Bounds for timely estimates of average household income

Domenico Depalo & David Loschiavo *

Abstract

We propose timely estimates of average household income based on survey data, which are typically affected by item non-response. We introduce a set-identification approach that exploits income brackets, exact declarations, and mild monotonicity assumptions to bound average income without the need for imputation or instruments. Using data from the Bank of Italy's Household Outlook Survey, we show that our bounds contain the official statistics while being available almost one year earlier. The economically plausible assumptions we impose tighten the initial bounds by 25%. Our approach is simple, broadly applicable, and well suited for contexts where timeliness and robustness to selection bias are crucial.

JEL classification: C14; C83; D31; I32

Keywords: Set identification; Partial identification; Household income; Income brackets; Timely statistics; Sample selection bias.

*We would like to thank Andrea Neri, Alfonso Rosolia and Federico Tullio. Replication files and additional results will be available at the webpage: http://sites.google.com/site/domdepalo/. The views expressed in this paper are those of the authors and do not imply any responsibility of the Bank of Italy. Corresponding address: Domenico Depalo, Banca d'Italia, Economics and Statistics Department, Via Nazionale, 91 - 00184 Roma, Tel.: 39-06-4792 5989, e-mail: domenico.depalo@bancaditalia.it.

1 Introduction

The availability of microdata on household income is essential for measuring policy-relevant dimensions of heterogeneity among economic agents. These data may be collected through administrative sources or surveys. Administrative data, like tax records, are subject to substantial delays, access limitations and restrictions, potential non-random measurement errors due, for instance, to tax evasion, and are not collected for analytical purposes. As surveys do not suffer from these drawbacks, for research purposes they are usually preferred to administrative data. However, the more detailed and precise the information gathered in the survey questionnaire, the higher the probability of systematic unit-/item-non-response and the greater the time lag between the collection and the release of the information.

To meet the growing demand for timely income statistics, statistical agencies increasingly rely on short surveys that collect income information through a single question or a small set of summary questions (Daikeler et al., 2019). An issue with summary questions is the systematic non-response to the item(s). To circumvent the bias that would arise as a consequence (Heckman, 1979; Korinek et al., 2006), in addition to —or instead of— the exact income level, sometimes surveys ask brackets, mainly adopting the unfolding bracket technique (Juster and Smith, 1997). This survey design naturally leads to a set-identification problem, where income is only partially observed and can be bounded rather than point-identified (Manski, 1990).

In this paper, we introduce set identification to provide a timely estimate of average household income, robust to non-response bias, using short surveys. Important features of our approach are its simplicity and general validity, which make it easily adaptable to various contexts. As a practical example, we focus on Italy in 2023. We mainly use the Household Outlook Survey (HOS) conducted by the Bank of Italy twice a year. The HOS collects timely information on the evolution of household economic conditions. For income, data referring to 2023 became available in May 2024, while the Italian National Statistical Institute released the official income statistics almost a year later (in March 2025). Three additional advantages are worth emphasizing for this paper. Firstly, the HOS asks both the income level and its bracket. Only the latter is compulsory. This feature

gives us the opportunity to explore a number of properties of our bounding approach. Secondly, the HOS respondents are drawn from the Survey on Income and Wealth (SHIW), a larger survey conducted by the Bank of Italy every three years. Hence, we are in a unique position to follow individuals over time. The panel structure of the data extends the array of possible bounds that we can identify. Finally, brackets are common to other important and widely used surveys, like Gallup World Poll, a survey which provides comprehensive and globally comparable income figures from 131 countries, and the Consumer Expectation Survey of the European Central Bank, which includes most of the euro area countries. Overall, the characteristics of our data span many of the circumstances typically encountered in empirical applications.

The idea of our approach is simple. We first consider households that declare their exact income level. If this variable were without gaps, the analysis would stop here. Unfortunately, this is not the case, as about half of the households in the HOS do provide only the bracket in which their income falls. Therefore, statistics based on the declared income level are unlikely to be representative of the entire population. This is the main motivation to apply set identification techniques. We start with bounds that always identify the correct income level because they rely on brackets only, but are too large to provide useful information. Nevertheless, this step is important because in the HOS there is no item non-response on income brackets. Then, we combine brackets and levels. This step remarkably narrows the width of the bounds, i.e. the distance between the upper and the lower bounds. Finally, we impose monotonicity assumptions that further shrink the identified set. Monotonicity provides constraints on the unobserved income, therefore reducing its admissible range of variation. For each assumption, we propose falsification tests.

In Italy, in 2023, the combination of brackets and levels has substantial identification power, as the width of the bounds shrinks by half compared to bounds using only income brackets. Monotonicity further narrows this width, even by 25%. Under our preferred assumption, that income increases with the educational attainment of household components, the average income of Italian households was between [28,602; 31,460] euros, consistent with the official data (31,356 euros).

The focus of this paper is on the average income. Extensions to other functionals are possible. Namely, the extension to quantiles is straightforward using the approach in Foresi and Peracchi

(1995) and Peracchi (2002). This is important to derive other indicators, such as the incidence of poverty. Yet, it is important to be aware of the limitations of the exercise, which, for instance, does not extend to interquantiles or other spread parameters (Manski, 1994; Stoye, 2010), e.g. the variance or the Gini index, at least under all the assumptions that we consider.

This paper makes three contributions. First, we show how income brackets alone can be used to construct bounds that are robust to item non-response and available with minimal delay. Second, we demonstrate how combining brackets with exact declarations and panel information substantially tightens these bounds without the need to rely on imputation or exclusion restrictions. Third, we illustrate how economically motivated monotonicity assumptions can further reduce uncertainty, while remaining transparent about their credibility and empirical support.

We contribute to the vast literature evaluating misreporting of income in household surveys. In particular, Crossley et al. (2023) have recently assessed the performance of summary income questions in rapid surveys using a measurement error model. They find systematic under-reporting of household income from single questions. However, measurement errors in single questions are substantially uncorrelated with true income. Using the HOS, Tullio (2025) extends Crossley et al. (2023) measurement error model to longitudinal settings and estimates the dynamics of household income. He finds that single questions allow for consistent inference about underlying income changes between two time periods. Both papers show that single questions preserve the relative ranking of household income along the distribution to a satisfactory extent. Leveraging these contributions, we provide a tool to narrow the width of values in which the mean income falls. This allows a timely and more precise measurement of both income dynamics between two time periods and a number of heterogeneity indices using income levels across the distribution.

