Within-household selection of respondents: The last step of sampling in household surveys

Achim Koch

(GESIS – Leibniz-Institute for the Social Sciences, Germany)

Timothy P. Johnson

(University of Illinois at Chicago)

Outline

- The problem

- Within-household selection methods

- Crossnational surveys: The ESS as an example

- Special challenges in developing countries

- A few basic recommendations

The problem (1)

- Countries where no population register available

  => Sample of households (hhs) as intermediary
       step to arrive at a sample of persons

- Within hh-selection of respondents in F-t-F-surveys:
  Interviewer + hh-informant involved

- Decentralized sampling operations during data collection

The problem (2)

Aim:
- Probability sampling from all eligible hh-members
- Avoiding coverage errors / misselection

Needed:
- Clear definition of a hh / residence rules
- Random sampling method

Challenge:
- Trade-off coverage vs. nonresponse error?

Selection methods: Kish (1)

- "Gold standard" for F-t-F surveys
- Listing of all eligible hh-members in a pre-specified order
- Random selection: selection tables or CAPI-software

Criticism:

- Listing burdensome, time-consuming, intrusive?
- Negative effect on rapport interviewer-respondent?
- Reducing coverage problems at the expense of increasing nonresponse problems?

Selection methods: Birthday (2)

- Less intrusive alternative; faster to implement

- Widely used in telephone surveys

- No listing of hh-members

- Ask for the number of eligible hh-members +
  "Which of them had most recent / has next birthday?"

Criticism:

- Not completely random

- Day/month of birth of all hh-members not always known

- Selection errors

- Difficult to validate

Selection methods: New approaches (3)

Aim:

- Minimize burden (restrict number of cases with full listing)
- Stick to random selection

Procedure:

- Differentiating selection schemes according to hh-size

Examples:

- Rizzo et al. 2004 (if average hh-size is small)
- Le et al. 2013 (if average hh-size is large)

Selection methods: non-probability (4)

Methods / procedures:

- Troldahl/Carter (1964) + variants
  (oldest/youngest male/female hh-member)

- Restricting selection on persons at home when selection is made

- Convenience method:
  select any eligible person who is available + willing

=> Not an option for surveys aiming at high quality standards!

Selection methods in crossnational surveys

A few examples:

- Programme for the Int. Assessment of Adult Competencies (PIAAC)
  Kish (birthday methods explicitly not allowed)

- European Quality of Life Survey (EQLS)
  (Next) birthday method
  (interviewers recommended to list day/month of birth of all hh-members)

- Afrobarometer
  Modified Kish
  (alternately interviewing a man and a woman; listing only for respective gender)

- Gallup World Poll
  Kish or (last) birthday method

- European Social Survey (ESS)
  Kish or (last/next) birthday method

ESS as an example (1)

- Fielded every 2 years since 2002

- Target population: residents in private households aged 15 years and older

- 36 countries participated (at least once) in the first six surveys rounds

- Pooling data across the first 6 rounds: 153 cases
                                                                     (= country-round-combinations)

- Sample of individuals from a register: 70 cases

- Sample of households or addresses: 83 cases
   => among which: 28 cases Kish
                                                                 41 cases last birthday
                                                                 14 cases next birthday

ESS as an example (1)

- Fielded every 2 years since 2002

- Target population: residents in private households aged 15 years and older

- 36 countries participated (at least once) in the first six surveys rounds

- Pooling data across the first 6 rounds: **153 cases**
                                       (= country-round-combinations)

- Sample of individuals from a register: 70 cases

- Sample of households or addresses: 83 cases
  => among which: 28 cases Kish
                                       41 cases last birthday
                                       14 cases next birthday

ESS as an example (1)

- Fielded every 2 years since 2002

- Target population: residents in private households aged 15 years and older

- 36 countries participated (at least once) in the first six surveys rounds

- Pooling data across the first 6 rounds: 153 cases
  (= country-round-combinations)

- Sample of individuals from a register: **70 cases**

- Sample of households or addresses: 83 cases
   => among which: 28 cases Kish
  41 cases last birthday
  14 cases next birthday

ESS as an example (1)

- Fielded every 2 years since 2002

- Target population: residents in private households aged 15 years and older

- 36 countries participated (at least once) in the first six surveys rounds

- Pooling data across the first 6 rounds: 153 cases
                                            (= country-round-combinations)

- Sample of individuals from a register: 70 cases

- Sample of households or addresses: **83 cases**
   => among which: 28 cases Kish
                                            41 cases last birthday
                                            14 cases next birthday

ESS as an example (1)

- Fielded every 2 years since 2002

- Target population: residents in private households aged 15 years and older

- 36 countries participated (at least once) in the first six surveys rounds

- Pooling data across the first 6 rounds: 153 cases
  (= country-round-combinations)

- Sample of individuals from a register: 70 cases

- Sample of households or addresses: 83 cases
  => among which: **28 cases** Kish
  41 cases last birthday
  14 cases next birthday

ESS as an example (2)

- Sample quality criteria:
  misrepresentation of gender among a subset of respondents (Kohler 2007)

- HHs with a gender heterogeneous couple:
  the male + female partner should have the same chance of being interviewed

- Among the respondents from such couples
  we should expect a proportion of 50% females (and 50% males)

- Bias = difference from 50% female, divided by the S.E. of the estimate

  Bias = (% female – 50) / sqrt [(50 * 50) / n],
  with n = number of respondents from gender heterogeneous couples

- Statistic follows normal distribution with critical values > |1.96|

ESS as an example (2)

