# Dynamics and measurement error in household income data collected with single questions

Tullio, Federico

Bank of Italy

2025

# DYNAMICS AND MEASUREMENT ERROR IN HOUSEHOLD INCOME DATA COLLECTED WITH SINGLE QUESTIONS*

FEDERICO TULLIO†

## ABSTRACT

I provide insights into the dynamics of income collected in surveys using single questions, by extending to longitudinal settings a measurement error model previously developed in the literature. In this framework, single-question income data are validated against a benchmark provided by detailed source-by-source questions, which are considered the best practice for measuring income in surveys. I outline the assumptions required to infer benchmark income changes between two time periods (e.g., two subsequent survey waves), both at the macro- and micro-levels, when income is collected using single questions. Potential heterogeneity in respondents' misreporting behaviour in single questions and its implications in longitudinal settings are also discussed. I apply the methodology to estimate income changes in Italy between 2022 and 2023, using data from a new web-survey conducted by the Bank of Italy.

*JEL codes*: C81, C83, I3

*Keywords*: Income data, Measurement error, Data quality

---

* The views expressed herein are mine and should not be attributed to the Bank of Italy.

† federico.tullio@bancaditalia.it. Bank of Italy, DG Economics, Statistics and Research, via Nazionale 91, 00184 Rome, Italy.

## 1. Introduction

An essential aspect of household survey questionnaire design is to minimise the response burden on households, thereby reducing non-response rates. To this aim, earnings and income data are often collected with single-summary questions covering all forms of earnings or income received by an individual or by households. On the one hand, compared to detailed source-by-source questions, which represent the best-practice approach for collecting earnings and income data in sample surveys (United Nations, 2011), single questions provide less guidance in helping respondents retrieve the desired income measure (i.e., understanding which income sources to include, recalling the information from memory, and aggregating each source), particularly when assessing household total income. On the other hand, single questions offer the clear advantage of reducing respondent burden and permitting greater focus on other survey priorities. Single questions on household and individual earnings and income are therefore common in surveys (see, e.g., Micklewright and Schnepf, 2010, and references therein; Institute for Social and Economic Research, 2021; and Guiso and Jappelli, 2024), making the reliability of data collected through this method a crucial issue. This matter is particularly relevant for web surveys, which suffer from low response rates compared to other interviewing modes (Daikeler et al., 2020), and its importance is likely to increase in the future. Indeed, the increasing demand for timely quantitative and qualitative information on household economic conditions is prompting data producers to adopt web surveys more frequently than traditional surveys, which are mainly performed by in-person interviews and are typically conducted annually or at longer intervals, with results often released a year after the conclusion of the data collection process.[1] Recent examples of web surveys are the new Bank of Italy's Conjunctural Survey of Italian Households (CSIH), and the ECB's Consumer Expectations Survey, among others.

Although there is a long-standing tradition of validating income (or earnings) measures against a benchmark, using measurement error models (Fuller, 1967), the benchmark is typically represented by administrative data. Bingley and Martinello (2017), for instance, propose a framework for validating income survey data that allows for measurement error in the validation data (represented by administrative records), building a system of moment restrictions to identify the characteristics of

[1] The spread of the Covid-19 pandemic in 2020 highlighted the importance of alternative data collection tools, as many field activities involving in-person interviews were suspended to protect the health of respondents and interviewers, and several new web surveys were launched to obtain timely insights into the economic situation of households during the pandemic-related crisis, such as the Bank of Italy's Special Survey of Italian Households (Neri and Zanichelli, 2020), the Understanding Society COVID-19 study (Institute for Social and Economic Research, 2021), and the Survey on COVID-19 and Consumption (Immordino et al., 2022), among others.

measurement error in survey-reported gross income, and estimating the model parameters using the Generalized Method of Moments (GMM; Hansen, 1982).[2]

A different strand of literature analyses whether single questions provide good measures of individual and household earnings and income compared to sets of source-by-source questions. The latter are assumed to be closer to underlying true income values and serve as a reasonable benchmark for what could be achieved with substantially more resources (both in terms of economic cost and respondents' burden) dedicated to income data collection. Micklewright and Schnepf (2010) investigate the reliability of single questions on individual and household income using UK data from a macro-level perspective (in the terminology of Bound et al., 2001), namely, they compare the two approaches implemented in separate samples.[3] The authors find that single questions on individual income provide a distribution that is very close to the distribution based on detailed questions, whereas the question on total household income induce lower item response rates and lead to mild income understatement on average. Crossley et al. (2023) evaluate earnings and income data collected with single questions through short web surveys fielded during the pandemic - the Understanding Society Covid-19 Study. The statistical method employed by Crossley et al. (2023) builds on the measurement error model developed by Bingley and Martinello (2017), making use of the data source with the detailed source-by-source questions (Understanding Society Main Study) as validation archive. The authors conclude that measures of household earnings and income from single questions are noisier than those from the validation archive, and that there is evidence of systematic under-reporting, even though measurement errors in single questions are substantially uncorrelated with true earnings and income values.

While these studies validate single-question measures of individual and household earnings and income in a cross-sectional perspective, insufficient attention has been devoted, to my knowledge, to providing insights into the dynamics of income and earnings collected through single questions compared to sets of source-by-source questions. This issue is of primary importance from a longitudinal perspective, as conjunctural analyses based on survey data often aim to capture the distributional implications of income evolution over time. To fill this gap in the literature, I employ the measurement error model developed by Bingley and Martinello (2017) within a framework that is similar to that of Crossley et al. (2023). I extend the model setting to a longitudinal framework, eliciting the further assumptions that are required to draw inference on the income change between

[2] Other examples include Kapteyn and Ypma (2007), Abowd and Stinson (2013), and Jenkins and Rios-Avila (2023), among others.

[3] In other words, households answering to single questions and those answering to the set of source-by-source questions are not the same.

two time periods (say, two subsequent editions of a survey), using single questions, both at the macro- and the micro-level. The implications of potential heterogeneity in respondents' misreporting behaviour are also discussed. I implement the proposed method to estimate household income changes in Italy between 2022 and 2023, employing the first two editions (fielded in August/September 2023 and March/April 2024) of a new web-survey conducted by the Bank of Italy, the Conjunctural Survey of Italian Households (CSIH), which was launched to track household finances in the periods when the main survey, the Survey on Household Income and Wealth (SHIW), is not conducted. I find that the average income of Italian households increased by 6.1 per cent between 2022 and 2023. Income change was higher for households in the upper half of the income distribution and for those residing in Southern Italy.

The remainder of this paper is organised as follows. The next section outlines the data employed in the analysis. Section 3 describes the adopted methodology, based on the longitudinal extension of a measurement error model developed by Bingley and Martinello (2017). Section 4 discusses the results of the empirical application using Italian data from 2022 and 2023. Concluding remarks are provided in the last section.

## 2. Data

I make use of the Bank of Italy's main household survey, the Survey on Household Income and Wealth (SHIW), and the new Conjunctural Survey on Italian Households (CSIH).

The SHIW is conducted by the Bank of Italy since 1965, and it collects information on demographics, income, real and financial assets, and loans for a representative sample of Italian households. It is conducted every two years starting from 1989, thus leaving an informative gap on the economic conditions of Italian households between two subsequent waves. Data are collected using the Computer-Assisted Personal Interviewing (CAPI) method, in which interviewers record responses on a tablet during the interview. Starting from 2020, significant methodological changes were introduced in the SHIW.[4] Specifically, the sample of households to be interviewed was selected based not only on traditional demographic variables but also on households' fiscal (administrative) income classes (Loschiavo et al., 2024). Moreover, nearly all household members interviewed are linked to fiscal records. In other words, for each individual interviewed, we have a comprehensive set of variables from the fiscal records, including total gross income, gross employee and self-employed earnings, etc. The latest edition of the survey, which we employ in our analysis, was fielded throughout 2023

[4] For an up-to-date list of survey designs employed at the European level, see the technical report of the Household Finance and Consumption Network (2023, Chapter 4.3)

and it refers to household conditions in 2022. In addition to the detailed item-by-item set of questions on each income source of each household component,[5] its questionnaire included a single question on total net household income for a random sub-sample of respondents. The question was asked with the *unfolding bracket technique* (Juster and Smith, 1997)[6] to avoid item non-response, and it was asked before the set of detailed questions to avoid helping respondents to recall each household-level income source.

The Conjunctural Survey of Italian Households (CSIH) is a new web-based survey that has been designed for two main purposes. The first is to track the evolution of household conditions throughout the economic cycle in the periods when the SHIW is not conducted, while the second is to have a flexible and prompt tool to meet current informational needs and analytical requirements of the Bank of Italy. The main data collection method is the CAWI (Computer-Assisted Web Interviewing), which is used with approximately 90 per cent of interviewed households, on average, in the two editions of the current analysis.[7] Households are selected from participants in the most recent SHIW survey; they are contacted via email and complete the on-line questionnaire independently, without assistance from an interviewer.

The first edition of the CSIH took place between August and September 2023, while the second edition took place in March and April 2024. The gross sample (i.e., the list of households contacted via e-mail for the interview) is represented by households that had participated in the 2022 edition of the SHIW by the end of July 2023, in the case of the first CSIH wave, and the entire 2022 SHIW sample in the case of wave two. 1,317 households were interviewed in both CSIH waves.

The CSIH questionnaires[8] included a large number of qualitative questions to capture the most significant phenomena during this particular economic phase, and they also included a single question on total household net income. To minimise the burden on households participating in the survey, a set of item-by-item questions on all household income sources was considered unfeasible. The question formulation is identical to that of the 2022 edition of the SHIW. Single-question income refers to 2022 in the first CSIH edition, and to 2023 in the second edition of the survey.