The paper is organized as follows. Section 2 describes the data. Section 3 introduces the methods that we apply in Section 4, where we derive the average income in Italy. Section 5 provides some concluding remarks.

2 Data

We use the Bank of Italy's main household surveys, the Survey on Household Income and Wealth (SHIW), and the new Household Outlook Survey (HOS).

Survey on Household Income and Wealth (SHIW)

By collecting data on all the economic variables comprising household balance sheets, the SHIW provides a comprehensive and accurate overview of the economic conditions of Italian households. Alongside detailed information on income, the survey provides a large array of individual and household characteristics, including age, educational attainment, marital and occupational status of each household member.

Starting with the 2025 edition, it will permanently switch from a biennial to a triennial frequency. The survey relies on a sample of about 10,000 households and is representative of the Italian population. To derive the population aggregates from the sample, a set of weights is available, and we always use them throughout the analysis (Faiella and Gambacorta, 2007; Loschiavo et al., 2024). The latest edition of the SHIW, which we employ in our empirical application, was fielded throughout 2023, and it refers to household conditions in 2022.

Importantly for this paper, in addition to the traditional detailed item-by-item set of questions on each income source of each household component, the SHIW 2022 questionnaire included a single question on total net household income for a 20% random sample of respondents. The question was asked before the set of detailed questions to avoid helping respondents recall each household-level income source and was asked with the unfolding bracket technique (Juster and Smith, 1997) to prevent item non-response. In particular, three questions were used. The first was about the net overall household income in 2022, distinguishing between 9 possible brackets: 0-10,000; 10,001-15,000; 15,001-20,000; 20,001-25,000; 25,001-35,000; 35,001-50,000; 50,001-100,000; 100,001-500,000; 500,001+. The second asked for the exact net overall household income with answers top-coded at 2,000,000 euros. The third, reserved for those who did not answer the exact income question, asked whether the exact income was closer to the middle, the lower, or the upper

half of the bracket chosen in the first question.

Household Outlook Survey (HOS)

The loss of information that would follow with the new triennial frequency of the SHIW is filled by a new survey, the Household Outlook Survey (HOS). The HOS is a survey run twice a year on a sample of 2,500 households selected from those who participated in the most recent edition of the SHIW (panel dimension). The HOS allows monitoring household economic conditions throughout the economic cycle, collecting information on their income and wealth in years when the SHIW is not conducted.

Compared with other timely surveys, the HOS benefits from a richer set of information on household characteristics, which can be retrieved from the previous SHIW wave. The HOS wave we use in this paper was conducted between March and April 2024. Its questionnaire included the set of three questions on total household net income in 2023 formulated in exactly the same way as in the 2022 wave of the SHIW. Indeed, to minimize the burden on households participating in the survey, a set of item-by-item questions on all household income sources was considered infeasible. Differently from the SHIW, this set of questions was asked to the entire sample instead of a random 20%.

3 Methods

In this section, we present our bounding approach for the average household income. Before introducing the formal notation, it is useful to describe the logic of our approach informally. If all interviewed households reported their exact income, estimation would be straightforward, as we would directly observe the quantity of interest. In practice, many households report only income brackets, which provide incomplete information but are available for the entire sample.

The width of the bounds, i.e. the distance between the upper and the lower bound, depends on the share of households declaring their income levels and on the assumptions that we are willing to impose. Therefore, our strategy is to start with bounds based solely on brackets—which are always

valid but potentially wide—and to progressively tighten them by exploiting additional information available for subsets of households or supported by economic reasoning. This approach is consistent with the "law of decreasing credibility" stating that stronger assumptions yield conclusions that are more powerful but less credible (Manski, 2011).

3.1 General bounds

To simplify notation, but without loss of generality, we illustrate the bounds using a simplified example with only two brackets. The first bracket ($B(1)$) includes income values between the lower limit $L(1)$ and the upper limit $U(1)$, i.e. $B(1) \in [L(1), U(1)]$. Similarly, $B(2) \in [L(2), U(2)]$. Although $B(1)$ and $B(2)$ are disjoint sets, $B(1) \cup B(2)$ covers the entire range of admissible variation of income. Under these conditions, the average income for the N households is $y \in [L, U]$, with

$$
\begin{array}{l|l}
\multicolumn{1}{c|}{L} & \multicolumn{1}{c}{U} \\
\hline
\frac{1}{N}\sum_{i=1}^{N}\{ & \frac{1}{N}\sum_{i=1}^{N}\{ \\
\quad L(1) \quad 1[y_i \in B(1)] & \quad U(1) \quad 1[y_i \in B(1)] \\
+ \quad L(2) \quad 1[y_i \in B(2)] & + \quad U(2) \quad 1[y_i \in B(2)] \\
\} & \}
\end{array}
\tag{1}
$$

where $L(j)$ and $U(j)$ for $j \in \{1, 2\}$ do not have subscript because they do not change with the i-th observation. Equation 1 says that each household reporting the first income bracket has an income at least equal to $L(1)$ and at most equal to $U(1)$. The overall average is weighted by the share of households in each bracket. Notice that the width (W) of these bounds is equal to $\frac{1}{N}\sum_{i=1}^{N}\{(U(1) - L(1))\, 1[y_i \in B(1)] + (U(2) - L(2))\, 1[y_i \in B(2)]\}$.

The HOS asks both the income level and bracket. Therefore, within each bracket some individuals provide the exact level y (which we flag with $E = 1$) instead of the bracket $B(j)$ (i.e. $E = 0$),

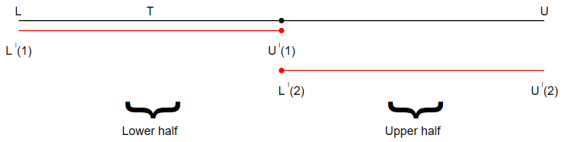
so that

$$
\begin{array}{l|l}
\multicolumn{1}{c|}{L} & \multicolumn{1}{c}{U} \\
\hline
\frac{1}{N}\sum_{i=1}^{N}\{ & \frac{1}{N}\sum_{i=1}^{N}\{ \\
\quad L(1)\ 1[y_i \in B(1)\ \&\ E=0] & \quad U(1)\ 1[y_i \in B(1)\ \&\ E=0] \\
+\quad y_i\ 1[y_i \in B(1)\ \&\ E=1] & +\ y_i\ 1[y_i \in B(1)\ \&\ E=1] \\
+\ L(2)\ 1[y_i \in B(2)\ \&\ E=0] & +\ U(2)\ 1[y_i \in B(2)\ \&\ E=0] \\
+\quad y_i\ 1[y_i \in B(2)\ \&\ E=1] & +\ y_i\ 1[y_i \in B(2)\ \&\ E=1] \\
\} & \}
\end{array}
\quad . \qquad (2)
$$

