OESTERREICHISCHE NATIONALBANK
EUROSYSTEM

Working Paper 285

Let's Talk Money:
Reducing Income Nonresponse in Surveys

Katharina Allinger, Melanie Koch

OeNB Working Paper Series
36 pages

Let's Talk Money: Reducing Income Nonresponse in Surveys

Survey participants frequently refuse to answer questions about their income. This is problematic as income is an important variable in social surveys and nonresponse can complicate and potentially bias results. We conduct a randomized experiment in four Central and Eastern European countries. In the survey respondents receive a question on the exact income amount followed by a categorical question if they refuse to answer. The experiment shows that a simplification of the follow-up question – reducing the number of brackets from more than twenty to three – raises response rates substantially.

Authors

Katharina Allinger
Oesterreichische Nationalbank,
Economic Analysis Division,
katharina.allinger@oenb.at

Melanie Koch
Oesterreichische Nationalbank,
Economic Analysis Division,
melanie.koch@oenb.at

JEL classification
C83; D91; G50

Keywords
Survey experiment; item nonresponse; household income

Simpler questions reduce income nonresponse

Our experiment shows that replacing a long list of more than twenty income brackets with only three broad income categories significantly and substantially reduces income nonresponse in a survey of individuals. The treatment is effective for all four countries examined.

Refusal more common than lack of knowledge

In the CESEE countries studied, people who do not answer the question on exact income overwhelmingly give "no answer" rather than saying they "don't know". Such refusals might be an indicator of privacy concerns when declaring one's income, and our treatment effectively reduces nonresponse in this case.

No positive spillover to personal income question

Simpler, less intrusive questions may increase trust in the survey and interviewer and thus potentially reduce nonresponse to subsequent questions. However, our treatment does not improve the respondents' willingness to answer a subsequent question on exact personal income.

The views expressed are those of the authors and do not necessarily reflect those of the Oesterreichische Nationalbank or the Eurosystem.

Let's Talk Money: Reducing Income Nonresponse in Surveys[a]

Katharina Allinger[b] Melanie Koch[c]

October 2026

Abstract

Item nonresponse is a common issue in surveys. We implement an experiment to reduce nonresponse to income questions in an international household survey, looking at four different countries. Survey respondents are asked to report their exact household income. We randomize those who refuse to answer into two groups. In a follow-up question, the control group is asked to choose their income from a very granular list of brackets, the treatment group receives only three options to choose from. In all four countries, the treatment leads to a significant decrease in nonresponse. Thus, when condensed income data are sufficient, fewer answer options are a cost-effective way to reduce nonresponse.

Keywords: survey experiment; item nonresponse; household income
JEL: C83; D91; G50

[a]We thank Elisabeth Beckmann, Joachim Winter, seminar participants at the DIW Berlin and several anonymous referees for very helpful comments. The paper also benefited from discussions with participants of the SABE/IAREP conference 2024, the EEA annual meeting 2024, the NOeG annual meeting 2024 and the ESRA Conference 2025. Further, we thank Katarina Rohsmann for proofreading the paper. Declarations of competing interest: none. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The opinions expressed do not necessarily reflect the official viewpoint of the Oesterreichische Nationalbank (OeNB) or the Eurosystem. This study uses data from the OeNB Euro Survey. The data and variables to replicate our study will be made available along with the study once it is published. Researchers can apply for free access to a larger dataset including more variables following the steps outlined here: https://www.oenb.at/en/Monetary-Policy/Surveys/OeNB-Euro-Survey/data-sharing.html.

[b]Corresponding Author. Oesterreichische Nationalbank, Otto-Wagner-Platz 3, 1090 Vienna, Austria;
Email: katharina.allinger@oenb.at
[c]Oesterreichische Nationalbank, Otto-Wagner-Platz 3, 1090 Vienna, Austria;
Email: melanie.koch@oenb.at

Non-technical summary

Income is one of the most common and important variables collected in social science surveys. It is used in a multitude of economic analyses either as an outcome, explanatory or control variable. At the same time, income questions are particularly prone to item nonresponse, meaning that survey participants decline to answer income questions. This can reduce the usefulness of survey data and may lead to biased results.

This paper studies whether a simple change to the questionnaire can reduce income nonresponse. We conduct a randomized survey experiment in the 2023 wave of the OeNB Euro Survey in Bosnia and Herzegovina, Bulgaria, Hungary and Poland, four countries with previously high nonresponse. Survey participants are first asked to report the exact amount of their monthly household income. Those who do not provide an answer receive a bracketed follow-up question. In the control group, respondents receive a long list of at least twenty income brackets, as in previous survey waves. In the treatment group, respondents are instead asked whether their income falls into the lower, middle or upper third of the national income distribution.

The treatment substantially reduces income nonresponse. Across all four countries, respondents who receive the simplified question are significantly more likely to provide income information – by between 11 and 24 percentage points. We also examine whether the simplified household income question raises the willingness to respond to a subsequent question on personal income, but find no supporting evidence. Additional analyses suggest that the treatment is particularly effective among respondents who refuse to answer the exact income question, rather than among those who do not know the answer. We also find some evidence that treatment effects differ by gender.

In summary, using very broad income categories can be a simple and cost-effective way to reduce income nonresponse. In surveys where income is mainly a control variable, the benefit of the substantial increase in response rates likely often outweighs the loss in precision. For other surveys, it would be necessary to investigate the trade-off between lower nonresponse and precision – an interesting avenue for further research.

1 Introduction

Surveys can be a convenient and cost-effective way to elicit microeconomic data otherwise not available. However, missing data, in particular nonresponse, are a persistent challenge. One aspect of nonresponse is item nonresponse, which means that persons are willing to participate in the survey but refuse to answer certain questions. The reasons for this are numerous, ranging from deep mistrust in authority to genuinely not knowing the answer to a question. Unsurprisingly, questions about monetary values such as income are especially prone to nonresponse.

Missing income information is a serious issue, given that income plays such an important role in the economic decisions and preferences of individuals. Therefore, missing values can severely compromise research outcomes. In the worst possible case – when income information is missing not at random – it can lead to the exclusion of a large, non-random portion of the sample from the analysis, and thus to biased results.

Indeed, income nonresponse is pervasive across surveys. For instance, Yan et al. (2010) collect income nonresponse figures for a selection of US surveys. In these surveys, income nonresponse to open-ended questions ranges from 14% to 35%. Nonresponse for follow-up bracketed questions ranges from 14% to 19%. Frick and Grabka (2010) note that in the German SOEP (1992-2004), the Australian HILDA (2001-2005) and the British BHPS (1991-2004) item nonresponse for annual labor income is around 14-16%.

Income nonresponse has also been a problem in the OeNB Euro Survey, which has been conducted in nine countries for more than 15 years. In the 2022 survey wave, nonresponse to household income questions ranged from around 10% in Croatia, Czechia and Romania to around 40% in Bulgaria and Bosnia and Herzegovina. Income nonresponse remains high, despite various attempts over time to reduce it.

We conducted an experiment in the 2023 OeNB Euro Survey wave, testing a new approach to reduce income nonresponse. The previous approach worked well in some countries but barely in others. We restrict the experiment to four countries in which nonresponse remained particularly high: Bosnia and Herzegovina, Bulgaria, Hungary

and Poland. In the survey, respondents are initially asked to provide exact amounts on monthly net household income. Until survey wave 2022, respondents who were not willing or not able to provide an exact amount (i.e. non-respondents) were offered to choose an answer from a list of very granular income brackets. In wave 2023, granular brackets are our control group. In contrast, we ask our treatment group if their income falls into the first, second or third income tercile, with boundaries predefined by us. In each country, respondents are randomly assigned to either control or treatment.

We have two pre-registered hypotheses, which are tested independently for each country. First, we expect that income nonresponse is significantly reduced in the group that gets the tercile question, as respondents are more likely to answer questions that are cognitively less demanding and less privacy-invading. Second, we expect spillover effects from our treatment to the answer behavior for the question on exact personal income, which follows the household income question for all respondents who live in multi-person households.

Our results show a significant reduction in nonresponse to household income questions. In comparison to the control groups, the nonresponse shares in the treatment groups are significantly lower in all four countries: by about 24 percentage points in Bosnia and Herzegovina, 11 percentage points in Bulgaria, 19 percentage points in Poland and 20 percentage points in Hungary. Thus, the evidence supports our first hypothesis but not our second. Respondents who receive the tercile question for household income are not more likely to provide an answer regarding the exact amount of their personal income.

Given the large effect sizes and power for hypothesis one, we also conduct supplementary analyses, which were not pre-registered in detail and are more exploratory. They contribute to a better understanding of what drives income nonresponse and how to mitigate it. First, we study whether our treatment is more effective for respondents who refused the exact income question or for those who answered with "don't know". This serves as a proxy for distinguishing between people who are unwilling and those who are unable to answer. The treatment is effective for the first group of people in all countries.

It is reassuring that the treatment is clearly effective for the more important type of nonresponse in our samples. "Don't know" responses were much less common, making our estimates on this issue very imprecise. They are also less consistent across countries. Related to this, in one country, the time spent on the treatment question was significantly shorter than for the control question. This might indicate a lower cognitive burden, which might also have contributed to the effectiveness of the treatment.

Second, we investigate differences in average treatment effects across gender, age and education. With the exception of gender, we find no significant differences in treatment effects. Men respond more strongly to the treatment than women, with the pooled treatment effect being more than ten percentage points larger. Finally, we analyze how the overall income distribution might be affected. There is no clear pattern but in two countries, the distribution between control and treatment clearly differs. This underlines the fact that income nonresponse does not occur randomly.

Literature Item nonresponse is closely linked to several strands of the survey literature which deal with cognitive processes and their implications. For instance, Yan and Williams (2022) review the literature and design a conceptual framework for response burden, which can trigger item nonresponse. Despite the importance of these wider aspects, we narrow our literature review down to the most closely related papers, those focusing on different strategies to reduce income nonresponse. For instance, Moore and Loomis (2001) conduct an experiment in which they test whether a brief statement before the income questions, designed to reassure the respondent, can mitigate the fact that income nonresponse tends to be higher in topic-based household interviews compared to person-based interviews. They find that for nine income items, nonresponse in topic-based interviews with the statement is not significantly different or, in some cases, even significantly lower than in person-based interviews without the statement. Other researchers have focused on the usefulness of bracket questions instead of open-ended questions or as a follow-up to open-ended questions. Bracketing techniques usually reduce nonresponse to monetary questions significantly (Juster and Smith, 1997; Heeringa and Suzman, 1995; Yan et al.,

2010). Juster and Smith (1997) compare standard bracketing techniques with unfolding brackets – where the ranges are obtained via a series of simple yes/no or higher/lower questions. They find that in their sample, unfolding brackets are somewhat more successful in reducing nonresponse than standard range cards. Wang (2010) studies how the switch from a categorical income question to unfolding bracket questions affects income nonresponse in a telephone survey in the US. He reports that, overall, few changes can be observed. However, he documents a decrease in answer refusals and an increase in "don't know" answers. Using brackets clearly also has its challenges. Besides getting less precise income data, it can lead to biases related to the number and distribution of brackets (Winter, 2002; Van Soest and Hurd, 2008). A related issue is whether respondents provide correct monetary values or whether they underreport (e.g., Meyer and Mittag, 2019) or round (e.g., Drechsler and Kiesl, 2016). To our knowledge, our paper is the first that provides experimental evidence on how much income nonresponse is reduced by drastically reducing the number of brackets.

Another strand of literature explicitly analyzes which personal characteristics are related to income nonresponse. For example, using British household surveys, Micklewright and Schnepf (2010) estimate that females have a higher nonresponse rate for household but not individual income, while older and self-employed persons have a higher nonresponse rate for individual but not household income. Our exploratory analyses suggest that men respond more strongly to reducing the number of household income brackets while we find no differences across age and education groups.