[5] It is important to note that income questions are mandatory in the SHIW. Respondents are required to report income amounts for any declared income sources, though some income sources may still be omitted (Neri and Zizza, 2010).
[6] The unfolding bracket technique requires indicating membership in value classes and subsequently specifying either the exact value or its approximate placement within the indicated class. The exact question formulation is provided in Appendix A1.
[7] The remaining households are interviewed by telephone. Telephone interviews are conducted to avoid excluding a non-negligible segment of the population (elderly individuals who are not connected to the internet).
[8] The data and its documentation will be available on the Bank of Italy’s website starting in 2025.

Table 1 - Data used in the analysis

| Dataset | Fieldwork | Household income reference period | Single question on household income | Source-by-source questions on household income | Gross sample |
|---|---|---|---|---|---|
| SHIW 2022 | Entire 2023 | 2022 | ✓† | ✓ | Households randomly extracted from the universe of Italian households |
| CSIH 2022 | Aug-Sep 2023 | 2022 | ✓ | | Households interviewed in SHIW 2022 up to July 2023 |
| CSIH 2023 | Mar-Apr 2024 | 2023 | ✓ | | All households interviewed in SHIW 2022 |

Notes: SHIW: Survey on Household Income and Wealth. CSIH: Conjunctural Survey on Italian Households. Household total income net of taxes and social contributions.
† For a random sub-sample of respondents only.

Table 1 summarises the data used in this study. Since the single questions on income refer to 2022 and 2023 in the first and second editions of the CSIH, respectively, I will refer to these datasets as CSIH 2022 and CSIH 2023 for simplicity.[9] Note that, because each household in the CSIH has also been interviewed in the SHIW, I have both a single-question and a source-by-source measure of income for each household in the CSIH 2022.

## 3. Methodology

### 3.1 Validating household income collected with a single question

Let $y_i$ be the unobserved variable of interest for household $i$, representing the logarithm of (a benchmark value of) household total income, with $E(y_i) = \mu_y$ and $Var(y_i) = \sigma_y^2$, $i = 1, \dots, n$.[10]

[9] This definition is somewhat imprecise, given the conjunctural nature of the survey and that most questions pertain to the time of the interview (e.g., respondent's perceived economic situation at the time of the interview, inflation expectations, etc.); however, it is the clearest in the present context.

[10] Alternatively, $y_i$ may represent individual or household earnings, other income sources (e.g., pension and capital income), or, more generally, any benchmark variable that can be measured using both single and detailed source-by-source questions.

Following Bingley and Martinello (2017) and Crossley et al. (2023), I assume that $y_i$ is linearly related with the observed survey log-income retrieved by the single question ($y_s$):

$$y_{is} = \mu_y + \kappa_s + (1 + \rho_s)(y_i - \mu_y) + \epsilon_{is}, \quad (1)$$

and the error term $\epsilon_{is}$ is independent of $y_i$, with $E(\epsilon_{is}) = 0$ and $E(\epsilon_{is}^2) = \sigma_s^2$. The parameter $\kappa_s$ reflects systematic under-reporting or over-reporting at the mean, while $\rho_s$ accounts for the correlation between the measurement error in $y_{is}$ and the actual benchmark value, $y_i$. As it happens, Equation (1) may also be expressed as:

$$y_{is} = y_i + \kappa_s + \rho_s(y_i - \mu_y) + \epsilon_{is},$$

which clearly indicates that $y_{is}$ is defined additively by $y_i$ and a compound error, consisting of the sum of a classical measurement error and a component capturing the systematic relation to $y_i$. Note also that if $\rho_s = \kappa_s = 0$, the measurement error in $y_{is}$ is classical. A negative $\rho_s$ indicates a pattern of mean reversion, where lower income earners report higher-than-actual values and higher earners report lower-than-actual values, thereby converging toward the mean. An opposite pattern is captured by a positive $\rho_s$.

I also assume that the measurement error in the benchmark/validation archive (the SHIW, in my case) is classical (i.e., independent of $y_i$):

$$y_{im} = y_i + \epsilon_{im}, \quad (2)$$

with $E(\epsilon_{im}) = 0$ and $E(\epsilon_{im}^2) = \sigma_m^2$, and *m* stands for *main*. In words, my aim is to validate household income as measured by single questions by comparing it to household income derived from source-by-source questions, which serves as our benchmark (i.e., what would be obtained if the main survey -the SHIW- was conducted), and in which $y_i$ is measured with error.

Combining Equations (1) and (2) we obtain:

$$y_{is} = \mu_y + \kappa_s + (1 + \rho_s)(y_{im} - \mu_y) + \epsilon_{is} - (1 + \rho_s)\epsilon_{im}. \quad (3)$$

Now let $\mu_{y_s} = \mu_y + \kappa_s$. Provided that there exists at least one instrument $z_i$ that is correlated to $y_{im}$ and orthogonal to the error terms $\epsilon_{im}$ and $\epsilon_{is}$, we can estimate the set of parameters $(\mu_y, \kappa_s, \rho_s, \sigma_y, \sigma_s, \sigma_m)$ by Generalized Methods of Moments (GMM), using the following system of moment restrictions:

$$
\begin{aligned}
E(y_{im} - \mu_y) &= 0, \\
E(y_{is} - \mu_{y_s}) = E(y_{is} - \mu_y - \kappa_s) &= 0, \\
E\left[(y_{im} - \mu_y)^2 - \sigma_y^2 - \sigma_m^2\right] &= 0, \\
E\left[(y_{is} - \mu_{y_s})^2 - (1+\rho_s)^2\sigma_y^2 - \sigma_s^2\right] &= 0, \\
E\left[(y_{is} - \mu_{y_s})(y_{im} - \mu_y) - (1+\rho_s)^2\sigma_y^2\right] &= 0, \\
E\left[z_i\left(y_{is} - \mu_y - \kappa_s - (1+\rho_s)(y_{im} - \mu_y)\right)\right] &= 0.
\end{aligned} \tag{4}
$$

### 3.2 Income change at the micro-level

For a given unit *i*, a straightforward extension of the model described by Equations (1)-(3) pertains the application to longitudinal settings. Indexing Equation (1) to a generic time period *t*:

$$
y_{is}^{(t)} = \mu_y^{(t)} + \kappa_s^{(t)} + \left(1 + \rho_s^{(t)}\right)\left(y_i^{(t)} - \mu_y^{(t)}\right) + \epsilon_{is}^{(t)}, \tag{5}
$$

and taking the difference between two time periods (such as two consecutive waves of a longitudinal survey) results in:

$$
\begin{aligned}
\Delta y_{is}^{(t)} = \Delta\mu_y^{(t)} + \Delta\kappa_s^{(t)} + \left(1 + \rho_s^{(t)}\right)\left(y_i^{(t)} - \mu_y^{(t)}\right) + \\
- \left(1 + \rho_s^{(t-1)}\right)\left(y_i^{(t-1)} - \mu_y^{(t-1)}\right) + \Delta\epsilon_{is}^{(t)},
\end{aligned} \tag{6}
$$

where $\Delta x^{(t)} = x^{(t)} - x^{(t-1)}$ is the difference-operator for a generic variable/parameter *x*. In this framework, I adopt the following set of assumptions:

*Assumption 1*: $\kappa_s^{(t)} = \kappa_s$, namely, under-/over-reporting at the mean is constant over time.

*Assumption 2*: $\rho_s^{(t)} = \rho_s$, namely, the systematic relation to the true value $y_i$ is constant over time.

The stability of $\kappa_s^{(t)}$ and $\rho_s^{(t)}$over time may be tested by the data producer, with repeated cross-sectional estimates of the system in Equation (4). For instance, in the case of a large (main) survey conducted at a low frequency (e.g. every two/three years), which implements the gold-standard set of source-by-source questions on household income, and a short survey conducted at a higher frequency to fill the informative gap between two subsequent editions of the main survey, the stability of these parameters may be tested over two consecutive years where the income reference period for the main and short surveys are the same (as is the case only for 2022 in my empirical application; see Table 1). Prior evidence on the stability of the parameters would enhance the reliability of Assumptions 1 and 2. Note also that different interview modes (e.g., face-to-face, web-based, etc.) are likely to be variably susceptible to misreporting behaviour (Angel et al., 2019). Therefore, to estimate income changes between two periods, it is recommended to employ the same interview mode to ensure credibility of Assumptions 1 and 2.

Under Assumptions 1 and 2, we have that:

$$\Delta y_{is}^{(t)} = \Delta\mu_y^{(t)} + (1+\rho_s)\left(\Delta y_i^{(t)} - \Delta\mu_y^{(t)}\right) + \Delta\epsilon_{is}^{(t)}. \tag{7}$$

A special case of Assumption 2 is that of no linear relation between $y_i^{(t)}$ and $y_{is}^{(t)}$, when $\rho_s$ is equal to zero:

*Assumption 2 (strong)*: $\rho_s^{(t)} = \rho_s = 0$, namely, no systematic relation to the true value over time.

This invariance can be tested over time as well. In this case, Equation (5) simplifies to:

$$y_{is}^{(t)} = y_i^{(t)} + \kappa_s + \epsilon_{is}^{(t)}, \tag{8}$$

and we have that:

$$\Delta y_{is}^{(t)} = \Delta y_i^{(t)} + \Delta\epsilon_{is}^{(t)}. \tag{9}$$

In both settings described by Equations (7) and (9), if we take the expectations we get:

$$E\left(\Delta_{y_{is}}^{(\mathrm{t})}\right) = E\left(\Delta_{y_i}^{(\mathrm{t})}\right) = \Delta\mu_y^{(t)}, \tag{10}$$

implying that the average change in the true values $y_i$ between two consecutive waves can be estimated using the sample counterpart of the average change in the values $y_{is}$.

