Income and Wealth Sample Estimates Consistent With Macro Aggregates: Some Experiments

Giovanni D'Alessio (Bank of Italy)
Andrea Neri (Bank of Italy)

Paper prepared for the 34th IARIW General Conference

Dresden, Germany, August 21-27, 2016

Session 2G: Research Forum I on Income Distribution Issues

Time: Monday, August 22, 2016 [Afternoon]

INCOME AND WEALTH SAMPLE ESTIMATES CONSISTENT WITH MACRO AGGREGATES: SOME EXPERIMENTS

Giovanni D'Alessio* and Andrea Neri*

Summary

The Bank of Italy's Survey of Household Income and Wealth (SHIW) is widely used to study the economic behavior of Italian households. Like most similar surveys, the SHIW is biased downward in its estimates by the lesser propensity of wealthy families to participate and by the tendency to underreport income and wealth. This work assesses the various techniques for correct the bias, applying them to the period 1995-2012. Calibration techniques, which produce estimates consistent with the macro-economic information available from other sources, are also employed.

Contents

* Bank of Italy, Economic and Statistics Department, Via Nazionale 91, 00184 Rome, Italy.

Non-technical summary

The measurement of household income and wealth through sample surveys is a daunting task. Both topics are very sensitive for respondents. As a result, some households refuse to participate in the survey. This is especially the case for wealthy families. Moreover, some respondents may be reluctant to truthfully report the amounts of wealth they hold or the income they earn. These facts may introduce bias in the survey-based estimates of household income and wealth.

In this paper we describe the experience accumulated over the years by the Bank of Italy with the survey of Italian household income and wealth (SHIW). The survey is the Italian component of the European Household finance and consumption survey (HFCS).

In the first part of the paper we review all the existing studies aiming at estimating and correcting these distortions in the SHIW survey.

We then apply the available adjustment methods to the 1995-2012 waves. We compute some statistics relating the household income and wealth distribution using different estimators. To this end we use all the external information available such as aggregate statistics coming from national accounts, information from administrative records and survey data which are considered to be more reliable than the SHIW survey on specific topics.

This exercise allows us to assess the robustness of the survey-based statistics and provides hints on how to better to use the survey data.

The main finding is that survey data provide very reliable information on the relative positions of households with given socio-demographics within the income and wealth distribution.

A second finding is that the level of concentration of both distributions is likely to be underestimated by survey data.

1. Introduction

The Survey of Household Income and Wealth (SHIW) conducted by the Bank of Italy every two years is widely used to analyse the economic behavior of Italian households. However, like those of all the surveys of this kind, the data are subject to various measurement errors, above all the tendency of wealthy families to decline participation and the unwillingness of respondents to state their full income and wealth.

Over the years, a good many studies have shown how the resulting downward bias is the main factor in the substantial differences between the sample estimates and other sources of data on households' budgets (both macroeconomic, such as the national accounts, and administrative, such as supervisory reporting and censuses).

This study first reviews the methods used over the years to adjust the SHIW data. We then explore the possibility of simultaneous application of some of them to the surveys carried out from 1995 to 2012. The aim is to assess the possibility of micro analysis on some of the main variables that determine the living conditions of Italian households (income, wealth and debt) through estimates that are consistent with the other macroeconomic information. Although the latter too is subject to measurement errors, we try to take advantage of the strengths of each kind of source. The paper finally discusses the extent to which these data can be used in microsimulation models.

2. A short review of the literature

Sample surveys inevitably have problems of measurement error and systematic non-participation. Notwithstanding substantial efforts to prevent and minimize these errors, ex-post adjustment is essentially unavoidable.

The correction methods set out in the literature fall into two broad categories (see Nicolini et. al, 2013). The first is the *design-based approach*, which serves chiefly to address the problem of non-response. Sample selection is taken as a two-phase process. The sample selected is the one obtained in the first phase, while the sample actually interviewed (respondents) is treated as the product of a second stage of sampling. Each unit in the population has a certain probability of participating in this second phase, which can be estimated in various ways and then used to construct estimators with better asymptotic properties. This is done by modifying the sampling weights.[1] [2]

---

[1] Deville and Särndal (1992) extend the calibration techniques by including the totals of quantitative variables. Fuller et. al (1994) first note that linear calibration implicitly adjusts for non-response if the model for non-response is linear. On this basis, other studies have introduced extensions. Folson and Singh (2000) find a general formulation that includes non-linear functions too in the calibration. Deville (2000) introduces the concept of generalized calibration, which allows inclusion of variables that explain the non-response but for which no external information is available at the population level (such as the information collected by the interviewers). Kott and Chang (2010), taking up an idea of Deville (2000), propose including the same variable of interest in the generalized calibration to correct the distortion due to the non-negligible non-response.

[2] For a more detailed description of the approach, see for instance Oh and Scheuren (1983). The statistical properties of these estimators are analyzed in various studies. For example, Little and Vartivarian (2005) show that if the variables used to construct the weights are associated both with non-participation and with the variable of interest, the bias and the variance of the estimators are reduced. More recently Kott and Liao (2012) present an estimator that allows a dual protection against non-response bias.

The second, *model-based* approach is characterized by two requirements: a model for the distribution of the measurement error and auxiliary information to estimate the parameters of the model. Among the various models found in the literature, those most suitable for our purposes are imputation methods. For a general description, see the seminal work of Rubin (1978, 1987). These methods are mainly used to address the issue of item non-response, but they can be readily generalized to the problem of measurement error. In fact, the variable affected by error may be deemed unrealistic for certain observations and a plausible value accordingly imputed[3].

In any case, the two approaches have some shared traits, so that clear separation is not always easy. For example, the weighting adjustment can also be seen as a method of value imputation consisting in compensating for the missing responses by using those of the respondents with the most similar characteristics; in the same way, the imputation of plausible in lieu of respondents' claimed values can be thought as a re-weighting method.

Further, within the design-based framework a model-assisted approach has recently been developed: the model describes the relationship between the variable of interest and one or more other variables for which external information is available in order to generate estimators with better asymptotic properties (which are always evaluated in a design-based framework).

That said, it is still possible to summarize the pros and cons of the two approaches. For more detailed discussions, see Gelman (2007) and Brick (2013).

One assumption generally made in both approaches is that the missing data are *missing at random*. By this assumption, the auxiliary variables available contain all the information necessary to make the adjustment.

The difference between the two methods emerges clearly when the corrections involve multiple variables. The model-based approach usually allows for a more flexible and tailored form of correction for each variable. For example, the under-reporting of financial investments is likely to be different from that of self-employment income (Neri and Zizza (2010)), so the use of imputation models specific to each variable would make for more effective correction.

Moreover, the imputation of one variable could require recalculating the derived variables, such as when some component of household wealth is imputed, which means modifying not only the aggregate wealth but also the financial income it generates.

Finally, imputation models modify the correlation among the variables associated with the one that is imputed, so careful study of the effects on associations is required[4].

In the case of weights-adjustments, the internal consistencies between the variables are preserved by definition. This represents a definite advantage, especially for micro analysis. On the other hand, a modification of the weights results in a modification of the distributions of all the surveyed variables, and should therefore be carefully monitored.

The model-based approach, working at the level of the single observation, generally yields estimates with smaller variance than would be obtained by modifying

[3] For a recent example of the use of these imputation models, see Peytchev (2012), who uses the technique to adjust jointly for non-response and measurement error.

[4] One solution is to impute according to a sequential scheme, to ensure consistency among the imputed variables.

the weights. Consider, for example, financial assets, which are heavily concentrated in the hands of a limited number of households and subject to significant under-reporting. This means that a part of the sample could be subject to very substantial weight adjustment, which could increase overall variability. And in these cases it is not uncommon for the method not to converge, as it fails to align the sample with both financial and socio-demographic external information.

Model-based methods afford greater opportunity to make adjustments that relax the *missing at random* assumption by giving researchers more flexibility in model specification. However, model-based estimators may have problems of robustness when the model's assumptions – which are not normally testable – are violated. Holt and Smith (1979) show instead that robustness (i.e protection from erroneous specification) is one of the strengths of the post-stratification.

According to Lohr (2007), model-based estimators are less desirable for the producers of official data and statistics, in that they entail more choices to be defended than design-based estimators. Further, the design-based approach is simpler to use and accessible to a wide variety of users.

The weight adjustment approach allows easy alignment of the survey findings to external sources like the census. The estimators obtained by these techniques generally have desirable statistical properties: in most cases the accuracy of the estimators can be increased, and if the variables used for calibration are also correlated with non-response, they also reduce bias (Little, Vartivarian 2005).

Yet it should be borne in mind that the choice of the method of adjustment is basically driven by the information that is available. If, for example, the only available auxiliary information is population totals, the design-based approach is preferable; but if auxiliary data are available at the individual level, then the model-based methods too may be employed.

In any event, the two approaches should not be considered as alternatives. This paper is intended as an instance of their joint use to align one particular survey (the Survey of Household Income and Wealth) to a variety of external sources.

3. Previous adjustments on SHIW data

The discrepancies between SHIW estimates and the corresponding macro aggregates have been public knowledge for decades. In the Bank of Italy *Bulletin* in 1970 Ulizzi, describing the findings of the 1968 survey, observed that "Among the mentioned errors [non-sampling errors], special reference is due to those attributable to the reticence of respondents about the financial assets held. The experience gained in numerous analyses, some of which are specific on the subject, has revealed considerable reluctance on the part of families to provide information on the ownership of financial assets (...). For savings and income, collaboration of respondents is generally better, being less the aversion to provide data on flows than on stocks. "

In those years, Ulizzi had worked on the under-reporting of financial assets using techniques of exact matching. He sought to interview about 900 persons whose true securities assets were known from other sources. Thirty per cent did not take part; the average value of the non-respondents' financial assets was only slightly higher than that of the respondents. This finding is most significant, as it suggests that the effect of non-participation on the overall estimates may be only marginal. But the average value of securities declared by the respondents was considerably lower (15 per cent) than their actual holdings. Most of the overall discrepancy was produced by non-reporting: that is, over 60 percent of the respondents denied any ownership of securities, and most of the others under-declared or refused to answer. Non-reporting and under-reporting were more common among the wealthiest households. This first study has been followed by many others focusing on non-participation and under-reporting[5].

