

Comparative Study of Electoral Systems (CSES) Module 5: Design Report (Sample Design and Data Collection Report)

September 14, 2016

Country: Great Britain
Date of Election: 8 June 2017

Prepared by: British Election Study
Date of Preparation: February 2021

NOTES TO COLLABORATORS:

- Where brackets [] appear, answer by placing an “X” within the appropriate bracket or brackets.
- If more space is needed to answer any question, please lengthen the document as necessary.

Collaborator(s):

Collaborators are the contact persons for election studies that appear in the CSES dataset - they are not necessarily the parties who collected the data. These collaborators and their contact information will be listed on the CSES website.

Name: Edward Fieldhouse Title: Professor Organization: University of Manchester Address: School of Social Sciences Humanities Bridgford Street M139PL E-Mail: ed.fieldhouse@manchester.ac.uk	Name: Jane Green Title: Professor Organization: Nuffield College, Oxford University Address: Nuffield College New Road OX11NF E-Mail: jane.green@nuffield.ox.ac.uk
Name: Geoffrey Evans Title: Professor Organization: Nuffield College, Oxford University Address: Nuffield College New Road OX11NF E-Mail: Geoffrey.evans@nuffield.ox.ac.uk	Name: Jonathan Mellon Title: Dr. Organization: University of Manchester Address: School of Social Sciences Humanities Bridgford Street M139PL E-Mail: jonathan.mellon@manchester.ac.uk

<p>Name: Chris Prosser Title: Dr. Organization: Royal Holloway, University of London</p> <p>Address: Egham Hill Egham TW20 0EX</p> <p>Email: chris.prosser@rhul.ac.uk</p>	
---	--

Data Collection Organization:

Organization that conducted the survey field work/data collection:

Organization: The data was collected by a consortium of three fieldwork agencies: Gfk, Kantar Public, and NatCen. Fieldwork was coordinated by Gfk:

Gfk UK Ltd

Address:
Level 18
25 Canada Square
Canary Wharf
London
E14 5LQ
UK

Telephone: +44 207 890 9830

E-Mail: nick.moon@gfk.com

Website: <http://www.gfk.com/en-gb/>

Funding Organization(s):

Organization(s) that funded the data collection:

Organization: Economic and Social Research Council, ES/K005294/1

Address:
Polaris House
North Star Avenue
Swindon
SN2 1UJ

Telephone: +44 1793 413000

Fax:

E-Mail: esrcenquiries@esrc.ac.uk

Website: www.esrc.ac.uk

Archiving Organization

If appropriate, please indicate the primary location where the full, original election study dataset (not just the CSES portion) will be archived:

Organization: UK Data Service
Address:
University of Essex
Wivenhoe Park
Colchester
Essex
C04 3SQ

Telephone: +44 1206 872001
Fax:
E-Mail: help@ukdataservice.ac.uk
Website: <https://www.ukdataservice.ac.uk/>

Please indicate the date when the study is expected to be available at this archive:

Study Design

1. Timing of the study that the CSES Module was included in:

- Post-Election Study (with interviewing starting within 6 months after the election)
- Post-Election Study (with interviewing starting more than 6 months after the election)
- Pre-Election/Post-Election Panel Study
- Between Rounds

2a. Date Post-Election Interviewing Began:

June 26th 2017

2b. Date Post-Election Interviewing Ended:

October 1st 2017

3a. Mode of interviewing for the post-election survey in which the CSES Module appeared:
(If multiple modes were used, please mark all that apply.)

- In person, face-to-face - using a questionnaire on paper
- In person, face-to-face - using an electronic/computerized questionnaire
- Telephone
- Mail or self-completion supplement
- Internet

3b. Was there a mode change *within* interviews (e.g., selected self-completion elements within the questionnaire)?

- No
- Yes; please provide details:

The main fieldwork for the survey was conducted in person using CAPI. Some CSES questions were included in the self-completion mail back survey. Respondents had the option of completing this section of the survey either via the internet or a mail-back paper questionnaire. See documentation deposited with UKDS for complete information.

4a. Was the survey part of a panel study?

- Yes
- No

4b. If the survey was part of a panel study, please describe the design of the panel study, including the date at which interviewing for each prior wave began and ended:

4c. If the survey was entirely or partly conducted via the Internet, please indicate whether it was based on an access panel (i.e. respondents were selected from a group of pre-screened panelists):

- Yes
- No

4d. If the survey was based on an Internet access panel, please describe the access panel (company, population [does it include persons without initial access to the Internet and how are they interviewed], method of recruiting members, total size of access panel, method of selecting survey respondents from the panel):

n/a

Translation

Please provide copies of questionnaires in all languages used as part of the election study deposit. For questionnaires in a language other than English, please also provide a version of each translated back into English. Note: Questions are based on those developed for the ISSP.