Also in longitudinal settings, when using $\Delta_{y_{is}}^{(\mathrm{t})}$ in place of $\Delta_{y_i}^{(\mathrm{t})}$ as a dependent or explanatory variable in linear models, the same considerations regarding the measurement-error-induced bias of the OLS and IV estimators discussed by Bingley and Martinello (2017) and Crossley et al. (2023) apply. We refer the reader to these references for further details. In the present context, I underline that estimating by OLS a linear model with exogenous covariates of the type $\Delta_{y_i}^{(\mathrm{t})} = \boldsymbol{\beta}'\boldsymbol{x}_i + u_i$, employing $\Delta_{y_{is}}^{(\mathrm{t})}$ as dependent variable in place of $\Delta_{y_i}^{(\mathrm{t})}$, results in: $\widehat{\boldsymbol{\beta}}_{ols} \xrightarrow{p} \boldsymbol{\beta}(1+\rho_s)$.

### 3.3 Income change at the macro-level: Change at the mean of the distribution

In the previous section I described the extension to longitudinal settings, at the micro-level, of the model presented in Section 3.1. Some implications at the macro-level can be derived as well, for comparing income averages over time with distinct cross-sectional archives.

In particular, under the set of assumptions discussed in the following, the relative change in the single-question average income between two time periods is equal to the relative change at the mean of the benchmark income distribution. This result holds under Assumptions 1 and 2, with the necessary condition that $\rho_s$ is equal to zero (i.e., the strong version of Assumption 2 is needed). Moreover, the following further assumption is needed:

*Assumption 3*: $\epsilon_{is}^{(t)} \sim N(0, \sigma_s)$,

namely, normality of the error term distribution with a time-constant standard deviation.

To derive the result, the model described by Equation (8) may be written in levels as:

$$Y_{is} = \exp(y_{is}) = \exp(y_i + \kappa_s + \epsilon_{is}) = kY_i \exp(\epsilon_{is}), \tag{11}$$

with $k = \exp(\kappa_s)$. Taking the expectation, we get:[11]

$$E(Y_{is}) = kE(Y_i)\, E(\exp(\epsilon_{is})), \tag{12}$$

and indexing to time *t* and taking the first difference, we get:

$$E\left(Y_{is}^{(t)}\right) - E\left(Y_{is}^{(t-1)}\right) = k\left[E\left(Y_i^{(t)}\right)E\left(\exp\left(\epsilon_{is}^{(t)}\right)\right) - E\left(Y_i^{(t-1)}\right)E\left(\exp\left(\epsilon_{is}^{(t-1)}\right)\right)\right], \tag{13}$$

with the samples at time *t* and *t-1* possibly consisting of different households. Then, by the properties of the log-normal distribution, we get:[12]

$$E\left(Y_i^{(t)}\right) - E\left(Y_i^{(t-1)}\right) = \frac{E\left(Y_{is}^{(t)}\right) - E\left(Y_{is}^{(t-1)}\right)}{k \exp\left(\frac{\sigma_{is}^2}{2}\right)}, \tag{14}$$

and

[11] Because of the non-linearity of the log-transformation, if the strong version of Assumption 2 were to fail (i.e., with $\rho_s$ different from zero), no implications for the change at the mean of the benchmark income distribution between two time periods can be drawn. As a matter of fact, with $\rho_s$ different from zero, Equation (11) must be re-written starting from Equation (5): $Y_{is} = \exp(\kappa_s - \rho_s\mu_y + (1+\rho_s)y_i + \epsilon_{is}) = \exp(\kappa_s - \rho_s\mu_y)\exp(y_i^{1+\rho_s})\exp(\epsilon_{is})$, entailing the unfeasibility of the passages starting from Equation (12) and leading to the result in (15).

[12] From Equation (14), note that the change in the average single-question income in levels is lower than the change in benchmark values when the factor k*exp$\left(\frac{\sigma_{is}^2}{2}\right) = \exp\left(\frac{\sigma_{is}^2}{2} + k_s\right)$ is less than 1, namely, when the argument of the exponential function is less than zero. Even in the case of $k_s < 0$ (i.e., when there is under-reporting at the mean), the argument is less than zero only if $|k_s| > \frac{\sigma_{is}^2}{2}$. In other words, even if there is under-reporting at the mean, the change in the average single-question income may be overstating the benchmark values' change if the classical measurement error variance $\sigma_{is}^2$ is large enough.

$$\frac{E\left(Y_i^{(t)}\right) - E\left(Y_i^{(t-1)}\right)}{E\left(Y_i^{(t-1)}\right)} = \frac{E\left(Y_{is}^{(t)}\right) - E\left(Y_{is}^{(t-1)}\right)}{E\left(Y_{is}^{(t-1)}\right)}, \tag{15}$$

implying that the relative change at the mean of the income distribution between two time periods can be consistently estimated using the relative change in the single-question average income estimates.

### 3.4 Heterogeneity of misreporting and implications in longitudinal settings

*3.4.1 Heterogeneous misreporting behaviour*

Now, suppose there is evidence of heterogeneous misreporting behaviour among respondents. This may be caused, for instance, by varying levels of household income complexity (i.e., the more income sources, the more likely it is that single-question income will be under-reported), or by certain respondent characteristics (e.g., their educational level). For simplicity, I assume there are two groups, with a grouping variable $G$ assuming two values, and that the model in (1) is defined with group-specific average true income value and under-/over-reporting at the mean:[13]

$$y_{isg} = \mu_{yg} + \kappa_{sg} + (1 + \rho_s)\left(y_{ig} - \mu_{yg}\right) + \epsilon_{is}, \quad g = 1{,}2. \tag{16}$$

The system of moment restrictions described in (4) is modified accordingly.[14] In this setting, Assumptions 1 and 2 must hold conditional on each group $g$ to draw any inference on the income change between two time periods.

*3.4.2 Implications at the micro-level*

In the case of heterogeneous misreporting behaviour among respondents, from a micro perspective, the result reported in Equation (10) holds conditional on group $g$. I emphasize that for modelling the conditional mean of $y_i^{(t)}$ in cross-sectional linear settings and to estimate it by OLS using $y_{is}^{(t)}$ as dependent variable, it is fundamental to include among the set of covariates all the characteristics contributing to the heterogeneity in misreporting behaviour. In longitudinal settings, a noteworthy situation occurs when households can switch group between two time periods, namely, when the

[13] A less restrictive, though less parsimonious parametrisation would define each parameter conditional on group $g$. To improve model identifiability, we prefer to retain the more parsimonious approach presented in the main text, particularly given the limited number of observations (a common characteristic of many web surveys) in the empirical application discussed in Section 4. A similar consideration applies to the choice of the number of groups $g$. We also note that both frameworks may be easily extended to g>2.

[14] The modified system of equations is provided in Appendix A2.

grouping variable is allowed to vary over time. Then, in such regression settings one must control for changes in these characteristics between the two time periods.

### *3.4.3 Implications at the macro-level*

It is straightforward to prove that, if also Assumption 3 holds conditional on $g$, the results in Equations (14) and (15) hold conditional on group $g$ as well. As discussed in Section 3.3, a crucial assumption is that $\rho_s = 0$. Then, if the relative share of each group remains constant within the two time periods (it may be the case that this can be directly verified in the data, and, if not, it must be assumed constant over time), the relative change in the mean of the distributions can also be retrieved. In the interest of space, a formal derivation is provided in Appendix A3.

# 4. Results

## 4.1 Descriptive comparisons

Because each household in the CSIH has already been interviewed in the SHIW, I have two single questions (SHIW 2022 and CSIH 2022) that may be validated against the best-practice source-by-source approach implemented in the 2022 edition of the SHIW (see Table 1).[15] To reduce the leverage of extreme observations, all data were winsorised at the $5^{th}$ and $95^{th}$ percentiles.

Even though validating the single question on income from the SHIW may seem sufficient, the single question referring to the same period was also asked in the subsequent CSIH 2022 wave. This choice was motivated by the interview mode differing between the two surveys. It is desirable to measure income change conditional on the same mode, to ensure credibility of Assumptions 1-3, as the interview mode is known to represent an important factor influencing respondents (Angel et al., 2019), and affecting measured income inequality (Fessler et al., 2018).

[15] The selection of an income class for single-question household income is mandatory in both the SHIW and CSIH. However, households are not required to indicate an exact total income amount, but rather its approximate placement within the selected class. Nonetheless, around 68 per cent of households in the SHIW and 63 per cent of households in the CSIH report the exact amount. Depending on respondents' indications, missing amounts are set equal to: (i) $lb + (ub - lb)/8$; (ii) $(ub + lb)/2$; (iii) $lb + (ub - lb) \times 7/8$; where $ub$ and $lb$ stand for upper and lower bound, respectively.