The survey is intended to be representative of the resident population. Since the selection of households is from municipal civic registers, which are not always perfectly accurate, some groups may be under-represented in the sample, such as recent immigrants, who do not always comply with the obligation to notify the authorities of changes of residence in Italy or departure from the country.

However, the main source of inaccuracy in the estimates is far more likely to be sample composition, as determined by the type of households that are not interviewed.[6] Whether the reason for non-interviewing is explicit refusal or unavailability (not at home at the time set), it represents a problem for statistical surveys, a selection bias that may produce samples in which those less willing to cooperate (or not reached) may be under-represented. Since the estimates draw information from respondents only, the bias increases with the share of non-response and with the difference between the average values of respondents and non-respondents.

The SHIW incorporates various procedures to limit the effects of non-participation (Bank of Italy, 2014). First, households that cannot be interviewed are replaced by others, randomly extracted, in the same municipality. This controls for the potential source of bias due to the relationship between the local and household characteristics. Second, post-stratification is performed on the basis of some individual characteristics, in order to balance the weights of the different population segments within the sample. This is done by raking techniques, which impose the alignment of the weighted distributions of the sample by sex, age, geographical area and size of municipality with those of entire population.

However, some bias can be presumed to remain, since particular groups of households (say, the wealthy) may be less likely than others to be interviewed. This is hard to gauge, because information on non-respondents is not generally available.

In an examination of panel attrition, Cannari and D'Alessio (1992) compared the households that ceased collaboration with those that continued to participate in the survey. The non-response behavior in the panel was then extrapolated to the entire sample, and the under-reporting of income due to attrition was estimated at 5 percent.

Other methods have also been applied to this question, and in particular procuring information on households that have never been interviewed, in whose regard studies like the foregoing are impossible. Analysis of the call attempts needed to get the interview (i.e., the number of visits or phone contacts to persuade families to participate) can indicate the kinds of households that are hardest to interview and thus help in correcting sample weights by estimating the actual probability of participation of each household interviewed. D'Alessio and Faiella (2002) showed that when these aspects are taken into account, income and wealth estimates increase; households' average income and wealth differ depending on how easily they make themselves available for interviewing. Respondents who are persuaded to participate after an initial

[5] A number of studies have compared the survey estimates with those derived from other sources. See for instance Brandolini, 1999, and Bonci Marchese and Neri, 2005. In what follows we refer only to works that suggest methods of adjustment of the sample estimates. For a review of the literature see D'Alessio and Ilardi, 2013.

[6] D'Alessio and Faiella (2002).

refusal show average income and wealth 20 and 30 percent higher than the overall average; those interviewed after not being found at home the first time, a few percentage points below the overall average.

D'Alessio and Faiella (2002) also study a sample of about 2000 households whose information had been matched anonymously with some banking information; in this case they show that non-response is not random but is more frequent among the wealthiest families. The bias detected was greater for financial assets (with adjusted estimates 15 to 30 percent higher than unadjusted ones) than for income (underestimation of 5 to 14 percent), probably because of the greater inequality of the distribution of wealth than of income .

Neri and Ranalli (2011), using the results of a telephone survey conducted on SHIW non-respondents, report greater difficulty obtaining interviews from the wealthiest households and propose a corresponding adjustment of sampling weights. The result is confirmed by a more recent work, D'Alessio and Iezzi (2014).

Another issue relevant to the adjustment of the sample estimates is under-reporting, i.e. the non-declaration or undervaluation of real estate and financial assets and income.

Cannari and D'Alessio (1990) inquired into the SHIW estimates of real estate wealth. They found that the number of residential properties was quite well estimated, but that the number of rented homes according to landlords' declarations was inconsistent with tenants' answers. And a comparison with census estimates showed that the survey also underestimated the number of vacation homes. The authors proposed a method for correcting the survey estimate of the number of dwellings according to owners' reports, namely imputing additional homes to the sample households on the basis of estimated probabilities of owning a second residential property. [7]

Cannari, D'Alessio, Raimondi and Rinaldi (1990) performed a statistical matching of the financial assets declared by SHIW respondents with data provided by a sample of commercial bank clients from a survey carried out by the bank. Assuming that there was no under-reporting in the latter, the authors used statistical models to estimate both the probability of holding the various types of asset and the true amounts that the various types of household should hold. Comparison with the SHIW estimates showed that non-reporting was more frequent among some types of household (the poorer and less educated), under-reporting among others types of household. On the whole, the primary factor in the SHIW's underestimate compared to the aggregate data was under-reporting. The adjusted estimates obtained by a design-based approach are about twice the standard SHIW estimates, but even so there is some difference from the macro data. Although in some instances the revisions are quite substantial, the relative proportions of assets held by the various categories of households are not greatly changed. Cannari and D'Alessio (1993), with a more complex model-based

[7] The distribution of the number of dwellings (excluding the primary residence) is modeled by a Poisson distribution whose mean depends on a vector of observable characteristics (age, education, gender of the household head, household income, municipality of residence, etc.). The survey data are used to estimate the probability of a household's owning second homes, which in turn are used to impute the missing dwellings (i.e. the difference between the more reliable census data and the survey data).

methodology, also showed that the Gini concentration index is not significantly affected by the adjustment.

Brandolini et al. (2004) study Italian wealth distribution after adjusting for the underestimation of real and financial assets.

The statistical matching between commercial bank data and SHIW data has been replicated more recently (D'Aurizio et al., 2006). The adjusted estimates of financial assets average more than twice the original figures, reaching 85 percent of the aggregate. The adjustment is larger for households whose head is old or poorly educated. The paper also adjusted financial liabilities, whose corrected values are on average about 40 percent higher.

Neri and Monteduro (2013) propose an adjustment of housing wealth based on the aggregate distributions of ownership from tax records. SHIW tends to underestimate both the number of taxpayers who own just one and those who own more than five units of housing. Correcting the SHIW data by aligning the sample data with the administrative data increases total housing wealth by about a quarter. The adjustment does not significantly affect the concentration of wealth or the association between wealth and some socio-demographic characteristics.

As to the under-reporting of income, Cannari and Violi (1995), on the pattern of by Pissarides and Weber (1989) using British data, applied a method of 'indirect' reconstruction of real income, positing that income is correctly detected for some population groups and that some components of consumption are measured without systematic error for all groups. Under these hypotheses, the relationship between consumption (food consumption) and income is estimated using the sub-sample for which income data are accurate. For the rest, the relationship can be reversed, reconstructing estimated income consistent with observed consumption.[8]

This approach was replicated by Neri and Zizza (2010) using the value of the household's primary residence (which can be assumed not to be under-reported, thanks to face-to-face interviewing), not food consumption. The relationship between the value of the dwelling and income is first estimated for civil servants and then applied to the self-employed, to derive a consistent amount of labour income: the adjustment of the estimates is substantial (about 36 percent of income). The authors then develop corrections for other income components, largely based on revisions of paper described above.

Cifaldi and Neri (2013) use the results of previous studies to correct the SHIW income and consumption data and discuss the effects of their differential under-reporting on the estimate of the household saving rate.

4. Adjusting for non-response and under-reporting

As we have seen, the SHIW sample estimates of income and wealth fall significantly short of the relevant macroeconomic estimates. The differences are due in part to non-response but mainly to under-reporting.

In this section we set out several possible methods for adjusting the sample data. Sometimes corrections are based on external information at individual level; in other cases, the procedure posits that the national statistics are available and correct and so align the sample data with them, by minimizing a distance function defined on sample

8 A similar procedure can be found in Hurst, Li and Pugsley (2010).

weights. Here, as noted, we discussion several adjustment methods. Comparative analysis of the various results is left to the subsequent section.

4.1 *Proportional adjustemt - C_1*

The most elementary adjustment procedure, which we take as a benchmark and denote by C_1, simply inflations the sample values y_i by the coefficient $k = Y^T / y^T$, the ratio of the total known population value to the total sample estimate.

This method is based on a very simple under-reporting model, assuming that for every individual the amount declared y_{id} is a constant fraction of the true amount y_i, plus an error term:

$$y_{id} = y_i/k + e_i \qquad (1)$$

Simple as it is, this model can be useful, especially to adjust single components of income and wealth. Income and wealth obtained as the sum of inflated components can offer helpful indications on how under-reporting affects averages and concentration indices. On income, for example, the method separately corrects the data on wage or salary income (YL), pensions and other transfers (YT), income from self-employment (YM) and income from capital (YC). In the same way, for wealth the method can be applied to each single component – real assets (AR), financial assets (AF), and financial liabilities (PF) – which immediately indicates the extent of the greater underestimation of financial than real assets.

Of course, this estimator absolutely cannot adjust for non-reporting, i.e. the failure to declare a certain asset or source of income, as only the declared amounts are inflated.

4.2 *Adjustment based on interviewer score – C_2*

To get information on possible under-reporting, the SHIW also collects some paradata, asking interviewers to judge the reliability of respondents' answers on income and wealth. The judgment is based on the correspondence between the answers and the other information available, such as area of residence, type of property, apparent standard of living (furniture, etc.). In the 1993 and in 1995 waves this information on reliability was only qualitative (totally unreliable, fairly unreliable, fairly reliable, totally reliable); from the 1998 survey onwards the opinions of the interviewers were expressed with a score from 1 (totally unreliable) to 10 (totally reliable).

On the whole, the truthfulness of the answers is deemed satisfactory for all the years examined (Table 1): in 1993 and 1995, between 85 and 90 per cent of the responses are judged to be satisfactory (fairly or totally reliable); for subsequent surveys, shares are similar if one considers as satisfactory all scores of 6 or better. The average increases in the last two years.

Nevertheless, the judgments are not homogeneous in the sample. The scores are regularly higher for employee households, better educated households and those in the Centre and North. This information seems to complement that obtained in advance and can serve to correct the sample estimates.

