BANCA D'ITALIA
EUROSISTEMA

Mind the mode: lessons from a web survey on household finances

Andrea Neri

Joint project

ISTAT

Giulio Barcaroli
Laura Capparucci
Massimiliano Degortes
Martina Lo Conte
Loredana Mazza
Manuela Murgia
Roberta Rizzi

Banca d'Italia

Patrizia De Troia
Marco Di Pietropaolo
Romina Gambacorta
Andrea Neri
Giuseppina Papadia
Francesca Zanichelli

Foreword

- Banca d'Italia runs a survey on households' income and wealth (SHIW) since 1960s
- All the main surveys on income and wealth use interviewers: *Survey of Consumer Finances* (SCF), *Household Finance and Consumption Survey* (HFCS), *The European Union Statistics on Income and Living Conditions* (EU-SILC)
- The only example of large scale web survey on income and wealth is the DNB household Survey (CentERpanel)

Research questions

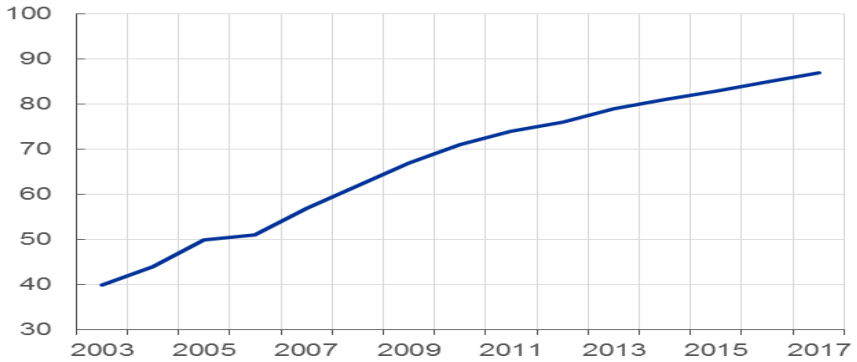
- Do we really need face-to-face interviews?

- Does the web represent a viable alternative to collect information on household income and wealth?

The use of web surveys

- Increasing use of Internet

Internet access of households
(percentage of all households, euro area)



- Appealing features:
  - cost reduction;
  - timeliness;
  - better quality answers to sensitive questions (Tourageau and Yan 2007).

The Web Survey on Italian Households (WEBIT)

- Probabilistic sample of about 10,000 hhs selected from population registers
- An invitation letter with a password was sent
- To access the website a valid email was required
- An incentive to respond was given to about 80% of the sample, to test the effect on survey participation
- Recalls by telephone and email by Istat
- Final sample of about 1,000 hhs

The Web Survey on Italian Households (WEBIT)

- Survey data are linked with tax records on income.
- The web survey was carried out in parallel to a CAPI survey in the same municipalities and using the same questionnaire.
- Randomized experiments.

How to assess data accuracy

- Coverage problems
- Self-selection bias
- Response behaviour
- We focus on Bias

Selection Bias in web surveys, Bethlehem 2010

The Science of Web Surveys, Tourangeau, Conrad, Couper 2013

Coverage

- Target population is usually different from Internet population
- Bias has two components:

$$Bias(\bar{y}_I) = E(\bar{y}_I) - \bar{Y} = \frac{N_{NI}}{N}(\bar{Y}_I - \bar{Y}_{NI})$$

$\frac{N_{NI}}{N}$	Share of population without access to Internet
$(\bar{Y}_I - \bar{Y}_{NI})$	Average difference between the 2 groups

Coverage: results

- 30% without Internet. Coverage increases with education, income, the presence of a young person in the household.

- Difference in the variable of interest (income from labour and transfers) between the two sub-groups:
  $\bar{Y}_I = €31.800 \;\; \bar{Y}_{NI} = € 18.300$

- $Bias(\bar{y}_I)$: positive bias $\approx 15\%$ of $\bar{Y}$

Self-selection

- Web surveys are known to have low unit response rates and are also affected by break-offs (Tourangeau, Conrad, Couper 2013).