In addition, there are several studies investigating whether income nonresponse is related to income itself and whether it is therefore missing not at random (MNAR). One important, early contribution comes from Lillard et al. (1986), whose results point towards the nonignorability of income nonresponse in the Current Population Survey (CPS) and issues with (then) implemented imputation procedures. Some more recent evidence comes from Bollinger and Hirsch (2013) and Bollinger et al. (2019), who also look at the CPS. Bollinger and Hirsch (2013) conclude that high-income men are less likely to report their

income. For women, the relationship between earnings and nonresponse is less clear-cut. Household (co-)heads are more likely to report their wages. Bollinger et al. (2019) find that income nonresponse is highest in the tails of the distribution and that imputation procedures cannot address this problem satisfactorily. They find income nonresponse to be nonignorable. As a result of these and similar papers, several researchers have tried to quantify the potential biases arising from missing income information for important economic topics, such as inequality (Hokayem et al., 2015), poverty (Meyer and Mittag, 2019) or wage gaps (Hirsch and Schumacher, 2004). Most papers use administrative data on top of survey data and find that income nonresponse in surveys and how it is addressed can have serious effects for research outcomes.

Therefore, finding ways to convince individuals to report monetary values stays of utmost importance for social science surveys. There is no one-fits-all solution to reducing income nonresponse. The outcome of a particular set of income questions may depend on many factors, inter alia, the cultural context, the target population, the mode of the survey or the remaining content of the survey (Essig and Winter, 2009; Moore et al., 2000; Schräpler et al., 2010). Nonetheless, our results offer useful insights when researchers and data providers decide how to elicit income in surveys. Finding similar results for four countries lends credibility to our results in terms of external validity.

The remainder of the paper is structured as follows: section 2 describes the setup of the experiment and provides details on income questions and nonresponse in previous OeNB Euro Survey waves. Section 3 outlines the hypotheses and how they are tested. Section 4 shows the results of the experiment. Section 5 includes some additional, exploratory analyses, and section 6 concludes.

2 Data and implementation

2.1 The OeNB Euro Survey

We test our treatment in four different country samples of an international survey of individuals, the OeNB Euro Survey. This survey is conducted annually by the Oesterreichische Nationalbank (OeNB) in nine countries in Central, Eastern and Southeastern Europe. The survey uses multi-stage stratified random route sampling and around 1,000 adult individuals per country are interviewed every year by local survey institutes. Our experiment was implemented in the 2023 wave of the survey. Our sample size was predetermined by the contract governing the conduct of the OeNB Euro Survey.

Questionnaires of the OeNB Euro Survey are ex ante harmonized such that income is elicited in the same manner, except for the currency and the number of brackets used (see appendix table 1).[1] To further ensure comparability, the survey also harmonizes research methodology and fieldwork practice. Some differences remain, e.g. the use of different CAPI software solutions, which affects the implementation of the experimental randomization (see subsection 2.4 and the methodological notes to the survey).

The main topics covered in the survey revolve around household finance issues and respondents are asked to report their household and personal net income. The selection of the individual in the household to be interviewed is random and knowledge about household income may vary. Usually, other household members are not present during the interview and not allowed to help to answer questions. However, around 90% of respondents report that they are in charge of managing household finances alone or jointly with their partner.

[1] The questionnaire is compiled in English by OeNB staff before each wave to maximize comparability across countries. This website provides more information on the questionnaire, including translations.

2.2 Income elicitation and nonresponse in previous waves

Questions regarding the exact amount of household and personal income have been asked in several previous OeNB Euro Survey waves. In the 2023 wave, the same questions for the exact amount (see appendix boxes 1 and 2) and the same order as for the 2022 wave were used. This means that income questions were the last questions respondents had to answer. Household income was asked first and personal income was asked only for multi-person households. The household member interviewed is randomly selected and not necessarily the most knowledgeable person on financial issues.

In all countries, questions on exact income amounts suffer from nonresponse and are often the questions with the highest nonresponse share in the survey. This is in line with results of Jabkowski and Piekut (2024), who report high income nonresponse in the European Social Survey especially in Southern and Eastern Europe. As in most international surveys, nonresponse shares in the OeNB Euro Survey differ considerably across countries. In 2022, nonresponse to exact household income ranged from around 20% in Croatia and Romania to more than 50% in Bulgaria and Bosnia and Herzegovina. The numbers are comparable or even higher for the case of personal income.

In the past, nonresponse has been partially tackled by asking follow-up questions. Instead of exact amounts, respondents are offered to choose their income from a list of income brackets. The number of brackets and the range of each bracket is different for each country. Amounts are guided by administrative income data from the respective countries and survey responses from previous years. Since income ranges were chosen to be relatively granular and the sampled countries generally experience strong nominal wage growth each year, at least twenty different brackets are given. This number of brackets is at the upper end compared to other surveys. Since the 2022 OeNB Euro Survey wave, the follow-up income brackets are asked with an introductory statement, which re-emphasizes that data are treated confidentially and are only used for research purposes. We keep this statement for our treatment and control groups, as it is also used in the countries that are not part of our experiment.

In some countries, the bracket questions decrease income nonresponse markedly – for instance, in Czechia. However, in some countries, the majority of the respondents also refused to answer the income brackets (see figures 2 and 3 in section 4.1). Based on household income nonresponse in past waves, we decided to conduct the experiment in four countries: Bosnia and Herzegovina, Bulgaria, Hungary and Poland. In these countries, nonresponse to exact income amounts is very high and asking about income brackets is a minor or no remedy. This also ensures that we have enough power to test our treatment. The nonresponse shares of exact, and categorical income questions for the four countries in wave 2022 can be seen in table 1.

Table 1: Share of nonresponse to household income questions in wave 2022

	BG		HU		PL		BA	
	Share	N	Share	N	Share	N	Share	N
Exact income	0.56	1009	0.48	1000	0.37	1010	0.53	1000
Categorical income	0.69	568	0.50	480	0.54	377	0.81	531
Both questions	0.39	1009	0.24	1000	0.20	1010	0.43	1000

Source: OeNB Euro Survey, wave 2022.

Note: Share of respondents who did not answer exact household income, who did not answer categorical income (as a share of those that did not answer household income) and joint (final) nonresponse to both questions in the full sample.

Although we also randomize groups for the personal income elicitation, our focus is on the randomization of the household income questions, which is asked first and directed to all respondents. Results on personal income can be found in appendix section D.

2.3 The survey experiment

We test whether reducing the number of brackets lowers income nonresponse. Our control and treatment groups consist of respondents who refused to answer the household income question which asks for exact amounts. Our control group receives the question as asked in previous waves (see box 1). As can be seen in appendix table 1, control respondents are

presented lists with at least 24 income brackets. The smallest brackets in each country capture a range as small as 50€.

Box 1: Control group

We know that people are not used to talking about their income, but we ask these questions to obtain overall statistics about your community and country, and not to obtain information about you personally. Please rest assured that your responses will be treated confidentially and will under no circumstances be used for commercial or marketing purposes.

I am now showing you a card with different amounts. Could you choose the range that best fits the amount of your monthly household income after taxes?

BRACKET LIST	[See appendix table 1 for brackets in each country]
No income	–77777
Don't know	–88888
No answer	–99999

In contrast, the treatment group is presented a much shorter and presumably easier question: whether their household income lies in the pre-defined first, second or third income tercile of a country's overall income distribution (see box 2). Hence, the treatment differs in one key aspect from the control. Instead of having to locate the correct bracket in a list of more than 24 granular brackets, respondents are offered only three brackets. We design the three brackets to approximate income terciles. As there is no way to know the exact future income distribution within a country in advance, we approximate the bounds. To do so, we use the bounds of the data from the previous survey wave but take two additional factors into account.

Box 2: Treatment group

We know that people are not used to talking about their income, but we ask these questions to obtain overall statistics about your community and country, and not to obtain information about you personally. Please rest assured that your responses will be treated confidentially and will under no circumstances be used for commercial or marketing purposes.

Could you tell me if your household income is

Below [Lower bound of the second tercile]	1
Between [Lower bound of the second tercile] and [Lower bound of the third tercile]	2
Above [Lower bound of the third tercile]	3
No income	–77777
Don't know	–88888
No answer	–99999

First, in all countries studied, there is substantial growth in average wages from year to year. For example, growth rates between 2021 and 2022, which were years with elevated inflation rates in all countries, were between 8-14%. We increase previous year bounds by around 10-15% to calculate tercile bounds for 2023. Second, we adjust the resulting values slightly to correspond to the closest available bracket boundary from our control group. This helps to later construct three groups from the control group. It makes direct comparisons more feasible and reliable as bounds are exactly the same. The final tercile bounds are presented in appendix tables 2 and 5. Appendix tables 3 and 6 show that the actual tercile bounds in 2023 – calculated from the exact income question and the control question – are not too far away from the pre-defined values. Appendix tables 4 and 7 show the growth rates of household and personal income from 2022 to 2023.

There is one caveat regarding the exact comparability of our control and treatment question. The bracket question for the control group was left unchanged compared to previous waves. It reiterates that the monthly, after-tax household income should be

stated. This repetition was not included for the treatment group, as the question was simplified as much as possible. If respondents in the treatment group were under the impression that they could report income on any scale (e.g., annual, gross), then this could have led to lower nonresponse. However, both groups were told to report monthly, after-tax income in the preceding exact income question, which harmonizes the framing and scope. Neither the survey institutes nor the interviewers reported that respondents were unsure about the scale in the treatment questions. Moreover, in 2024, the treatment approach was rolled out to all countries, including the "monthly, after-tax" addition in the tercile question. If the reduction in nonresponse between treatment and control in our experiment was driven by this, we would assume to see a sharp and systematic increase in nonresponse in 2024 and 2025 compared to 2023. However, nonresponse to the tercile question was roughly the same as in 2023 or even lower in three of the four countries and somewhat higher in one. Thus, there is little evidence to assume that the omission of the income scale from the treatment question has any consequence for our experiment.

2.4 Technical procedures

As in previous years, the field phase of the OeNB Euro Survey was aligned across countries and took place between 29 September and 9 December 2023. In total, 946 respondents were interviewed in Bosnia and Herzegovina, 1,001 in Bulgaria, 1,000 in Hungary and 1,012 in Poland. All interviews in the OeNB Euro Survey are interviewer-led and conducted face-to-face. Almost all interviews are computer assisted using tablets. In each country, local polling institutes are responsible for conducting the survey, which leads to some differences in technical procedures. For example, the software used to conduct the interviews differs and 22% of the interviews in Poland were conducted paper-based. As noted in the pre-analysis plan, the technique to randomize respondents into treatment and control group differs as well. This is described in more detail in appendix subsection B.1. Nevertheless, in each country, randomization is stratified by interviewer.

3 Hypotheses and empirical approach

We have two main hypotheses regarding the effect of our treatment. The study, including these hypotheses, is registered in the AEA RCT Registry (ID: 0012215). We assume that asking household income in terciles instead of more granular brackets (after respondents already refused to give an exact amount) decreases different kinds of nonresponse. First and foremost, we expect that nonresponse to the question itself is reduced. The first hypothesis H1-0 and its alternative, one-sided hypothesis H1-a are:

H1-0: The share of nonresponse to household income is equal in the treatment and in the control group.

(H1-a: The share of nonresponse to household income is lower in the treatment than in the control group.)

We attribute nonresponse to two main factors: ignorance of the exact amount or refusal to disclose the exact amount for privacy reasons. The assumption that easier questions with lower response burden have lower nonresponse is frequently discussed in the studies in our literature review but has barely been tested rigorously. For both reasons, asking in broader brackets should be a remedy. Respondents are not required to know the exact amount (an approximation is enough) and the interviewer does not learn the exact amount, which preserves at least some privacy. In our sample, "no answer" makes up a much larger share of nonresponse than "don't know" (see the exploratory analyses in section 5). Related to privacy concerns, individuals might not report behavior that is socially undesirable or even illegal, like tax evasion. This might be especially relevant in economies with higher levels of informality (Gërxhani and Schram, 2006). In extreme cases, asking respondents in very broad terms might not reduce nonresponse. Instead, researchers rely on indirect questioning techniques (see, for example, Yu et al., 2008; Korndörfer et al., 2014). Although it is the social norm in several countries we study

not to talk about personal earnings, we do not believe that income constitutes such an extreme case that indirect questioning techniques are necessary.