Equation 2 is identical to Equation 1 but exploits the fact that some households declare their exact income beyond brackets. For them, we already observe the outcome of interest and there is no need of brackets. This improvement provides three relevant results. Firstly, if there is at least one individual providing the income level y, the width of bounds in eq. 2 is strictly smaller than the width of bounds in eq. 1. To see this, in each of the two equations take the difference between the upper and lower bounds (i.e. the column-difference) for each row. If the household declares the income level, her observation has a null contribution to the width of the bounds in eq. 2 (because $y_i - y_i = 0$ regardless of the income bracket). The other two results are a consequence of the former property. First, the larger (smaller) the share of households declaring income level the narrower (larger) the bounds. In particular, if everybody declares the income level, the width is zero; if everybody declares the income bracket, the width of bounds in eq. 1 is equal to the width of bounds in eq. 2. Second, the narrower the brackets, i.e. the larger the number of brackets, the narrower the width of the bounds. For instance, suppose that the true income of a generic household declaring brackets is equal to T in Figure 1. Under the original bracket definition, the contribution to the width of the bounds is $\frac{1}{N}(U-L)$. Under a narrower definition of the bracket, e.g. if we split the segment between U and L in two equal halves, e.g. $[L(1), U(1)]$ and $[L(2), U(2)]$ as divided by the circles in Figure 1, the width of bounds becomes correspondingly smaller (by half in this special case), namely the bounds go from $U-L$ to $U'-L'$ (where we use superscript for the new definitions) and the width shrinks from $W(U-L)$ to $W(U'-L') = \frac{1}{2}W(U-L)$. However, one should be wary of concluding in favour of the largest possible number of brackets—which is

the income level, at the limit— because, as brackets become narrower, individuals may become more reluctant to answer. In this respect, the solution of the HOS that first asks for the larger bounds definition, thus securing the highest response rate, and then asks either the level or, as a second best, whether the true value is in the lower or in the upper half of the original bracket, is reasonable.

Figure 1: Narrower definition of brackets reduces the width of bounds



3.2 Monotonicity

To further narrow the bounds, we consider additional monotonicity-based assumptions. The first exploits the panel structure of the data and is rooted in the survey design. The second relies on economic theory and considers the marginal distributions of income. Although these assumptions are well motivated, they are not innocuous. Rather than treating them as maintained hypotheses, we view these assumptions as optional restrictions whose credibility can be empirically assessed.

3.2.1 Within-households (panel) monotonicity

We take advantage of repeated interviews of the same households that we follow between the SHIW (time $t = 0$) and the HOS (time $t = 1$), henceforth "within-households" or "panel" monotonicity. This monotonicity exploits the idea that typically (nominal) income does not decrease from period $t = 0$ to $t = 1$, if all the other characteristics ($x \in X$ that we omit —unless necessary— to simplify notation) remain identical. However, we caution that this hypothesis may not hold in all years or

for all household types, e.g. may not hold for those who receive income support which undergoes a revision of the rules, or for self-employed whose labour income may decrease during downturns. Below, we propose a check of this restriction.

Formally, within-households monotonicity implies that the income function of each household with given observable characteristics is weakly increasing over years, i.e. $t_1 \geq t_0 \Rightarrow y_i(t_1) \geq y_i(t_0), \forall i$ (Manski, 1997; Manski and Pepper, 2000).

Under this restriction, if household income belongs to $B(j)$ and at time $t = 0$ the household declared the level $y_{t=0}$ —which we index with $(S = 1)$ because the reference survey is the SHIW at that time—, then in $t = 1$ $B(j) \in [y_{t=0}, U(j)] \subseteq [L(j), U(j)] \, \forall j$ where $y_{t=0} \geq L(j)$. Upon appropriate substitution, the new bounds are

$$
\begin{array}{l|l}
\multicolumn{1}{c|}{L} & \multicolumn{1}{c}{U} \\
\hline
\frac{1}{N}\sum_{i=1}^{N}\{ & \frac{1}{N}\sum_{i=1}^{N}\{ \\
\quad L(1) \;\; 1[y_i \in B(1) \;\&\; E=0 \;\&\; S=0] & \quad U(1) \;\; 1[y_i \in B(1) \;\&\; E=0] \\
+ \;\; y_{i,t=0} \;\; 1[y_i \in B(1) \;\&\; E=0 \;\&\; S=1] & \\
+ \;\; y_i \;\; 1[y_i \in B(1) \;\&\; E=1] & + \;\; y_i \;\; 1[y_i \in B(1) \;\&\; E=1] \\
+ \;\; L(2) \;\; 1[y_i \in B(2) \;\&\; E=0 \;\&\; S=0] & + \;\; U(2) \;\; 1[y_i \in B(2) \;\&\; E=0] \\
+ \;\; y_{i,t=0} \;\; 1[y_i \in B(2) \;\&\; E=0 \;\&\; S=1] & \\
+ \;\; y_i \;\; 1[y_i \in B(2) \;\&\; E=1] & + \;\; y_i \;\; 1[y_i \in B(2) \;\&\; E=1] \\
\} & \}
\end{array}
\tag{3}
$$

This set is smaller than that of eq. 2. The degree to which the width of bounds reduces depends on how many households declare their income level in the SHIW at $t = 0$ $(S = 1)$ and on the difference between $y(j)$ and $L(j)$.

3.2.2 Between-households monotonicity

A different restriction follows from between-households monotonicity, which exploits the marginal distributions of income along the desired dimension, instead of the longitudinal structure of the data. The desired dimensions may include, for example, educational attainment -that in this section

we use as a running example - or age. In Section 3.3, we discuss the role of these dimensions more thoroughly.

Under between-households monotonicity, households with better/higher-valued characteristics have higher mean income functions than those with worse/lower-valued characteristics, i.e. $E[y(j) \mid X_1] \geq E[y(j)|X_0]$ where $X_1 \geq X_0$ (Manski and Pepper, 2000). For instance, using education as an example, households whose components have higher education are on a higher mean income trajectory than households whose components have a lower education. A motivation behind this assumption is that better-educated individuals are more likely to work and enjoy better-paid labour positions (Mincer, 1974; Card, 1999), on average. As labour income is the largest portion of overall income, having a better-paid job may imply a higher income level.