  - Bias has 2 components (Bethlehem 1998):

$$Bias(\bar{y}_{I,r}) \approx \frac{Cov(p, Y)}{\bar{p}}$$

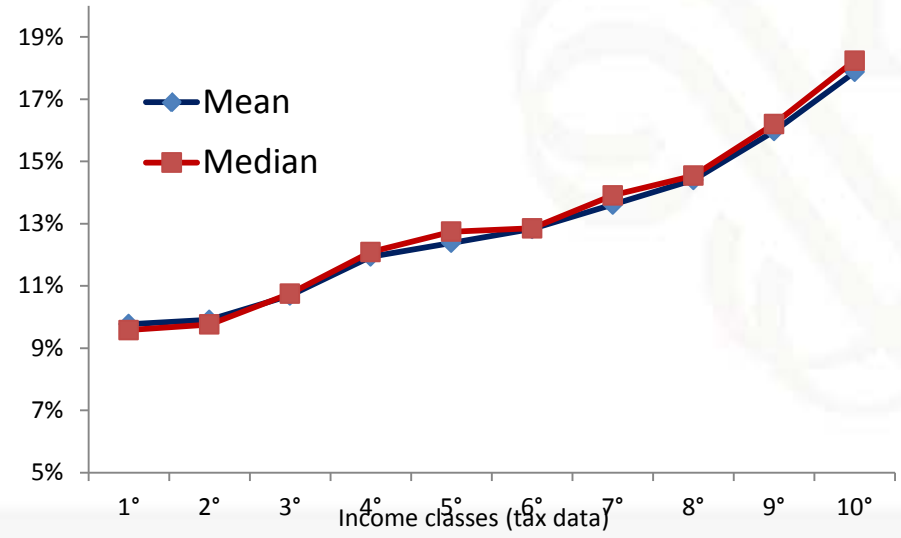
$Cov(p, Y)$ Association between response probability and Y.

$\bar{p}$ Average response rate.

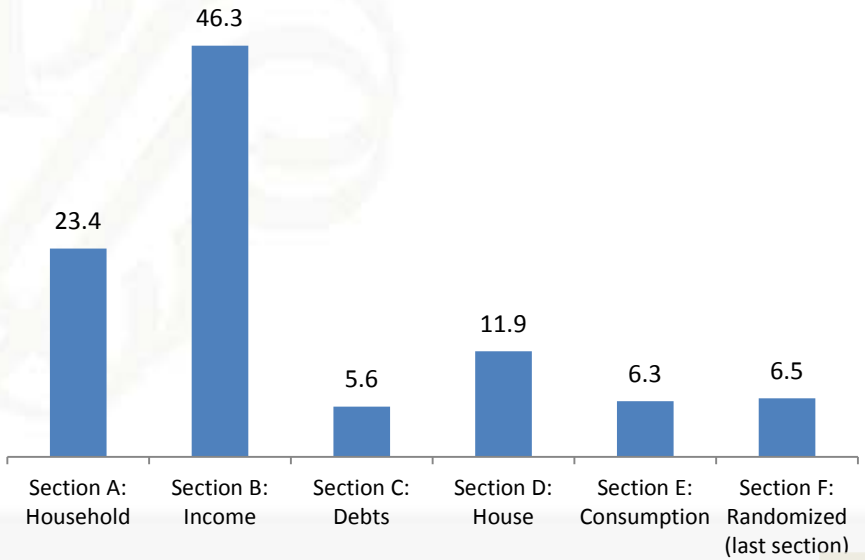
Selection bias: results

- Average response rate ($\bar{p}$): 13 %
- Positive and significant covariance between response probabilities and income

Response probabilities by household income classes



Break-offs by section



Selection bias: results

- *Y* = household income from labour and transfers (tax data)
- Association between *Y* and $\hat{p}$: rho=0.43

$$Bias\left(\bar{y}_{I,r}\right) \approx \frac{Cov(p,Y)}{\bar{p}} \approx € 4.500 = 17\%$$

Measurement error

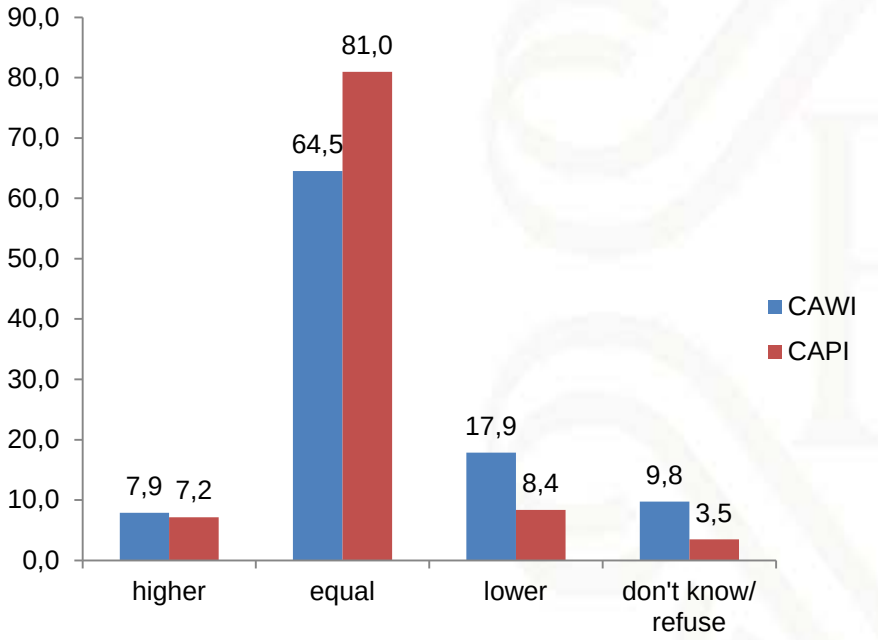
- Ways to assess the response behaviour in the WEB survey:

  1) a comparison with the CAPI survey :
     - selecting the CAPI respondents more similar to those of the WEB survey;
     - aligning the weights of the two samples.

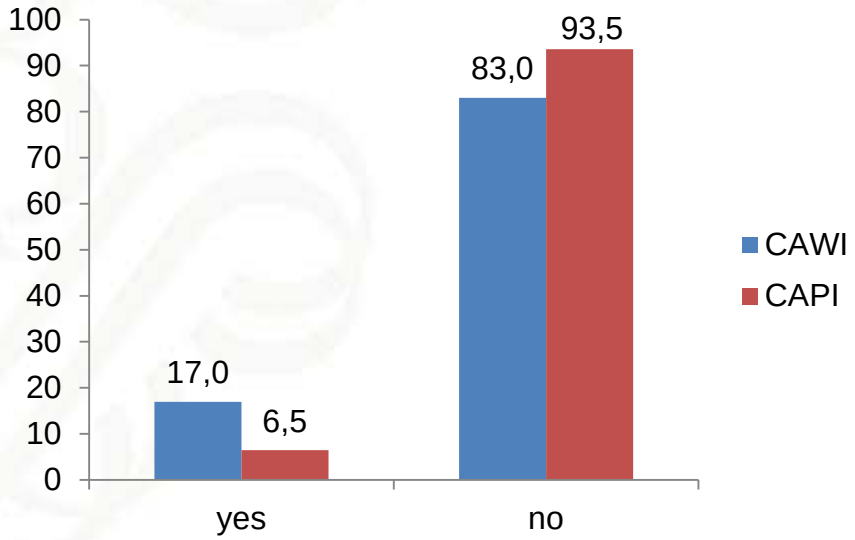
  2) a comparison with the tax registers.

Response behaviour: results

Do you expect that your hh's total income in 2016 will be higher, equal or lower than the one you had in 2015?



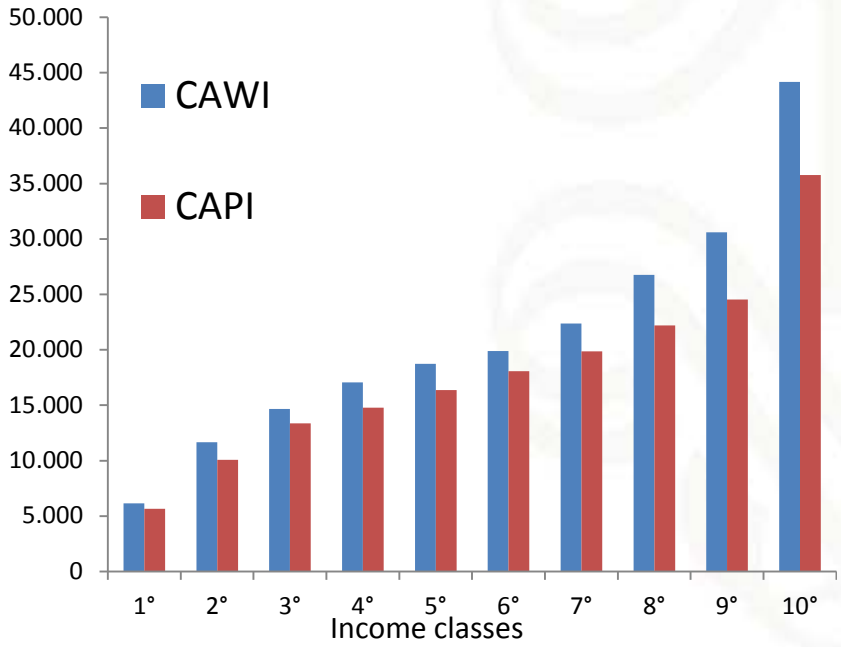
In the last 3 years has your hh received any assistance or non-economic aid from relatives or friends ?