In addition, we hypothesize that there might be spillover effects to subsequent income questions. Given that respondents from multi-person households are first asked about household and then personal income, they might "learn" from the household question and change their behavior in the subsequent question. On the one hand, people might learn that a simpler question will follow the exact question and nonresponse to the exact personal income question might increase. On the other hand, the tercile question is less sensitive. This might increase trust in the interviewer and the questionnaire. Moreover, the broad brackets may make it clearer to the respondent that a very approximate answer is acceptable, if they are unsure about the true amount of their income. These effects might increase the willingness of respondents to answer the exact personal income question. Juster and Smith (1997) report that respondents who used unfolding brackets for wealth items in earlier parts of the survey are more likely to give answers on exact amounts later in the survey. We also consider positive spillover effects more likely, and hypothesize that the treatment might lower nonresponse in personal income questions.
H2-0: The share of nonresponse to the exact amount of personal income is equal in the treatment and in the control group.

(H2-a: The share of nonresponse to the exact amount of personal income is lower in the treatment than in the control group.)

As we treat each country as independent and look at the country samples separately, we test eight one-sided hypotheses using bivariate and multivariate approaches. Let *treat* be an indicator variable that equals 1 if a respondent receives the tercile questions (treatment group) and 0 if the respondent receives the question with many brackets (control group). Thus, *treat* is missing if a respondent gives an exact amount on household income. For individual $i = 1, .., N_c^{treat}$ in country $c \in \{BA, BG, HU, PL\}$, y_{hi} is the binary

variable that equals 0 if the respondent gives an answer to either the bracket list or the tercile question and 1 if the respondent refuses to answer the question. Analogously, y_{pi} equals 0 if the respondent gives an exact amount on personal income and 1 in case of nonresponse. Please note again that y_{pi} will be missing for single-person households by definition.

Then, for each country c, we perform Fisher's exact test, chi-square tests and two proportions tests for the equality of proportions of y_{hi} by $treat$ and y_{pi} by $treat$.[2] Given that interviewer and regional effects can be strong (see Olbrich et al., 2025) in the OeNB Euro Survey, we additionally run logistic regressions for each country c, in which we control for the region (Z_i) using dummy variables, and for interviewer (J_i) effects using interviewer characteristics (gender, age, education and interview experience)[3]. Moreover, given the imbalances in the covariates, controlling for them in a regression serves as an additional robustness check.

$$\ln \frac{P(y_{hi}=1|treat_i, Z_i, J_i)}{P(y_{hi}=0|treat_i, Z_i, J_i)} = \alpha + treat_i'\beta + Z_i'\gamma + J_i'\delta + \epsilon_i, \quad \forall c \in \{BA, BG, HU, PL\} \tag{1}$$

$$\ln \frac{P(y_{pi}=1|treat_i, Z_i, J_i)}{P(y_{pi}=0|treat_i, Z_i, J_i)} = \alpha + treat_i'\beta + Z_i'\gamma + J_i'\delta + \epsilon_i, \quad \forall c \in \{BA, BG, HU, PL\} \tag{2}$$

We conduct one-sided, bivariate tests. For the regressions, we report two-sided significance levels. It was our intention in the pre-analysis plan to report one-sided significance tests for both, but two-sided significance levels are more common in regressions. If we detect significant differences for the proportions within a country and if the $treat$ coefficient is significant in regressions (1) or (2) respectively, we interpret this as evidence supporting our research hypothesis (H1-a or H2-a) for the specific country. As robustness, we adjust our regression results for multiple hypothesis testing with sharpened q-values as used in Anderson (2008).

[2] Note that the Fisher's exact test is a very conservative test as it does not test if the average treatment effect is zero, but if the treatment is zero for all units.

[3] In the pre-analysis plan, we stated that we would control for interviewer effects using interviewer dummies. However, this approach leads to a drop in observations, given that some interviewers conducted only few interviews and have no variation in the outcome variable. We therefore use interviewer characteristics rather than interviewer dummies.

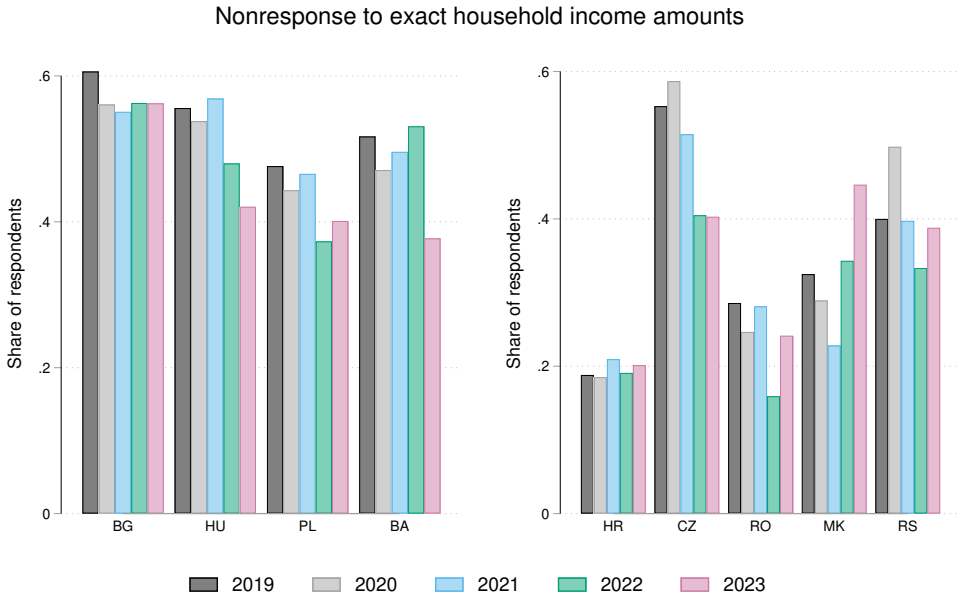
4 Main results

4.1 General descriptive statistics

We first look at the evolution of nonresponse to exact, categorical and both household income questions across all countries of the OeNB Euro Survey over the last years, including wave 2023.

Nonresponse shares for exact household income in 2023 look comparable to previous waves – see figure 1. In the countries included in the experiment (the left panel), nonresponse shares are roughly the same as the 2019-2022 average in Bulgaria and Poland and a bit lower in Hungary and Bosnia and Herzegovina. In Bosnia and Herzegovina, we see the largest drop compared to a year ago. Still, nothing suggests that there are significant changes in trends in general response behavior that interfere with our experiment.

Figure 1: Respondents who refuse to report the exact amount of household income

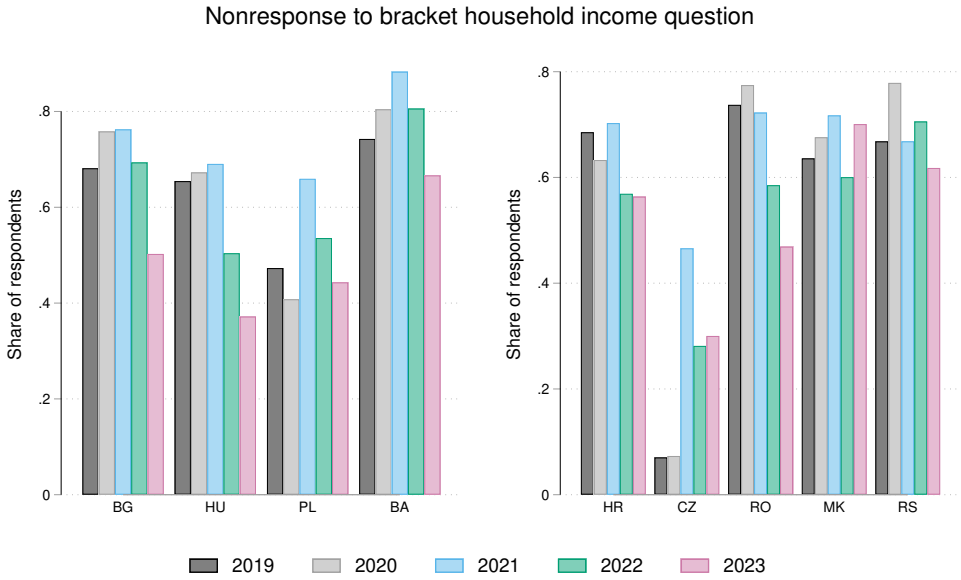


Source: OeNB Euro Survey, waves 2019–2023.
Note: 2021 might not be comparable as income was asked differently.

For the categorical income questions (control and treatment group for wave 2023) in figure 2, the aggregate shares of nonresponse have declined in 2023 in all countries included in the experiment (see again left panel).

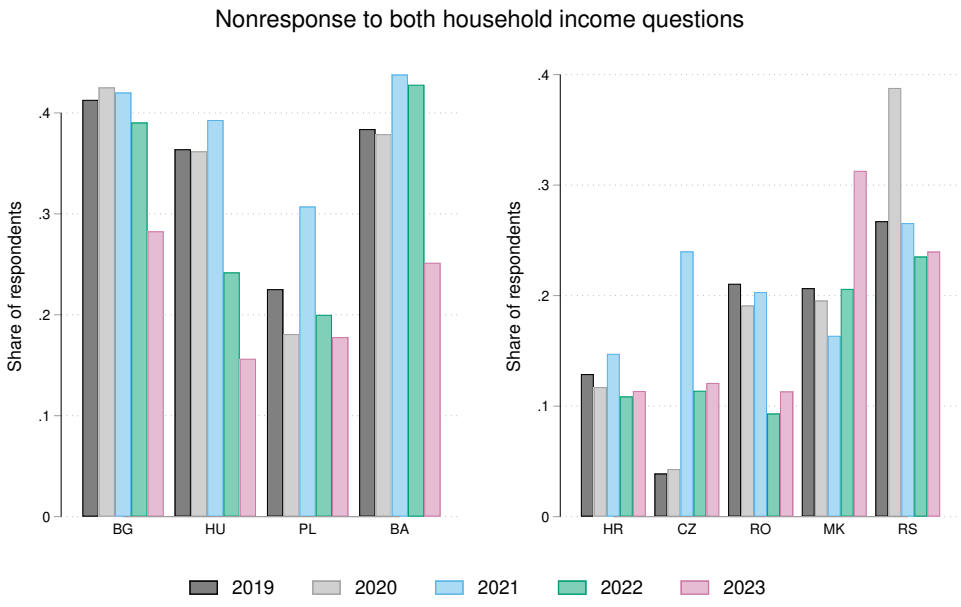
Figure 3 shows overall income nonresponse to both questions in the full sample. In most countries that are not part of the experiment, joint nonresponse stayed roughly unchanged in 2023 compared to previous waves.

Figure 2: Respondents who refuse to report household income in categories



Source: OeNB Euro Survey, waves 2019–2023.
Note: 2021 might not be comparable as income was asked differently. In BG, HU, PL and BA, respondents are randomized into receiving two different kinds of bracket questions (many vs few) in 2023.

Figure 3: Respondents who did not answer any income questions



Source: OeNB Euro Survey, waves 2019–2023.
Note: 2021 might not be comparable as income was asked differently.

Having found that overall nonresponse strongly decreased only in the countries that are part of the experiment, we continue to look at experimental results in more detail.

4.2 Descriptive statistics for the treatment groups

In Hungary, respondents are assigned to either control or treatment group if they do not answer the question on exact income amounts. In the three other countries, all respondents are assigned to either control or treatment group because treatment assignment is determined in the beginning of the survey. We calculate some descriptive statistics for two different samples in these countries. In the main text, we present statistics on the sample that refuse to give an amount for household income and then receive either the control or the treatment question. In subsection B.2 in the appendix, we show the same tables for all respondents interviewed.