Table 1

Truthfulness of answers on income and wealth, 1993-2012

(percentages, scores in tenth)

Year	Qualitative judgment on the reliability of the income and wealth answers provided by respondents (interviewers' opinions)				
	Totally unreliable	Farily unreliable	Fairly reliable	Totally reliable	Total
1993	0.9	9.4	50.5	39.2	100.0
1995	1.0	11.7	53.3	34.1	100.0

Score from 1 (totally unreliable) to 10 (totally reliable) on the truthfulness of respondents' answers on income and wealth (interviewers' opinions)

Year	1	2	3	4	5	6	7	8	9	10	Total	Average score
1998	1.5	1.3	1.7	2.7	6.5	12.3	16.5	22.0	17.5	18.1	100.0	7.6
2000	0.6	0.7	1.3	3.1	6.7	11.8	16.6	20.0	19.7	19.5	100.0	7.7
2002	0.7	1.2	1.3	2.2	6.3	12.3	17.2	21.1	18.2	19.6	100.0	7.7
2004	1.0	1.4	1.2	2.6	7.0	12.1	17.8	22.0	16.9	18.0	100.0	7.6
2006	0.3	0.7	1.1	2.4	6.3	13.1	18.7	23.5	17.8	16.1	100.0	7.6
2008	0.7	0.8	1.0	2.3	6.1	13.4	18.8	23.8	19.7	13.5	100.0	7.6
2010	0.6	0.5	0.7	1.6	4.0	8.6	15.7	22.8	26.0	19.6	100.0	8.0
2012	0.3	0.4	0.6	1.0	3.7	8.8	13.3	22.0	25.6	24.3	100.0	8.2

We therefore estimate the following model:

$$\log(y_{id}) = x_i \beta + \gamma v_i + e_i \qquad (2)$$

where x_i is a vector of control variables and v_i is the interviewer's truthfulness score on income and wealth answers. Once the contribution of component V is estimated, we can estimate the income and wealth that the household should have declared to get the maximum truthfulness score (v_i).

This model suggests that the interviewers' judgments do capture some elements of under-reporting. For instance, the revaluations of income and wealth are greater for the self-employed than for pensioners and employees. Nevertheless, the average adjusted values remain quite distant from the totals known from aggregate sources.

One alternative estimator (which we can designate C_2) takes interviewers' scores into account and totally aligns survey and aggregate figures:

$$y_{id} = y_i / k_i + e_i \qquad (3)$$

where k is an inverse function of the interviewer's score v_i

$$k_i = 1 + (10 - v_i)\,\alpha \qquad (4)$$

When v_i is maximum (v_i=10) there is no correction; when it is lower the adjustment is proportional to the distance from peak score. The coefficient α is calibrated so that the sample estimate of the total y^T is equal to the total drawn from the macro source Y^T.

As above, the estimator does not correct for non-reporting.

4.3 The adjustment of single phenomena – C_3

External information can sometimes improve estimation. Below we present the adjustments for non-response and under-reporting of income by self-employed workers, of real estate assets (other than primary residence) and of financial assets. These corrections are designated respectively as C_{3A}, C_{3B}, C_{3C}, C_{3D}; together, as C_3.

4.3.1 Non-response – C_{3A}

The adjustment for non-response is based on Neri and Ranalli (2011). The methodology corrects sampling weights as follows:

$$w_c^{(NR)} = w_c^{(DES)} \alpha_c \tag{5}$$

where $w_c^{(NR)}$ is the weight adjusted for non-response of households in the class c, $w_c^{(DES)}$ is the design weight, and α_c is the correction factor (defined as the inverse of the estimated participation probability of this class of households.

For panel households we use the information available from the past survey combined with contact attempts by the interviewers. The probability of participation is estimated by a logistic model, using as covariates the geographical area and the size of the municipality, the income and wealth brackets, and the interviewer's judgment on the climate in which the interview was carried out. In order to avoid outliers, the probabilities estimated are then grouped into deciles, and each household is assigned the relevant decile's average probability of participation.

For non-panel households, instead, we use data collected on a sample of non-respondents[9]. In the 2008, 2010 and 2012 waves, the main, face-to-face survey was followed by telephone survey of a sample of about 500 non-respondents whose telephone numbers could be found and who agreed to a brief interview. In total, across all the surveys, 863 not-panel households provided data. For each survey, this sample is appended to that of the regularly interviewed households. We then estimated a logistic model to obtain the probability of belonging to the group of non-respondents. The covariates were geographical area and size of the municipality, age, employment status, education, home ownership, number of household members and number of income earners.

The correction method depends on some simplifying assumptions. First, the non-response is assumed to be a function of the observed variables only (missing at random). Second, the non-response and measurement errors described below are assumed independent of each other. Consequently, the adjustment described here is made independently of all the other adjustments.

4.3.2 Adjustment of self-employment income – C_{3B}

As we have seen, the under-reporting of a group of respondents can be estimated by using a benchmark group in whose regard the absence of under-reporting is plausible (say, employees). If for the entire sample we have some income-related indicators that are not affected by measurement error, they can be used to estimate income indirectly.

In what follows we take the value of the primary residence as the pivotal variable to correct the under-reporting of self-employment income. As the interviews are conducted in person and at home, this value cannot be easily concealed from the

[9] This information is not currently used in constructing the official weights for the survey. A similar correction is also used for the panel families. On this point see the methodological appendix of the report on the 2012 survey.

interviewer, so we imagine that it is not systematically underestimated, or at least less so than income.

The extent of under-reporting by households whose head is self-employed can be estimated by the following model:

$$\log(V) = \alpha + \beta \log(Y_d) + \gamma A + \theta X \qquad (4)$$

where it is assumed that the logarithm of the indicator V is a function of a constant α, the logarithm of the declared income Y_d (which in the case of the control group coincides with the actual income Y), other characteristics (sex, age etc.) collected in the matrix X, and a dummy A for self-employed households. Assuming that the two sets of households behave in the same way with respect to V, the portion of income declared by the self-employed π can be estimated from equation (4) as:

$$\pi = Y_d/Y = \exp(-\gamma/\beta) \qquad (5)$$

The coefficient π is not theoretically restricted to the interval 0-1, although in the estimates computed it always did fall there.

The first column of Table 2 gives the estimated coefficients π for the three geographical areas and for the whole sample. The coefficients indicated under-reporting of about 35 percent, slightly more in the South.

To compensate for possible measurement errors in the independent variables, we made an instrumental variables estimate; by these new estimates income under-reporting by the self-employed was reduced to between 10 and 20 percent, and the greater under-reporting in the South disappeared.[10]

In the following we use a single adjustment factor at national level, which we estimate at 20 percent.

Table 2

Reporting coefficients

Value of primary residence	Logarithm	Log (IV)
North..	0.7369	0.8438
Center...	0.7873	0.8717
South and Islands	0.6276	0.9087
Italy..	**0.6761**	**0.8709**

4.3.3 Adjustment of real estate other than primary residence – C_{3C}

A significant share of Italian households' wealth consists in real estate. Most of these properties are primary residences, whose SHIW estimate is close to that resulting from other surveys such as EU-SILC or from census data. Dwellings other than the primary residence, however, are underestimated. The first evidence of this came from consistency checks between some SHIW estimates (Cannari and D'Alessio, 1990). The number of dwellings that the owners declare they rent to other households can be

10 Neri and Zizza (2010), with a slightly different method, re-value self-employed earnings by about 36 percent; Cannari and Violi (1995) estimate an increase of about 25 percent.

compared with the number of tenants interviewed, i.e. those who say their home is owned by someone else.

If there were no under-reporting the two estimates wouldbe equal, save for sampling fluctuations. Actually, however, the number of houses declared by the owners is substantially underestimated at between 1 and 1.5 million, while the number of tenant households comes to 3 million. In other words, only 30 or 40 per cent of rental homes are reported by their owners (Table 3). [11]

Table 3

Houses declared by owners and leaseholders, 1991-2012

(percentages)

Year	Tenant households (a)	Dwellings that owners report renting (b)	Share (b) / (a)
1991	3,291,258	983,777	29.9
1993	3,220,253	1,391,772	43.2
1995	3,360,512	1,533,344	45.6
1998	3,255,218	1,112,374	34.2
2000	3,182,180	1304,149	41.0
2002	2,970,913	978,709	32.9
2004	3,304,629	967,758	29.3
2006	3,360,706	861,826	25.6
2008	3,320,834	1,529,607	46.1
2010	3,646,078	1,205,595	33.1
2012	3,683,863	1,210,284	32.9
Average	-	-	35.8

Comparing the interviewees' reports on housing with census data reveals about the same level of under-reporting (Table 4). According to the SHIW, in 1991 there were about 15.3 million homes owned by households, whereas the census put the number at 22.9 million[12]. Considering that there were some 12.4 million primary residences, we can estimate that the share of houses reported – excluding first homes, which are presumably not unreported – is less than 30 percent. Comparing the 2002 SHIW with the 2001 census, we find that 35 per cent of second homes are reported in the survey. Such substantial under-reporting requires adequate treatment.[13]

Drawing on this evidence, Cannari and D'Alessio (1990) developed a method for imputing missing properties to their most likely owners.[14]

[11] The breakdown of this indicator by region shows the highest values for North and the Centre compared to South and Islands. Since according to survey data about 90 percent of the properties owned by families is located in the same geographic area of residence (the share rises to 98 percent for housing rented to families), it is likely that the observed gap is due to the higher level of under-reporting that characterizes southern families.

[12] Part of the gap is *likely due* to the presence of dwellings in usufruct or in free use.

[13] See for example Cannari and D'Alessio (1990) and Brandolini, Cannari, D'Alessio and Faiella (2004).

[14] The method assumes that the number of dwellings follows a Poisson distribution.