To make the assumption operational, suppose that there are two households with the same characteristics ($x \in X$, omitted for notational convenience), but one with a low educational level (LE) and the other with a high educational level (HE).[1] According to this restriction, $E[y(j)|HE] \geq E[y(j)|LE] \equiv \tilde{y}(j)$. As a consequence, this assumption has no identification power for the lowest educational attainment, whereas it provides an improvement to the lower bound of households with higher educational attainment:

$$
\begin{array}{l|l}
\multicolumn{1}{c|}{L} & \multicolumn{1}{c}{U} \\
\hline
\frac{1}{N}\sum_{i=1}^{N}\{ & \frac{1}{N}\sum_{i=1}^{N}\{ \\
\quad L(1)\ 1[y_i \in B(1)\ \&\ E=0\ \&\ LE] & \quad U(1)\ 1[y_i \in B(1)\ \&\ E=0] \\
+\ \tilde{y}(1)\ 1[y_i \in B(1)\ \&\ E=0\ \&\ HE] & \\
+\ y_i\ 1[y_i \in B(1)\ \&\ E=1] & +\ y_i\ 1[y_i \in B(1)\ \&\ E=1] \\
+\ L(2)\ 1[y_i \in B(2)\ \&\ E=0\ \&\ LE] & +\ U(2)\ 1[y_i \in B(2)\ \&\ E=0] \\
+\ \tilde{y}(2)\ 1[y_i \in B(2)\ \&\ E=0\ \&\ HE] & \\
+\ y_i\ 1[y_i \in B(2)\ \&\ E=1] & +\ y_i\ 1[y_i \in B(2)\ \&\ E=1] \\
\} & \}
\end{array}
\tag{4}
$$

[1] In this section, we assume that all individuals in a household achieve the same educational attainment. While this simplifying assumption is useful for the exposition of the method, it may be restrictive in the empirical application. Therefore, we do not impose it in our application. In Section 4 we explain how we make it operational with our data.

Three considerations are in order. Firstly, whilst the within-household monotonicity must be satisfied for all households, the between-household monotonicity should be satisfied on average. It follows that the former assumption is more restrictive than the latter. Secondly, a direct verification of the conditions pertaining to the within-households and between-households monotonicity assumptions is not directly testable. However, we can build appropriate quasi-falsification tests. For both restrictions, the income function should be non-decreasing in the appropriate dimension - i.e. time for the case of within-household monotonicity and educational attainment for the example of between-household monotonicity. To this aim, we suggest to use the sample of households reporting income levels in both waves or across educational levels to gain intuition about the validity of the assumption. A limitation of this approach is that neither of the two assumptions would apply to the households used in the test because for them we already have the target variable, i.e. income level. This explains why this is a "*quasi*-falsification" test and not a proper "falsification test". Yet, the rationale for this test is that, had these households not declared their income in the HOS, the assumption would apply to them. Therefore, although evidence against the restriction does not directly falsify the restriction, it jeopardizes its credibility. Thirdly, (from a theoretical viewpoint) if we refuse the monotonicity assumptions, we can still gain insights on a possible correction to restore them. For instance, consider the within-households monotonicity and compare $y(j)_{t=1} - y(j)_{t=0}$ for all households. Suppose that for some households that declare exact incomes in bracket j in both the SHIW and the HOS we obtain $y(j)_{t=1} - y(j)_{t=0} = \delta(j) < 0$. Following Manski and Pepper (2018) and Rambachan and Roth (2023) we can use δ to inflate $y(j)_{t=1}$. This approach simply requires to use $y(j)_{t=0} + \delta(j)$ instead of $y(j)_{t=0}$ in the lower bound of eq. 3.

3.3 The role of covariates

The monotonicity assumptions exploit observable characteristics ($x \in X$). Their role is remarkably different between the set identification approach of this paper and the more standard point identification approach. Pepper (2000, p.475) clarifies that in set identification "covariates are merely used to define subpopulations of interest, not to control for spurious effects. It is important to recognize that this use of covariates to define subpopulations of interest differs from the conventional use of

covariates to account for nonrandom treatment assignment [in point identification... With point identification], researchers attempt to 'correctly' choose a set of covariates or 'control variables' such that the exogenous selection assumption applies. If certain variables are omitted, the estimated [...] effects will be biased. Inevitably, however, these analyses are controversial with much debate about whether or not the researcher omitted 'important' explanatory variables." Hence, the concept of correctly specified model specification does not apply to set identification where covariates flag the population to which the restrictions apply (Manski and Nagin, 1998).

We conduct inference using the methods presented in Imbens and Manski (2004), for which we applied the Stata code of McCarthy et al. (2015).

4 An application to the Italian data

In this section, we estimate the average income of Italian households in 2023, with the data described in Section 2. In the HOS, 55% of the households report their exact income. However, the response behaviour changes systematically with observable economic, socio-demographic, and labour market characteristics (Table A.1). Reporting the income level is less frequent in the lower half and for the highest bracket of the income distribution —although estimates are not very precise— whereas it is more common for individuals who are older and better educated. Finally, employees and retirees are more likely to report the exact level of income than self-employed. Using valid answers, the average income was 30,701 euros (95% confidence intervals: [28,527; 32,875]).

To recover the full sample, we may impute missing values. Without a reference model, we opt for a flexible approach. We start with all the possible interactions used to explain the response behaviour in Table A.1 and let LASSO (Belloni and Chernozhukov, 2013; Lucchetti et al., 2025) select the most appropriate set of covariates (the full list is available upon request).[2] With this approach, the average income was 31,740 euros (95% confidence intervals: [29,543; 33,939]). Two considerations are worth stressing. Firstly, there is no guarantee that the LASSO imputation is correct, because the set of available characteristics may still miss important drivers, i.e. may

[2]We also tried the standard parametric sample selection model (Heckman, 1979), invoking the normality assumption for identification, but without exclusion restrictions, which are not available to us. This model failed to converge.

suffer from an omitted variable bias. Even if the model is correctly specified, it is estimated on a selected sample. As a consequence, the coefficients may not be valid beyond the estimation sample. Secondly, the range of variation of the two estimates is sizeable. For reference, the estimate consistent with the official release of Istat (2025), based on EU-SILC, is 31,356 euros. Therefore, the point estimate based on the valid answers is too low whereas the one based on imputation is too high. This feature emphasizes that point-identification methods rely on strong assumptions, such as a correct model specification for imputation, that, on the one hand, are difficult to verify, and, on the other hand, may substantially affect the results. This is a strong motivation in favour of an alternative approach.