Table 2 depicts the share of respondents (who refuse to disclose their exact income) by assigned treatment. The shares are relatively balanced, with half of the respondents assigned to control and the other half to treatment. The only exception is Bosnia and Herzegovina, where the shares are 55% and 45% respectively. Beckmann and Koch (2025) note that probably the randomization procedure in Bosnia and Herzegovina leads to these imbalances as the randomization is not fully computerized which can lead to stronger variation in treatment assignment by chance (see appendix subsection B.1).

Table 2: Distribution of treatment groups by country

	BG	HU	PL	BA
Control: many brackets	0.49	0.48	0.52	0.55
Treatment: terciles	0.51	0.52	0.48	0.45
Observations	562	414	406	357

Source: OeNB Euro Survey, wave 2023.

Note: Share of respondents who are assigned to control and treatment group respectively. Control is the group in which respondents are asked about income in many brackets and treatment is the group in which they are asked in terciles.

In table 3, we show mean comparisons for selected individual characteristics across treatment groups. Based on previous literature, these are the characteristics likely important for predicting nonresponse to income questions (Jabkowski and Piekut, 2024).

Table 3: Differences between control and treatment group

	BG	HU	PL	BA
Gender (f/m)	−0.02	0.07	0.03	−0.04
Age (in years)	0.33	−2.10	−0.59	−2.29
Education (in categories)	0.17	−0.22	−0.05	0.01
Main earner in household (0/1)	−0.02	0.10*	0.05	−0.03
Manages household finances (0/1)	0.05	−0.06*	−0.02	0.04
Size of Household	−0.02	0.25*	0.22*	−0.03
Apprehensive before survey (0/1)	0.03	0.00	−0.03	0.03
Observations	562	414	406	357

Source: OeNB Euro Survey, wave 2023.

Note: Difference between control and treatment group for each variable and country. Positive (negative) numbers indicate that the value is larger (smaller) for the control group than for treatment group. Control is the group in which respondents are asked about income in many brackets and treatment is the group in which they are asked in terciles.

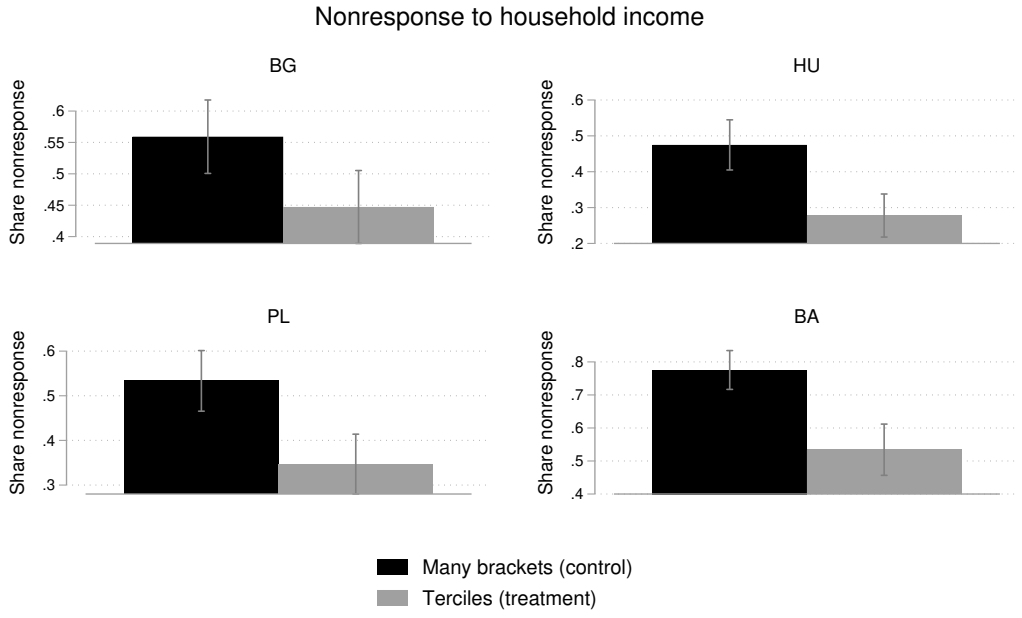
* $p < .05$, ** $p < .01$, *** $p < .001$

Sociodemographic variables like gender, age, education, whether the respondent is the main earner of the household, whether the respondent is involved in managing household finances and whether the respondent was apprehensive before the survey are mostly well balanced between the two groups. In Poland, there is an imbalance with respect to the number of household members. This is also the case in Hungary, where there are further imbalances with respect to being the main earner and person to manage household finances. An F-test on joint orthogonality shows that all characteristics together predict treatment assignment (p=0.059). We do not have any reason to believe that the computerized randomization in Hungary did not work properly but we will control for these covariates in our regression analysis. The F-test for Poland is not significant (p=0.738).

4.3 Treatment effects

First, we look at the effect the treatment has on nonresponse in household income. As a visual aid, figure 4 shows the nonresponse shares for the control and treatment groups, including confidence intervals based on t-tests. However, as specified in the pre-analysis plan, we use three different tests to see if the shares are significantly different from each other. Table 4 presents the p-values for these one-sided tests. In all four countries, the share of nonresponse for the tercile treatment question is significantly lower than for the control group under all three tests. The magnitude of the treatment effects is impressive and economically large: the nonresponse share is 11 percentage points smaller in Bulgaria, 19 percentage points in Poland, 20 percentage points in Hungary, and 24 percentage points in Bosnia and Herzegovina.

Figure 4: Mean comparison household income nonresponse between brackets and terciles



Source: OeNB Euro Survey, wave 2023.
Note: Bars indicate the share of people who refused to report their household income under control and treatment respectively. Confidence intervals are obtained using two–sided t–tests.

Table 4: Household income – p-values for smaller nonresponse shares in treatment group

	BG	HU	PL	BA
One-sided Fisher's exact	0.004	<0.001	<0.001	<0.001
Chi-squared	0.007	<0.001	<0.001	<0.001
One-sided two proportions	0.003	<0.001	<0.001	<0.001

Multivariate regressions in table 5 also show significant average treatment effects. Columns (1), (3), (5) and (7) show the unconditional treatment effects, while we control for region, interviewer characteristics and sociodemographic variables in columns (2), (4), (6) and (8). In all estimations, treatment effects are highly significant, even when adjusting for multiple hypothesis testing. On average, the share of nonresponse is reduced between 11 and almost 23 percentage points.

We reject H1-0 and thus find support for our hypothesis H1-a that the treatment lowers refusal rates. It should also be noted that we have sufficient power and all the estimated effects exceed the minimum detectable changes we calculated in our pre-analysis plan (see also subsection B.3 in the appendix).

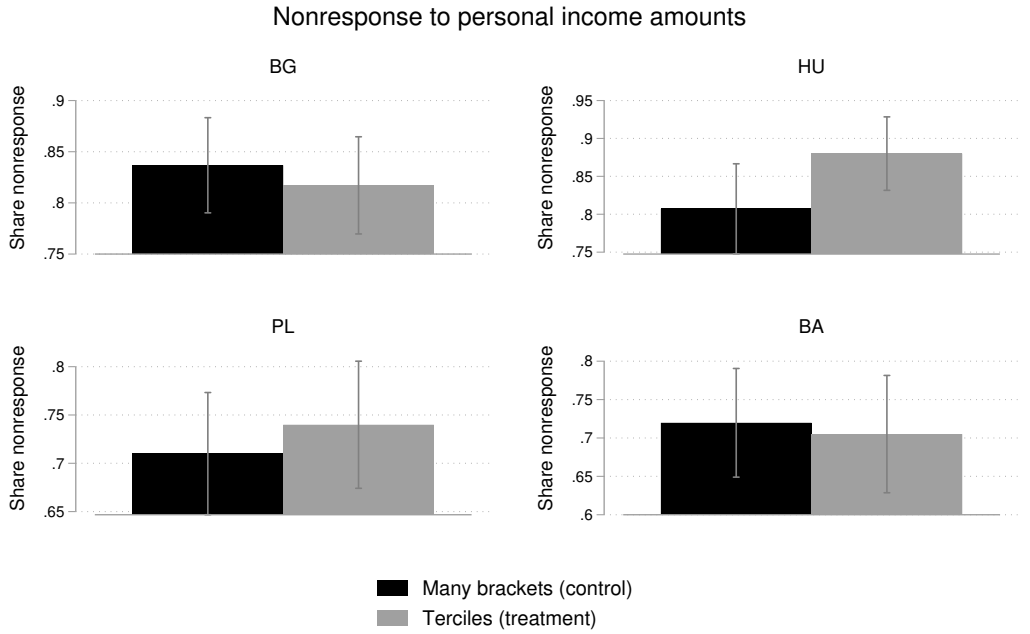
Table 5: Household income nonresponse share – logit regressions with controls

	BG		HU		PL		BA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment: terciles	−0.111**	−0.119**	−0.191***	−0.193***	−0.182***	−0.140**	−0.229***	−0.238***
	(0.039)	(0.041)	(0.040)	(0.039)	(0.042)	(0.040)	(0.047)	(0.046)
DK/NA control	0.56	0.56	0.47	0.47	0.53	0.53	0.78	0.78
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-386.7	-348.9	-264.6	-239.3	-271.6	-221.7	-215.6	-168.6
McFaddens R^2	0.01	0.05	0.03	0.12	0.03	0.13	0.05	0.23
Observations	563	532	414	414	406	372	357	341

* p < .05, ** p < .01, *** p < .001. To adjust for multiple hypothesis testing, sharpened q-values were calculated (two sets of eight hypotheses: hypothesis 1 and 2 without controls as well as hypothesis 1 and 2 with controls). q-values lie between 0.001 and 0.01 for all columns.
DK/NA control shows the nonresponse share in the control group. Control variables include age, gender, education, main earner in household, manages household finances, size of household, whether respondent was apprehensive before the survey, region dummies, gender of interviewer, age of interviewer, interviewer education and interviewer experience.

Next, we want to see if the treatment on household income has some spillovers on the willingness to answer the question on exact personal income amounts. We again run the three bivariate tests. As suggested by figure 5, the treatment does not have the intended spillover effects on exact personal income. The share of nonresponse is mostly not significantly different (see table 6). In addition, effect sizes are very small, being below 2 percentage points in Bulgaria and Bosnia and Herzegovina, and 3 percentage points in Poland. An exception is Hungary, where it is 7 percentage points in the specification without control variables. Like in Poland, the effect runs in the opposite direction, and the share of nonresponse after respondents answered the tercile question is higher than the share of nonresponse after the many brackets question. However, the effect is only significant under the one-sided Fisher's exact test. The overall insignificant results might either be because respondents do not change their response behavior – after all, we ask only one tercile question – or because the theoretical effects discussed in section 3 offset each other.

Figure 5: Mean comparison personal income nonresponse between brackets and terciles



Source: OeNB Euro Survey, wave 2023.
Note: Confidence intervals are obtained using two–sided t–tests.

Table 6: Personal income – p-values for smaller nonresponse shares in treatment group

	BG	HU	PL	BA
One-sided Fisher's exact	0.330	0.042	0.299	0.440
Chi-squared	0.576	0.061	0.520	0.780
One-sided two proportions	0.288	0.969	0.740	0.390

Regression results in table 7 provide further evidence that the treatment effect is insignificant and the effect size small in all countries. To put the coefficient sizes into perspective, the standard deviation of the nonresponse share for the exact personal income question between 2019 and 2022 is 1.8 in Bulgaria, 8.8 in Hungary, 5 in Poland and 2.8 percentage points in Bosnia and Herzegovina. Thus, the results do not support our hypothesis 2a that there is a positive spillover effect on the willingness to give an exact amount for personal income.

