names of all supporting parties
    [ ] Yes, candidate's name appears as many times as there are different parties endorsing
        him/her, each time with the name of the endorsing party
    [ ] Yes, other; please explain: _____

<u>**Data on Electoral Institutions**</u>

If possible, please supplement this section with copies of the electoral law, voters' handbooks, electoral commission reports, and/or any other relevant materials.

Questions 15a through 21d must be repeated for each electoral tier (segment) of each directly elected house of the legislature.

Electoral Tier (Segment) and House

15a. In your answers for questions 15a through 21d, which electoral tier (segment) is being referred to? (Note: Countries with only one tier may skip this question.)

15b. In your answers for questions 15a through 21d, which house is being referred to (lower or upper)? (Note: Countries with only one tier may skip this question.)

Questions about Voting

16a. How many votes do voters cast or can cast? In systems where voters rank order the candidates, if there are 10 candidates (for example), the response to this question should be 10.

16b. Do they vote for candidates (not party lists) or party lists? (Note: Collaborators may select multiple answers, if appropriate.)
(Definition: Party bloc voting is used in multi-member districts where voters cast a single party-centered vote for their party of choice; the party with the most votes wins all of the district seats.)
    [ ] Candidates
    [ ] Party Lists
    [ ] Party Bloc Voting
    [ ] Other; please explain: _____

16c. How many rounds of voting are there?

16d. If there are lists, are they closed, open, flexible, or is there party bloc voting?
    [ ] Closed (order of candidates elected is determined by the party and voters are unable to
        express preference for a particular candidate)
    [ ] Open (voters can indicate their preferred party and their favored candidate within that
        party)
    [ ] Flexible (voters can allocate votes to candidates either within a single party list or across
        different party lists as they see fit)

17. Are the votes transferable?
(Definition: In systems with preferential voting, a voter can express a list of preferences. E.g., votes can be cast by putting a '1' in the column next to the voter's preferred candidate, a '2' beside their second favorite candidate and so on. Votes are counted according to the first preferences and any candidates who have achieved the predetermined quota are elected. To decide which of the remaining candidates are elected the votes are *transferred* from candidates who have more than the necessary number to achieve the quota and from the candidate with the least number of votes. An example of this is the election in Ireland in 2002.)

      [ ] Yes
      [ ] No

18. If more than one vote can be cast, can they be cumulated?
(Definition: Cumulative voting refers to systems in which voters are allowed to cast more than one vote for a single candidate.)

      [ ] Yes
      [ ] No

19. Is voting compulsory?
(Definition: Voting is compulsory if the law states that all those who have the right to vote are obliged to exercise that right.)

      [ ] Yes; Strictly Enforced Sanctions
      [ ] Yes; Weakly Enforced Sanctions
      [ ] Yes; Without Sanction for Violation
      [ ] No

20. Please list and describe any other features of voting that should be noted.

Questions about Converting Votes into Seats

21a. Are there legally mandated thresholds that a party must exceed before it is eligible to receive seats?
    [ ] Yes
    [ ] No

21b. If YES in Question 21a, what is the threshold?

21c. If YES in Question 21a, what is the unit for the threshold mentioned in Question 21b?
    [ ] Percent of total votes
    [ ] Percent of valid votes
    [ ] Percent of the total electorate
    [ ] Other; please explain: _____

21d. If YES in Question 21a, please specify to what House/ Electoral Tier (Segment) the threshold(s) apply.

Please repeat questions 15a through 21d for each electoral tier (segment) of each directly elected house of the legislature. Countries with only one tier should proceed to Question 22.

References

22. Please provide an official source for district-level election results. English language sources are especially helpful. Include website links or contact information if applicable.

23. Please list any resources that were consulted in the preparation of this report, or that the CSES community may find especially helpful in understanding the political system described. Include website links if applicable.

Appendix D

Data Bridging: Possibilities for accommodating merging of CSES data with other macro data sources

Memorandum prepared by the CSES Secretariat for the M5 PC Macro and District subcommittee

24 September 2015

Objective

This memorandum is for the CSES Module 5 PC macro and district variable subcommittee. The objective of this memo is to ascertain which country identifiers are used in five cross-national datasets namely the ACE Electoral Knowledge Network (provided by the IDEA), the Constituency-Level Elections Archive (CLEA), the Quality of Government Institute (QoG), the UNESCO, and the World Bank Database. While there are a multitude of alternative sources available, we chose five major sources of data for existing CSES data as test cases to explore the issue. Our mission is to provide information on which country identifiers are used in each dataset and how the CSES might accommodate bridging macro data not included in the CSES with these other sources.