Table 2 - Descriptive statistics for the variables of interest in 2022

| | SHIW 2022 | | | | CSIH 2022 | | |
|---|---|---|---|---|---|---|---|
| | Source-by-source income | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | Entire sample | Sub-sample† | Single question† | Ratio (3)/(2) | Source-by-source income | Single question | Ratio (6)/(5) |
| Percentile | | | | | | | |
| 5 | 7,674 | 8,177 | 7,000 | 0.86 | 7,908 | 5,000 | 0.63 |
| 10 | 9,750 | 10,803 | 10,000 | 0.93 | 9,854 | 8,750 | 0.89 |
| 25 | 15,522 | 15,616 | 14,880 | 0.95 | 17,005 | 15,000 | 0.88 |
| 50 | 24,185 | 24,249 | 22,400 | 0.92 | 25,628 | 21,000 | 0.82 |
| 75 | 37,945 | 37,024 | 33,750 | 0.91 | 38,416 | 35,000 | 0.91 |
| 90 | 55,156 | 52,002 | 48,125 | 0.93 | 55,101 | 49,000 | 0.89 |
| 95 | 68,509 | 67,612 | 65,000 | 0.96 | 67,612 | 70,000 | 1.04 |
| Mean | 28,533 | 28,286 | 25,942 | 0.92 | 29,371 | 26,103 | 0.89 |
| Gini index | 0.325 | 0.314 | 0.318 | 1.01 | 0.312 | 0.339 | 1.09 |
| Number of households | 9,641 | 1,509 | 1,509 | | 1,924 | 1,924 | |

Notes: Weighted estimates. SHIW: Survey on Household Income and Wealth. CSIH: Conjunctural Survey on Italian Households. Household total income net of taxes and social contributions. All distributions are winsorised at the 5th and 95th percentile. Ratio of households providing point values to single questions to total number of households: 0.68 in the SHIW 2022 and 0.63 in the CSIH 2022 (1,023 and 1,217 households, respectively).
† Randomised sub-sample of respondents who were asked the single question on household total income.

Supporting this statement, the degree of under-reporting of the single question compared to the source-by-source approach is larger in the CSIH than in the SHIW across the entire distribution, with the only exception of the 95th percentile (Table 2), in which the CSIH single-question income is overstating the corresponding source-by-source amount.[16] Moreover, while the Gini index of the single-question income distribution of the SHIW is practically identical to that of the source-by-source income distribution, in the case of the CSIH inequality is slightly overstated.

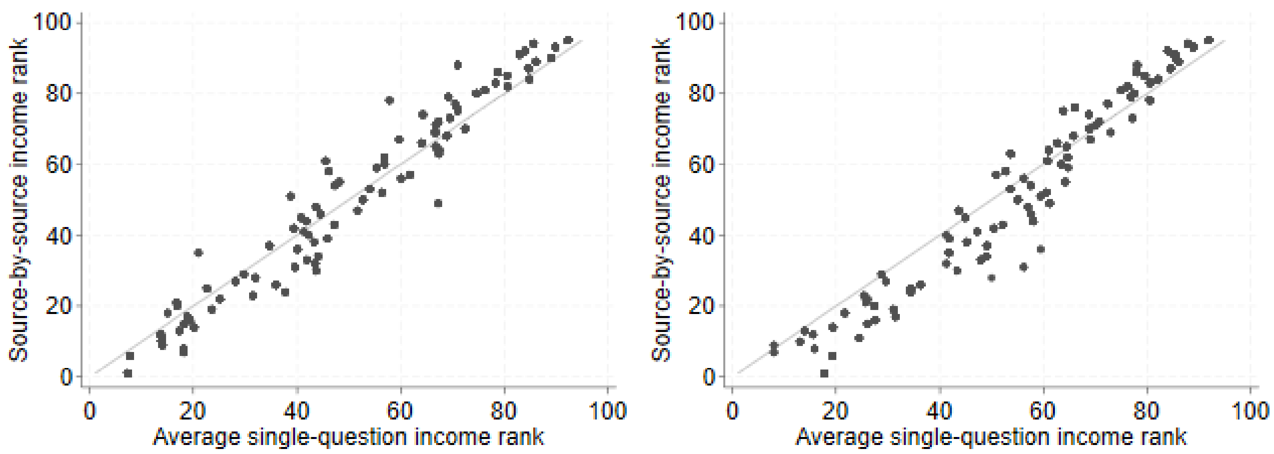
Although the values are underestimated compared to those obtained through the source-by-source method, the reconstruction of income using a single question preserves the relative ranking of households along the distribution to a satisfactory extent. In Figure 1, the average conditional percentile rank retrieved from the single question is plotted against the percentile rank in the validation source. Most points in the rank-rank plots cluster around the 45-degree lines, suggesting an acceptable level of agreement between the income distributions.

[16] This may be the result of a spurious measurement error in my data. The ratio of the 99th percentiles in the same dataset (not shown in Table 2) is again less than one, specifically 0.95.

Figure 1 - Average single-question income percentile rank conditional on source-by-source income percentile rank

(a) SHIW 2022
(Spearman's rho = 0.90)

(b) CSIH 2022
(Spearman's rho = 0.83)




Notes: Weighted income ranks.

However, there are notable deviations, especially within the CSIH 2022. In the latter case, the Spearman's rank-rank correlation (0.83) is equal to that found in Crossley et al. (2023) for household total income in the UK.

### 4.2 GMM estimates of the measurement error model

I estimate the measurement error model described in (4), using both SHIW and CSIH data, and employing three different sets of instruments to identify $\rho_S$. I recall that for the instruments to be valid, they must be correlated to the SHIW income and uncorrelated to measurement errors in both surveys. Following Crossley et al. (2023), my first set of instruments makes use of well-measured SHIW variables. It includes household food consumption expenditure, and expenses related to the household main residence (including condominium fees, utilities, etc.) reported in SHIW, and it represents my preferred specification.[17] In fact, it is known that food expenditure and main residence expenses are well estimated compared to those sources providing reliable external benchmarks (Cifaldi and Neri, 2013; Donatiello et al., 2025). In the second specification, the only instrument is represented by the sum of gross fiscal employee and pension income, which are known to be highly correlated with the corresponding survey counterparts (Barcaroli et al., 2021).[18] In this case, the

[17] For comparison, Crossley et al. (2023) employ the following survey variables as instruments: the number of cars the household owns or has access to, and the number of rooms in the home. As and additional instrument, they use council tax liability (from administrative records).

[18] Fiscal records were missing for approximately 4 per cent of the interviewed individuals for whom personal IDs were not available. Missing fiscal records were imputed from similar individuals, based on their survey-reported income,

model described in (4) is just-identified. The last set of instruments also includes gross fiscal total household income, in addition to the sum of gross fiscal employee and pension income.

In practice, I have two single-question measures (SHIW 2022 and CSIH 2022) that may be validated against the benchmark SHIW 2022 income, and I estimate the model described in (4) for each of the two data pair, employing each of the three sets of instruments, resulting in a total of six specifications. Estimated parameters of the different specifications of the model are presented in Table 3 (columns 1 to 6). Due to zero tax incomes, the number of observations is smaller for the models employing fiscal records as instruments (models (B) and (C)). All estimates of interest are stable across specifications when using the two data sources. Importantly, in each specification, the estimate of $\rho_s$ is low and statistically not different from zero, which indicates that measurement error in the single-question income measure is not related to the benchmark income value, and it simplifies the interpretation of the income changes discussed in the following sections.

Estimates of $\kappa_s$ capture a significant under-reporting ranging from approximately 9 per cent to almost 16 per cent, depending on the model specification. Conditional on a given specification, the CSIH data exhibits a larger degree of under-reporting, compared to the SHIW data, consistent with previous descriptive evidence. Moreover, the ratio $\sigma_s^2/\sigma_y^2$ increases from 0.102 in the SHIW data to 0.415 using the CSIH data (model (A)), indicating a substantially larger measurement error in the CSIH. These results point to the importance of estimating income changes in two subsequent periods conditional on the same interview mode, to ensure the reliability of Assumptions 1 and 3.

In the last three columns of Table 3, I report the estimates of the model using the CSIH data and restricting the sample to those households whose respondents were the same than in the SHIW, to ensure that the validation source and the single-question income are provided by the same interviewed individuals. Ideally, the interest of the researcher is being able to recover a meaningful signal on the household income regardless of which household component is responding to the questionnaire. When restricting the sample to households with the same respondents, the sample size of the main specification (which employs the entire sample) reduces by 225 units to 1,699 households. In other words, respondents have changed between the two surveys for only approximately 10 per cent of the sample. This value is not sizeable, and, as a result, the estimated parameters reported in the last three columns of Table 3 are only marginally different from those reported in columns 4 to 6.

geographical location, gender, and occupation type. The employed imputation method was nearest-neighbour matching (D'Orazio et al., 2006).