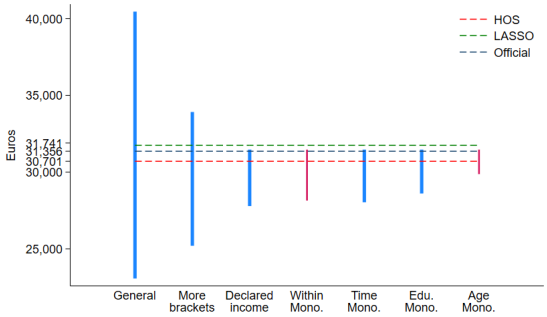
Bounds

The non-random pattern of the item non-response implies that the income recovered from the reported or the imputed levels may be biased (Heckman, 1979; Little, 1995). A solution to avoid sample-selection bias exploits the bounds. Since our bounds use also brackets, an advantage of the HOS is that all the observations report this information. As a consequence, we do not need to model the missingness behaviour.[3]

We apply the bounds of Section 3, imposing increasingly more restrictive assumptions. The main results are in Figure 2. Each solid vertical line is a bound. The dotted lines represent the average income estimated using valid answers on income level (HOS) or imputation (LASSO), as well as the official release (Official). Table 1 reports the exact numbers of the upper and the lower bounds, their 95% confidence intervals based on Imbens and Manski (2004), and the width.

[3]If this were the case, a solution could be along the lines of Horowitz and Manski (2000).

Figure 2: Bounds on households income under various assumptions



Notes: Bound are in red if the underlying assumptions are questionable for the reasons discussed in the text.

Table 1: Bounds on households income - exact values

Variable	General	More brackets	Declared income	Monotonicity				EU-SILC income
				Within	Between			
					Time	Education	Age	
C.I. 95% Lower	21,504	23,812	26,352	26,706	26,629	27,073	28,345	30,893
Lower	23,068	25,198	27,784	28,145	28,031	28,602	29,866	31,356
Upper	40,449	33,911	31,460	31,460	31,460	31,460	31,460	31,356
C.I. 95% Upper	42,542	35,473	32,957	32,957	32,957	32,916	32,916	31,818
Width	17,381	8,713	3,676	3,315	3,429	2,858	1,594	0

Note: Standard errors are from Imbens and Manski (2004). For reference, we report the EU-SILC average income in 2023, net of imputed rents and adjusted to be comparable to the SHIW (see below).

We start with the 'General' bounds based on eq. 1 and 9 brackets. With these bounds, household income on average is between [23,068; 40,449] euros. Although these bounds solve the statistical issue of selection-bias in response behaviour, the width (almost 17,400 euros) is very large for policy making analysis. However, the HOS (and the SHIW as well) has two following questions on income: one is whether the exact level is closer to the lower or the upper half of the declared bracket; one is the precise amount. The former question doubles the number of brackets.[4] Like in the 'General' bounds, there is no missing value on this follow-up question. Using 'More

[4]To be precise, the question has three possible answers: closer to the lower half (29%); broadly in the middle (41%); closer to the upper half (30%). In principle, we could triple the number of brackets. We did not proceed in this way because different respondents have different opinions on the quantitative meaning of 'middle'. Therefore, we avoid introducing discretion on this issue by using our numerical interpretation of the word. However, we run the exercise twice: in the main exercise we work as if 'middle' means the lower half, but we checked that the conclusions remain untouched if we would assume it means the upper half (results are available upon request). A clarification of the definition of "middle" represents an area of improvement for the SHIW and the HOS. Alternatively, providing only two alternatives would simplify the analysis. For technical considerations on similar issues, see Dominitz and Manski (2017).

brackets', household income on average is between [25,198; 33,911] euros, with a width that shrinks by half, as expected using the logic of Figure 1. This result suggests that the approach adopted by the HOS provides a good balance between the need of avoiding systematic non-response while providing useful information for policy making.

When asked for the exact amount of income, the household response rate is slightly above 50% - as described at the beginning of the section. We combine this information with the brackets, using eq. 2. The 'Declared income' bounds are much smaller than the previous, ranging in the set [27,784; 31,460] euros, with a remarkably small width of 3,676 euros, i.e. less than half that of 'More brackets' bounds and almost 5 times smaller than the 'General' bounds. Therefore, without invoking any assumption, we are able to substantially shrink the initial set, while at the same time avoiding the statistical drawbacks that may jeopardize the validity of the point identification analysis, namely the sample selection bias. Related to this issue, note that the point estimates of the LASSO-imputed values are inconsistent with the interviews, as the estimated average income is above the upper bound. However, because of the large uncertainty surrounding these imputations, the confidence intervals overlap.

It is instructive to understand the source of the gain in width because the logic applies to all the following bounds. Define $\rho = \Pr(E = 1)$ the share of observations for which we observe the exact income level. Going from the 'More brackets' to the 'Declared income' bounds the width (W) becomes

$$\begin{aligned} W(\text{More brackets}) &\equiv \rho \times W(\text{More brackets}|E=1) &+ (1-\rho) \times W(\text{More brackets}|E=0) \\ &> \\ W(\text{Declared income}) &\equiv \rho \times 0 &+ (1-\rho) \times W(\text{More brackets}|E=0) \end{aligned}$$

which shrinks by more than half, consistent with the response rate, as predicted in Section 3. If the response rate of the income level were higher, the width of the bounds would be even smaller. Increasing the response rate for this item is a challenge for the next waves of the survey (Stantcheva, 2023).

Monotonicity

From now on, we impose the monotonicity assumptions. They narrow the bounds, i.e. have identification power, but, if falsified, they may undermine the credibility of the analysis (Manski, 2011). Therefore, it is important to support their validity.

