An experimental web survey on income and wealth among the Italian Households

Giulio Barcaroli*


(*) Italian National Institute of Statistics (ISTAT)
(**) Bank of Italy
Outline

- Istat - Bank of Italy research project
- The Web Survey on Italian Households
- Research question: is response behavior relating to sensitive topics different when collected with different modes?
- Some results
- Conclusions
Istat - Bank of Italy research project

- Bank of Italy carries out the Survey of Italian Household Income and Wealth (SHIW) since the Sixties using face to face interviews (now using CAPI technique).

- It is well known how the presence of an interviewer can affect the propensity to respond to a survey, especially when dealing with sensitive topics such as income and wealth.

- For this reason, Istat and Bank of Italy have collaborated in a research project aimed at evaluating whether the adoption of CAWI can enhance response propensity and improve data quality.
The Web Survey on Italian Households

- As part of the project, an experimental web survey was carried out in 2016, parallel to a CAPI survey run by the Bank of Italy on a subset of households participating to the SHIW.

- A sample of 10,030 households was selected from the Population Register in the same municipalities used for the CAPI survey.

- An incentive to respond was given to about 80% of the sample, in order to assess the effect on survey participation (to 80% of the sample was proposed to participate to 5 lotteries to win an iPad each time; the remaining households were not given any incentive).
The Web Survey on Italian Households

In the web survey we also included some experiments:

- The presence/absence of ‘don’t know’ option
- Inverse order for long list of items
- Different scale graphical layout
- Different text wording for the same question
- A usability test for an automatic coding system for the variable Occupation
Aim of the presentation

In this presentation we illustrate some results on whether response behavior relating to sensitive topics is different when collected with different modes.

Furthermore, web survey data was also linked to administrative tax records on income in order to study potential underreporting.
Response rates and break offs

The incentive had a slight effect in boosting survey participation and in reducing break-offs.

Maybe other forms of incentive should be evaluated.
Response rates and break-offs

Break-offs happened mostly in income section

Response rate per section

Web data collection: challenges and future perspectives
ITACOSM 2017 – The 5th ITAlian Conference on Survey Methodology
Does response behavior relating to sensitive topics differ in the CAPI and CAWI surveys?
Comparing CAPI and CAWI data

The comparison is made among variables that might have been affected by the “social desirability” effect.

The presence of an interviewer may influence answers to questions relating to:

- savings
- debts
- non economic aids
- income
- expectations on hh income variations
Comparing CAPI and CAWI data

Differences between CAPI and WEB results can derive from:

• **Selection effects**
  Due to differences in sample compositions (since individuals differ in the likelihood to which they respond to different modes of data collection, e.g. by age, level of education..)

• **Mode effects**
  refer to differences caused by the method used to collect the data (people may respond differently to survey questions depending on the different interviewing technique and on the different cognitive processes taking place when respondents answer to a survey question)
Comparing CAPI and CAWI data

In order to handle the selection effect, so to make CAPI and CAWI data more comparable from the point of view of the mode effect, we use:

• **Adjusted weights for CAWI data**
  Weights were created in two steps:
  1. to adjust for population totals (*by sex, age, geographical area and education level*)
  2. to align CAWI respondents to the subset of the population using Internet (*estimates obtained from the CAPI survey*)

• **Subgroup of CAPI sample**
  Comparisons are made among CAWI data and the subgroup of CAPI respondents that navigate in Internet using a computer or a smartphone (*potential CAWI respondents*)
Expectations on income variation 2015-2016

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

Web data collection: challenges and future perspectives
ITACOSM 2017 – The 5th ITAlian Conference on Survey Methodology
Debts and non economic aids

By the end of 2015 did your household have debts to banks and financial companies?

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

Statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>1</td>
<td>145.082</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>1</td>
<td>60.2034</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Savings/debts and expenditure variation 2014-15

Expenditure in 2015 was higher, equal to or lower than your hh total annual income?

Expenditure for all consumption in 2015 was higher, equal to or lower than in 2014?

<table>
<thead>
<tr>
<th>Statistics</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>2</td>
<td>68.708</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistics</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>4</td>
<td>432.295</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Web data collection: challenges and future perspectives
ITACOSM 2017 – The 5th ITAlian Conference on Survey Methodology
Comparing CAPI and CAWI data on income

**Employment Yearly Net Income (Head of the family)**

<table>
<thead>
<tr>
<th>euro values</th>
<th>CAWI</th>
<th>CAPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>22,745</td>
<td>18,400</td>
</tr>
<tr>
<td>Median value</td>
<td>20,500</td>
<td>18,000</td>
</tr>
<tr>
<td>Modal value</td>
<td>20,000</td>
<td>15,000</td>
</tr>
</tbody>
</table>

CAPI survey seems to show lower levels of employment income.
Income ADmin archive (IAD) as a benchmark

Are respondents confident with the web mode? Do they feel comfortable in declaring their income?

**Employment income**

<table>
<thead>
<tr>
<th></th>
<th>Don’t declare an income</th>
<th>Declare an income</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IAD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don’t declare an income</td>
<td>30,6</td>
<td>7,5</td>
<td>38,1</td>
</tr>
<tr>
<td>Declare an income</td>
<td>4,7</td>
<td>57,2</td>
<td>61,9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>35,3</td>
<td>64,7</td>
<td>100,0</td>
</tr>
</tbody>
</table>

*4,7 Potential Under reporting*

Web respondents seem quite confident with the web mode: low level of income underreporting

*time lag: IAD 2014 and CAWI 2015*
Income ADmin archive (IAD) as a benchmark

Comparison of employment income distributions*

<table>
<thead>
<tr>
<th>euro values</th>
<th>CAWI</th>
<th>IAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>24,825</td>
<td>25,395</td>
</tr>
<tr>
<td>Median value</td>
<td>21,357</td>
<td>22,688</td>
</tr>
<tr>
<td>Modal value</td>
<td>20,000</td>
<td>20,939</td>
</tr>
</tbody>
</table>

* Data are not weighted

Income underestimation in web survey data seems to be small
Income Admin Archive (IAD) as a benchmark

The Web Survey seems to slightly underestimate the hhs’ income. Differences are mostly between -20% and 5%.

If IAD as benchmark → CAPI might underestimate income even more
Some conclusions

CAWI seems to be a good technique to collect sensitive data on income and wealth. However, there are pros and cons:

- **Pros**: more confidentiality and privacy for the respondent (no interviewer effect)
  
  Respondents seem more likely to reveal more difficult economic conditions and to declare an income more in line with other sources (administrative data).

- **Cons**: low response rates (non response)
  
  An incentive seems slightly help in improving response rates and break-offs, but other forms of incentive should be evaluated, together with an effective reminder strategy.
Thank you for your attention

barcarol@istat.it
loconte@istat.it
murgia@istat.it
andrea.neri@bancaditalia.it
giuseppina.papadia@bancaditalia.it
francesca.zanichelli@bancaditalia.it