Table 4

Houses reported to SHIW and census data, 1991-2012

Year	SHIW estimates				Census data (*)	Percentage of owned homes declared
	Primary residence owned (a)	Other homes owned by households (b)	Total homes owned by households (c) = (a) + (b)	of which: usufruct or free use	Homes owned by households (d)	(c) / (d)
1991	12,791,339	3,181,017	15,972,357	2,020,510	22,958,865	69.6
2002	14,825,485	3,823,484	18,648,969	2,151,803	25,257,775	73.8

(*) The share of total unoccupied houses owned by the households is assumed equal to the share of occupied houses.

The method (C_3) imputes the difference between the number of houses declared in SHIW and those resulting from the census, suitably interpolated for the years between censuses (Bank of Italy, 2012). The imputation model comprises various characteristics and different average value of primary residences and other homes.

In valuing houses, the C_3 adjustment takes account of respondents' tendency to overestimate their actual market value, ignoring the usual difference between the price asked by the seller and the price paid by the buyer. According to the survey of the housing market (Bank of Italy, 2013) this gap averages between 10 and 15 percent; we take 12 percent.[15]

4.3.4 Adjustment of financial assets – C_{3D}

A detailed comparison between the Financial Accounts and the SHIW estimates of financial wealth was made by Bonci, Marchese and Neri (2005), quantifying the discrepancies between the two sources and attributing them to the various possible factors: differences in definition, measurement errors, sampling and non-sampling errors. A more recent comparison (Bank of Italy, 2012) indicates that the sample estimate of financial assets and liabilities comes to between 30 and 40 percent of the aggregate.

The adjustment procedure proposed here is based on an extension of the method described in D'Aurizio et al. (2006), which compared the 2004 SHIW data with those of a 2003 survey of a commercial bank's customers and corrected the SHIW accordingly. For effective comparison, the sampling and other operating procedures for this external survey had been made as similar as possible to those of SHIW.

The sample of clients, stratified according to brackets of financial wealth, geographical area and size of the municipality of residence, was made up of 1,834 households. Before the matching experiment, a post-stratification was performed in order to reproduce the main socio-demographic characteristics of the population of bank customers in Italy.

The adjustment of the SHIW data was in two steps. First, reticence was measured by comparing the customers' declarations with the real data on the stocks they held, as a function of the amounts declared and the socio-economic characteristics of households. Second, these estimated reatios were applied to the SHIW sample to obtain adjusted financial wealth for the entire population of Italian banking customers.

---

[15] The comparison between the survey data and the administrative data on house prices confirms that respondents tend to overestimate the market value of the homes they own.

The methodology here proposed amends that described only to extrapolate the adjusted estimates for subsequent years. For the years before 2010, instead, we use the adjustment method of Cannari and D'Alessio (1993).

4.4 *Calibrations – C_4 / C_8*

Sample surveys quite commonly incorporate auxiliary information from external sources in the weights. A typical use is post-stratification, or raking, techniques that are used in the SHIW. For instance, this method aligns the socio-demographic composition of the sample with some distributions known from the census, so as to reduce (in general) the standard errors of estimates of the variables that are related to socio-demographic composition (for example, income). These treatments also provide samples for which the known characteristics (say, composition by sex or age) exactly reproduce the data known from other sources.

Starting with Deville and Särndal (1992), the calibration techniques have been generalized to include, in the a priori information set, not only the distributions of qualitative or ordinal variables but also the totals of quantitative variables. Using numerical algorithms, this method finds adjustment weights that are as close as possible to the design weights (by a distance criterion), and at the same time satisfy the constraints on sample composition (as in traditional raking) and the totals of certain variables (e.g. total income). In what follows we refer to the calibration techniques implemented in the SAS macro Calmar (Sautory, 1993).[16]

The strategy was to impose the alignment of distributions of the socio-demographic characteristics of the household head resulting from SHIW as well as total income by source or type of wealth, as described in Table 5.

The alignment of the sample with the totals of the four sources of income (employment YL, pensions and other transfers YT, self-employment YM, and capital YC) and total net wealth W is obtained with an increase in the deviation standard of the weights that, on average in the years considered, from 1.02 to 1.87.[17]

Aligning the sample estimates of totals to the known values of the various forms of wealth is more difficult. The calibrations that take account of the totals of the main categories of real assets (AR), financial assets (AF) and financial liabilities (PF), in addition to income (Y), converge only in some years, and with a significant increase in the variability of the weights. Imposing additional constraints, such as that of total risky assets (AF3) or the distribution of housing other than the primary residence (OTHERW), the algorithm does not converge. Imposing constraints regarding both income and wealth does not appear feasible.

In short, this first block of calibrations shows that if income convergence is attained with a set of weights whose variability is not too great compared with the initial weights, for wealth convergence is attainable only with much more highly variable weights and with a limited set of variables. Presumably this reflects the greater under-

---

16 The Calmar macro furnishes four criteria to search for solutions: linear, raking, logistic, and linear truncated. we use linear truncation, which in most cases produces a solution and avoids negative weights.

17 According to some estimates based on the 2010 survey, an increase in the standard deviation of the weights due to calibration produces an increase of the same magnitude in the standard errors of the estimates. For example, if the standard error of average income is about €500 in 2010, with an average of €35,000 euro, then doubling the variability of the weights would produce a standard error of €1,000 euro. This is obviously an approximation, but it does allow us to assess, roughly, the impact of calibration on the variability of the estimates.

reporting and greater concentration of wealth than income. Another factor could be some inconsistency between SHIW data and the constraints used in the calibrations.

The calibration of total wealth was replicated (for 2010 only) with an enlarged sample that combines SHIW households with 198 households identified by the Italian Private Banking Association (AIPB) with a sample scheme and a questionnaire similar to those of the SHIW. These households, selected among AIPB bank customers, all hold more than €500,000 worth of financial assets, although, as in SHIW, they do not necessarily declare the full amount possessed.

The integration of the two samples was done by post-stratification, computing in SHIW the share of households with that amount of wealth and reproducing the same share in the combined SHIW-AIPB sample.

The higher frequency of wealthy families in the combined sample produces a smaller increase in the standard deviations of the weights (2.60) when control of totals of the forms of wealth is imposed. The adjustment of the sample weights remains problematic when alignment with the number of properties owned (other than the primary residence) is also required.

The results thus far suggest the difficulty of applying the calibration methods to substantially under-reported data. Therefore, we repeated the calibration experiments on SHIW data whose weights take account of non-responses and whose data on real estate, financial assets and income of the self-employed were adjusted beforehand by the procedures described above (C_3).

Calibrations on adjusted SHIW data on sources of income and total wealth (C_6) have weights of relatively low variability (the standard deviation of the final weights, on average across years, is 1.91). And taking total real assets (AR), financial assets (AF) and financial liabilities (PF), and total income (Y) – (C_7) – the calibrated weights have a variability (1.35) only slightly higher than the design weights (Table 5).

The alignment with the total of types of both income and wealth (C_8), applied to already corrected data, yields weights whose standard deviation is significantly greater (2.77).

Various hypotheses could be evaluated, adding or eliminating constraints. In any case, we believed the material was sufficient for a comparative assessment of the results generated by the foregoing corrections of SHIW data.

Table 5

Result of the calibrations

(Standard deviation of the calibration weights *)

Year	SHIW (C_0)	Non-response weight (C_3)	Controls on totals**					
			YL YM YT YC W (C_4)	AR AF PF AF3 Y (C_5)	YL YM YT YC AR AF PF	YL YM YT YC W (C_6)	AR AF PF Y (C_7)	YL YM YT YC AR AF PF (C_8)
			SHIW data		SHIW + AIPB Data	Adjusted SHIW data***		
1995	0.94	1.04	1.85	No convergence	-	1.99	1.47	2.74
1998	0.98	1.16	1.97	2.76	-	2.10	1.01	3.10
2000	0.94	1.19	1.98	No convergence	-	1.98	1.10	2.84
2002	1.04	1.48	2.12	No convergence	-	2.02	1.57	3.18
2004	1.05	1.34	1.70	No convergence	-	1.75	1.47	2.61
2006	1.04	1.34	1.50	No convergence	-	1.61	1.36	2.48
2008	1.03	1.12	1.64	No convergence	-	1.61	1.46	2.89
2010	1.06	1.21	1.97	2.96	2.60	1.98	1.38	2.26
2012	1.07	1.14	2.08	2.82	-	2.13	1.33	2.83
Mean	**1.02**	**1.24**	**1.84**	**2.85**	**2.60**	**1.91**	**1.35**	**2.77**

(*) The standard deviation of weights in adjustments C_1 and C_2 is equal to that in C_0. (**) Includes the marginal distribution of sex, age and profession of household head, number of household members, size of municipality, and geographical area. (***) Adjustment for non-response, number of houses other than primary residence, the value of houses, financial assets, and income from self-employment.

5. Assessment of the 2012 estimates

Tables A1, A2, A3 and A4 in the Appendix show the average values of income and net worth, by household characteristics, calculated both on SHIW data and on the adjustments considered above.

In the proportional correction (C_1), the greater appreciation of self-employment than salaried income changes the relative position of entrepreneur households with respect to managers, whose incomes are modified only marginally. The other self-employed workers also have larger than average corrections, employees less than average. The average profiles for the other characteristics are not greatly altered by this adjustment. The ratios between the initial and final values of households residing in the various geographic areas, for example, are almost identical.

For net wealth, the procedure tends to the values for the North more than for the Center or South. The wealth of the elderly and the better educated also change more than the average.

Adjustment C_2, which incorporates interviewers' judgments, does not differ greatly from C_1; income and wealth of entrepreneurs and university graduates are revalued somewhat less than C_1, those of other persons and residents in the South a bit more.

Among the corrections denoted as C_3, that for non-response (C_{3A}) yields average revaluations of 9 per cent for income and 15 per cent for wealth. The revaluation is greater for entrepreneurs and other self-employed workers, less for executives and managers.

The correction of self-employment income (C_{3B}), which is increased by 25 per cent, results in a revaluation of total income of 3.9 percent; obviously the increase in total income is greater for the self-employed.

The correction of properties other than the primary residence (C_{3C}), which increases the number of properties owned but decreases their market value, increases income by 3.8 per cent and wealth by 3.1 per cent.