We first exploit the panel structure of the data between the SHIW ($t = 0$) and the HOS ($t = 1$).[5] According to the within-households monotonicity assumption, nominal income should not decrease over time, for each household. Although this assumption seems plausible most of the time, we emphasize that there may be cases where it is falsified. For instance, in our application, the average income of the lowest income brackets decreased between 2022 and 2023 (Table 2). This pattern reflects the new rules concerning the minimum income scheme ('Reddito di Cittadinanza'), introduced to support low-income households, which changed over the two years. Until 2022, 'Reddito di Cittadinanza' was granted to households with low income (defined as almost 10,000 euros per year, increased with the number of the household's components), for 18 months; the benefit was up to 780 euros per month, including rent. Since 2023, the length of the benefit decreased to 7 months and the eligibility conditions became tighter; also, the in-pocket benefit no longer included rent, which was paid directly to the landlord (for further details, see Bovini et al., 2023). For brackets not affected by the minimum income scheme, the average household income increased over time (with the only, very small exception between 50,000-100,000 euros).

[5] In principle, we could recover the total income for all the respondents to the SHIW rather than focusing on the random 20% who received the question (Section 2). However, we decided to fully preserve the panel structure of the data, namely to use the same questions between the SHIW and the HOS. This approach is motivated by the design of the two surveys and the exact wording of the questions. For further technical details, see Tullio (2025).

Table 2: Inspection of the within-household monotonicity assumption

Bracket	$y_{t=0}$	$y_{t=1}$	$y_{t=1} - y_{t=0}$
0-10,000	13894.4	3735.2	-10159.3
10,001-15,000	14242.6	14049.9	-192.7
15,001-20,000	17500.0	17195.3	-304.7
20,001-25,000	22944.8	23234.8	290.0
25,001-35,000	30189.7	30332.2	142.5
35,001-50,000	39223.2	42511.7	3288.5
50,001-100,000	69707.3	67603.4	-2104.0
100,001-500,000	116444.1	149802.0	33357.8
500,001+	500000.0	680037.4	180037.4

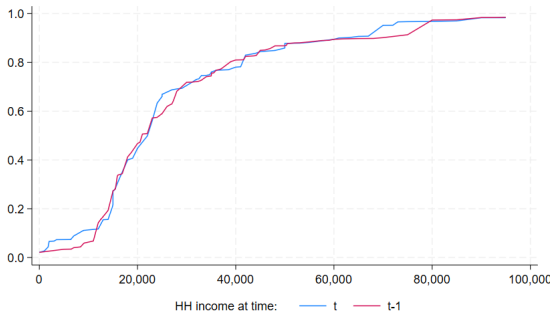
Strictly speaking, these statistics constitute evidence against the within-household monotonicity assumption, which is required to hold for all observations. To emphasize the potential violation of the assumption, we plot the resulting bounds in red in Figure 2. Yet, we believe it is useful to show how the method works for at least two reasons. Firstly, the observed violations are economically small and consistent with well-identified institutional changes affecting specific income components. Therefore, the Italian case illustrates an important feature of our framework: monotonicity assumptions may fail, e.g. for institutional or policy-driven reasons. Rather than invalidating the approach, this highlights the value of a framework that makes assumptions explicit and empirically assessable. Secondly, the income level in the SHIW is available only for a random 20% of the sample. In the HOS all respondents received the question. The share of households responding in the SHIW but not in the HOS is about 5%. Therefore, the design of the SHIW makes the assumption not very binding. Imposing the within-households monotonicity, the bounds are [28,145; 31,460] euros, with a width of 3,315 euros, not much smaller than that based on the 'Declared income'.

We then move to between-households monotonicity assumptions. We present three possible versions: 1) over time; 2) over education; 3) over age classes.[6] When we consider time, we use the

[6]In principle, it is possible to combine these dimensions, e.g. the combination of age and education. The theoretical requirements are the same as those presented here. However, similar combinations are seldom used in set identification, given the different role of covariates with respect to point identification strategy (see Manski and Pepper, 2018, and Section 3.3) and the weaker theoretical support. We are not aware of a theory suggesting whether being 45 years old

SHIW and the HOS, without exploiting the panel structure of the data. The difference between the two assumptions relies on how they exploit time (Manski and Pepper, 2000). The within-household monotonicity must be valid for each household, whereas the between-household monotonicity must be valid on average. Moreover, under the current restriction, households need not preserve their ranking position in the income distribution. Like in the within-households monotonicity, we support the assumption using households for whom we observe income level in both surveys. We compare the cumulative distribution functions of income over time. Given the nature of the between-households monotonicity, we consider the marginal distributions of income in the two periods (Doksum, 1974). The two distributions are largely overlapping (Figure 3), apart from the extreme values, and, according to the Kolmogorov-Smirnov test, they are not different ($KSstat. = 0.067; P - val. = 0.776$). Based on this result, we do not expect a large gain from this assumption. The new bounds

Figure 3: Household income stochastic dominance



are [28,031; 31,460] euros, which confirm the theoretical prediction of a small gain in terms of width (only 250 euros less than the 'Declared income' assumption and 100 euros larger than the within-monotonicity assumption, which imposes a stronger restriction).

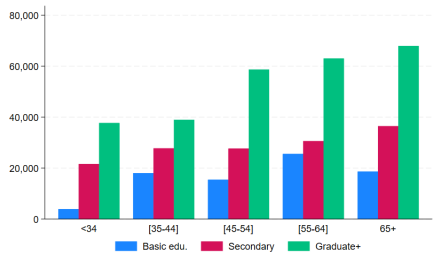
The second between-households assumption that we consider is over educational attainment. We allow different individuals in the household to achieve different educational attainments. Therefore, we need a synthetic measure of household education. To this aim, we proceed in three steps, limiting

with 15 years of education delivers higher or lower income than being 40 years old with 20 years of education, and solving the issue for all possible {age,education} combinations. Also, from a statistical point of view, the larger the set of covariates the smaller the sample size of each cell defined by the combination of the dimensions.

the attention to the head of the household and his/her partner (if any). In the first step, we fix the age class of all household components. In the second step, within each age class, we order the school achievement of the head of household and of the cohabiting partners (if any), regardless of the formality of the union, i.e. cohabiting partners are equivalent to married partners. In the final step, we compare pairwise the school achievement of the components, i.e. head of household i vs head of household j, and partner of household i vs partner of household j. In this way, we compare households, i.e. we impose the restriction, within the same age class and family structure. We do not control for gender of the head of the household, because in the greatest majority is a man and sample size considerations suggest to avoid this control. We view this approach as a balance between the combination of age and education and the focus on education only.

