

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

September 14, 2016

Country: Italy
Date of Election: March 4, 2018

Prepared by: Federico Vegetti (with the help of Demetra Opinioni.net)
Date of Preparation: May 20, 2018

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: Paolo Segatti Title: Professor Organization: Dipartimento di Scienze Sociali e Politiche, Università degli Studi di Milano Address: Via Conservatorio, 7, 20122 Milano MI Telephone: Fax: E-Mail: paolo.segatti@unimi.it Website:	Name: Federico Vegetti Title: Postdoctoral Research Fellow Organization: Dipartimento di Scienze Sociali e Politiche, Università degli Studi di Milano Address: Via Conservatorio, 7, 20122 Milano MI Telephone: Fax: E-Mail: federico.vegetti@unimi.it Website: federicovegetti.github.io
Name: Title: Organization: Address: Telephone: Fax: E-Mail: Website:	Name: Title: Organization: Address: Telephone: Fax: E-Mail: Website:

Data Collection Organization:

Organization that conducted the survey field work/data collection:

Organization: Demetra Opinioni.net
Address: Via Andrea Costa 34/c Venice (VE), Italy

Telephone: 041982429
E-Mail: info@opinioni.net
Website: www.opinioni.net

Funding Organization(s):

Organization(s) that funded the data collection:

Organization:
Address:

Telephone:
Fax:
E-Mail:
Website:

Organization:
Address:

Telephone:
Fax:
E-Mail:
Website:

Organization:
Address:

Telephone:
Fax:
E-Mail:
Website:

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: Address: Telephone: Fax: E-Mail: Website:
--

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:

March 8, 2018

2b. Date Post-Election Interviewing Ended:

May 2, 2018

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:

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):

The web panel of Demetra Opinioni.net includes 13734 panelists with access to internet, recruited by CATI mode (Computer Assisted Telephone Interviewing).

The panelists were selected by constituency; they were invited by e-mail and, after clicking the link to the questionnaire in the e-mail, they could answer the questions. The questionnaire was self-administered.

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:
Italian

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:

Q05a: The concept of "ethnic minority" in Italy commonly refers to Italian groups living near the borders to Austria, France, Slovenia. They are usually referred to as "linguistic minorities". Otherwise, the concept of "ethnic minority" is not much used in Italy --the most used term in common language would be "foreigner", even though it might apply to people who are by all

means Italian citizens. We nevertheless used the term "ethnic minority" ("minoranze etniche") for consistency with other CSES waves.

Q09: The reference to the capital city in the English version of the questionnaire was dropped as it might have been mistaken as a question about the city government in Rome (which has been under the spotlight during the campaign).

Sample Design and Sampling Procedures

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

Italian residents of voting age (18 or more)

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:

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? 1.53%

If yes, please explain:

The region Valle d'Aosta/Vallée d'Aoste, and the province of Bolzano/Bozen were excluded from the sample as the citizens there voted with a slightly different electoral system. Additionally, given the 2-stage nature of the sampling procedure (we sampled electoral districts at stage 1 and respondents within districts at stage 2 excluding large cities; within the large cities we directly sampled respondents), people living in districts which included a large city but who did not live within the municipality of the large city ended up excluded from the sample. All in all we only ended up excluding 1.5% of the population. With "large cities" we refer to all cities which included more than 1 electoral district within their municipality (Bari, Bologna, Florence, Genova, Milan, Naples, Palermo, Rome, and Turin).

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? _____ %

If yes, please explain:

10c. Were military personnel excluded from the sample?

Yes

No

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? 2.5%

Please explain:

Source: <https://www4.istat.it/it/files/2012/04/primocapitolo.pdf>

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? 38.7 %

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:

The questionnaire was partly administered by CATI (Computer Assisted Telephone Interviewing) and CAMI (Computer Assisted Mobile Interviewing) mode

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: _____ %

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.

At the first stage, a sample of single-member districts was drawn, stratified by the population in 7 macro-areas, excluding 9 large cities which had more than one district in their territory. At the second stage, a random sample of residents was drawn from within each district and from the 9 cities excluded from the first-stage. The sampling at the first stage produced 92 clusters, 83 districts and 9 large cities. The number of districts drawn from each macro-area is proportional to the population of voting age in 2017, based on estimates from the Italian National Institute of Statistics, and divided as follows: 7 districts from the North-West regions (consisting of the regions Liguria and Piedmont, excluding Valle d'Aosta/Vallée d'Aoste); 14 districts from Lombardy; 11 districts from the North-East regions (Friuli-Venezia Giulia, Trentino-Alto Adige/Südtirol without the province of Bolzano/Bozen, and Veneto); 16 districts from the "Red" regions (Emilia-Romagna, Marche, Tuscany, and Umbria); 8 districts from the Central regions (Abruzzo, Lazio, Molise); 18 districts from the Southern regions (Basilicata, Calabria, Campania, Puglia); 9 districts from the Islands (Sardinia, Sicily). The region Valle d'Aosta/Vallée d'Aoste and the province of Bolzano/Bozen were excluded from the sampling because the citizens there voted under a slightly different electoral system. The 9 large cities excluded from the first-stage sampling are Bari, Bologna, Florence, Genova, Milan, Naples, Palermo, Rome, and Turin. At the second stage, 20 observations were drawn from within each district and within each large city, with the exception of Milan (where 60 observations were drawn), Naples (40 observations), Rome (100 observations) and Turin (40 observations).

12a. What were the primary sampling units?

Districts in all country except for the 9 large cities, individuals within the 9 large cities

12b. How were the primary sampling units selected?