The adjustment of financial assets (C_{3D}) increases the wealth by an average of 18 per cent; and income is indirectly revalued by 3.7 per cent as a result of the allocation of the corresponding earnings. Here, in contrast to the previous two corrections, the appreciation of the wealth of the self-employed and managers is smaller than the average.

Altogether, the four adjustments C_{3A}-C_{3D} result in an increase of 19 per cent in average income and 37.7 per cent in net wealth. Even so, the sample estimates of the totals are lower than the National Accounts figures. The income of the self-employed increases significantly (due to the specific adjustment C_{3B}), but their wealth is revalued by less than the average.

The calibration of the income sources (C_4) involves appreciable revaluations of both income (30 per cent) and net wealth (23 percent), aligning the means to those derived from the national accounts. The revaluations are greater for the self-employed, for larger households, for residents in smaller municipalities (up to 40,000 inhabitants), and in the South. For employees (particularly production workers and teachers), the revaluations are modest. For net wealth – the method only controls for consistency with the aggregate total – the holdings of self-employed workers and entrepreneurs are revalued very substantially, while those of executives, workers and retirees are decreased with respect to the interview figures.

The calibration of the different components of wealth, while controlling for total income (C_5), is quite unstable. Overall, the calibration confirms the indications of correction C_4, with a greater revaluation of both income and wealth of households in the North.

The revaluations generated by calibrations C_6, C_7 and C_8 (applied to the data already adjusted by corrections C_{3A}-C_{3D}) are not always fully concordant. All in all, the greater appreciation of the income of the self-employed and of university graduates is corroborated. But on wealth the results of these household types are mixed, above average in some cases and below in others.

Figures 1 and 2 give an overview of the corrections (the thicker line indicates the unadjusted SHIW estimates). The profiles of income by sex, age and educational attainment show some stability. The results for the northern and central regions have some variability, while the South remains permanently below the average, fluctuating around 30 percentage points below the North. Most of the estimates confirm the higher figures for larger cities.

Overall, the income profiles produced by correction C_3 have the closest correlation with the unadjusted SHIW data, both for income and for wealth. This correction rather faithfully preserves the picture furnished by the unadjusted data.

Among the adjusted estimates of income, the greatest variability is that connected with professional qualification. The estimates of the income of executive, entrepreneur and other self-employed households are quite variable, which may be due in part to their low sample weights.

On the whole the estimates of net wealth confirm this pattern, albeit with sharper revaluations owing to the greater variability of wealth than of income estimates. Note that some of the adjusted estimates of net wealth are smaller than the unadjusted estimates, mainly because of a reduction that takes account of the problems of valuation of properties in survey data.

The corrections frequently produce weaker correlations between the adjusted variables (Tables A5 and A6). The correlation between income and wealth, which in the unadjusted data is 0.57, falls to 0.50 and 0.44 with corrections C_1 and C_2; but with C_3, which imputes houses and financial assets and the incomes generated by these assets, it increases (0.62). The calibrations show no common pattern: in some cases they strengthen the correlation (C_5 and C_7), in others they weaken it. Other indices of the degree of concordance between the variables provide similar indications.

The index of concentration of adjusted income (both absolute and equivalent) is always higher than that of unadjusted income, especially when calibrations are applied. The index of wealth concentration based on the C_3 correction is slightly lower than that computed on unadjusted data. All the corrections applied to adjusted data (C_3) gave higher values than those on unadjusted data (Tables A7 and A8). Overall, these findings appear to suggest that the survey may underestimate the concentration of both aggregates.

Figure 1 – Profiles of household income: comparison among corrections

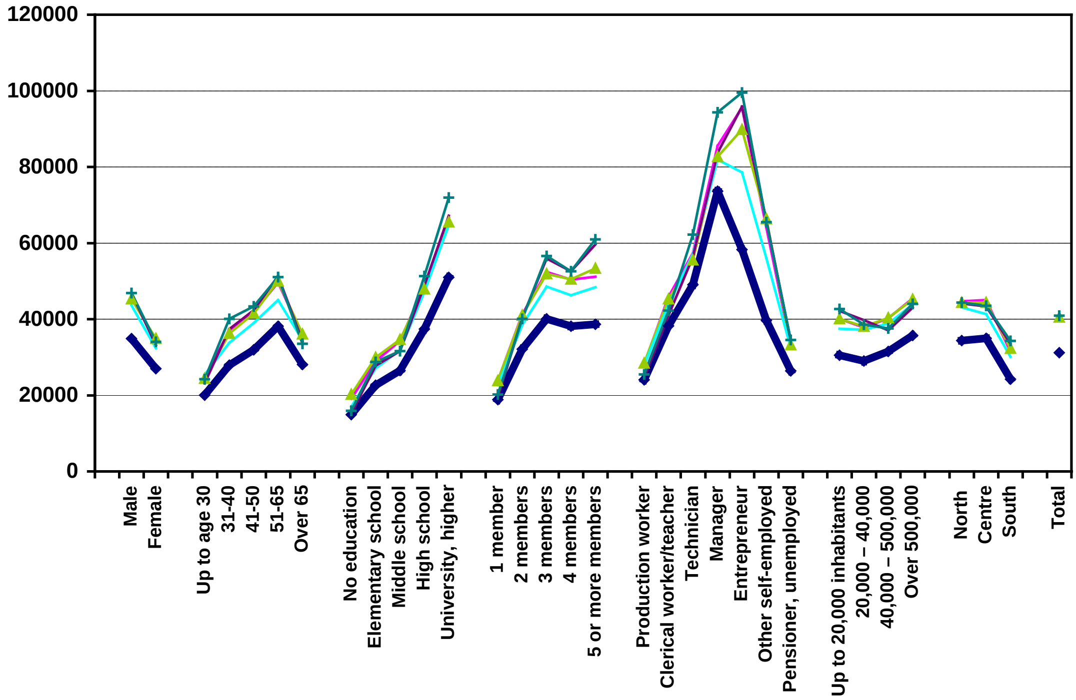

Figure 2 – Profiles of household net wealth: comparison among corrections

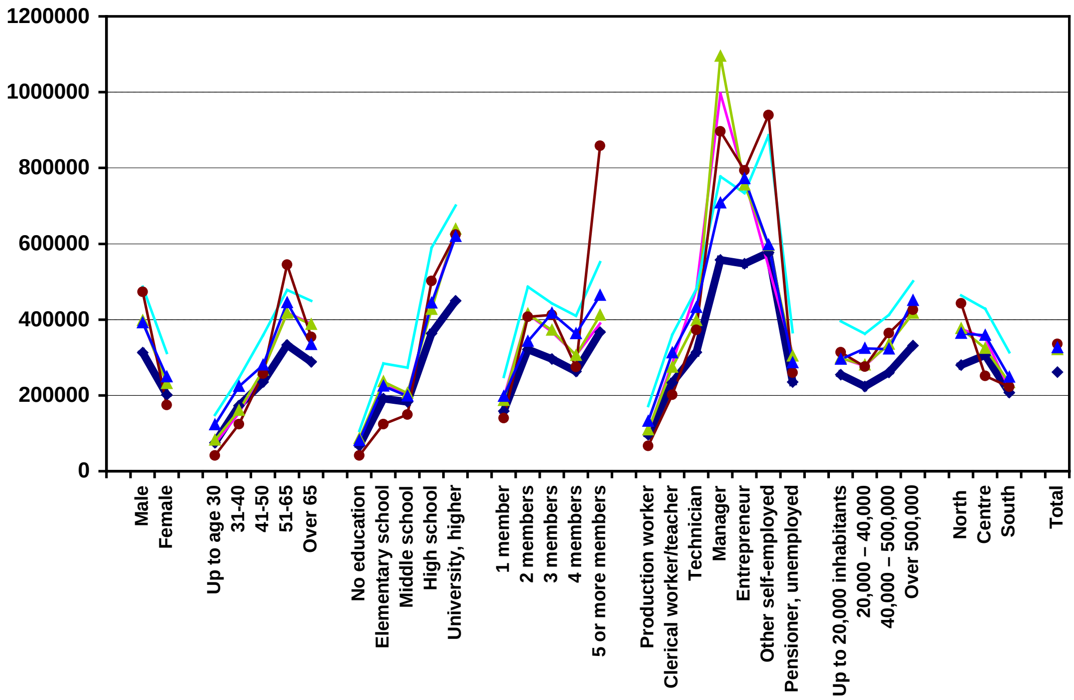

Table A9 shows the variability of these estimates and an estimate of their distance from the National Accounts values – distance which we simply call "bias", although clearly the aggregate estimates too are subject to errors of various kinds.

The SHIW estimates of both income and wealth have low standard error but high bias. All the other estimators have less bias (or none), although the result is obtained by an increase in variance.

These two aspects can be assessed jointly by mean square error. Overall, excluding the estimators that simply re-proportion the values (C_1 and C_2), the C_7 estimator performs best for both income and wealth.

Comparing the distribution of the number of properties (other than the primary residence) as estimated from tax data[18] with our various corrections (Table 6), we see that the SHIW substantially underestimates real estate holdings (85 per cent of the survey households claim they own no other properties, as against 68.2 per cent in the tax data). All the corrections reduce this gap except for C_4, which mainly corrects income rather than wealth, and C_9, which refers to the SHIW-AIPB sample.[19]

The most satisfactory corrections are those of previously adjusted data (C_3), in particular C_6, C_7 and C_8, which are also calibrated.

18 See Neri and Monteduro (2013).

19 Households in the AIPB sample have considerable financial wealth, which makes it possible to align total net wealth without increasing the number of properties.