Table 7: Personal income nonresponse share – logit regressions with controls

	BG		HU		PL		BA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment: terciles	−0.020	−0.016	0.073	0.035	0.030	0.028	−0.015	−0.001
	(0.033)	(0.030)	(0.039)	(0.034)	(0.049)	(0.048)	(0.044)	(0.045)
DK/NA control	0.84	0.84	0.81	0.81	0.71	0.71	0.72	0.72
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-231.3	-166.1	-148.1	-110.0	-219.6	-173.8	-177.4	-134.9
McFadden's R^2	0.00	0.22	0.01	0.27	0.00	0.15	0.00	0.21
Observations	502	466	346	346	373	336	296	281

* $p < .05$, ** $p < .01$, *** $p < .001$. To adjust for multiple hypothesis testing, sharpened q-values were calculated (two sets of eight hypotheses: hypothesis 1 and 2 without controls as well as hypothesis 1 and 2 with controls). q-values lie above 0.05 in all columns.

DK/NA control shows the nonresponse share in the control group. Control variables include age, gender, education, main earner in household, manages household finances, size of household, whether respondent was apprehensive before the survey, region dummies, gender of interviewer, age of interviewer, interviewer education and interviewer experience.

In summary, asking income in tercile bounds instead of brackets reduces overall nonresponse considerably. Extrapolating from our sample of 1,000 respondents per country, applying the treatment to all respondents in future waves could generate income infor-

mation for around 55–85 additional respondents per country compared to a control group approach – assuming that overall nonresponse to the exact household income question stays unchanged. In contrast, our treatment does not have any positive spillover effects on subsequent questions on exact income amounts.

5 Heterogeneous treatment effects

Given the large effect sizes and power for hypothesis one, we are able to conduct analyses supplementary to the questions outlined in the pre-analysis plan. These analyses, although not pre-registered in detail, were mentioned in the plan. They provide insights into what drives income nonresponse and how to mitigate it.

5.0.1 Not knowing the answer or not willing to answer

The natural starting point is to understand why respondents do not give an exact income amount and which reason for refusal is the more prevalent: not knowing the answer or not wanting to answer. While we cannot say for sure why people chose not to respond, we do have a reasonable proxy. The survey allows respondents to choose between "Don't know" or "No answer". It seems plausible that people who lack knowledge to answer the question would pick the former, while people who want to refuse would pick the latter – but we cannot test this assumption. As previously mentioned, the treatment should be a remedy for both motives, but it is worth exploring if one group responds more strongly than the other.

Table 8 shows the distribution of answers by country. "No answer" was the more common response in all countries. This is in line with privacy concerns as a main determinant of nonresponse for income questions and with Jabkowski and Piekut (2024). They find for the European Social Survey that refusals are high in Czechia, Hungary and Poland, while "don't know" is more common in several Northern and Western European countries. Similar to their paper, we find that "Don't know" responses are more common among

respondents living in larger households, in line with the argument of task complexity, but unlike their paper, there is no significant correlation between trust in other people and "Don't know" versus "No answer" responses.

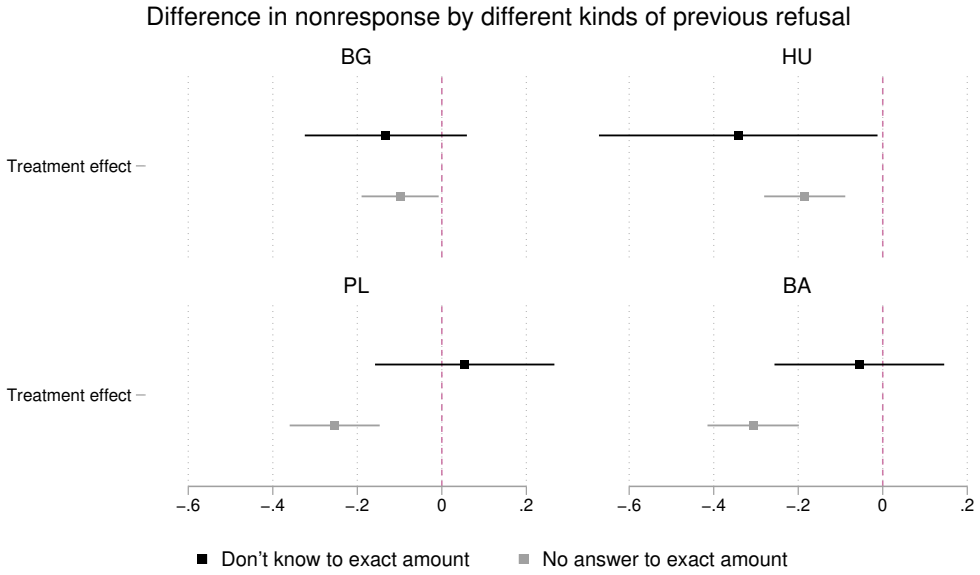
Table 8: Types of nonresponse to exact household income in wave 2023

	BG		HU		PL		BA	
	n	Share	n	Share	n	Share	n	Share
Don't know	103	0.18	33	0.08	88	0.22	93	0.26
No answer	459	0.82	381	0.92	318	0.78	264	0.74
No answer and don't know	562	1.00	414	1.00	406	1.00	357	1.00

Source: OeNB Euro Survey, wave 2023.

Figure 6 shows the results of t-tests on the difference between treatment and control group by country and type of nonresponse. The estimates are more precise for the respondents unwilling to give an answer, as expected given the higher number of observations. For this group, the differences are significant in all countries.

Figure 6: Average treatment effects by type of nonresponse

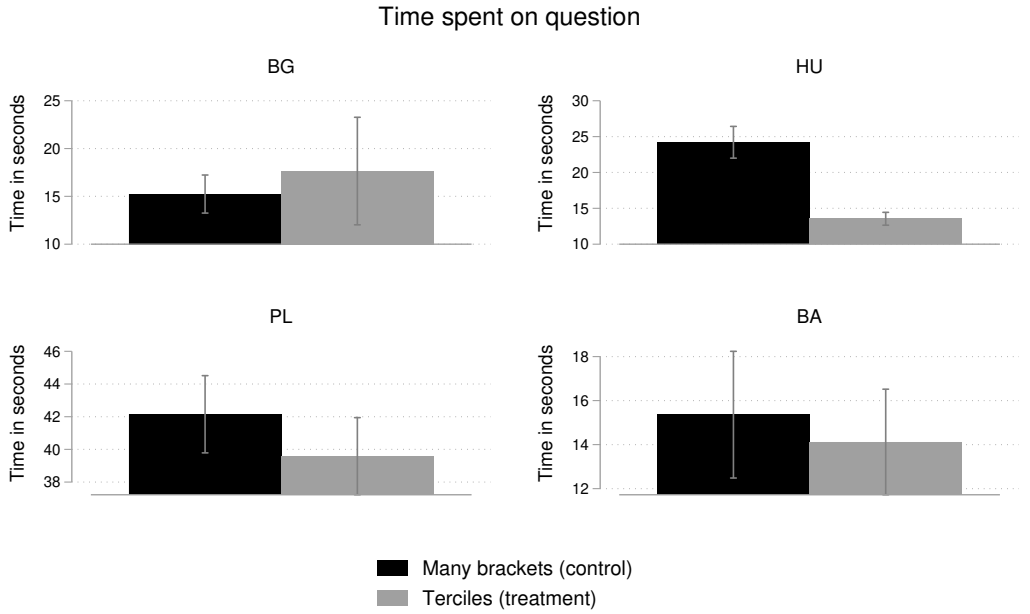


Source: OeNB Euro Survey, wave 2023.
Note: Graphs plot the estimates from t–tests on the difference in nonresponse between treatment and control, separately for those who previously answered "don't know" or "no answer" when asked about the exact income amount. Negative coefficients indicate that nonresponse was lower in the treatment group.

The average treatment effect for "don't know" respondents is significant only in Hungary – but it is based on very few observations. In Poland and Bosnia and Herzegovina, the difference in the treatment effect between types of nonresponse is rather large. It suggests that the treatment is more effective for those unwilling to answer.

One additional aspect we study is the time respondents spend on the control or treatment question. Spending less time on a question can be an indicator that the question is less cognitively demanding. It seems plausible that especially respondents who are fatigued or impatient, would assess the time it takes to answer a question and decide to skip very cognitively demanding ones by answering "don't know." Figure 7 illustrates that in three out of four countries, the time spent on the tercile question is shorter than on the bracket question, while in Bulgaria it is longer. Hungary has the largest and the only significant difference, which is in line with the significant treatment effect for "don't know" respondents observed above. Therefore, at least in one country, cognitive load might provide an additional explanation for the effectiveness of our treatment.

Figure 7: Time spent (in seconds) on control or treatment question



Source: OeNB Euro Survey, wave 2023.
Note: Confidence intervals are obtained using two−sided t−tests. BA is not perfectly comparable because the time partially includes the whole time spent on the income questions.

5.0.2 Heterogeneous effects by sociodemographic characteristics

We further study whether the effects of the treatment differ by sociodemographic characteristics which were shown to be correlated with (income) nonresponse. The most studied characteristic is gender, with studies finding that women are less likely to respond to income questions (Yan et al., 2010; Schräpler et al., 2010; Juster and Smith, 1997). This is not the case in our survey, as in three out of four countries, men are slightly less likely to respond. Yan et al. (2010), Riphahn and Serfling (2005), Jabkowski and Piekut (2024) investigate age effects, reporting differing results. Similarly, these studies also find ambiguous effects for education. Differences across studies might also be the result of different ways to elicit income. Our results show for all four countries that persons below 55 and persons who have tertiary education are less likely to answer the open-ended income question than older persons and persons without tertiary education.

We test whether our treatment works especially well for some sub-groups. Given the low number of observations when the sample is further split, we show the results pooled over the countries. Even pooled, heterogeneous treatment effects may be difficult to detect given insufficient power. According to table 9, the treatment is more effective for men than women. However, the difference between genders is only weakly significant and disappears if either Hungary or Bulgaria is excluded from the pooled sample – this may partially be the case due to a decrease in power. Younger people have a higher effect size, but the difference to older people is not significant. People with tertiary education have a similar effect size as those without. When we run regressions by country, effects always go in the same direction as for the pooled sample, except for education – but none of the differences in treatment effects between sociodemographic groups are significant.

Table 9: Average treatment effect for different sociodemographic groups

	Gender			Age			Tertiary education		
	Male	Female	Diff.	<55 years	>=55	Diff.	No	Yes	Diff.
Treatment: terciles	-0.240***	-0.137***		-0.207***	-0.144***		-0.185***	-0.182***	
	(0.034)	(0.032)		(0.029)	(0.040)		(0.028)	(0.045)	
Difference			-0.103*			-0.063			-0.002
			(0.048)			(0.050)			(0.052)
Observations	803	936	1739	1120	619	1739	1258	478	1736

* p < .05, ** p < .01, *** p < .001

Source: OeNB Euro Survey, wave 2023.

Note: The sample consists of respondents who said "don't know" or "no answer" when asked about exact household income. Estimates are obtained via linear regressions for each sub-sample pooled over all four countries. Differences between groups and their significance are calculated using STATA's *nlcom* command.

The appendix includes regression tables similar to tables 5 and 7 but pooled for all countries and including interaction effects of the treatment with age, gender and education. For hypothesis 1 the conclusions are unchanged to those from table 9, while for hypothesis 2 the variables and their interactions are insignificant.