5. Was the questionnaire translated?

- Yes, translated by member(s) of research team
- Yes, by translation bureau
- Yes, by specially trained translator(s)
- No, not translated

6. Please list all languages used for the fielded module:

English

7a. If the questionnaire was translated, was the translated questionnaire assessed/checked or evaluated?

- Yes, by group discussion
- Yes, an expert checked it
- Yes, by back translation
- Other; please specify: _____
- No
- Not applicable

7b. If the questionnaire was translated, was the questionnaire pre-tested?

- Yes
- No
- Not applicable

7c. If the questionnaire was translated, were there any questions which caused problems when translating?

- Yes
- No
- Not applicable

7d. If the questionnaire was translated, please provide a list of all questions which caused problems when translating. For each question listed, describe what problems were encountered and how they were solved:

n/a

Sample Design and Sampling Procedures

8. Please describe the population that your sample is meant to be representative of:

The sample is representative of the population of 18+ living in Great Britain who were eligible to vote in the 2017 election.

Eligibility Requirements

9a. Must a person be a certain age to be interviewed?

Yes

No

If yes, what ages could be interviewed?

18

9b. Must a person be a citizen to be interviewed?

Yes

No

9c. Must a person be registered to vote to be interviewed?

Yes

No

9d. Please list any other interviewing requirements or filters used:

Only those eligible to vote in the 2017 election were interviewed. Those who are eligible to vote are British citizens, Citizens of the Republic of Ireland or citizen of a Commonwealth country with the right to remain in the UK. Interviewers had a card on the doorstep with these 3 conditions and asked the person on the doorstep whether there was anyone in the household who fit these criteria.

Sample Frame

10a. Were any regions of the country excluded from the sample frame?

Yes

No

If yes, what percent of the total eligible population did this exclude from the sample frame? Less than 0.25 %

If yes, please explain:

For practical reasons the northern islands were not included.

10b. Were institutionalized persons excluded from the sample?

Yes

No

If yes, what percent of the total eligible population did this exclude from the sample frame? ~2 %

If yes, please explain:

The sample is drawn from the small user Postcode Address file and some sections of the population fall outside this sampling frame e.g. elderly people in residential care, prisoners and military personnel living in defense establishments

10c. Were military personnel excluded from the sample?

Yes

No [but see above]

If yes, what percent of the total eligible population did this exclude from the sample frame? _____ %

If yes, please explain:

10d. If interviews were conducted by telephone, what is the estimated percentage of households without a phone? _____ %

Please explain:

10e. If interviews were conducted by telephone, were unlisted telephone numbers included in the population sampled?

Yes

No

If no, what percent of the total eligible population did this exclude from the sample frame? _____ %

10f. If interviews were conducted via the Internet, what is the estimated percentage of households without access to the Internet? _____ %

10g. If interviews were conducted via the Internet, were provisions taken to include members of the population without access to the Internet? And if so, which?

Yes

No

If "Yes", please explain:

If "No", what percent of the total eligible population did this exclude from the sample frame? _____ %

10h. Were other persons excluded from the sample frame?

Yes

No

If yes, what percent of the total eligible population did this exclude from the sample frame? _____ %

If yes, please explain:

10i. Please estimate the total percentage of the eligible population excluded from the sample frame: ____~2____ %

Sample Selection Procedures

11. Please describe, in your own words, how the sample for the study was selected. If the survey is part of a panel study and/or based on an Internet access panel, please also describe the original sample, from the beginning of the study.

The sample was selected based on a multi-stage design, summarised as:

- 1) Stratified random sample of 234 Parliamentary constituencies
- 2) Two Lower Layer Super Output Areas (LSOAs) per constituency were selected with probability proportional to size
- 3) Selection of addresses from the Small user Postcode Address File (PAF)
- 4) One individual randomly selected per address by the interviewer

12a. What were the primary sampling units?

Parliamentary constituencies

12b. How were the primary sampling units selected?

The study was based in 234 Parliamentary constituencies, sampled from the whole of Great Britain (from 630 constituencies - excluding the seat held by The Speaker, and (for practical reasons) Orkney and Shetland).