Table 6

Distribution of number of houses in addition to primary residence in fiscal data and in SHIW original and adjusted data

	Houses other than primary residence							
	0	1	2	3	4	5	6 and more	Total
Fiscal data	68.2	23.0	6.8	0.5	0.5	0.4	0.6	100.0
C_0	85.0	11.7	2.3	0.6	0.2	0.1	0.0	100.0
C_3	66.0	22.9	7.0	2.7	0.8	0.4	0.1	100.0
C_4	87.7	9.5	2.0	0.4	0.2	0.0	0.1	100.0
C_5	82.5	12.6	3.3	0.9	0.6	0.1	0.0	100.0
C_6	72.3	18.9	6.0	2.1	0.3	0.4	0.1	100.0
C_7	68.4	21.5	6.5	2.4	0.6	0.5	0.0	100.0
C_8	73.8	17.9	5.6	1.9	0.3	0.4	0.1	100.0
C_9	90.9	7.2	1.4	0.3	0.1	0.0	0.1	100.0

We have seen that the calibrations can increase the variability of the weights and produce unstable estimators. To get more robust estimates we can take data from contiguous surveys, on the assumption that these represent the structural characteristics of the population, and use the calibration techniques to bring the estimates to the year we want. For example, to get more robust estimates for 2012, we can use data from the 2008, 2010 and 2012 surveys jointly and then calibrate applying the constraints for 2012. The procedure can be replicated for the other two years.[20] Inflation is quite low during our period, but monetary variables can in any case be readjusted on the level of the single-year estimates. Since the estimates of the three years are derived from the same dataset and since they only differ in the constraints used, this method yields information on changes in the profiles induced by changes in the constraints themselves.

Table A4 shows how the C_6 corrections for 2008-2012 compare with these robust estimators (C_6^R). The C_6^R estimators are much less variable from year to year than the C_6 estimators, since as noted they express only the effect of the change of constraints. We find no excessive, implausible changes in these estimators, like that, for example, for the C_6 estimator for households whose head has lower secondary education.

This method can also be used to evaluate forecast or simulated scenarios, using constraints for years subsequent to those to which the micro data refer. To assess this practice, we have applied the 2012 constraints for C_6 to the 2010 SHIW data, after re-proportioning the mean of income and wealth.

The percentage changes between the 2010 and the 2012 estimates based on the C_6 correction show some consistency with those obtained by comparing the C_6^R robust estimates for 2010 and 2012 (the correlation is 0.54), indicating that a significant part of the information contained in the constraints is transferred to the estimates.

20 As is shown in Cannari D'Alessio (2003), given a panel component and phenomena that are correlated across time, this problem should be taken into account at the weighting stage. Panel households that are interviewed twice should be weighted by a function that is inversely proportional to the correlation of income and wealth across time: $(1+\rho)$, where ρ is the correlation. The weight of panel households that are interviewed three times should be adjusted by $[1+ (4/3)\rho + (2/3)\rho^2]$. For simplicity, in the present paper none of these corrections is used.

6. Conclusion

We have examined various methods of correction for non-sampling errors (mainly selectivity bias and under-reporting) in the data of the Bank of Italy's Survey of Household Income and Wealth. Corrections based on specific knowledge of the phenomena are costly, require many assumptions, and do not totally fill the gaps between the estimates and information from other sources. Calibrations appear to be an interesting instrument, but when applied to the SHIW data they are effective only in conjunction with the model-based corrections. When the estimates are very distant from the constraints, in fact, the calibrations do not converge, and even when they do they produce very unstable estimators.

In practice, application of corrections based on SHIW data indicates that a single, all-purpose correction is hard to conceive of. Adjusting the various sources of income may entail greater difficulty in obtaining adequate estimates for the components of wealth, and vice versa.

All in all, the various corrections yield quite similar profiles of the main demographics, profiles that are also similar to unadjusted SHIW data (the correction of individual phenomena, C_3, is the most conservative with respect to the SHIW estimates). In other words, very often the adjustment does not significantly affect the relative positions of the different groups of households. However, mean square error analysis shows that the calibrations that perform best are those that correct the various components of the variables examined (corrections C_4 and C_7 for income and C_7 for wealth).

Our inquiry suggests that the unadjusted SHIW data underestimate the Gini concentration indexes of both income and wealth.

The calibration-based correction methods appear to be promising both for interpretation and for designing forecasting scenarios. When the variance of the calibrated estimators seems to be too great, more robust results can be obtained by aggregating successive waves of the survey.

Appendix A –Statistical tables

Table A1

Mean income in unadjusted SHIW data (C_0) and in the adjustments (C_1-C_8), 2012

(euros)

	SHIW	Adjustments			Calibrations				
	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
Gender									
male	34,896	45,656	45,312	43,619	46,918	49,028	46,891	46,016	46,938
female	26,982	34,478	34,878	32,348	33,196	30,897	34,043	34,224	33,920
Age									
30 and under	20,058	24,634	24,413	25,397	23,123	22,849	24,218	25,386	25,373
31 - 40	27,917	36,885	36,232	33,820	37,434	33,033	40,101	36,315	41,318
41 - 50	31,912	41,830	41,511	38,957	42,388	42,279	43,321	43,801	43,146
51 - 65	38,118	49,683	50,010	45,072	51,295	51,384	51,105	50,746	50,335
over 65	28,129	35,761	36,051	34,428	34,242	36,569	33,536	33,872	33,479
Educational qualification									
none	14,962	19,160	20,179	17,011	15,252	13,331	15,930	16,434	16,339
primary school certificate	22,658	28,901	29,965	27,079	27,796	21,683	28,752	26,212	27,057
lower secondary school certificate	26,488	34,282	34,595	31,895	31,644	29,631	31,635	32,126	31,123
upper secondary school diploma	37,439	48,698	47,866	47,203	48,782	48,537	51,303	48,740	52,681
university degree	50,947	66,634	65,482	64,726	67,189	69,742	71,923	71,873	70,271
Household size									
1 member	18,888	23,812	23,851	22,676	19,456	21,488	20,236	22,159	20,687
2 members	32,131	41,227	40,966	38,483	40,309	42,907	40,068	39,176	39,111
3 members	40,082	52,371	51,940	48,585	55,952	60,461	56,639	56,197	59,884
4 members	38,129	50,399	50,456	46,261	52,618	42,328	52,549	52,937	49,870
5 members or more	38,686	51,154	53,339	48,425	59,624	53,845	60,966	50,290	60,281
Work status									
Employee	24,039	28,195	28,386	27,874	24,618	25,107	25,506	27,764	25,165
blue-collar worker	38,275	45,901	45,244	43,832	41,005	42,625	42,244	45,490	42,085
office worker	49,085	57,328	55,542	57,071	56,378	58,941	62,258	60,127	61,721
manager, executive	73,602	85,523	82,673	81,945	83,671	95,031	94,366	92,316	91,391
Self-employed - business-owner, member of profession	58,320	95,637	89,847	78,566	95,844	83,940	99,524	97,868	104,631
other self-employed	39,675	64,108	66,348	56,038	65,913	61,880	65,533	59,794	63,722
retired and other	26,455	33,058	33,214	32,423	34,272	31,694	34,538	31,952	34,330
Town size									
up to 20,000 inhabitants	30,554	40,087	40,063	37,456	42,148	39,808	42,721	38,559	43,474
20,000 - 40,000	29,033	37,779	38,059	37,244	39,748	35,985	38,560	40,092	38,286
40,000 - 500,000	31,506	40,284	40,301	38,842	37,026	41,791	37,582	40,761	36,580
more than 500,000	35,760	45,488	45,232	43,886	42,994	46,753	44,038	48,398	43,319
Geographical area									
North	34,400	44,704	44,375	43,138	44,109	51,592	44,351	46,598	43,818
Centre	34,971	45,021	44,377	41,379	43,336	38,437	43,499	43,530	43,458
South and Islands	24,247	31,430	32,307	30,079	33,577	25,353	34,270	29,675	35,031
TOTAL	**31,236**	**40,487**	**40,487**	**38,602**	**40,579**	**40,663**	**40,964**	**40,563**	**40,932**

(*) Individual characteristics refer to the head of household, defined as the member with the highest income.

Table A2

Mean income in the single adjustments of C_3 (C_{3A}-C_{3D}), 2012

(euros)

	Adjustments				
	C_{3A}	C_{3B}	C_{3C}	C_{3D}	C_3
Gender					
male	38,300	36,391	36,318	36,190	43,619
female	28,691	27,883	27,920	27,991	32,348
Age					
30 and under	23,357	20,501	20,345	20,969	25,397
31 - 40	30,186	29,346	28,553	28,690	33,820
41 - 50	34,342	33,564	32,897	32,848	38,957
51 - 65	39,822	39,760	39,547	39,425	45,072
over 65	30,016	28,702	29,680	29,547	34,428
Educational qualification					
none	15,735	15,095	15,188	15,331	17,011
primary school certificate	24,210	23,108	23,437	23,690	27,079
lower secondary school certificate	28,739	27,433	27,263	27,172	31,895
upper secondary school diploma	40,729	39,153	39,083	39,054	47,203
university degree	56,365	53,503	53,418	53,000	64,726
Household size					
1 member	19,838	19,433	19,569	19,802	22,676
2 members	33,520	33,124	33,949	33,341	38,483
3 members	43,362	41,794	41,271	41,258	48,585
4 members	40,727	40,054	39,263	39,698	46,261
5 members or more	43,451	40,611	39,770	39,625	48,425
Work status					
Employee	25,852	24,203	24,635	24,824	27,874
blue-collar worker	39,833	39,031	39,516	39,426	43,832
office worker	51,991	49,765	50,828	50,307	57,071
manager, executive	75,530	74,443	77,338	74,132	81,945
Self-employed - business-owner, member of profession	65,930	66,848	60,356	59,325	78,566
other self-employed	44,584	45,235	41,162	41,630	56,038
retired and other	28,836	26,855	27,662	27,611	32,423
Town size					
up to 20,000 inhabitants	33,068	31,799	31,518	31,619	37,456
20,000 - 40,000	32,673	30,183	30,024	30,458	37,244
40,000 - 500,000	34,138	32,643	33,053	32,619	38,842
more than 500,000	38,834	37,143	37,353	37,105	43,886
Geographical area					
North	37,557	35,781	35,821	35,857	43,138
Centre	37,011	36,366	35,997	36,059	41,379
South and Islands	26,892	25,121	25,210	25,004	30,079
TOTAL	**34,022**	**32,457**	**32,435**	**32,398**	**38,602**