5.0.3 Effects on the overall income distribution

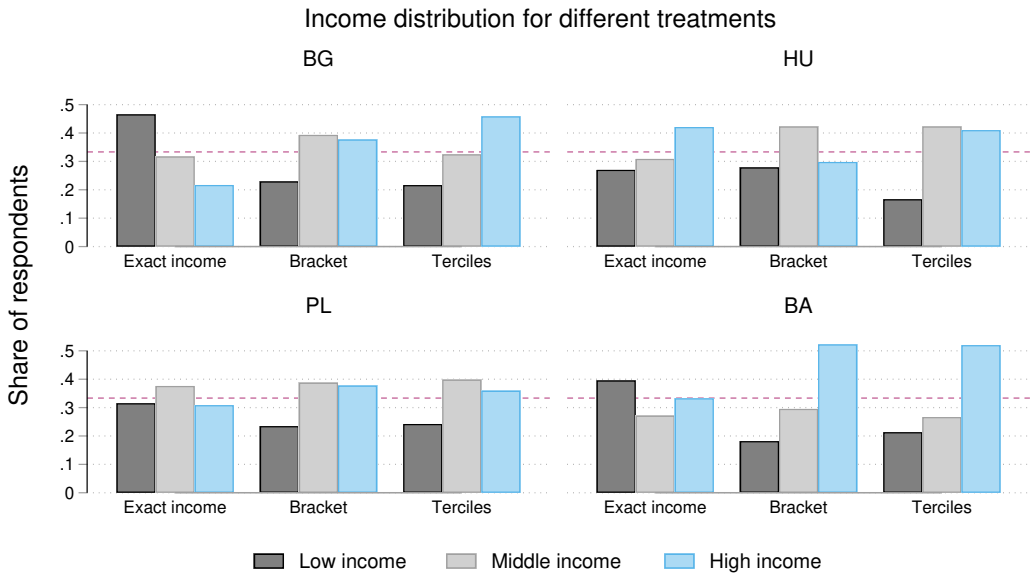
Finally, we study how nonresponse depends on income itself. We compare the resulting income distributions between the question on exact household income, the control group with many brackets and the treatment with terciles. For this purpose, we convert information from all three questions into terciles.

Figure 8 shows that in all countries, the distribution of terciles differs between the three groups. The share of low income respondents is higher for the exact income question than for the terciles in all countries. It is also higher than for the brackets in all countries except Hungary. This points towards the well-known issue that individuals with top incomes are often underrepresented in surveys due to a combination of unit and item nonresponse

(Bollinger and Hirsch, 2013; Bollinger et al., 2019; Ledić et al., 2024). It highlights the importance of a follow-up question on exact income to reduce asymmetric nonresponse across the income distribution.

Between the tercile and bracket question, the distributions are clearly distinct in two out of four countries. Bulgaria shows a higher share of high income versus middle income respondents for the tercile question. Hungary shows a higher share of high income and a lower share of low income for the tercile question. It seems that in these countries, the tercile question may not only reduce nonresponse, but simultaneously affect the collected distribution. The observed pattern suggests that the tercile question might motivate high income people more than the bracket question to provide an answer.

Figure 8: Income distribution in terciles for exact income, control group and treatment group



Source: OeNB Euro Survey, wave 2023.
Note: Bars indicate the share of people in each pre–defined income tercile when using only information on exact household income, only for the control group (many brackets) and only for the treatment group (terciles). Necessarily, only those who eventually provided an income are included.

In line with the above discussion, the overall income distribution shifts a little from low to medium or high income in all countries. The strongest differences occur in Bulgaria, where the distribution with only exact income is strongly skewed towards the lower income tercile. In the other countries, the differences in the tercile distributions are small and

indicate a balanced distribution of the terciles.

So far, we have focused on terciles. What if we wanted to use exact income? In this case, the treatment is associated with a trade-off between lower nonresponse and lower precision of the obtained categorical information. If we tried to impute missing data, it is *a priori* unclear whether asking only brackets or only terciles would lead to a better approximation of the true, continuous income distribution, and less biased estimates of relationships of interest. This point is discussed in D'Elia and Martelli (2003), who use Italian survey data for non-experimental simulations and show that the optimal bracket size to accurately estimate average income indicators lies between 5 and 9. Longmuir and Grabka (2024) estimate the optimal number of brackets to predict the true underlying distribution of income and wealth variables using the G-SOEP. They conclude that 5-7 brackets are optimal, thus three would be not enough. However, in their simulations, the share of nonresponse is taken as given, meaning the trade-off between more responses but less precision is not considered.

The severity of this trade-off also depends on the statistic of interest. For headcount measures such as poverty rates, the loss from coarser brackets seems small, provided the cut-offs are chosen accordingly. For inequality measures, three categories are clearly insufficient. These measures usually depend on the shape of the distribution, and with three groups the result is largely determined by assumptions about the within-bracket distribution. However, the tercile question appears to recover relatively more high-income respondents, who are typically underrepresented. This in turn could improve distributional statistics.

Empirically evaluating the trade-off between reduced nonresponse and less precision is beyond the scope of this paper, but a great avenue for future research. Nonetheless, we make some general remarks on the difficulties associated with studying this aspect. The most promising avenue for an empirical analysis is probably simulations – which always rest on a number of assumptions. One of the most important assumptions would be whether income is missing at random (MAR) or not at random (MNAR). In the former

case, imputations would reduce the bias in a hypothetical, estimated relationship between income and another variable. In the latter case, imputation could increase the bias. There are many reasons to assume that income in our data is MNAR. Several studies find that income nonresponse is related to income itself (e.g., Lillard et al., 1986; Riphahn and Serfling, 2005; Bollinger and Hirsch, 2013; Bollinger et al., 2019), like we do in at least two of our countries. This would make the use of imputations questionable and the results of simulations difficult to interpret.

Moreover, the outcome of any simulation would depend on the relationship of interest investigated. Given that income is an important control or explanatory variable for many questions of interest, it would be difficult to pick a topic for a simulation exercise. Outcomes would also likely depend on how well-specified both the outcome and imputation model are. Finally, one should be aware that imputation with multiple chained equations is focused on identifying the correct relationships between variables. Predictions may very well be wrong on the individual level.

6 Conclusion

Item-nonresponse in surveys is common and poses challenges for statistical inference and representativeness of the data. Especially questions about monetary values, like income, are prone to nonresponse. We use a randomized experiment to test if simplifying income elicitation reduces income nonresponse in an international survey. We look at four countries of the survey, which exhibited income refusal rates between 20-40% in previous survey waves. The previous approach for respondents who did not provide exact income amounts was to ask income using a list of very granular numerical brackets and serves as our control group. The treatment group, in contrast, is simply asked if their income falls into the first, second or third tercile of a country's household income distribution.

The treatment significantly reduces the nonresponse to household income. In all of the countries, refusal rates decrease by at least 11 percentage points (the average across

countries is about 18 percentage points). We do not find evidence for positive spillover effects on the willingness to subsequently answer questions about exact personal income. In exploratory analyses, the treatment works especially well for persons who refuse to report an exact income amount in comparison to those who say that they do not know the exact amount, and for men in comparison to women. These heterogeneous effects, in combination with the resulting overall income distributions being different between treatment and control group in at least two countries, highlight that income nonresponse is not random.

Our results are informative for other institutions and researchers conducting surveys. Clearly, our approach of aggregating information in exchange for less nonresponse does not work for all surveys. Its usefulness crucially depends on the purpose of the collected data. If income information is not one of the main objectives of the survey but serves as control variable, it can suffice to group households into low, middle and high income. Our study shows that in such cases, it is worthwhile to consider asking for income with three answer options.

Even if exact income information must necessarily be obtained, asking income terciles as follow-up question could be a more fruitful approach in combination with multiple imputation routines than using more granular brackets. The implications of the trade-off between getting less precise data but less overall nonresponse for imputation is an empirical question which is beyond the scope of this paper. Given the substantial reduction of nonresponse achieved by asking income in terciles, this question is worth exploring.

Another natural avenue for future research is to look for a middle ground between three and more than twenty categories, by testing an intermediate number of brackets or unfolding brackets more thoroughly. We would caution, however, that part of the appeal of three categories lies in the question format rather than the number of categories as such: three options can be administered as a single question without a show card. Having to present a list likely increases response burden.

References

Anderson, Michael L., 2008, "Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, *103* (484), 1481–1495. https://doi.org/10.1198/016214508000000841.

Beckmann, Elisabeth and Melanie Koch, 2025, "Perceived sustainable minimum wages - a test of randomization and question scales across countries." *Mimeo.*

Bollinger, Christopher R. and Barry T. Hirsch, 2013, "Is earnings nonresponse ignorable?" *Review of Economics and Statistics*, *95* (2), 407–416. https://doi.org/10.1162/REST_a_00264.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak, 2019, "Trouble in the tails? What we know about earnings nonresponse 30 years after Lillard, Smith, and Welch." *Journal of Political Economy*, *127* (5), 2143–2185. https://doi.org/10.1086/701807.

D'Elia, Enrico and Bianca Martelli, 2003, "Estimation of Households Income from Bracketed Income Survey Data." *Documento di Lavoro ISAE*, (35).

Drechsler, Jörg and Hans Kiesl, 2016, "Beat the heap: An imputation strategy for valid inferences from rounded income data." *Journal of Survey Statistics and Methodology*, *4* (1), 22–42. https://doi.org/10.1093/jssam/smv032.

Essig, Lothar and Joachim K. Winter, 2009, "Item non-response to financial questions in household surveys: An experimental study of interviewer and mode effects." *Fiscal Studies*, *30* (3-4), 367–390. https://doi.org/10.1111/j.1475-5890.2009.00100.x.

Frick, Joachim R. and Markus M. Grabka, 2010, "Item nonresponse and imputation of annual labor income in panel surveys from a cross-national perspective." *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, pp. 355–372. https://doi.org/10.1002/9780470609927.ch19.

Gërxhani, Klarita and Arthur Schram, 2006, "Tax evasion and income source: A comparative experimental study." *Journal of Economic Psychology*, *27* (3), 402–422. https://doi.org/10.1016/j.joep.2005.08.002.

Heeringa, Steven G. and Richard Suzman, 1995, "Unfolding brackets for reducing item nonresponse in economic surveys." *PSID Technical Series Paper, No. 95-01.* https://psidonline.isr.umich.edu/publications/Papers/tsp/1995-01_Reducing_Item_Nonresponse.pdf.

Hirsch, Barry T. and Edward J. Schumacher, 2004, "Match bias in wage gap estimates due to earnings imputation." *Journal of Labor Economics*, *22* (3), 689–722. https://doi.org/10.1086/383112.

Hokayem, Charles, Christopher Bollinger, and James P Ziliak, 2015, "The role of CPS nonresponse in the measurement of poverty." *Journal of the American Statistical Association*, *110* (511), 935–945. https://doi.org/10.1080/01621459.2015.1029576.

Jabkowski, Piotr and Aneta Piekut, 2024, "Between task complexity and question sensitivity: nonresponse to the income question in the 2008–2018 European Social Survey." *Survey Research Methods*, *18* (2), 113–135. https://doi.org/10.18148/srm/2024.v18i2.8117.

Juster, F. Thomas and James 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. https://doi.org/10.1080/01621459.1997.10473648.

Korndörfer, Martin, Ivar Krumpal, and Stefan C Schmukle, 2014, "Measuring and explaining tax evasion: Improving self-reports using the crosswise model." *Journal of Economic Psychology*, *45*, 18–32. https://doi.org/10.1016/j.joep.2014.08.001.

Ledić, Marko, Ivica Rubil, and Ivica Urban, 2024, "Using tax records to correct for underrepresentation of top income sources in surveys." *Fiscal Studies*, *45* (4), 521–541. https://doi.org/10.1111%2F1475-5890.12363.

Lillard, Lee, James P. Smith, and Finis Welch, 1986, "What do we really know about wages? The importance of nonreporting and census imputation." *Journal of Political Economy*, *94* (3, Part 1), 489–506. https://doi.org/10.1086/261386.

Longmuir, Maximilian and Markus Grabka, 2024, "How many brackets should we ask for to derive adequate metric information for income and wealth?" *Survey Research Methods*, *18* (3), 251–261. https://doi.org/10.18148/srm/2024.v18i3.8187.