To check the credibility of the assumption, we focus on households that declared their income level and compare the average income by educational attainment. For a clearer representation, we show the results only for the head of the household. However, they are confirmed if we consider jointly the heads of the household and their partner (results available upon request). From Figure 4, income monotonically increases with educational attainment, within each age-class. This finding supports the validity of the restriction. Evidence in favour of a similar assumption is fairly common in the literature (e.g. Manski and Pepper, 2000; Okumura and Usui, 2014, in the set identification literature). Under education monotonicity, the income bounds are [28,602; 31,460] euros, a width

Figure 4: Houshold income and education monotonicity

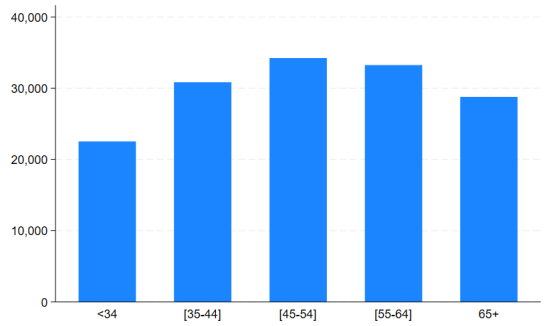


Note: The figure relies on households that declared their income level. For a clearer representation, we show the results only for the head of the household. However, they are confirmed if we focus on partners or if we consider jointly the heads of the household and their partner. These results are available upon request.

equal to 2,858, or 5-600 euros less than the time-based assumptions.

Finally, we consider age monotonicity. To show the pitfalls of using an inappropriate conditioning set, here we do not consider other characteristics beyond age, namely educational attainment. To see the implications of this approach, consider an (extreme) example of two individuals, one 50 years old with a PhD, and the other 55 years old with basic education. Under this assumption, the latter should have a higher income than the former, on average. Whether this is possible or not is an empirical issue that we check using the households that declared their exact income level. From Figure 5, the pattern of income is a reverse-U-shaped in age. Therefore, the imposed assumption is falsified. If we impose the age monotonicity assumption, bounds are remarkably small, at [30,057; 31,460] (a width smaller than 1,500 euros). However, the evidence against the restriction suggests that this monotonicity might not be credible. For this reason, we plot these bounds in red in Figure 2. This example shows the role of covariates in set identification, which identify subpopulations of interest to which assumptions apply, instead of controlling for spurious effects (Section 3).

Figure 5: Houshold income and age monotonicity



Note: The figure relies on households that declared their income level.

Our preferred bounds of the average income in 2023 rely on education monotonicity, which deliver the set [28,602; 31,460] euros. This includes the official EU-SILC figure, which is 31,356 euros (last column of Table 1) **metterei in commento ciò che segue fino alla fine della frase, inclusa la nota; se ce lo chiedono lo tiriamo fuori. ciò che è nella nota sta già nella footnote che accompagna la tavola 1. se non ti è super chiara la proposta, chiamami**, after adjusting for methodological differences with the SHIW.[7] Although our method produces a

[7]EU-SILC increases households' declared self-employment income when this is lower than that recorded in tax register data. Social transfers and other income sources are also added to the declared income when not reported.

set rather than a point estimate, it delivers results nearly a year ahead of the official release and with high credibility, as the EU-SILC value lies within the identified range.

5 Conclusions

This paper introduces a new approach based on set identification (Manski, 1990) to estimate the average income of households, using surveys. Official statistics remain superior to estimates based on survey data. However, they become available with a long delay that makes their information less useful for timely policy making analysis.

To the best of our knowledge, set identification is new to this literature. This is an important improvement over standard approaches based on point identification, because the latter typically must rely on imputation or on some instruments to purge potential sample selection bias (Heckman, 1979). Imputation requires a correct model specification, which in some cases cannot be estimated, e.g. because not all the necessary variables are collected. Instruments are always difficult to find and, when available, generally identify the desired outcome only for specific sub-samples (Vytlacil, 2002). Bounds overcome these limitations, upon a proper definition of the relevant sample to which assumptions apply (Pepper, 2000). However, they identify a set of admissible values instead of a single point.

We apply our method to Italy. Using our preferred bounds, household income in 2023 was in the set [28,602; 31,460] euros, consistent with the official estimates, which became available only a year later. Interestingly, in our application, point identification methods based on imputation, namely LASSO, yield an estimation that lie outside our bounds, i.e. are inconsistent with interviews. This feature shows how sensitive such approaches can be to specification choices and, more broadly, to the underlying modeling assumptions.

We propose restrictions that are broad enough to apply to a wide range of empirical settings. The key requirements of our framework—income brackets, partial response, and short panels—are increasingly common in international surveys. At the same time, specific sample designs may justify

As a consequence, SHIW average income has been consistently around 80% of EU-SILC average income. In the text, we use 83.5%, which represents the average since 2004. The results are remarkably stable if we consider the last five waves (81.7%) or even the last wave (81.9%). Further details on EU-SILC are available in Istat (2025).

additional or alternative assumptions. Rather than advocating a one-size-fits-all approach, we view set identification as a disciplined starting point when point identification would rely on hard-to-verify assumptions, leaving researchers free to incorporate stronger restrictions when warranted by their institutional context or survey designs.

Overall, our results suggest that partial identification provides a credible and transparent tool, specially when timeliness and robustness are central concerns. While bounds necessarily trade point precision for credibility, they offer a rigorous way to incorporate additional information as it becomes available. We view this framework as particularly well suited to the growing use of rapid surveys.

References

Belloni, A. and V. Chernozhukov (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli 19*, 521–547.

Bovini, G., E. D. Carlo, and A. Tomasi (2023). The revision of the measures against poverty in italy. Questioni di Economia e Finanza (Occasional Papers) 820, Bank of Italy, Economic Research and International Relations Area. In Italian.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 30, pp. 1801–1863. Elsevier.

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

Daikeler, J., M. Bošnjak, and K. Lozar Manfreda (2019). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology 8*(3), 513–539.

Doksum, K. A. (1974). Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case. *The Annals of Statistics 2*(2), 267–277.

Dominitz, J. and C. F. Manski (2017). More Data or Better Data? A Statistical Decision Problem. *The Review of Economic Studies 84*(4), 1583–1605.

Faiella, I. and R. Gambacorta (2007). The Weighting Process in the SHIW. Temi di discussione (Economic working papers) 636, Bank of Italy, Economic Research and International Relations Area.

Foresi, S. and F. Peracchi (1995). The Conditional Distribution of Excess Returns: An Empirical Analysis. *Journal of the American Statistical Association 90*(430), 451–466.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*(1), 153–161.