At the first stage the constituencies were stratified by country and then (within England) by Region.

Within each country/region, constituencies were classified by party competition, defined as a combination of winning party and party competition from the 2010 election.

The final stage of stratification was to sort the constituencies within each cell from the least to the most marginal. The constituencies were then selected with probability proportional to population size.

12c. Were the primary sampling units randomly selected?

- Yes
 No

Please explain how the units were randomly selected. If the units were not randomly selected, please provide a justification for why the units were not randomly selected.

See 12b

13. Were there further stages of selection?

Yes

No

13a. If there were further stages of selection, what were the sampling units at each of the additional stages?

Two LSOAs per constituency were selected with probability proportional to size

Addresses were selected within each LSOA from the Small user Postcode Address File (PAF)

13b. If there were further stages of selection, how were the sampling units selected at each of the additional stages?

LSOAs

LSOAs were used as secondary sampling units in each sampled constituency. Because there is not a perfect match between LSOAs and constituencies (some LSOAs straddle two constituencies) the LSOAs were treated as being part of the constituency in which the majority of its population live.

All LSOAs were ranked in each constituency by their Index of Multiple Deprivation (IMD) score, and then divided into quartiles. Within each quartile the LSOAs were listed from lowest to highest population density.

Two LSOAs were selected with probability proportional to size. Rather than used a sampling interval, two random numbers – one between 1 and the total population of the top two quartiles, and the other between 1 and the total population of the lower two quartiles were used for selection.

Addresses

Addresses were sampled from the latest version of the PAF using a fixed sampling interval and random start.

13c. If there were further stages of selection, were units at each of these stages randomly selected?

Yes

No

Please explain how the units were randomly selected. If the units were not randomly selected, please provide a justification for why the units were not randomly selected.

See 13b

14a. How were individual respondents identified and selected in the final stage?

At each address the interviewer established the number of dwellings, then households, and finally people aged 18 or over who are eligible to vote in the general election. At each of these levels, if there was more than one present, one was selected by the interviewer using a Kish grid, randomised for each address.

14b. Could more than one respondent be interviewed from a single household?

Yes

No

If yes, please explain:

15. Did the sample design include clustering at any stage?

Yes

No

If yes, please describe:
as noted in 12b, the sample was clustered in 300 of 630 constituencies.

16. Did the sample design include stratification?

Definition: Stratification involves the division of the population of interest according to certain characteristics (for instance: geographic, political, or demographic). Random selection then occurs within each of the groups that result.

Yes

No

If yes, please describe (please include the list of characteristics used for stratification, and in the case of multi-stage selection processes the stage[s] at which stratification occurred):

Geographic stratification, for details see answer to 12b and 13b

17. Was quota sampling used at any stage of selection?

Yes

No

If yes, please describe:

18. Was substitution of individuals permitted at any stage of the selection process or during fieldwork?

Yes

No

If yes, please describe:

19. Under what circumstances was a household designated non-sample? Please check all that apply:

Non-residential sample point

All members of household are ineligible

Housing unit is vacant

No answer at housing unit after _____ callbacks

Other (Please explain):

20. Were non-sample replacement methods used?

Yes

No

Please describe:

21a. For surveys conducted by telephone, was the sample a random digit dial (RDD) sample?

Yes

No

21b. For surveys conducted by telephone, was the sample a listed sample?

Yes

No

21c. For surveys conducted by telephone, was the sample a dual frame sample?

Yes

No

If yes, what % list frame _____ and what % RDD _____

22. For surveys conducted by mail, was the sample a listed sample?

Yes

No

Please describe:

23. For surveys conducted on the Internet, did respondents self-select into the survey, at any stage?

Yes

No

Please explain:

Incentives

24a. Prior to the study, was a letter sent to the respondent?

Yes

No

(If yes, please provide a copy of the letter.)

24b. Prior to the study, was a payment sent to the respondent?

Yes

No

If yes, please describe (including amount of payment):

24c. Prior to the study, was a token gift sent to the respondent?

Yes

No

If yes, please describe:

24d. Did respondent receive an additional payment after their participation? (Do not include any payment made prior to the study.)

Yes

No

If yes, please describe (including amount of payment):

A gift voucher: £25 for 18-24s, £20 for addresses in London with respondents aged 25+ and £10 for addresses outside of London with respondents aged 25+

24e. Were any other incentives used?

Yes

No

If yes, please describe:

Interviewers

25. Please describe the interviewers (e.g., age, level of education, years of experience):

Interviewing was carried out by fully trained and experienced GfK interviewers: in total 208 interviewers were used to conduct the study.