Meyer, Bruce D. and Nikolas Mittag, 2019, "Using linked survey and administrative data to better measure income: Implications for poverty, program effectiveness, and holes in the safety net." *American Economic Journal: Applied Economics*, *11* (2), 176–204. https://doi.org/10.1257/app.20170478.

Micklewright, John and Sylke V. Schnepf, 2010, "How reliable are income data collected with a single question?" *Journal of the Royal Statistical Society Series A: Statistics in Society*, *173* (2), 409–429. https://doi.org/10.1111/j.1467-985X.2009.00632.x.

Moore, Jeffrey C. and Laura Loomis, 2001, "Reducing income nonresponse in a topic-based interview." *Research Report Series*, *2002* (06). https://www.census.gov/library/working-papers/2002/adrm/rsm2002-06.html (last accessed 28 May 2025).

Moore, Jeffrey C., Linda L. Stinson, Edward J. Welniak et al., 2000, "Income measurement error in surveys: a review." *Journal of Official Statistics*, *16* (4), 331–362. https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/income-measurement-error-in-surveys-a-review.pdf (last accessed 28 May 2025).

Olbrich, Lukas, Elisabeth Beckmann, and Joseph W. Sakshaug, 2025, "Multivariate assessment of interviewer errors in a cross-national economic survey and the role of fieldwork institutes." *Journal of the Royal Statistical Society Series A: Statistics in Society*, *qnaf006*. https://doi.org/10.1093/jrsssa/qnaf006.

Riphahn, Regina T. and Oliver Serfling, 2005, "Item non-response on income and wealth questions." *Empirical Economics*, *30*, 521–538. https://doi.org/10.1007/s00181-005-0247-7.

Schräpler, Jörg-Peter, Jürgen Schupp, Gert G. Wagner et al., 2010, "Changing from PAPI to CAPI: introducing CAPI in a longitudinal study." *Journal of Official Statistics*, *26* (2), 239–269. https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/changing-from-papi-to-capi-introducing-capi-in-a-longitudinal-study.pdf (last accessed 28 May 2025).

Van Soest, Arthur and Michael Hurd, 2008, "A test for anchoring and yea-saying in experimental consumption data." *Journal of the American Statistical Association*, *103* (481), 126–136. https://doi.org/10.1198/016214507000000266.

Wang, Kevin, 2010, "Effects of unfolding brackets on the quality of income data in a telephone survey." *Survey Practice*, *3* (3). https://doi.org/10.29115/SP-2010-0014.

Winter, Joachim, 2002, "Bracketing effects in categorized survey questions and the measurement of economic quantities." *Sonderforschungsbereich 504, Discussion Paper*, *02-35*. https://epub.ub.uni-muenchen.de/19729/ (last accessed 28 May 2025).

Yan, Ting and Douglas Williams, 2022, "Response burden–review and conceptual framework." *Journal of Official Statistics*, *38* (4), 939–961. https://doi.org/10.2478/jos-2022-0041.

Yan, Ting, Richard Curtin, and Matthew Jans, 2010, "Trends in income nonresponse over two decades." *Journal of Official Statistics*, *26* (1), 145–164. https://www.scb.se/contentassets/

ca21efb41fee47d293bbee5bf7be7fb3/trends-in-income-nonresponse-over-two-decades.pdf (last accessed May 2025).

Yu, Jun-Wu, Guo-Liang Tian, and Man-Lai Tang, 2008, "Two new models for survey sampling with sensitive characteristic: design and analysis." *Metrika*, *67*, 251–263. https://doi.org/10.1007/s00184-007-0131-x.

A Additional descriptions and data

A.1 Additional information on income questions

Box A.3: Household income question on exact amounts

What is the total monthly income of your household after taxes? If you cannot provide an exact amount an approximate answer would also be helpful.

_____ [CURRENCY of your COUNTRY]

No income	–77777
Don't know	–88888
No answer	–99999

[INTERVIEWER: If the respondent says "Zero" (no income) please ask again if the household really has no income or if the respondent doesn't want to reveal the amount. If the respondent doesn't want to reveal the amount, please choose "no answer" (99999).]

Box A.4: Personal income question on exact amounts (only if more than one hh-member)

And, what is your personal total monthly income after taxes? If you cannot provide an exact amount, an approximate answer would also be helpful.

_____ [CURRENCY of your COUNTRY]

No income	–77777
Don't know	–88888
No answer	–99999

[INTERVIEWER: Income that is received irregularly should be converted to a monthly amount. The reference period for income received is the last 12 months.]

Table A.1: Income brackets respondents can choose from in control group

	BG	HU	PL	BA
1	1-400 BGN	1-80,000 HUF	1-1,000 PLN	1-100 KM
2	401-600 BGN	80,001-100,000 HUF	1,000-1,499 PLN	101-200 KM
3	601-800 BGN	100,001-120,000 HUF	1,500-1,999 PLN	201-250 KM
4	801-1,000 BGN	120,001-140,000 HUF	2,000-2,249 PLN	251-300 KM
5	1,001-1,200 BGN	140,001-160,000 HUF	2,250-2,499 PLN	301-350 KM
6	1,201-1,400 BGN	160,001-180,000 HUF	2,500-2,749 PLN	351-400 KM
7	1,401-1,600 BGN	180,001-200,000 HUF	2,750-2,999 PLN	401-500 KM
8	1,601-1,800 BGN	200,001-220,000 HUF	3,000-3,249 PLN	501-600 KM
9	1,801-2,000 BGN	220,001-240,000 HUF	3,250-3,499 PLN	601-700 KM
10	2,001-2,250 BGN	240,001-260,000 HUF	3,500-3,749 PLN	701-800 KM
11	2,251-2,500 BGN	260,001-280,000 HUF	3,750-3,999 PLN	801-900 KM
12	2,501-2,750 BGN	280,001-300,000 HUF	4,000-4,399 PLN	901-1,000 KM
13	2,751-3,000 BGN	300,001-320,000 HUF	4,400-4,799 PLN	1,001-1,100 KM
14	3,001-3,250 BGN	320,001-340,000 HUF	4,800-5,199 PLN	1,101-1,200 KM
15	3,251-3,500 BGN	340,001-360,000 HUF	5,200-5,699 PLN	1,201-1,300 KM
16	3,501-3,750 BGN	360,001-380,000 HUF	5,700-6,199 PLN	1,301-1,500 KM
17	3,751-4,000 BGN	380,001-400,000 HUF	6,200-6,999 PLN	1,501-1,750 KM
18	4,001-4,250 BGN	400,001-425,000 HUF	7,000-7,999 PLN	1,751-2,000 KM
19	4,251-4,500 BGN	425,001-450,000 HUF	8,000-8,999 PLN	2,001-2,250 KM
20	4,501-4,750 BGN	450,001-475,000 HUF	9,000-10,999 PLN	2,251-2,500 KM
21	4,751-5,000 BGN	475,001-500,000 HUF	11,000-12,999 PLN	2,501-2,750 KM
22	5,001-5,500 BGN	500,001-550,000 HUF	13,000-14,999 PLN	2,751-3,000 KM
23	5,501-6,000 BGN	550,001-600,000 HUF	15,000-17,500 PLN	3,001-3,250 KM
24	Over 6,000 BGN	600,001-650,000 HUF	Over 17,500 PLN	3,251-3,500 KM
25		650,001-700,000 HUF		3,501-3,750 KM
26		700,001-750,000 HUF		3,751-4,000 KM
27		750,001-800,000 HUF		Over 4,001 KM
28		800,001-900,000 HUF		
29		900,001-1,000,000 HUF		
30		1,000,001-1,200,000 HUF		
31		1,200,001-1,400,000 HUF		
32		Over 1,400,000 HUF		

Table A.2: Household income tercile bounds used for the treatment group

	BG	HU	PL	BA
1. Tercile, lower bound	0 BGN	0 HUF	0 PLN	0 KM
2. Tercile, lower bound	1,601 BGN	320,001 HUF	4,800 PLN	901 KM
3. Tercile, lower bound	2,751 BGN	500,001 HUF	8,000 PLN	1,501 KM

Table A.3: Household income tercile bounds calculated from 2023 data

	BG	HU	PL	BA
1. Tercile, lower bound	0 BGN	0 HUF	0 PLN	0 KM
2. Tercile, lower bound	1,400 BGN	370,000 HUF	5,000 PLN	850 KM
3. Tercile, lower bound	2,500 BGN	560,000 HUF	7,500 PLN	1,600 KM

Source: OeNB Euro Survey, wave 2023.

Note: Bounds calculated from the data collected in 2023 including the exact income question and the control group treatment of granular brackets. For bracketed questions, mid-points of the bracket are used.

Table A.4: Household income in local currency, 2022 versus 2023

	BG		HU		PL		BA	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Income 2022	1,946	1,900	404,669	390,001	5,686	5,200	1,122	1,000
Income 2023	2,061	1,901	485,301	450,000	6,669	6,250	1,396	1,151
Growth rate (%)	5.9	0.0	19.9	15.4	17.3	20.2	24.4	15.1

Source: OeNB Euro Survey, waves 2022 and 2023.

Note: Values are shown in local currency, as reported by the respondents. Includes answers to exact income question and control group bracketed question.

	BG	HU	PL	BA
1. Tercile, lower bound	0 BGN	0 HUF	0 PLN	0 KM
2. Tercile, lower bound	801 BGN	200,001 HUF	2,500 PLN	601 KM
3. Tercile, lower bound	1,401 BGN	300,001 HUF	3,750 PLN	1,001 KM

Table A.5: Personal income tercile bounds used for the treatment group

Table A.6: Personal income tercile bounds calculated from 2023 data

	BG	HU	PL	BA
1. Tercile, lower bound	0 BGN	0 HUF	0 PLN	0 KM
2. Tercile, lower bound	800 BGN	200,000 HUF	2,700 PLN	500 KM
3. Tercile, lower bound	1,250 BGN	290,000 HUF	4,000 PLN	900 KM

Source: OeNB Euro Survey, wave 2023.

Note: Bounds calculated from the data collected in 2023 including the exact income question and the control group treatment of granular brackets. For bracketed questions, mid-points of the bracket are used.

Table A.7: Personal income in local currency, 2022 versus 2023

	BG		HU		PL		BA	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Income, 2022	1,043	900	217,823	195,000	2,991	2,900	642	530
Income, 2023	1,178	1,000	264,903	250,000	3,475	3,200	720	600
Growth rate (%)	13.0	11.1	21.6	28.2	16.2	10.3	12.1	13.2

Source: OeNB Euro Survey, waves 2022 and 2023.

Note: Values are shown in local currency, as reported by the respondents. Includes answers to exact income question and control group bracketed question.

A.2 Question texts and coding of control variables

Table A.8: Coding and question text of sociodemographic control variables

Variable name	Question text and coding
Age (in years)	*"What is your age?"*
	-Exact age in years
	-No answer
Gender (f/m)	*"Gender:"*
	-Female
	-Male
	-Gender diverse
	Dummy variable that is 1 for female respondents and 0 for male respondents.
Size of household	*"How many people, including yourself, live permanently in this household?"*
	-Number of people
	-No answer
Education (in categories)	*"What is your highest educational attainment?"*
	The number of answer categories varies by country. In Hungary and Bosnia and Herzegovina, ten categories, in Poland and Bulgaria, eleven categories are given. National categories are eventually mapped into the the ISCED-2011 classification for educational attainment, which has eight categories.
Main earner in household (0/1)	*"Are you the person who earns the most in your household?"*
	-Yes
	-No
	-Don't know or No answer
Manages household finances (0/1)	*"Who is in charge of managing household finances?"*
	-I am
	-I am together with my partner
	-I am together with somebody else
	-Somebody else
	-Don't know or No answer
	The first three categories are coded as 1, the fourth as 0.
Apprehensive before survey (0/1)	*"Does the respondent seem apprehensive about the survey?"*
	-No, not at all
	-Yes, a little
	-Yes, very
	The yes-categories are coded as 1, the no-category as 0.