Horowitz, J. L. and C. F. Manski (2000). Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data. *Journal of the American Statistical Association 95*(449), 77–84.

Imbens, G. and C. F. Manski (2004). Confidence Intervals for Partially Identified Parameters. *Econometrica 72*(6), 1845–1857.

Istat (2025). Condizioni di vita e reddito delle famiglie - anni 2023-2024.

Juster, F. T. and J. P. Smith (1997). Improving the quality of economic data: Lessons from the hrs and ahead. *Journal of the American Statistical Association 92*(440), 1268–1278.

Korinek, A., J. Mistiaen, and M. Ravallion (2006). Survey nonresponse and the distribution of income. *Journal of Economic Inequality 4*, 33–55.

Little, R. J. A. (1995). Modeling the Drop-Out Mechanism in Repeated-Measures Studies. *Journal of the American Statistical Association 90*(431), 1112.

Loschiavo, D., G. M. Mariani, M. Moscatelli, A. Neri, E. Porreca, and F. Tullio (2024). Enhancing the SHIW weighting methodology with external data sources. Methodological note, Bank of Italy, Economic Research and International Relations Area.

Lucchetti, L., P. Corral, A. Ham, and S. Garriga (2025). An application of lasso and multiple imputation techniques to income dynamics with cross-sectional data. *Review of Income and Wealth 71*(1), e12693.

Manski, C. F. (1990). Nonparametric Bounds on Treatment Effects. *American Economic Review 80*(2), 319–323.

Manski, C. F. (1994). *The selection problem*, pp. 143–170. Econometric Society Monographs. Cambridge University Press.

Manski, C. F. (1997). Monotone Treatment Response. *Econometrica 65*(6), 1311–1334.

Manski, C. F. (2011). Policy Analysis with Incredible Certitude. *Economic Journal 121*(554), F261–F289.

Manski, C. F. and D. S. Nagin (1998). Bounding disagreements about treatment effects: A case study of sentencing and recidivism. *Sociological Methodology 28*, 99–137.

Manski, C. F. and J. V. Pepper (2000). Monotone Instrumental Variables : With an Application to the Returns to Schooling. *Econometrica 68*(4), 997–1010.

Manski, C. F. and J. V. Pepper (2018). How Do Right-to-Carry Laws Affect Crime Rates? Coping with Ambiguity Using Bounded-Variation Assumptions. *The Review of Economics and Statistics 100*(2), 232–244.

McCarthy, I., D. L. Millimet, and M. Roy (2015). Bounding treatment effects: A command for the partial identification of the average treatment effect with endogenous and misreported treatment assignment. *Stata Journal 15*(2), 411–436.

Mincer, J. (1974). *Introduction to Schooling, Experience, and Earnings*. NBER Chapters. National Bureau of Economic Research, Inc.

Okumura, T. and E. Usui (2014). Concave-monotone treatment response and monotone treatment selection: With an application to the returns to schooling. *Quantitative Economics 5*(1), 175–194.

Pepper, J. V. (2000). The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis. *Review of Economics and Statistics 82*(3), 472–488.

Peracchi, F. (2002). On estimating conditional quantiles and distribution functions. *Computational Statistics & Data Analysis 38*(4), 433–447.

Rambachan, A. and J. Roth (2023). A More Credible Approach to Parallel Trends. *Review of Economic Studies 90*(5), 2555–2591.

Stantcheva, S. (2023). How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible. *Annual Review of Economics 15*(1), 205–234.

Stoye, J. (2010). Partial identification of spread parameters. *Quantitative Economics 1*(2), 323–357.

Tullio, F. (2025). Dynamics and measurement error in household income data collected with single questions. Forthcoming.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica 70*(1), 331–341.

A Additional results

Table A.1: Probability of reporting income level as function of observable characteristics

Variable		Income	+ Socio-demo.	+ labour	+Area
Income class	10001-15000	-0.032	-0.026	-0.031	-0.024
		(0.037)	(0.037)	(0.037)	(0.037)
	15001-2000	0.043	0.038	0.013	0.027
		(0.037)	(0.037)	(0.038)	(0.038)
	20001-25000	-0.066 *	-0.082 **	-0.104 ***	-0.088 **
		(0.036)	(0.036)	(0.037)	(0.037)
	25001-35000	-0.011	-0.040	-0.057	-0.038
		(0.035)	(0.036)	(0.037)	(0.037)
	35001-50000	-0.013	-0.017	-0.033	-0.015
		(0.034)	(0.036)	(0.037)	(0.037)
	50001-100000	0.087 **	0.068	0.050	0.070
		(0.040)	(0.044)	(0.044)	(0.045)
	100001500000	0.011	-0.009	-0.011	0.008
		(0.088)	(0.091)	(0.091)	(0.091)
	500001+	-0.328	-0.255	-0.206	-0.192
		(0.217)	(0.216)	(0.216)	(0.214)
# components			-0.034 ***	-0.028 ***	-0.029 ***
			(0.009)	(0.009)	(0.009)
Age class	[35-44]		0.024	0.025	0.022
			(0.047)	(0.047)	(0.047)
	[45-54]		-0.034	-0.036	-0.025
			(0.045)	(0.045)	(0.045)
	[55-64]		-0.004	0.022	0.028
			(0.045)	(0.047)	(0.046)
	65+		0.089 *	0.110 **	0.114 **
			(0.046)	(0.052)	(0.052)
Edu. level	Primary		0.007	0.011	0.014
			(0.095)	(0.095)	(0.096)
	Lower secondary		0.053	0.071	0.087
			(0.093)	(0.093)	(0.094)
	Upper secondary		0.145	0.160 *	0.168 *
			(0.093)	(0.093)	(0.094)
	College		0.159	0.187 *	0.191 *
			(0.097)	(0.097)	(0.098)
	College+		0.225 *	0.257 **	0.249 **
			(0.122)	(0.123)	(0.123)
Employee				0.100 ***	0.095 ***
				(0.027)	(0.026)
Retired				0.073 **	0.083 **
				(0.033)	(0.033)
Area	Center				0.150 ***
					(0.026)
	South				0.046 *
					(0.024)
Intercept		0.543 ***	0.497 ***	0.408 ***	0.339 ***
		(0.024)	(0.102)	(0.104)	(0.107)

Notes: Standard errors are robust to heteroscedasticity. *(**)[***] indicate significance at 10(5)[1]% confidence level.