Table A3

Mean net wealth in unadjusted SHIW data (C_0) and in the adjustments (C_1-C_8), 2012

(euros)

	SHIW	Adjustments			Calibrations				
	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
Gender									
male	313,142	395,878	396,845	486,200	445,548	473,063	448,867	391,087	441,453
female	201,529	232,328	231,204	311,990	179,729	174,803	206,160	248,854	211,101
Age									
30 and under	75,172	72,694	81,676	148,506	61,198	41,535	85,978	122,669	88,106
31 - 40	173,029	153,789	160,344	246,223	191,824	124,017	236,168	223,768	228,159
41 - 50	235,283	266,314	266,435	358,123	276,510	257,570	276,423	279,588	289,036
51 - 65	332,956	417,878	416,406	478,319	510,957	544,980	509,564	444,326	515,117
over 65	288,292	389,963	386,753	449,206	295,862	354,265	317,270	334,303	301,416
Educational qualification									
none	67,374	82,588	86,006	106,114	51,395	41,383	57,295	80,881	56,333
primary school certificate	191,477	226,259	236,060	284,143	240,571	123,928	247,801	223,896	199,091
lower secondary school certificate	182,845	205,294	205,946	273,627	191,150	149,166	202,385	197,959	172,152
upper secondary school diploma	363,070	436,489	426,455	589,566	479,851	501,753	529,689	443,403	532,151
university degree	449,685	635,338	638,474	700,396	502,181	624,223	558,869	618,646	653,561
Household size									
1 member	158,900	193,970	186,534	248,368	122,666	139,735	146,073	197,712	161,318
2 members	321,158	413,888	414,602	486,939	341,427	407,243	388,339	342,226	324,848
3 members	296,399	367,437	371,548	442,330	391,145	412,504	427,316	416,429	471,691
4 members	263,033	306,232	305,178	410,031	377,224	275,107	309,114	362,359	318,121
5 members or more	366,466	389,185	411,193	551,597	817,117	858,798	797,243	463,868	820,275
Work status									
Employee	96,471	105,360	108,990	172,406	58,246	66,587	80,327	131,460	71,349
blue-collar worker	233,888	285,337	274,269	360,461	199,005	201,665	236,049	312,029	252,138
office worker	314,046	480,278	397,120	480,560	249,325	372,497	372,771	431,929	401,548
manager, executive	557,100	996,609	1,094,885	777,241	579,579	896,072	864,959	707,923	941,308
Self-employed - business-owner, member of profession	547,060	770,382	754,938	733,743	920,098	793,847	895,116	771,769	1,101,719
other self-employed	576,900	538,551	595,709	885,983	1,111,552	939,068	962,781	596,785	836,193
retired and other	235,546	310,730	303,212	367,175	240,909	259,850	271,715	285,579	271,825
Town size									
up to 20,000 inhabitants	254,598	298,791	298,746	395,293	382,388	313,550	399,380	295,236	377,534
20,000 - 40,000	223,650	277,055	280,066	362,644	286,335	275,675	237,035	323,598	263,947
40,000 - 500,000	260,980	336,078	335,093	412,585	229,688	363,926	249,370	321,792	267,063
more than 500,000	331,793	417,698	416,598	500,789	331,267	425,749	393,168	450,713	395,700
Geographical area									
North	279,878	375,508	375,954	463,468	319,016	442,376	339,924	363,447	310,132
Centre	306,401	351,747	323,652	428,075	353,078	251,373	369,361	358,237	409,332
South and Islands	207,352	217,977	233,808	314,043	310,512	222,490	313,040	248,190	329,613
TOTAL	**261,529**	**320,248**	**320,248**	**408,649**	**322,745**	**335,459**	**336,906**	**325,315**	**335,169**

Table A4

Mean net wealth in the single adjustments of C_3 (C_{3A}-C_{3D}), 2012

(euros)

	Adjustments				
	C_{3A}	C_{3B}	C_{3C}	C_{3D}	C_3
Gender					
male	366,284	313,142	315,556	365,248	486,200
female	222,716	201,529	216,537	243,892	311,990
Age					
30 and under	97,007	75,172	75,087	105,659	148,506
31 - 40	186,304	173,029	173,554	202,327	246,223
41 - 50	273,881	235,283	238,080	273,849	358,123
51 - 65	363,390	332,956	336,312	383,954	478,319
over 65	316,253	288,292	309,303	350,482	449,206
Educational qualification					
none	75,513	67,374	67,858	87,724	106,114
primary school certificate	212,423	191,477	186,786	240,180	284,143
lower secondary school certificate	210,510	182,845	183,575	210,603	273,627
upper secondary school diploma	428,140	363,070	389,557	430,034	589,566
university degree	510,320	449,685	463,376	520,822	700,396
Household size					
1 member	174,855	158,900	162,895	199,569	248,368
2 members	345,679	321,158	348,223	374,183	486,939
3 members	331,435	296,399	296,618	345,811	442,330
4 members	307,291	263,033	263,323	317,080	410,031
5 members or more	479,719	366,466	357,804	398,954	551,597
Work status					
Employee	118,037	96,471	103,454	127,393	172,406
blue-collar worker	260,801	233,888	240,228	276,238	360,461
office worker	338,169	314,046	324,740	367,381	480,560
manager, executive	578,530	557,100	589,716	574,189	777,241
Self-employed - business-owner, member of profession	621,936	547,060	540,922	578,972	733,743
other self-employed	714,447	576,900	566,758	660,781	885,983
retired and other	260,338	235,546	248,913	284,463	367,175
Town size					
up to 20,000 inhabitants	299,057	254,598	253,454	300,698	395,293
20,000 - 40,000	262,940	223,650	222,438	276,282	362,644
40,000 - 500,000	291,127	260,980	287,812	306,111	412,585
more than 500,000	380,728	331,793	347,308	384,794	500,789
Geographical area					
North	333,286	279,878	284,053	342,881	463,468
Centre	331,216	306,401	307,657	352,443	428,075
South and Islands	238,414	207,352	225,852	232,522	314,043
TOTAL	**302,374**	**261,529**	**269,767**	**309,130**	**408,649**

Table A5

Consistency between income and wealth in SHIW data (C_0) and in the adjustments (C_1-C_8)

(means 1995-2012)

	SHIW	Adjustments			Calibrations				
		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
Correlation between income and wealth.............	0.573	0.499	0.438	0.619	0.462	0.729	0.499	0.616	0.530
Cronbach alpha (*).............	0.768	0.760	0.721	0.786	0.779	0.828	0.760	0.799	0.746
Variance explained by the first principal component (*)	0.420	0.415	0.373	0.454	0.436	0.504	0.417	0.467	0.415

(*) Variables: Y, YL, YT, YM, YC, AR, AF, PF, W.

Table A6

Consistency between income and wealth in SHIW data (C_0) and in the single adjustments of C_3 (C_{3A}-C_{3D})

(means 1995-2012)

	Adjustments				
	C_{3A}	C_{3B}	C_{3C}	C_{3D}	C_3
Correlation between income and wealth	0.593	0.584	0.640	0.565	0.438
Cronbach alpha (*)...	0.777	0.770	0.774	0.760	0.721
Variance explained by the first principal component (*).....................	0.432	0.427	0.433	0.413	0.373

(*) Variables: Y, YL, YT, YM, YC, AR, AF, PF, W.

Table A7

Concentration index in SHIW data (C_0) and in the adjustments (C_1-C_8)

(Gini index, means 1995-2012)(*)

	SHIW	Adjustments			Calibrations				
	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
Household income	0.350	0.383	0.389	0.383	0.403	0.439	0.401	0.409	0.403
Equivalent income	0.319	0.351	0.357	0.359	0.367	0.422	0.364	0.383	0.377
Net wealth	0.596	0.632	0.637	0.577	0.680	0.693	0.656	0.621	0.702

(*) *Winsorized estimates (1st and 99th percentile).*

Table A8

Concentration index in SHIW data (C_0) and in the single adjustments of C_3 (C_{3A}-C_{3D})

(Gini index, means 1995-2012)(*)

	Adjustments				
	C_{3A}	C_{3B}	C_{3C}	C_{3D}	C_3
Household income	0.356	0.360	0.367	0.350	0.383
Equivalent income	0.325	0.329	0.319	0.319	0.359
Net wealth	0.576	0.596	0.595	0.590	0.577

(*) *Winsorized estimates (1st and 99th percentile).*

Table A9

Variability of estimators of income and net wealth in SHIW (C_0) and in the adjustments (C_1-C_8)

(euros)

	SHIW	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
	Income								
Mean	31,236	40,487	40,487	37,384	40,579	40,672	40,945	40,564	40,766
Std.err.	364	581	608	552	1,291	1,617	1,327	1,118	1,563
Bias	9,251	0	0	3,103	92	185	458	77	279
MSE	9,258	581	608	3,152	1,295	1,628	1,404	1,121	1,587
	Net wealth								
Mean	261,529	320,248	320,248	360,107	322,749	335,648	341,268	325,434	329,100
Std.err.	10,087	10,519	11,727	17,811	36,587	37,902	38,903	15,412	35,552
Bias	58,719	0	0	39,859	2,501	15,400	21,020	5,186	8,852
MSE	59,579	10,519	11,727	43,657	36,672	40,911	44,219	16,261	36,637

Table A4

Mean income – Adjustment C_6 and simulations, 2008 - 2012

(euros)