Table 3 - GMM estimates of the measurement error model

| | SHIW 2022† | | | CSIH 2022‡ | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Entire sample | | | Same SHIW 2022 respondent | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Model specification | (A) | (B) | (C) | (A) | (B) | (C) | (A) | (B) | (C) |
| $\mu_y$ | 10.086*** | 10.117*** | 10.114*** | 10.120*** | 10.162*** | 10.161*** | 10.096*** | 10.135*** | 10.134*** |
| | (0.024) | (0.025) | (0.026) | (0.027) | (0.027) | (0.028) | (0.029) | (0.030) | (0.030) |
| $k_s$ | -0.090*** | -0.100*** | -0.100*** | -0.150*** | -0.157*** | -0.157*** | -0.147*** | -0.153*** | -0.153*** |
| | (0.012) | (0.013) | (0.013) | (0.020) | (0.020) | (0.020) | (0.022) | (0.022) | (0.022) |
| $\sigma_y^2$ | 0.293*** | 0.296*** | 0.274*** | 0.313*** | 0.277*** | 0.280*** | 0.316*** | 0.290*** | 0.292*** |
| | (0.018) | (0.021) | (0.023) | (0.022) | (0.022) | (0.021) | (0.023) | (0.023) | (0.022) |
| $\sigma_m^2$ | 0.045*** | 0.026* | 0.048*** | 0.045*** | 0.050*** | 0.048*** | 0.051*** | 0.050*** | 0.048*** |
| | (0.008) | (0.014) | (0.016) | (0.011) | (0.009) | (0.008) | (0.012) | (0.010) | (0.009) |
| $\rho_s$ | 0.029 | -0.036 | 0.041 | -0.024 | -0.017 | -0.026 | -0.002 | -0.02 | -0.025 |
| | (0.028) | (0.041) | (0.058) | (0.038) | (0.045) | (0.040) | (0.041) | (0.047) | (0.042) |
| $\sigma_s^2$ | 0.030*** | 0.050*** | 0.028* | 0.130*** | 0.126*** | 0.127*** | 0.126*** | 0.129*** | 0.130*** |
| | (0.009) | (0.011) | (0.016) | (0.017) | (0.017) | (0.016) | (0.019) | (0.019) | (0.018) |
| Number of households | 1,509 | 1,357 | 1,357 | 1,924 | 1,766 | 1,766 | 1,699 | 1,559 | 1,559 |

Notes: Significance levels: * 10% ** 5% *** 1%.

† Using single-question SHIW 2022 income. ‡ Using single-question CSIH 2022 income. In both cases benchmark values are given by the source-by-source SHIW 2022 income. Instruments employed for model specification (A): food consumptions expenditure and expenses related to the household main residence; (B): logarithm of the sum of gross fiscal employee and pension income; (C): logarithm of the sum of gross fiscal employee and pension income, and logarithm of total household income. The number of observations is smaller for models (B) and (C) due to zero tax incomes. Model (B) is just-identified.

### 4.3 Income change at the macro-level in Italy between 2022 and 2023

To estimate the change at the mean of the income distribution (i.e., at the macro-level), I employ the entire CSIH 2022 and 2023 waves. In both waves' questionnaire the single question on household total income was included to provide an estimate of the income change of Italian households in the reference period (see Table 1).

In Table 4, I provide the percentage change at the mean of the income distribution, and the percentage change conditional on income quartiles. I emphasize that, while the percentage change at the mean of the single-question income distribution provides an estimate of the percentage change at the mean of the *true* income variable, the percentage change of the mean conditional on quartiles cannot be interpreted within the framework of the measurement error model defined in Section 3.1. Nonetheless, they are likely to provide a reliable indication of the heterogeneity of the change across the distribution. Between 2022 and 2023 the mean income of Italian households increased by 6.1 percentage points (first row) in nominal terms. For comparison, in the same period of analysis, the disposable income of Italian households reported by the National Accounts increased by 4.9 percentage points. According to Equation (14), the change in the level of average income that would have been observed if the main survey (the SHIW) was run is equal to 1,743 euros, whereas the change in the single-question average income is equal to around 1,600 euros. Note that the number of households in each income quartile is strongly unbalanced, with most of the households belonging to the top of the distribution, and less than 300 households belonging to the bottom quartile in both waves. This result, driven by the higher weights of households in the bottom tail of the distribution and jointly determined by the oversampling of the most affluent households in the SHIW (Loschiavo et al., 2024)[19] and their higher propensity to participate in the CSIH,[20] suggests that the CSIH may not represent the optimal setting for analysing the lower end of the income distribution. With this caveat in mind, I note that the income change in Italy between 2022 and 2023 was not uniform across the income distribution. It was more pronounced for the upper half of the distribution, whereas average income for the first quartile reduced. Furthermore, the Gini index increased by approximately 2 percentage points, from 33.9 to 36.2, in the considered period.[21]

[19] Indeed, the SHIW 2022 is also characterised by a similar feature, albeit less pronounced. Out of a total of 9,641 households, 1,870 belong to the bottom quartile, whereas 3,650 belong to the upper quartile.

[20] For example, participation rates in the CSIH 2022 range from a minimum of 9.5 per cent for households in the bottom SHIW 2022 income quartile to a maximum of 30 per cent for those in the top quartile.

[21] Fluctuations in the Gini index over time may be overstated when using single questions, as is the point value in a given period (Table 2). Additional data points over time with both income measures are needed to analyse these aspects, which are unavailable. We leave the topic for future research.

These results are qualitatively consistent (in the direction of the changes, though not in their magnitude) with those obtained using a question that directly asks for the percentage change over the period considered, which was applied to the source-by-source 2022 income.[22] According to this measure of income change, mean income increased by 1.5 percentage points. While the average income of each quartile changed in the same direction as that identified by single-summary questions, the magnitude of these changes is significantly attenuated (with the exception of the second quartile), particularly at the distribution's tails. In previous experiments conducted at the Bank of Italy with this type of question, it was argued that the high number of respondents reporting a zero per cent change in total household income may be attributed to opportunistic respondent behaviour aimed at minimising the effort required to complete the questionnaire (a phenomenon known as *satisficing behavior*; Krosnick, 1991), thereby potentially explaining, at least in part, this difference.

Table 4 - Macro-level: Single-question average income by quartile in 2022 and 2023
(cross sectional estimates; entire CSIH 2022 and 2023 samples)

| | CSIH 2022 | | CSIH 2023 | | | |
|---|---|---|---|---|---|---|
| | Estimate | Number of households | Estimate | Number of households | Perc. change | Perc. change (comparison method)† |
| Average income | 26,103<br>(24,787 27,420) | 1,924 | 27,704<br>(26,382 29,025) | 2,513 | 6.1<br>(2.9 9.4) | 1.5<br>(0.7 2.3) |
| Average income by income quartile | | | | | | |
| First | 10,040<br>(9,378 10,702) | 250 | 8,848<br>(8,341 9,356) | 270 | -11.9<br>(-19.2 -4.6) | -0.7<br>(-3.1 1.7) |
| Second | 18,590<br>(18,252 18,927) | 194 | 18,704<br>(18,198 19,210) | 367 | 0.6<br>(-1.6 2.8) | 1.4<br>(-0.2 3.0) |
| Third | 27,698<br>(27,035 28,361) | 446 | 29,573<br>(28,949 30,198) | 584 | 6.8<br>(4.1 9.4) | 1.5<br>(0.5 2.5) |
| Fourth | 50,798<br>(48,967 52,629) | 1,034 | 54,301<br>(52,517 56,086) | 1,292 | 6.9<br>(3.9 9.9) | 2.3<br>(1.5 3.0) |

Notes: Weighted estimates. 95% confidence intervals in parentheses. † Computed on the CSIH 2023 sample using the question directly asking for the percentage change over the period considered, which was applied to the source-by-source 2022 income. CSIH: Conjunctural Survey on Italian Households. Household total income net of taxes and social contributions. Both distributions are winsorised at the 5th and 95th percentile. Standard errors for percentage changes have been computed using a first-order Taylor approximation, and the correlation between the estimators at *t* and *t-1* is estimated on the panel component. Ratio of households providing point values to single questions to total number of households: 0.63 in the CSIH 2022 and 0.60 in the CSIH 2023 (1,217 and 1,520 households, respectively).

[22] Consequently, the change at the mean of the distributions is retrieved as: $(\overline{(1+c_i) \times y_{im}^{2022}} - \overline{y_{im}^{2022}})/\overline{y_{im}^{2022}}$ , where $c_i$ is the percentage change provided by respondents.

### 4.4 Heterogeneity analysis

I now discuss the estimated model accounting for heterogeneous misreporting behaviour described in Section 3.4, to investigate differences in misreporting across subgroups. To test whether varying levels of household income complexity affects respondents' behaviour, I divide the sample in two subgroups depending on the number of income earners among the household components (single-/multiple-earner households). I further divide the sample based on the following respondent characteristics: gender (male/female), level of education (university degree/no degree), and income earner status in SHIW 2022 (main income earner/other household component).

Results are reported in Table 5. The correlation between the measurement error in the single-question income and the actual benchmark value, $y$, is very close to (and statistically not different from) zero in each specification (Panel B), with the exception of the one based on the number of income-earner components. In the latter case, as expected, a lower value of $\kappa_{sg}$ in absolute terms is found for the single-earner group of households, which is statistically not different from zero, whereas for multiple-earner households is estimated in -0.253. Men are on average more likely to underreport than women (although the difference is not significant), as opposed to what is found by Micklewright and Schnepf (2010) and Crossley et al. (2023) for British households. Close values of $\kappa_{sg}$ are also found for respondents with and without a university degree. Respondents not representing the main income earner (as observed in the 2022 edition of the SHIW) are more likely to provide underreported income values, albeit the limited number of observations (381) suggests caution in interpreting the results.

Overall, the degree of heterogeneity observed in our data indicates the need to control for different factors of misreporting behaviour when using the single-question income as a dependent or independent variable in linear models for the conditional mean, in cross-sectional settings, and for changes in these factors in longitudinal settings.