26. Please provide a description of interviewer training. If possible please differentiate between general interviewer training and study-specific components:

Interviewers received extensive study specific training. All interviewers were provided with written instructions which gave a detailed explanation of all aspects of the study. A master briefing session was held via webinar by GfK researchers and attended by researchers from Kantar Public and NatCen, SLT members and experienced GfK interviewers. The session covered the background to the survey and further important aspects of interviewers' job; this session also included a run through of the questionnaire. Kantar Public and NatCen then conducted their own interviewer briefings.

26a. Please provide a description of the content, structure and time used for general training of interviewers:

26b. Please provide a description of the content, structure and time used for training interviewers in the specifics of the study within which CSES was run:

Webinar:

- Background
- Achieving high response rate
- Making contact/eligibility/ respondent selection
- Key questionnaire issues
- Questionnaire run through

Written instructions:

- Provides further detail on issues explored in the webinar
- Interviewer completed a practice interview on their CAPI machine in conjunction with the written notes about the questions

Contacts

27a. What was the average number of contact attempts made per household, for the entire sample?

Not available

27b. For households where contact was made, what was the average number of contact attempts prior to first contact?

Mean = 4.17, median = 3

27c. During the field period, how many contacts were made with the household before declaring it a **non-sample**?

This varied significantly - a vacant lot would be found at first call, but one that seemed occupied but was later found to be empty could take many calls

28d. During the field period, how many contacts were made with the household before declaring it a **non-interview**?

Again this varied depending on the circumstances, a refusal could be at first call or 10th call, and a non-contact must be on at least the 8th call

28e. During the field period, what were the maximum number of days over which a household was contacted?

28f. During the field period, did interviewers vary the time of day at which they re-contacted the household?

Yes

No

If yes, please describe:

Minimum number of two calls had to be made at each of weekday daytime, weekday evenings and weekends

Refusal Conversion

29a. Were efforts made to persuade respondents who were reluctant to be interviewed?

Yes

No

Please describe:

Interviewers were given training on how to respond to reluctance on the doorstep, and given various response types to use to answer respondent objections to taking part.

29b. Were respondents who were reluctant to be interviewed sent a letter persuading them to take part?

Yes

No

(If yes, please provide a copy of the letter or letters.)

If yes, please describe:

29c. Was payment offered to respondents who were reluctant to take part?

Yes

No

If yes, how much?

Initial non-contacts were reissued to a second interviewer and the incentives increased, from £5 to £15 depending on age and location

29d. Were respondents who were reluctant to take part turned over to a more experienced interviewer?

Yes

No

29e. What was the maximum number of re-contacts used to persuade respondents to be interviewed?

29f. Were any other methods used to persuade respondents reluctant to be interviewed to take part?

Yes

No

If yes, please describe:

Up to 3 reminder letters/emails were sent to those who agreed to complete the mail back portion of the survey

Interview/Survey Verification

Definition: Interview/survey verification is the process of verifying that an interview was conducted and that the survey was administered to the correct respondent, for quality control purposes.

30. Was interview/survey verification used?

Yes

No

If yes, please describe the method(s) used:

If yes, please indicate the percent of completed surveys that were verified: _____ %

Response Rate

Note: If multiple modes of interviewing were used for the post-election survey in which the CSES Module appeared, please repeat the following questions as appropriate for each of the modes used.

31. What was the response rate of the survey that the CSES Module appeared in? Please show your calculations. (If the CSES Module appeared in a panel study, please report the response rate of the first wave of the study, even if the CSES Module did not appear in that wave.)

32. Please provide the following statistics for the survey that the CSES Module appeared in. (If the CSES Module appeared in a panel study, please report the statistics for the first wave of the study, even if the CSES Module did not appear in that wave.)

	N
Total issued addresses	4866
Not eligible	
Out of scope properties	400
No eligible respondents in household i.e. not eligible to vote in the general election	170
Unknown eligibility, non-interview	
Refused before screening stage	744
Non contacts	388
Other unproductive before screening stage	32
<i>Total unknown eligibility, non-interview cases</i>	<i>1164</i>
<i>Total unknown eligibility, non-interview cases who are expected to have someone in the household eligible for the study (A)</i>	<i>1123</i>
Eligible, non-interview	
Refused	609
Non contacts	238
Other unproductive	319
<i>Total eligible, non-interview cases (B)</i>	<i>1166</i>
Full interviews (C)	1966
Total eligible (A+B+C)	4255
Main study response rate (C/(A+B+C))	46.20%
Issued re-contact sample (D)	249
Interviews from re-contact sample (E)	228
Response rate with re-contact sample included ((C+E)/(A+B+C+D))	48.71%
Completed CSES module cases (F)	984
CSES module response rate (F/(C+E))	44.85%

33. If the CSES Module appeared in a panel study, how many waves were conducted prior to the wave that included the CSES Module?