- Sample quality criteria:
  misrepresentation of gender among a subset of respondents (Kohler 2007)

- HHs with a gender heterogeneous couple:
  the male + female partner should have the same chance of being interviewed

- Among the respondents from such couples
  we should expect a proportion of 50% females (and 50% males)

- Bias = difference from 50% female, divided by the S.E. of the estimate

  Bias = (% female – 50) / sqrt [(50 * 50) / n],
  with n = number of respondents from gender heterogeneous couples

- Statistic follows normal distribution with critical values > |1.96|

ESS as an example (2)

- Sample quality criteria:
  misrepresentation of gender among a subset of respondents (Kohler 2007)

- HHs with a gender heterogeneous couple:
  the male + female partner should have the same chance of being interviewed

- Among the respondents from such couples
  we should expect a proportion of 50% females (and 50% males)

- Bias = difference from 50% female, divided by the S.E. of the estimate

  Bias = (% female – 50) / sqrt [(50 * 50) / n],
  with n = number of respondents from gender heterogeneous couples

- Statistic follows normal distribution with critical values $> |1.96|$

ESS as an example (3)

Over-/underrepresentation of females, by type of sample + within hh-selection method (ESS 1 – 6; 153 country-round-combinations)

$$Bias = (\%female - 50)/sqrt[(50*50)/n]$$
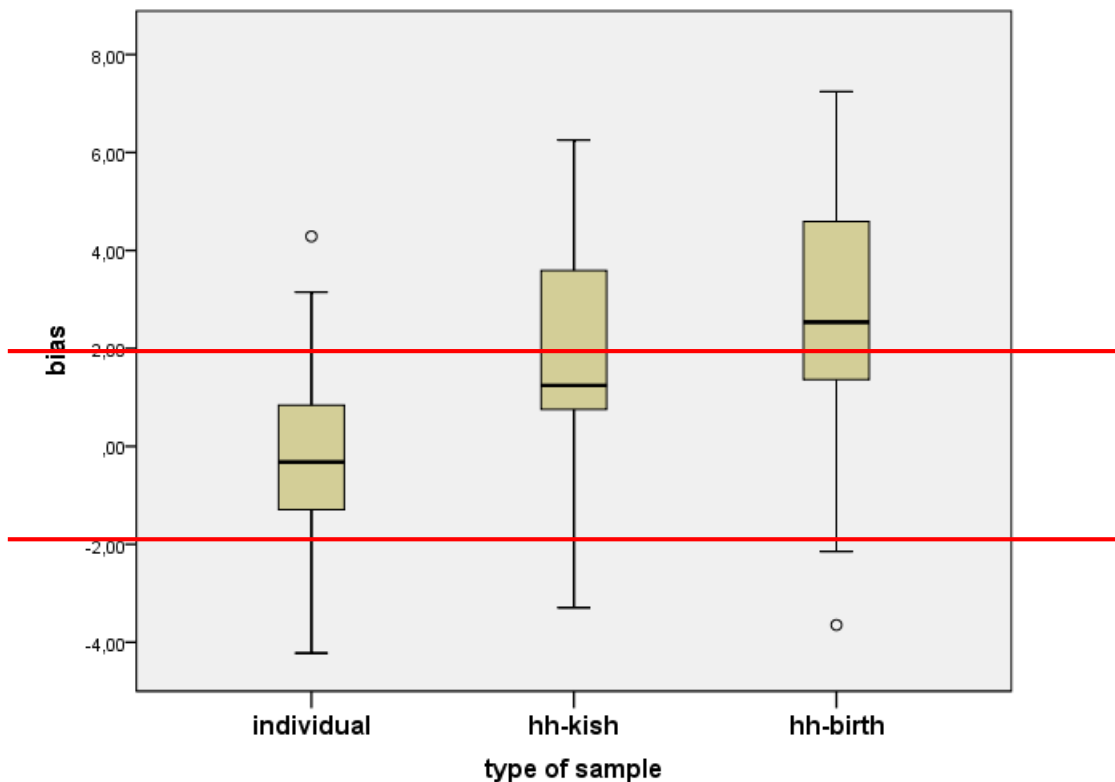
Sample of individuals:
n = 70 cases from 15 different countries;
13 cases = 18.6%: bias > |1.96|

Sample of hhs - Kish:
n = 28 cases from 10 different countries;
11 cases = 39.3%: bias > |1.96|

Sample of hhs - Birthday:
n = 55 cases from 19 different countries;
38 cases = 69.1%: bias > |1.96|

Beyond Western countries

- Selection techniques developed in/for Western countries
- Needs to be checked whether they can be applied (smoothly) in other parts of the world

Specific issues in developing/emerging countries
- Applicability of household / household membership definitions
- Larger average household sizes
- Availability of information on household members
- Concept of random selection understood?

=> Devising special selection methods appropriate for developing countries with large household sizes meritorious (Le et al. 2013)

Basic reccommendations for cross-national surveys (1)

- Use Kish, if possible.

  If birthday method is to be used:
  make sure that verification is possible
  (collect day/month of birth of all hh-members)

- Provide uniform definitions (hh, residence rules, etc.).
  Check whether national explanations/adaptations are necessary

- Coordinating center to check national procedures + materials
  (before fieldwork)

Basic recommendations for cross-national surveys (2)

- Train interviewers in method to be used

- Verify correct application of method used (during/after fieldwork)

- Document procedures + materials

Finally:
- More research on effects of different selection methods in F-t-F surveys still needed!
- Experimental comparisons desirable (effects on coverage – nonresponse – sample composition – costs)

Questions? Comments?

achim.koch@gesis.org