Table 5 - GMM estimates: Heterogeneity analysis

Panel A: Subgroup-specific parameters

| | Number of income-earner-components (1) | | Respondent's gender (2) | | Respondent's level of education (3) | | SHIW 2022 respondent's status (4) | |
|---|---|---|---|---|---|---|---|---|
| | single-earner | multiple-earner | male | female | university degree | no university degree | main income earner | other member |
| Number of households | 686 | 1,238 | 1,371 | 553 | 1,508 | 416 | 1,543 | 381 |
| $\mu_{yg}$ | 9.741*** | 10.459*** | 10.289*** | 9.866*** | 10.294*** | 9.898*** | 10.052*** | 10.355*** |
| | (0.032) | (0.027) | (0.027) | (0.045) | (0.032) | (0.041) | (0.030) | (0.048) |
| $k_{sg}$ | -0.037 | -0.253*** | -0.164*** | -0.129*** | -0.156*** | -0.142*** | -0.109*** | -0.298*** |
| | (0.030) | (0.025) | (0.026) | (0.031) | (0.026) | (0.032) | (0.021) | (0.047) |

Panel B: Common parameters among subgroups

| | Number of income-earner-components (1) | Respondent's gender (2) | Respondent's level of education (3) | SHIW 2022 respondent's status (4) |
|---|---|---|---|---|
| $\sigma_y^2$ | 0.189*** | 0.271*** | 0.276*** | 0.299*** |
| | (0.015) | (0.018) | (0.020) | (0.020) |
| $\sigma_m^2$ | 0.041*** | 0.044*** | 0.044*** | 0.044*** |
| | (0.009) | (0.011) | (0.011) | (0.010) |
| $\rho_s$ | 0.142** | -0.021 | -0.030 | 0.004 |
| | (0.057) | (0.042) | (0.044) | (0.039) |
| $\sigma_s^2$ | 0.121*** | 0.131*** | 0.132*** | 0.126*** |
| | (0.017) | (0.017) | (0.017) | (0.016) |

Notes: CSIH 2022 sample. Significance levels: * 10% ** 5% *** 1%. Instruments: food consumption expenditure and expenses related to the household main residence.

### 4.5 Income change at the micro-level using the panel sample

I now turn to the micro-level analysis of the determinants of income changes of Italian households between 2022 and 2023, employing the CSIH 2022-2023 panel sample consisting of 1,317 households and using the CSIH 2023 weights. In more than 92% of these households, the respondent was the same in both surveys.

The average income change, measured by the average of the differences in the log-values of the single-question income, $\Delta y_{is}^{(t)}$, is equal to 0.031,[23] whereas the median of the distribution is equal to zero (i.e., for half of the weighted sample, the income change is less than or equal to zero). To analyse whether the income change was heterogeneous among the population, for the conditional mean of $\Delta y_i^{(t)}$ I employ a linear model including among the covariates $\boldsymbol{x}_i$ a set of informative variables from the survey. As noted in Section 3.2, the average partial effects (APEs) of the $\boldsymbol{x}_i$'s on $\Delta y_i^{(t)}$ are consistently estimated by the OLS estimator, employing $\Delta y_{is}^{(t)}$ as dependent variable in place of the unobserved $\Delta y_i^{(t)}$, as long as $\rho_s$ is equal to zero. Evidence presented in Section 4.2 points in this direction. In other words, we would obtain the same average partial effects of the $\boldsymbol{x}_i$'s if we were able to collect, at a higher cost, an income measure similar to that in the SHIW. Importantly, to address potential bias due to households switching misreporting-behaviour group from 2022 to 2023, I include in my specification a set of labour market outcomes to control for changes in the number of income earners, to the extent allowed by our data. Moreover, since nearly the entire sample consists of households with the same respondent in both waves (1,214 out of 1,317 households), I perform a complementary analysis restricting the sample to these households rather than controlling for changes in respondent heterogeneity-related characteristics as covariates, to avoid sparseness and multi-collinearity issues.

Regression results are reported in Table 6. First, I regress the dependent variable on a set of household-level labour market outcomes (model (1)), which is used as a proxy for changes in the number of income earners. This set of variables also serves to assess the quality of our dependent variable $\Delta y_{is}^{(t)}$, as discrepancies between the two would cast doubt on the reliability of our measure of income change. As one would expect, income changes were higher for households with at least one component who started working, experienced an increase in labour income,[24] gained an income-earning member or reported an increase in the returns on financial capital, rental earnings and other sources of income. Conversely, income changes were lower for households where a member quit working, experienced a decrease in labour income, lost an income-earning member or reported a decrease in the returns on financial capital, rental earnings and other sources of income. In model (2), we regress $\Delta y_{is}^{(t)}$ on a set of determinants of the income change. The income change between 2022 and 2023 was higher for households whose main income source was payroll employment, compared

[23] From a macro-level perspective, the average of the income distribution of the panel sample increased by 4.8 percentage points. This value is also computed weighting panel households with 2023 cross-sectional weights, and it is fairly close to that reported in Section 4.3.
[24] Due to the question formulation, I cannot distinguish between households with a component who started/quit working and those with a component who experienced an increase/decrease in their labour income.

to all other households, although these differences are not statistically significant.[25] The change was lower for households in the lower half of the income distribution in 2023 (APE of -0.201), while it was higher (though not statistically significant) for those residing in Southern Italy. These results remain robust when labour market outcomes are included in the specification (model (3)), with the APE of the indicator variable for the South becoming significant at the 5% level. The results remain stable also when we include the set of misreporting-behaviour-related characteristics discussed in the previous section among the covariates, and when we restrict the sample to households with the same respondents in both waves to address the issue of group switching. The parameters for the indicators related to income-earning individuals joining or leaving the household are less robust, due to the small number of households reporting such occurrences (only 8 and 26 households, respectively).

Finally, note that the CSIH does not provide longitudinal weights for panel analyses. As a robustness check, I replicated the analysis using a tailored set of longitudinal weights that I computed, which can be reproduced using the disseminated data.[26] Using these weights, all regression results remained essentially unchanged (for further details, see Appendix A4).[27]

[25] The CSIH does not provide information on the main household-level income source. This information is thus retrieved from the 2022 edition of the SHIW.

[26] The process follows a similar approach to that used for the SHIW (Loschiavo et al., 2024). First, the CSIH 2022 cross-sectional weights were adjusted for unit non-response using a logit model estimating the conditional probability of panel retention in the CSIH 2023. Then, the non-response-adjusted weights were further adjusted using a raking technique to align with known population-level frequency distributions. Further details are provided in Appendix A4.

[27] Moreover, using longitudinal weights the average income in the panel sample increased by 6.4 percentage points, nearly identical to the increase shown in Table 4.

Table 6 – Micro-level: Regression analysis using the CSIH 2022 - CSIH 2023 panel sample

| | Entire panel sample | | | | Same respondent | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| Labour market outcomes | | | | | | | | |
| At least one component began working or experienced an increase in labour income | 0.135*** | | 0.080** | 0.090** | 0.134*** | | 0.077* | 0.087** |
| | (0.035) | | (0.038) | (0.036) | (0.038) | | (0.041) | (0.038) |
| At least one component ceased working or experienced a decrease in labour income | -0.153** | | -0.151** | -0.176** | -0.109 | | -0.121* | -0.139** |
| | (0.070) | | (0.069) | (0.069) | (0.070) | | (0.071) | (0.070) |
| An income-earning individual joined the household | 0.315*** | | 0.285** | 0.158*** | 0.211*** | | 0.171*** | 0.140*** |
| | (0.115) | | (0.124) | (0.054) | (0.062) | | (0.057) | (0.050) |
| An income-earning individual left the household | -0.367** | | -0.470*** | -0.478*** | -0.373** | | -0.482** | -0.528*** |
| | (0.150) | | (0.156) | (0.162) | (0.188) | | (0.198) | (0.201) |
| Other non-employment income† increased | 0.161** | | 0.095 | 0.104* | 0.139** | | 0.064 | 0.079 |
| | (0.062) | | (0.068) | (0.063) | (0.066) | | (0.073) | (0.067) |
| Other non-employment income† decreased | -0.065 | | -0.072 | -0.052 | -0.03 | | -0.038 | -0.013 |
| | (0.080) | | (0.083) | (0.088) | (0.065) | | (0.069) | (0.069) |
| Main household income source (SHIW 2022) | | | | | | | | |
| Self-employment | | 0.022 | 0.011 | 0.035 | | 0.024 | 0.011 | 0.04 |
| | | (0.159) | (0.157) | (0.155) | | (0.164) | (0.163) | (0.159) |
| Pensions | | 0.035 | 0.027 | 0.013 | | 0.046 | 0.036 | 0.031 |
| | | (0.144) | (0.144) | (0.143) | | (0.147) | (0.148) | (0.147) |
| Payroll employment | | 0.083 | 0.078 | 0.082 | | 0.102 | 0.094 | 0.106 |
| | | (0.141) | (0.138) | (0.138) | | (0.144) | (0.142) | (0.141) |
| Geographical area: Centre | | 0.036 | 0.051 | 0.069 | | 0.051 | 0.061 | 0.086 |
| | | (0.056) | (0.053) | (0.052) | | (0.058) | (0.056) | (0.055) |
| Geographical area: South | | 0.073 | 0.110** | 0.120** | | 0.067 | 0.098* | 0.115** |
| | | (0.053) | (0.052) | (0.053) | | (0.055) | (0.054) | (0.055) |
| Housheold income below the median in 2023 | | -0.201*** | -0.182*** | -0.202*** | | -0.192*** | -0.181*** | -0.210*** |
| | | (0.044) | (0.042) | (0.045) | | (0.045) | (0.043) | (0.047) |
| Constant | 0.023 | 0.053 | 0.045 | -0.028 | 0.027 | 0.046 | 0.043 | -0.033 |
| | (0.028) | (0.143) | (0.144) | (0.157) | (0.029) | (0.146) | (0.148) | (0.158) |
| Controls for misreporting heterogeneity | | | | ✓ | | | | ✓ |
| Observations | 1,317 | 1,317 | 1,317 | 1,313(‡ | 1,214 | 1,214 | 1,214 | 1,214 |

Notes: Weighted estimates. Significance levels: * 10% ** 5% *** 1%. † Including returns on financial capital, rental earnings, etc. Controls for misreporting heterogeneity: CSIH 2023 respondent's gender, educational level, and main income earner status in SHIW 2022. ‡ In four households, the CSIH 2023 respondent was not a household member in 2022; therefore, their main income-earning status in SHIW 2022 is unavailable. Baseline category for: (i) Main household income source: Other income; (ii) Geographical area: North.