We sampled districts randomly from 7 macro areas, with the number of districts sampled within each macro area being proportional to the population of voting age (18yo or more).

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.

An R routine was developed to sample randomly N district names from within each macro area (N being proportional to the population of age 18+ within the macro area).

To select respondents from within the 9 large cities, random digit dialing based on the

national phone book was used for landline telephone interviews, and random digit dialing was used for mobile interviews.

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?

Individuals

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

For telephone interviews we used random digit dialing or the phone book.

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 answers above

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

See answers above

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:

Being a 2-stage sample, individuals are clustered within district by design

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):
The population of age 18 or more in 7 macro areas (see question 11).

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 12 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 66% and what % RDD 34%

23. For surveys conducted on the Internet, did respondents self-select into the survey, at any stage?
 Yes
 No

Please explain:

They opted in in the panel in the first place

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):

1€ (in the form of a telephone top-up or Amazon voucher) for the CAWI respondents

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):

The interviewers were 21 years or more, with 2 or more years of experience in the field of survey administration, and their level of education was at least high school diploma

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

There is a general interviewing training that includes 19 videos of 15 minutes each with questions at the end of each videos.

There is also a study-specific training that includes 2 hours of training on the questionnaire: in this session all questions are explained to the interviewers.

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

See point 26

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

See point 26

Contacts

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

2.67

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

0.83

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

0.96 for being classified as "non-quota" (e.g. businesses phone numbers), 12 for being classified as "not reachable."

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

1.71

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

10 days

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:

Respondents could choose the time when they wanted to be called back.

Refusal Conversion

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

- Yes
 No

Please describe:

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?

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:

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.)

CATI

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.)

A. Total number of households in sample:
 25180

B. Number of valid households:
 23433

C. Number of invalid (non-sample) households:
Non-quota 1252

D. Number of households of unknown validity:
 495

E. Number of completed interviews:	1000
F. Number of partial interviews:	846
G. Number of refusals and break-offs:	14672
H. Number non-contact (never contacted):	0
I. Other non-response:	6915

The sum of B+C+D should equal the value of A. If not, please describe why:

If statistic D (number of households of unknown validity) has a value greater than zero (0), please estimate the proportion of households of unknown validity that are valid:
Can't estimate

The sum of E+F+G+H+I should equal the value of B. If not, please describe why:

If statistic I has a value greater than zero (0), please describe what cases fall into this category:

All call-backs at the end of the survey: manual call backs, non response phone, busy line, answering machine.

CAMI

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.)

A. Total number of households in sample:
14533

B. Number of valid households:
13013

C. Number of invalid (non-sample) households:
Non-quota: 407

D. Number of households of unknown validity:
1113

E. Number of completed interviews:	500
F. Number of partial interviews:	689
G. Number of refusals and break-offs:	6001
H. Number non-contact (never contacted):	0
I. Other non-response:	5823

The sum of B+C+D should equal the value of A. If not, please describe why:

If statistic D (number of households of unknown validity) has a value greater than zero (0), please estimate the proportion of households of unknown validity that are valid:
Can't estimate

The sum of E+F+G+H+I should equal the value of B. If not, please describe why:

If statistic I has a value greater than zero (0), please describe what cases fall into this category:

All call-backs at the end of the survey: manual call backs, non response, busy line, answering machine.

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 populated being studied?

Yes

No

If yes, please explain:

Design weights + post-stratification weights + political weights are included.

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:

The design weight is calculated in order to correct for the different probabilities of selection of clusters within different macro-areas, and for the different probabilities of selection of individuals within different clusters (see explanation of the sampling procedure above).

The demographic weight is computed in order to match the sample to the population frequencies by gender (male, female), age group (18-44, 45-59, 60+), geographical area (north, center, south), and education (low, middle, high). Population frequencies have been obtained from the Italian national institute of statistics (ISTAT, see [here](#)).

The political weight is computed in order to match the sample to the population frequencies by vote shares of the major coalitions (plus non-voters) at the regional level. Population frequencies are based on the official election results.

The design weight was calculated by hand, the demographic and political weights were computed using the "survey" package for R. All the syntax to produce the weights is available upon request.

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:

The design weight is calculated in order to correct for the different probabilities of selection of clusters within different macro-areas, and for the different probabilities of selection of individuals within different clusters (see explanation of the sampling procedure above).

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:

The demographic weight is computed in order to match the sample to the population frequencies by gender (male, female), age group (18-44, 45-59, 60+), geographical area (north, center, south), and education (low, middle, high). Population frequencies have been obtained from the Italian national institute of statistics (ISTAT, see [here](#)).

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:

The political weight is computed in order to match the sample to the population frequencies by vote shares of the major coalitions (plus non-voters) at the regional level. Population frequencies are based on the official election results.

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

Characteristic	<u>Completed Interviews</u>		
	<u>Population Estimates</u>	<u>Unweighted Distribution</u>	<u>Weighted Distribution</u>
<u>Age</u>			
18-44	38.5%	33.9%	38.5%
45-59	27.7%	30.7%	27.7%
60+	33.9%	35.4%	33.9%
<u>Education</u>			
Low (up to lower secondary)	50.0%	50.7%	50.0%

Middle (upper secondary, post-secondary non-tertiary)	35.5%	26.1%	35.5%
High (university degree or higher)	14.5%	23.2%	14.5%
<u>Gender</u>			
Male	48.1%	47.9%	48.1%
Female	51.9%	52.1%	51.9%

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. See: <http://dati.istat.it/Index.aspx?lang=en&SubSessionId=0adcaa71-7e36-4588-a620-b1a5a27a88d5>

The data on the ISTAT website come from the Labor Force Survey, and refer to the year 2017.