34. If the CSES Module appeared in a panel study, what was the total panel attrition between the first wave of the study and the wave that included the CSES Module? Please show your calculations.

35. If the CSES Module appeared in a panel study, please provide the number of completed interviews for the wave that included the CSES Module:

36. If the CSES Module appeared in a panel study, please provide the following statistics for panel attrition by age and education. In each cell, indicate the percent of all completed interviews in each category for the indicated wave.

Age	First wave of study	Wave that included CSES
18-25	%	%
26-40	%	%
41-64	%	%
65 and over	%	%

Education	First wave of study	Wave that included CSES
None	%	%
Incomplete primary	%	%
Primary completed	%	%
Incomplete secondary	%	%
Secondary completed	%	%
Post-Secondary Trade/Vocational	%	&
University incomplete	%	%
University degree	%	%

Post-Survey Adjustment Weights

37. Are weights necessary to make the sample representative of the population being studied?

Yes

No

If yes, please explain:

To ensure that the respondents who took part in the study represent the views of the population (18+ adults in Great Britain who are eligible to vote) the data collected were weighted. There were two weights which were applied: initially selection weights to correct for unequal selection probabilities and secondly post-stratification weights which account for differing levels of response from various groups in the population.

38. Are weights included in the data file?

Yes

No

39. If weights are included in the data file, please describe in detail how the weights were constructed:

See answers at 40a and 40b

40a. If weights are included in the data file, are the weights designed to compensate for disproportionate probability of selection?

Yes

No

If yes, please describe:

These weights need to be applied to correct for unequal selection probabilities; during the selection process this happened at the following points:

1. If a selected address on PAF contains a number of separate dwellings (typically flats) and the interviewer has to select one of the dwellings for interview
2. If a dwelling contained more than one household (a household is defined as people who share a living room or who have common catering for at least one meal a day) and one of these households has to be selected
3. If a selected household contains more than one eligible person and one person has to be randomly selected for interview

At all these levels, people living at addresses with multiple dwellings/households/people have less of a chance of selection than a person living alone, and weighting is needed to compensate for this. To calculate a person's chance of being interviewed: the number of number of

dwellings was multiplied by the number of households within the selected dwelling which is in turn multiplied by the number of adults in the selected household. The probability of selection is the inverse of this number, and so to correct for it we simply need to weight by the result of the multiplication.

It was decided to cap the selection weight at 5 – a range of possible caps were tested and this was felt to have the least impact on results for key questions whilst also increasing overall effective size and reducing the impact of any individual with extreme weights. Only 14 cases were affected by this cap. After the selection weight was capped it was rescaled to arrive at the original sample size.

40b. If weights are included in the data file, are the weights designed to match known demographic characteristics of the population?

Yes

No

If yes, please describe:

A number of demographics were considered for the non-response weighting, and it was decided that the demographics that should be corrected were age, gender and region. The targets for these were taken from the Office for National Statistics (ONS) mid-year population estimates and weights were calculated after the selection weights had been applied.

40c. If weights are included in the data file, are the weights designed to correct for non-response?

Yes

No

If yes, please describe:

40d. If weights are included in the data file, are the weights designed to correct to the official election results?

- Yes
 No

If yes, please describe:

41. Comparison of Completed Interviews to Population (please provide as percentages of the total):

Characteristic	Completed Interviews		
	Population Estimates	Unweighted Distribution	Weighted Distribution
<u>Age</u>			
18-25	%	%	%
26-40	%	%	%
41-64	%	%	%
65 and over	%	%	%
<u>Education</u>			
None	%	%	%
Incomplete Primary	%	%	%
Primary Completed	%	%	%
Incomplete Secondary	%	%	%
Secondary Completed	%	%	%
Post-Secondary Trade/ Vocational	%	%	%
University Incomplete	%	%	%
University Degree	%	%	%
<u>Gender</u>			
Male	%	%	%
Female	%	%	%

42. Please indicate the source of the population estimates in the prior question. English language sources are especially helpful. Include website links or contact information if applicable.