ACE Electoral Knowledge Network[1]

The country identifier used by the ACE Electoral Knowledge Network is a variable 'country' containing the country name written out (e.g.: "Australia"). To accommodate data bridging, CSES might introduce another identification variable into its data detailing the official name of a country written out in the English language form to allow for easy merging with the ACE database.

Constituency-Level Elections Archive (CLEA)[2]

The country identifier used by the CLEA dataset is variable CTR_N' – 'Country Name' which contains the written out names of the countries. Furthermore, the variable 'CTR' – 'Country Code' contains the country codes assigned by the International Organization for Standardization (ISO).[3] The CSES similarly uses ISO codes (in Module 4 see variable D1006 – polity identifier). However, the CSEs adds another digit to the code to indicate the number of sampling components used within one country (or '0', if only one sampling component exists). To accommodate bridging CSES data with CLEA data, one could think about dropping this last digit of the current CSES variable, which would ensure the ISO codes in each dataset exactly matched. Alternatively, the CSES might just have a three digit ISO code identifier to ensure the data can be merged easily.

QoG – Quality of Government Database[4]

The Quality of Government datasets contain a range of different country identifiers. First, the numeric ISO Code is used in variable 'ccode'. Second, an alphabetic, three-letter country code based on the ISO-3166-1 alpha3 standard is used in variable 'ccodealp'. Third, country codes from the Correlates of War Project are used in variable 'ccodecow'. These have two to three digits and do not correspond to ISO Codes. Fourth, country codes from the World Bank are used in variable 'ccodewb'. These are the same codes as the ISO Codes, except that

---

[1] *http://aceproject.org/epic-en*
[2] *http://www.electiondataarchive.org/*
[3] *http://unstats.un.org/unsd/methods/m49/m49.htm*
[4] *http://qog.pol.gu.se/data*

Serbia and Montenegro are not assigned separate codes.[5] The fifth country identifier consists of the country name written out verbatim in the variable 'cname'. The above holds for all QoG datasets, except their dataset on EU Regional Data. In this dataset, only two variables identifying the country are used, namely both country and region names are spelled out in the variable 'Name'. Second, a number is assigned to each country and each region in the dataset in the variable 'number'. This classification does not seem to be based on any recognized standard.

Considering that CSES and the QoG uses the same ISO classifications, the scope here for further means of merging are limited. But as mentioned with the CLEA data, the PC might consider about dropping this last digit of the current CSES variable (D1006 in Module 4), which would ensure the ISO codes in each dataset exactly matched. Alternatively, the CSES could add a three digit ISO code identifier to ensure the data can be merged easily. Alternatively, the CSES also makes use of the alphabetic three letter country code based on the ISO-3166-1 alpha3 standard that is used in this dataset (for Module 4 see variable D1004). It is usually used in combination with the country's year of election study (for e.g. ALB_2005: ALBANIA (2005). The PC might consider creating a separate variable that contains only the three letter country code (e.g.: ALB in the case of Albania).

UNESCO[6]

The UNESCO Institute for Statistics provides a number of datasets on different themes. All of them use only a single country-variable, named 'Country', which contains the country name written out. As with the CLEA and the QoG datasets, there would seem to be merit in creating a variable within the CSES which just had the country name written out verbatim in English.

World Bank Databank[7]

The World Bank offers a wide range of data on a multitude of topics. Each slightly differs in which country identifiers are used but all include a variable that contains the written out country name verbatim in English. Furthermore, two alphabetical codes are used: a) the 3 letter ISO 3166-1 alpha-3 code and b) the 2 letter ISO 3166-1 alpha-2 code. However, the World Bank also states that it deviates from these codes for around ten countries.[8]

As with the above cases, a potential means of allowing easy bridging would be for the CSES to have a variable with the country name written out verbatim in English and/or to have a variable name with the 3-letter ISO classification of the country.

Summary

This probe has explored the country identifiers used by five leading sources of cross-national macro data. The key finding of the analysis has identified that most data sources contain a country identifier where the name of the country is written out verbatim in English. Many also contain a variable using the ISO 3 letter classification of a country, which the CSES also party employs. To allow easy bridging with the other datasets, the CSES PC might consider ensuring that a) there is a country identifier based on the name of the country and b) solely on their 3-letter ISO classification. Such a strategy would not add any additional burden to the CSES Secretariat.

---

[5] http://wits.worldbank.org/wits/wits/witshelp/Content/Codes/Country_Codes.htm
[6] http://data.uis.unesco.org/
[7] http://databank.worldbank.org/data/home.aspx
[8] http://data.worldbank.org/developers/api-overview/country-queries).