## 5. Conclusions

Minimising the response burden on households is a crucial aspect of designing household survey questionnaires, as it helps increase response rates. To this end, single-summary questions are often used to collect earnings and income data. For instance, the growing demand for timely household economic data has prompted data producers to adopt web surveys, which commonly rely on single questions for household income and earnings to reduce respondent burden and eventually to allocate more resources to other survey priorities. However, single-summary questions provide less detailed guidance compared to source-by-source questions, which are considered the best-practice income data collection strategy. Recent studies comparing these methods suggest that while single questions can yield reasonable income distributions, they tend to produce noisier data with some underreporting (Micklewright and Schnepf, 2010; Crossley et al., 2023).

The contribution of this paper to the validation literature on assessing the quality of single-question income, is to provide insights into the dynamics of income and earnings collected through single questions, compared to detailed question sets, extending to a longitudinal framework a measurement error model first developed by Bingley and Martinello (2017). I show that, to draw micro-level inference on the income change between two time periods using single-question income data, it is essential to assume that respondents' misreporting behaviour remains constant over time. Additionally, the assumption of a normally distributed error term with a time-constant standard deviation is necessary to estimate changes in the distribution's mean. I also discuss potential heterogeneity in respondents' misreporting behaviour and its implications in longitudinal settings.

I then implement the proposed method to estimate income changes in Italy between 2022 and 2023, employing data from the new web-survey conducted by the Bank of Italy, the Conjunctural Survey of Italian Households. According to my findings, the average income of Italian households increased by 6.1 per cent between 2022 and 2023. Income changes were higher for more affluent households and for those residing in Southern Italy.

As web-based surveys will continue to play a crucial role in economic data collection, refining methodologies for analysing single-question measures in a dynamic perspective will remain an essential area of research. Future work could build on the proposed framework by exploring alternative approaches to account for misreporting behaviour, such as adopting a multidimensional perspective using latent class analysis (Collins and Lanza, 2010). Additionally, direct validation of single-question income changes against source-by-source income changes could be investigated in settings where both types of measures are available at two time points.

## References


Abowd, J. M. and Stinson, M. H. (2013). Estimating Measurement Error in Annual Job Earnings: A Comparison Of Survey And Administrative Data. *The Review of Economics and Statistics,* 95:1451–1467.

Angel, S., Disslbacher, F., Humer, S. and Schnetzer, M. (2019). What did you really earn last year?: Explaining measurement error in survey income data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182:1411-1437.

Neri, A. and Zanichelli, F. (2020). The main results of the special survey of Italian households in 2020 (only in Italian). Note Covid-19, Banca d'Italia.

Barcaroli, G., Ilardi, G., Neri, A. and Tuoto, T. (2021). Optimal sampling design for household finance surveys using administrative income data. In *Rivista di Statistica Ufficiale,* 2. Istituto Nazionale di Statistica. Rome, Italy.

Bingley, P. and Martinello, A. (2017). Measurement Error in Income and Schooling and the Bias of Linear Estimators. *Journal of Labor Economics*, 35:1117-1148.

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. In: *Handbook of Econometrics*, Vol. 5. Elsevier, pp. 3705–3843.

Cifaldi, G. and Neri, A. (2013). Asking income and consumption questions in the same survey: what are the risks. *Banca d'Italia Temi di Discussione (Workin papers)*, n. 908.

Collins, L. M. and Lanza, S. T. (2010). Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral and Health Sciences. Wiley, New York.

Crossley, T., Fisher, P. and Hussein, O. (2023). Assessing data from summary questions about earnings and income. *Labour Economics*, 81, 102331.

Daikeler, J., Bošnjak, M. and Manfreda, K. L. (2020). Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. *Journal of Survey Statistics and Methodology*, 8:513–539.

Donatiello, G., D'Orazio, M., Loschiavo, D., Neri, A. and Tullio, F. (2025). Filling the data gap in Italy on household multidimensional poverty by income, consumption and wealth. IAOS-ISI 2024 Conference Paper. Mexico City, Mexico.

D'Orazio, M., Di Zio, M. and Scanu, M. (2006). Statistical Matching: Theory and Practice. John Wiley & Sons, Chichester (UK).

Fessler, P., Kasy, M. and Lindner, P. (2018). Survey mode effects on measured income inequality. *Journal of Economic Inequality*, 16:487–505.

Fuller, W. A. (1987). Measurement Error Models. Wiley, New York.

Guiso, L., and Jappelli, T. (2024). The Italian Survey of Consumer Expectations: Statistical Bulletin. CESF Working Paper N.772.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029-1054.

Household Finance and Consumption Network (2023). The Household Finance and Consumption Survey: Methodological report for the 2021 wave. *ECB Statistics Paper Series*, N. 45, ECB.

Immordino, G., Jappelli, T., Oliviero, T. and Zazzaro, A. (2022). Fear of COVID-19 contagion and consumption: Evidence from a survey of Italian households. *Health Economics*, 31:496-507.

Institute for Social and Economic Research (2021). Understanding society COVID-19 user guide. Version 10.0. University of Essex, Colchester.

Jenkins, S. and Rios-Avila, F. (2023). Reconciling reports: modelling employment earnings and measurement errors using linked survey and administrative data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 186:110–136.

Juster, F. T. and Smith, J. P. (1997). Improving the quality of economic data: lessons from the HRS and AHEAD. *Journal of the American Statistical Association*, 92:1268-1278.

Kapteyn, A. and Ypma, J. Y. (2007). Measurement Error and Misclassification: A Comparison of Survey and Administrative Data. *Journal of Labor Economics*, 25:513-551.

Krosnick, J. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5:213- 236.

Loschiavo, D., Mariani, G. M., Moscatelli, M., Neri, A., Porreca, E. and Tullio, F. (2024). Enhancing the SHIW weighting methodology with external data sources. Available at https://www.bancaditalia.it/statistiche/tematiche/indagini-famiglie-imprese/bilanci-famiglie/metodologia-ibf/index.html?com.dotmarketing.htmlpage.language=1.

Micklewright, J. and and Schnepf, S. V. (2010). How reliable are income data collected with a single question? *Journal of the Royal Statistical Society Series A: Statistics in Society*, 173: 409-429.

Neri, A. and Zizza, R. (2010). Income reporting behaviour in sample surveys. *Temi di discussione (Economic working papers)*, n. 777, Bank of Italy.

United Nations (2011). *Canberra Group Handbook on Household Income Statistics*, second edition. United Nations, New York and Geneva.

## Appendix

### A1 Single question on household total income

(Mandatory question)

1. What was your household's total net income* for the year 202X?
   Please select one of the income brackets provided.

| |
|---|
| *** Consider the income received by all household members, net of taxes and social security contributions. The following income sources should be included in your calculation:**<br>• Payroll employment income<br>• Self-employment income<br>• Work-related pensions, disability pensions, old-age pensions, survivor pensions, social pensions, or private pensions (annuities) from insurance<br>• Temporary lay-offs, mobility or unemployment benefits, or severance payments<br>• Other forms of financial assistance (e.g., basic income schemes)<br>• Rental income<br>• Income from financial assets (dividends, interest, coupons, and other sources of financial income) |

- Up to 10.000 euros .................................... 1
- 10.001 - 15.000 euros................................ 2
- 15.001 - 20.000 euros................................ 3
- 20.001 - 25.000 euros................................ 4
- 25.001 - 35.000 euros................................ 5
- 35.001 - 50.000 euros................................ 6
- 50.001 - 100.000 euros.............................. 7
- 100.001 - 500.000 euros............................ 8
- Over 500.000 euros ................................... 9

2. Could you provide a specific or approximate value of your household’s total net income for 202X?

   |_|_|_|_|_|_|_| (limits to feasible answers from boundaries in Question 1)

   |_| No answer

3. (Mandatory question if respondent does not provide an answer to Question 2)
   Could you at least indicate whether your household’s total net income was closer to ... (lower limit), closer to ... (upper limit), or approximately halfway between these two values?