	Adjustment C_6			Simulation(*)			Forecast(**)
	2008	2010	2012	2008	2010	2012	2012
Gender							
male	48,820	46,868	46,890	47,746	47588	47,692	47,668
female	33,316	34,816	34,041	35,195	33900	33,082	33,003
Age							
30 and under	23,932	26,889	24,219	24,837	25648	23,716	28,046
31 - 40	36,839	38,620	40,099	38,149	39245	37,748	33,105
41 - 50	44,651	46,048	43,319	46,314	44995	43,521	43,269
51 - 65	55,256	52,756	51,105	55,099	51341	50,742	52,957
over 65	38,296	32,749	33,536	36,598	34522	34,876	34,476
Educational qualification							
none	18,160	16,929	15,931	17,083	17388	16,219	15,940
primary school certificate	29,396	24,586	28,752	28,543	28254	27,330	22,933
lower secondary school certificate	42,632	39,776	31,634	38,904	38389	37,205	38,565
upper secondary school diploma	48,774	46,360	51,301	50,286	48778	46,905	46,827
university degree	71,673	72,270	71,920	72,832	72231	69,996	68,427
Household size							
1 member	23,661	23,000	20,236	23,272	22216	21,329	21,082
2 members	43,377	40,165	40,066	42,361	39729	41,138	43,295
3 members	56,156	50,358	56,635	55,645	53551	53,488	49,229
4 members	54,449	56,953	52,542	55,822	54597	54,838	55,755
5 members or more	54,968	51,832	60,968	59,317	54140	54,691	55,743
Work status							
Employee	26,441	27,261	25,507	26,777	27199	24,964	25,506
blue-collar worker	41,964	46,722	42,244	44,590	45071	42,093	43,899
office worker	66,172	53,491	62,259	61,112	63035	57,987	48,261
manager, executive	98,933	82,508	94,367	93,087	93716	89,551	78,889
Self-employed - business-owner, member of profession	82,033	85,744	99,510	90,127	85231	85,911	87,241
other self-employed	75,390	71,755	65,528	73,978	69262	70,000	66,239
retired and other	36,290	33,375	34,537	35,717	34075	34,689	35,035
Town size							
up to 20,000 inhabitants	46,783	39,641	42,719	44,605	42983	42,224	40,656
20,000 - 40,000	36,145	39,799	38,558	37,981	38130	37,236	36,476
40,000 - 500,000	41,604	42,572	37,581	42,082	40422	40,022	40,893
more than 500,000	41,130	47,994	44,038	46,009	42050	42,252	46,947
Geographical area							
North	49,148	44,274	44,348	47,289	45409	44,818	44,633
Centre	44,352	49,353	43,499	46,995	45721	45,616	48,605
South	32,953	32,271	34,271	34,074	32858	32,303	30,679
TOTAL	**43,149**	**41,520**	**40,963**	**43,149**	**41520**	**40,963**	**40,963**

(*) The simulation proceeds by the following steps: 1) a single database is constructed using the 2008, 2010 and 2012 waves; 2) average household income is aligned to the value resulting in the year specified; 3) correction C6 is then applied.

(**) The forecast is obtained by applying the 2012 external information to the 2010 sample. We use the C6 correction. The average household income for the 2010 sample is aligned to the figure resulting in the 2012 wave.

References

Banca d'Italia (1970), *Risparmio e struttura della ricchezza delle famiglie italiane nel 1968*, in A. Ulizzi, (a cura di), Bollettino, Banca d'Italia, n. 1, gennaio-febbraio, pp. 103-167.

Banca d’Italia (2012), *La ricchezza delle famiglie italiane - 2011*, Supplementi al Bollettino statistico n. 65, dicembre.

Banca d’Italia (2013), *Sondaggio congiunturale sul mercato delle abitazioni in Italia*, Luglio 201, Supplementi al Bollettino statistico n. 41, agosto.

Banca d’Italia (2014), *I bilanci delle famiglie italiane nell'anno 2012*, a cura di F. Carta, R. Gambacorta, G. Ilardi, A. Neri, C. Rondinelli, Supplementi al Bollettino Statistico (nuova serie), Banca d'Italia, n. 5, Gennaio

BCE (2013), The Eurosystem Household Finance and Consumption Survey - Methodological Report for the First Wave. Statistics Paper Series, N.1, April.

Bonci, R., G. Marchese, A. Neri (2005), *La ricchezza finanziaria nei conti finanziari e nell’indagine sui bilanci delle famiglie italiane*, Temi di Discussione n. 565 - Novembre.

Brandolini A. (1999), *The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality*, Giornale degli Economisti e Annali di Economia, vol. 58, n. 2, pp. 183-239.

Brandolini A. L. Cannari, G. D'Alessio, I. Faiella, (2004), *Household wealth distribution in Italy in the 1990s*, Working papers, The Levy Economics Institute, No. 414.

Brick, J. M. (2013), *Unit Nonresponse and Weighting Adjustments: A Critical Review*, Journal of *Official* Statistics, n. 29(3): 329-469.

Cannari L., G. D’Alessio (1990), *Housing Assets in the Bank of Italy's Survey of Household Income and Wealth*, in Dagum e Zenga (a cura di), “Income and Wealth Distribution, Inequality and Poverty”, Springer Verlag, Berlino, p. 326-334.

Cannari L., G. D’Alessio (1992), *Mancate interviste e distorsione degli stimatori*, Banca d'Italia, Temi di discussione, n.172.

Cannari L., G. D’Alessio (1993), *Non-reporting and Under-reporting Behavior in the Bank of Italy's Survey of Household Income and Wealth*, in “Bulletin of the International Statistical Institute”, vol. LV, n. 3, Pavia, p. 395-412.

Cannari L., G. D'Alessio (2003), La distribuzione del reddito e della ricchezza nelle regioni italiane,Temi di Discussione n. 482, Banca d’Italia, Roma, Giugno.

Cannari L., G. D’Alessio, G. Raimondi, A.I. Rinaldi (1990), *Le attività finanziarie delle famiglie italiane*, Banca d'Italia, Temi di discussione, n. 136.

Cannari L., R. Violi (1995), *Reporting Behaviour in the Bank of Italy's Survey of Italian Household Income and Wealth*, Research on Economic Inequality, vol. 6, JAI Press Inc., pp. 117-130

Cifaldi G., A. Neri (2013), *Asking income and consumption questions in the same survey: what are the risks?*, Temi di discussione, n. 908 - Aprile.

D'Alessio G., I. Faiella (2002), *Nonresponse behaviour in the Bank of Italy's Survey of Household Income and Wealth*, Banca d'Italia, Temi di discussione, n. 462.

D'Alessio G., S. Iezzi (2014), *How the time of interviews affects estimates of income and wealth*, mimeo, Banca d'Italia.

D'Alessio G., G. Ilardi (2012), *Non sampling errors in sample surveys: the Bank of Italy's experience*, in C. Davino e L. Fabbris (Eds), *Survey data collection and integration*, Springer-Verlag.

D'Aurizio L., I. Faiella, S. Iezzi, A. Neri (2006), *L'under-reporting della ricchezza finanziaria nell'indagine sui bilanci delle famiglie*, Temi di discussione n. 610.

Deville, Jean-Claude (2000), *Generalized calibration and application to weighting for non-response*, COMPSTAT, Physica-Verlag HD: 65-76.

Deville, J., C. Särndal, (1992), *Calibration estimators in survey sampling*, Jour. Amer. Statist. Assoc. n. 87, 376-382.

Folsom R.E, Singh, A.C. (2000*). The Generalized Exponential Model for SamplingWeight Calibration for Extreme Values, Nonresponse, and Poststratification*, ASA Proceedings of the Section on Survey Research Methods, 598-603.

Fuller, W.A., Loughin, M.M., Baker, H.D, (1994), *Regression Weighting for the 1987-88 National Food Consumption Survey*, Survey Methodology, n. 20: 75-85.

Gelman, A., (2007) *Struggles with Survey Weighting and Regression Modeling*, Statistical Science 22, n. 2: 153--164.

Hurst E., G. Li, B. Pugsley (2010), *Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self-Employed*, NBER WP 16527, Ca,bridge MA.

Kott, P.S, Chang, T, (2010) *Using calibration weighting to adjust for non-ignorable unit nonresponse,* Journal of the American Statistical Association n. 105(491):1265–1275.

Kott P.S, Liao D, (2012), *Providing double protection for unit nonresponse with a nonlinear calibration-weighting routine*, Survey Research Methods, 6 n. 2: 105-111.

Little, R J., Vativarian S. (2005), *Does Weighting for Nonresponse Increase the Variance of Survey Means?,* Survey Methodology n. 31(2):161–68.

Lohr, S. (2007) Comment*: Struggles with Survey Weighting and Regression Modeling*, Statistical Science 22, no. 2: 175--178.

Sautory O. (1993), *Le macro CALMAR, Redressement d'un échantillon par calage sur marges*, Document n. F 9310, INSEE.

Neri A., T. Monteduro (2013), *La ricchezza immobiliare delle famiglie italiane: un confronto fra dati campionari e censuari*, Questioni di Economia e Finanza, n. 146 - Gennaio.

Neri A., M.G. Ranalli (2011), *To misreport or not to report? The measurement of household financial wealth*, Statistics in transition new series, 12, 2, 281-300.

Neri A., R. Zizza (2010), *Income reporting behaviour in sample surveys*, Banca d'Italia, Temi di discussione, n. 777.

Oh, H. L., Scheuren, F. J. (1983). Weighting adjustments for unit non-response. In W. G. Madow, I. Olkin, and D. B. Rubin (Eds.), *Incomplete data in sample surveys* (Vol. 2): Theory and bibliographies, pp. 143 184. Academic Press (New York; London).

Nicolini, G., Marasini, D., Montanari, G.E., Pratesi, M., Ranalli, M.G., Rocco, E, (2013), *Metodi di stima in presenza di errori non campionari*, UNITEXT, Collana di Statistica e Probabilità Applicata, Springer Science & Business.

Pissarides C.A., G. Weber (1989), *An Expenditure Based Estimate of Britain's Black Economy*, Journal of Public Economics, 39(1), Giugno.

Rubin, D B, (1978) *Multiple Imputations in Sample Surveys—A Phenomenological Bayesian Approach to Nonresponse*, Proceedings of the Survey Research Methods Section, American Statistical Association.

Rubin, D.B, (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley.

Statistics Canada, (2009), *Quality Guidelines*, Fifth Edition.