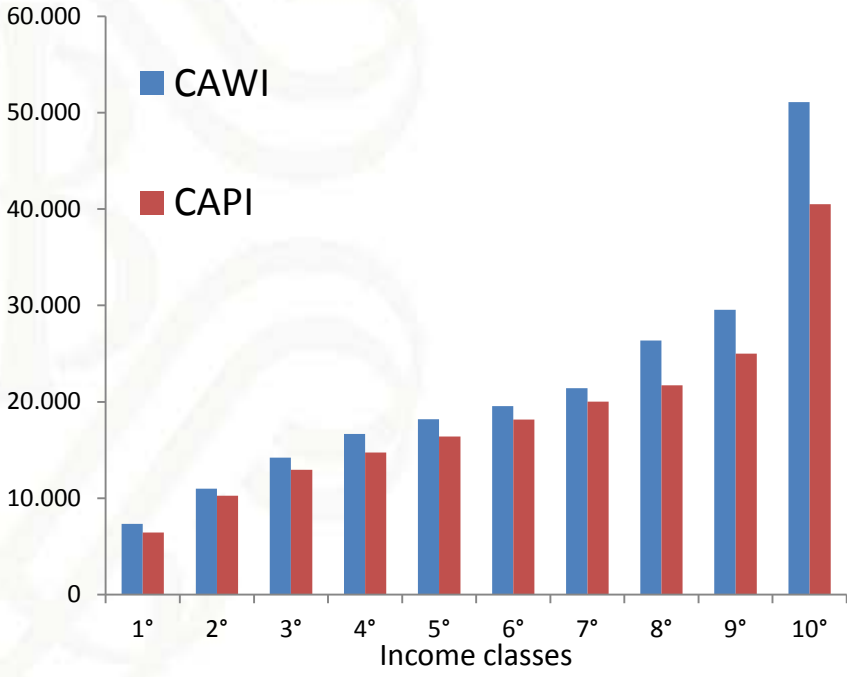
Response behaviour: results

Percapita income by income class (euro)

Income from employment

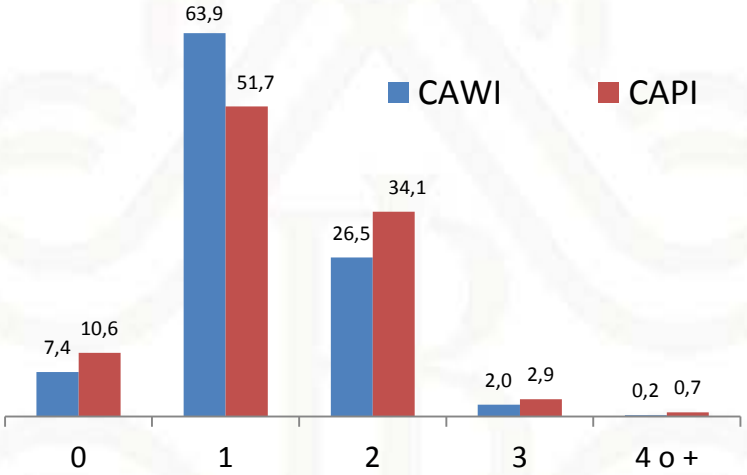


Income from pensions

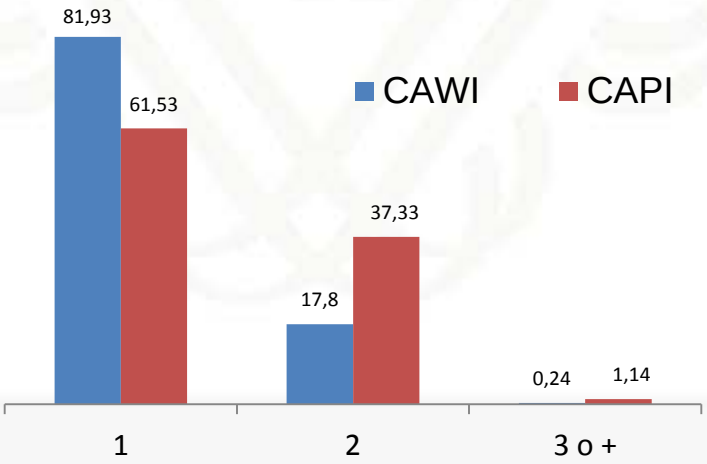


Response behaviour: results

Distribution of recipients of income from employment



Distribution of recipients of income from pensions



Conclusions

- Does the web represent a viable alternative to collect information on household income and wealth?

Probably not yet,
unless powerful auxiliary information is available.

Conclusions

- Self-selection is most challenging issue to overcome.
- The role of interviewers in enrolling households is difficult to replace.
- Good results for qualitative questions (economic conditions, expectations, saving decisions).
- As to income the evidence is mixed: interviewers are needed to prone respondents.