## A2 Modified system of moment restrictions for heterogeneous-misreporting behaviour

With two groups of households, define an indicator variable equal to one if household $i$ belongs to the first group: $d_i = \mathbb{1}\{g_i = 1\}$. The modified system of moment restrictions of the model accounting for heterogeneous respondents' misreporting behaviour is given by:

$$
\begin{aligned}
E\left(y_{img} - \mu_{yg}\right) &= 0, \quad g = 1,2, \\
E\left(y_{isg} - \mu_{y_{sg}}\right) = E\left(y_{isg} - \mu_{yg} - \kappa_{sg}\right) &= 0, \quad g = 1,2, \\
E\left[d_i\left(y_{im1} - \mu_{y1}\right)^2 + (1 - d_i)\left(y_{im2} - \mu_{y2}\right)^2 - \sigma_y^2 - \sigma_m^2\right] &= 0, \\
E\left[d_i\left(y_{is1} - \mu_{y_{s1}}\right)^2 + (1 - d_i)\left(y_{is2} - \mu_{y_{s2}}\right)^2 - (1 + \rho_s)^2\sigma_y^2 - \sigma_s^2\right] &= 0, \\
E\left[d_i\left(y_{is1} - \mu_{y_{s1}}\right)\left(y_{im1} - \mu_{y1}\right) + (1 - d_i)\left(y_{is2} - \mu_{y_{s2}}\right)\left(y_{im2} - \mu_{y2}\right) - (1 + \rho_s)^2\sigma_y^2\right] &= 0, \\
E\Big[z_i\Big( d_i\left(y_{is1} - \mu_{y1} - \kappa_{s1} - (1 + \rho_s)\left(y_{im1} - \mu_{y1}\right)\right) & \\
+ (1 - d_i)\left(y_{is2} - \mu_{y2} - \kappa_{s2} - (1 + \rho_s)\left(y_{im2} - \mu_{y2}\right)\right)\Big)\Big] &= 0.
\end{aligned}
$$

## A3 Change at the mean of the income distribution with heterogeneous-misreporting behaviour

To retrieve the relative change in the means of the income distributions, in a setting with two groups of households ($g = 1,2$) with heterogeneous misreporting, I take Equations (12) and (14) conditional on group g, as follows:

$$E\left(Y_{isg}\right) = k_g E\left(Y_{ig}\right) E\left(\exp(\epsilon_{is})\right), \tag{A1}$$

and

$$E\left(Y_{ig}^{(t)}\right) - E\left(Y_{ig}^{(t-1)}\right) = \frac{E\left(Y_{isg}^{(t)}\right) - E\left(Y_{isg}^{(t-1)}\right)}{k_g \exp\left(\frac{\sigma_{is}^2}{2}\right)}, \tag{A2}$$

for $g = 1,2$, where the suffix $g$ indicates group-specific variables and parameters. I recall that in such setting the parameter $\rho_s$ is assumed to be time-constant and equal to zero. If we let $\alpha^{(t)}$ denote the share of households belonging to the first group, $g = 1$, at time $t$, we have that

$$E\left(Y_i^{(t)}\right) = \alpha^{(t)} E\left(Y_{i1}^{(t)}\right) + \left(1 - \alpha^{(t)}\right) E\left(Y_{i2}^{(t)}\right), \tag{A3}$$

and

$$E\left(Y_i^{(t)}\right) - E\left(Y_i^{(t-1)}\right) = \alpha^{(t)} E\left(Y_{i1}^{(t)}\right) + \left(1 - \alpha^{(t)}\right) E\left(Y_{i2}^{(t)}\right) + \\ - \alpha^{(t-1)} E\left(Y_{i1}^{(t-1)}\right) - \left(1 - \alpha^{(t-1)}\right) E\left(Y_{i2}^{(t-1)}\right). \qquad \text{(A4)}$$

Further assume that the relative share of each group remains constant within the two time periods. It may be the case that this can be directly verified in the data, if the characteristics from which the heterogeneity stems are observed in both time periods. In this case, Equation (A4) can be re-written as:

$$E\left(Y_i^{(t)}\right) - E\left(Y_i^{(t-1)}\right) \\ = \alpha \left[E\left(Y_{i1}^{(t)}\right) - E\left(Y_{i1}^{(t-1)}\right)\right] + (1 - \alpha) \left[E\left(Y_{i2}^{(t)}\right) - E\left(Y_{i2}^{(t-1)}\right)\right], \qquad \text{(A5)}$$

with $\alpha$ being the time-constant relative share of the first group.

Then, to estimate of the relative change in the means of the income distributions it is sufficient to take the ratio of the quantities described by Equations (A5) and (A3), where $\Delta E\left(Y_{ig}^{(t)}\right)$ is replaced by the empirical counterpart of the right-hand-side of Equation (A2), making use of the estimated parameters for group $g$, and $E\left(Y_{ig}^{(t-1)}\right)$ is estimated starting from Equation (A1).

## A4 Longitudinal weighting process and regression results

The longitudinal weighting process follows a similar approach to that used for the SHIW (Loschiavo et al., 2024).

First, the CSIH 2022 cross-sectional weights ($w_i^{2022}$) were adjusted for unit non-response using a logit model for the conditional probability of panel retention in the CSIH 2023:

$$\log \frac{p_i}{1 - p_i} = \boldsymbol{x}_i' \boldsymbol{\beta},$$

where $p_i = (S_i = 1 \mid \boldsymbol{X}_i = \boldsymbol{x}_i)$, $S_i$ is an indicator function equal to one for households in the CSIH 2022 that were also interviewed in the CSIH 2023 (identifying the panel sample), and with the following characteristics included among the covariates $\boldsymbol{X}_i$: an indicator for households interviewed in the 2020 SHIW (to account for long-term panel households), homeownership status, geographical area of residence, municipality size, household size, quintiles of SHIW 2022 wealth and CSIH 2022 single-question income, and respondent’s main occupation, educational level, and citizenship. Non-response adjusted weights are thus computed using the inverse of the estimated probability of retention: $w_i^{nr} = w_i^{2022} \times 1/\widehat{p_\iota}$.

Then, the $w_i^{nr}$'s are adjusted using an iterative weight rebalancing technique (raking) to align the weighted frequency distribution of individuals with specific characteristics in the CSIH 2023 with actual shares obtained from external sources:

$$w_i^{2023} = w_i^{nr} \times \gamma_i,$$

where $\gamma_i$ is the adjustment factor for household $i$. External benchmarks include socio-demographic characteristics such as gender, age groups, geographical area of residence, municipality size, household size, educational level, and household income quintile in 2022 (retrieved from the SHIW).

Regression results obtained using this set of weights are reported in Table A1.

Table A1 - Regression analysis using longitudinal weights

| | Entire panel sample | | | | Same respondent | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| Labour market outcomes | | | | | | | | |
| At least one component began working or experienced an increase in labour income | 0.135*** | | 0.073 | 0.081* | 0.147*** | | 0.087* | 0.092** |
| | (0.042) | | (0.046) | (0.043) | (0.045) | | (0.049) | (0.047) |
| At least one component ceased working or experienced a decrease in labour income | -0.072 | | -0.09 | -0.115 | -0.023 | | -0.057 | -0.075 |
| | (0.088) | | (0.084) | (0.085) | (0.091) | | (0.090) | (0.090) |
| An income-earning individual joined the household | 0.364*** | | 0.339** | 0.190*** | 0.225*** | | 0.181*** | 0.162*** |
| | (0.138) | | (0.154) | (0.058) | (0.054) | | (0.052) | (0.060) |
| An income-earning individual left the household | -0.147 | | -0.241 | -0.247 | -0.038 | | -0.134 | -0.173 |
| | (0.199) | | (0.215) | (0.211) | (0.214) | | (0.241) | (0.247) |
| Other non-employment income† increased | 0.211*** | | 0.118* | 0.127* | 0.198*** | | 0.094 | 0.109 |
| | (0.063) | | (0.069) | (0.065) | (0.069) | | (0.076) | (0.070) |
| Other non-employment income† decreased | -0.062 | | -0.075 | -0.051 | -0.033 | | -0.05 | -0.021 |
| | (0.069) | | (0.075) | (0.083) | (0.058) | | (0.066) | (0.072) |
| Main household income source (SHIW 2022) | | | | | | | | |
| Self-employment | | -0.055 | -0.059 | -0.023 | | -0.058 | -0.065 | -0.022 |
| | | (0.182) | (0.177) | (0.177) | | (0.188) | (0.186) | (0.184) |
| Pensions | | -0.039 | -0.036 | -0.054 | | -0.032 | -0.028 | -0.035 |
| | | (0.158) | (0.155) | (0.156) | | (0.163) | (0.163) | (0.163) |
| Payroll employment | | 0.009 | 0.011 | 0.017 | | 0.025 | 0.021 | 0.036 |
| | | (0.158) | (0.152) | (0.152) | | (0.163) | (0.160) | (0.160) |
| Geographical area: Centre | | 0.084 | 0.093* | 0.119** | | 0.098* | 0.105* | 0.138** |
| | | (0.056) | (0.055) | (0.055) | | (0.060) | (0.060) | (0.059) |
| Geographical area: South | | 0.093 | 0.119** | 0.126** | | 0.089 | 0.110* | 0.124** |
| | | (0.057) | (0.059) | (0.058) | | (0.060) | (0.061) | (0.062) |
| Housheold income below the median in 2023 | | -0.223*** | -0.198*** | -0.212*** | | -0.218*** | -0.197*** | -0.216*** |
| | | (0.047) | (0.045) | (0.048) | | (0.049) | (0.048) | (0.052) |
| Constant | 0.019 | 0.142 | 0.114 | 0.005 | 0.021 | 0.141 | 0.112 | 0.001 |
| | (0.034) | (0.158) | (0.157) | (0.174) | (0.035) | (0.164) | (0.166) | (0.183) |
| Controls for misreporting heterogeneity | | | | ✓ | | | | ✓ |
| Observations | 1,317 | 1,317 | 1,317 | 1,313‡ | 1,214 | 1,214 | 1,214 | 1,214 |

Notes: Significance levels: * 10% ** 5% *** 1%. Longitudinally weighted estimates. † Including returns on financial capital, rental earnings, etc. Controls for misreporting heterogeneity: CSIH 2023 respondent's gender, educational level, and main income earner status in SHIW 2022. ‡ In four households, the CSIH 2023 respondent was not a household member in 2022; therefore, their main income-earning status in SHIW 2022 is unavailable. Baseline category for: (i) Main household income source: Other income; (ii) Geographical area: North.