Note: Where not otherwise indicated, "Don't know" and "No answer" are set to missing. All questions are translated into local languages by professional translators.

Table A.9: Coding of regional and interviewer-level control variables

Variable name	Question text and coding
Age of interviewer (in years)	*"Exact age of interviewer"* - Exact age in years
Gender of interviewer (f/m)	*"Gender of interviewer:"* -Female -Male -Gender diverse Dummy variable that is 1 for female respondents and 0 for male respondents.
Education of interviewer (dummy)	*"What is your highest educational attainment?"* Same as for respondents' education. Dummy is 1 if interviewer has completed at least Post-secondary non-tertiary education or higher and 0 otherwise.
Interviewer experience	*"Has the interviewer already conducted interviews for the OeNB Euro Survey in the past?"* -Yes, in wave 2021 -Yes, before 2021 - No The first two categories are coded as 1, the fourth as 0.
Region	We use the OeNB Euro Survey's region variable, which is harmonized over the waves. The number of regions varies across countries, totaling twelve in Bosnia and Herzegovina, nine in Bulgaria, seven in Hungary and sixteen in Poland.

B Additional methodological notes

B.1 Randomization procedure

As aforementioned, randomization mechanisms are not unified across countries. In Hungary and for the CAPI-interviews in Poland, computerized randomization was used, meaning the tablet randomizes treatment and control group independently of the interviewer. In Bosnia and Herzegovina and Bulgaria, this approach was not feasible as the software differs. Therefore the randomization relied on some interviewer input. In the Polish PAPI-interviews, questionnaires were prepared beforehand such that treatment and control were randomized.

As it is possible that interviewers conduct only a few interviews, in each randomization approach, a strict alternating order between control and treatment group was enforced. This ensures that interviewers had a balanced mix between the two groups. The computerized approach automatically switched between groups. In Bosnia and Herzegovina and Bulgaria, group assignment depended on whether the interviewer types in an odd or even number. This number was determined by how many interviews the interviewer had already conducted in the specific primary sampling unit (PSU). Thus, in contrast to the computer, interviewers restarted the counter for each PSU and interrupted interviews were counted in. In the PAPI-case, questionnaires received running numbers and were stacked in alternating order.

Randomization was tested in all countries in the 2022 wave of the survey, using the same methods as for this experiment. We did not experience any major problems with any randomization approach. Both the fully computerized and the interviewer-input method yielded well-balanced treatment assignments. However, in 2022, all questionnaires in a country were randomized. Due to time and budget constraints, Bosnia and Herzegovina, Bulgaria and Poland kept this procedure instead of randomizing if income nonresponse occurred. Therefore, we also show some data on randomization for the overall samples in these countries in the following section. Hungary implemented the randomization only if

nonresponse in the exact household income question occurred.

B.2 Descriptive statistics for all respondents in BA, BG and PL

Table A.10: Distribution of treatment groups by country – all respondents

	BG	PL	BA
Control: many brackets	0.52	0.50	0.54
Treatment: terciles	0.48	0.50	0.46
Observations	1001	1012	946

Source: OeNB Euro Survey, wave 2023.

Note: Share of respondents who are assigned to control and treatment group respectively. Control is the group in which respondents are asked about income in many brackets and treatment is the group in which they are asked in terciles. Note that statistics are calculated on all respondents in a country, not only on those who refuse to give an amount for household income. In Hungary, it was not randomized across all respondents.

Table A.11: Differences by treatment across countries: all respondents

	BG	PL	BA
Gender (f/m)	−0.03	−0.02	0.02
Age (in years)	1.10	−1.16	−1.84
Education (in categories)	0.11	−0.04	0.05
Main earner in household (0/1)	0.04	−0.00	0.03
Manages household finances (0/1)	0.05**	−0.02	0.00
Size of Household	−0.10	0.24***	0.01
Observations	1001	1012	946

Source: OeNB Euro Survey, wave 2023.

Note: Difference between control and treatment group for each variable. Positive (negative) numbers indicate that the value is larger (smaller) for the control group than for treatment group. Control is the group in which respondents are asked about income in many brackets and treatment is the group in which they are asked in terciles. Note that statistics are calculated on all respondents in a country, not only on those who refused to give an amount for household income. In Hungary, it was not randomized across all respondents.

* p < .05, ** p < .01, *** p < .001

B.3 Power

This sections shows the calculations of minimum detectable effect (MDE) sizes, which were included in the pre-analysis plan. For the proportion power tests, we use nonresponse shares for exact household income from the 2022 wave of the OeNB Euro Survey as benchmark. We focus on minimum detectable effect (MDE) sizes rather than power given a lack of empirical papers to guide the choice of effect size and because our sample size is fixed by the contract governing the survey.

Table B.3 reports the MDE for both hypotheses separately, given the sample sizes from 2022, a power of 80% and an alpha of 5%. Our MDE is the minimum detectable change in percentage points, which is often used in the case of proportions. We use one-sided tests as our hypotheses state that the shares of nonresponse decrease with the treatment.

The reason why sample sizes to test hypothesis 2 are smaller is that not at all respondents answer the question on personal income. For both hypotheses in all countries, we should be able to detect effects that are at least around 10-13 percentage points.

Table A.12: Summary power statistics

	BG	HU	PL	BA
Hypothesis 1				
Sample size	568	480	377	531
Sample size control, expected	284	240	189	266
Sample size treatment, expected	284	240	189	266
Share of nonresponse, control	69%	50%	54%	81%
Minimum detectable change	10 ppt	11.25 ppt	12.77 ppt	9.19 ppt
Hypothesis 2				
Sample size	515	377	324	412
Sample size control, expected	258	189	162	206
Sample size treatment, expected	258	189	162	206
Share of nonresponse, control	82%	80%	70%	81%
Minimum detectable change	9.15 ppt	11.16 ppt	13.27 ppt	10.47 ppt

Sample size and proportion of the control group are based on data from the 2022 Euro Survey wave. For the minimum detectable change, power of 80% and alpha of 5% assumed.

C Additional analyses on heterogeneous treatment effects

Table A.13: Treatment effects for household income nonresponse – OLS regressions with controls and interactions for pooled sample

	(1)	(2)	(3)	(4)	(5)	(6)
Terciles (treatment)	-0.181***	-0.181***	-0.137***	-0.177***	-0.195***	-0.145***
	(0.023)	(0.023)	(0.032)	(0.028)	(0.029)	(0.042)
HU	-0.164**	-0.157**	-0.155**	-0.156**	-0.157**	-0.154**
	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)
PL	-0.116**	-0.113*	-0.111*	-0.112*	-0.112*	-0.111*
	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)
BA	0.068	0.073	0.072	0.073	0.073	0.073
	(0.050)	(0.050)	(0.050)	(0.050)	(0.051)	(0.051)
Gender	0.022					
	(0.029)					
Age (in years)	-0.001					
	(0.001)					
Education	-0.011					
	(0.011)					
Male		0.026	0.074*	0.026	0.026	0.075*
		(0.029)	(0.037)	(0.029)	(0.029)	(0.037)
Tertiary education		-0.014	-0.014	-0.007	-0.014	-0.005
		(0.030)	(0.030)	(0.040)	(0.030)	(0.040)
Age >55		0.003	0.002	0.003	-0.016	-0.017
		(0.028)	(0.028)	(0.028)	(0.036)	(0.036)
Treatm. x Male			-0.096*			-0.097*
			(0.049)			(0.049)
Treatm. x High educ.				-0.013		-0.017
				(0.051)		(0.052)
Treatm. x Age >55					0.040	0.038
					(0.049)	(0.049)
Adjusted R^2	0.09	0.09	0.09	0.09	0.09	0.09
Observations	1661	1661	1661	1661	1661	1661

* p < .05, ** p < .01, *** p < .001.

Other control variables (not shown) are main earner in household, manages household finances, size of household, whether respondent was apprehensive before the survey, gender of interviewer, age of interviewer, interviewer education and interviewer experience.

Table A.14: Hypothesis 2: Treatment effects for personal income nonresponse – OLS regressions with controls and interactions for pooled sample

	(1)	(2)	(3)	(4)	(5)	(6)
Terciles (treatment)	0.010	0.010	0.008	0.016	-0.007	-0.000
	(0.020)	(0.020)	(0.029)	(0.024)	(0.024)	(0.034)
HU	-0.051	-0.059*	-0.059*	-0.058*	-0.059*	-0.058*
	(0.030)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
PL	-0.159***	-0.162***	-0.162***	-0.161***	-0.160***	-0.160***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
BA	-0.145***	-0.147***	-0.147***	-0.146***	-0.146***	-0.146***
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
Gender (f/m)	-0.031					
	(0.029)					
Age (in years)	0.001					
	(0.001)					
Education (in categories)	0.010					
	(0.008)					
Male		-0.032	-0.033	-0.032	-0.032	-0.031
		(0.029)	(0.035)	(0.029)	(0.029)	(0.035)
Tertiary education		0.018	0.018	0.030	0.018	0.029
		(0.023)	(0.023)	(0.034)	(0.023)	(0.034)
Age >55		-0.023	-0.023	-0.023	-0.048	-0.048
		(0.024)	(0.024)	(0.024)	(0.033)	(0.033)
Treatm. x Male			0.003			-0.001
			(0.040)			(0.040)
Treatm. x High educ.				-0.023		-0.021
				(0.045)		(0.045)
Treatm. x Age >55					0.052	0.051
					(0.043)	(0.043)
Adjusted R^2	0.15	0.15	0.14	0.15	0.15	0.14
Observations	1439	1439	1439	1439	1439	1439

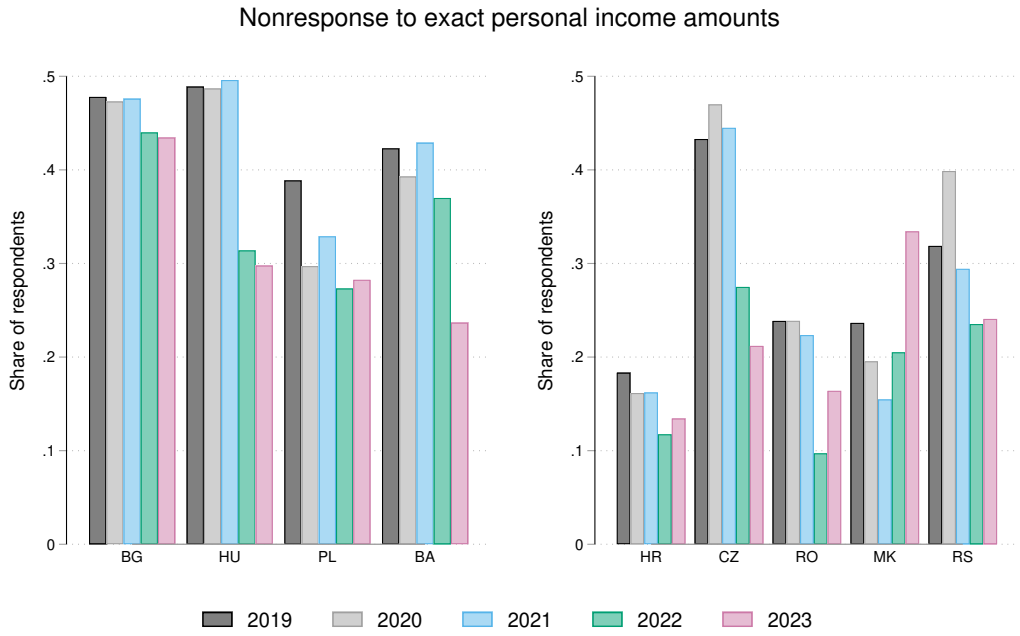
* p < .05, ** p < .01, *** p < .001.

Other control variables (not shown) are main earner in household, manages household finances, size of household, whether respondent was apprehensive before the survey, gender of interviewer, age of interviewer, interviewer education and interviewer experience.

D Analyses on personal income questions

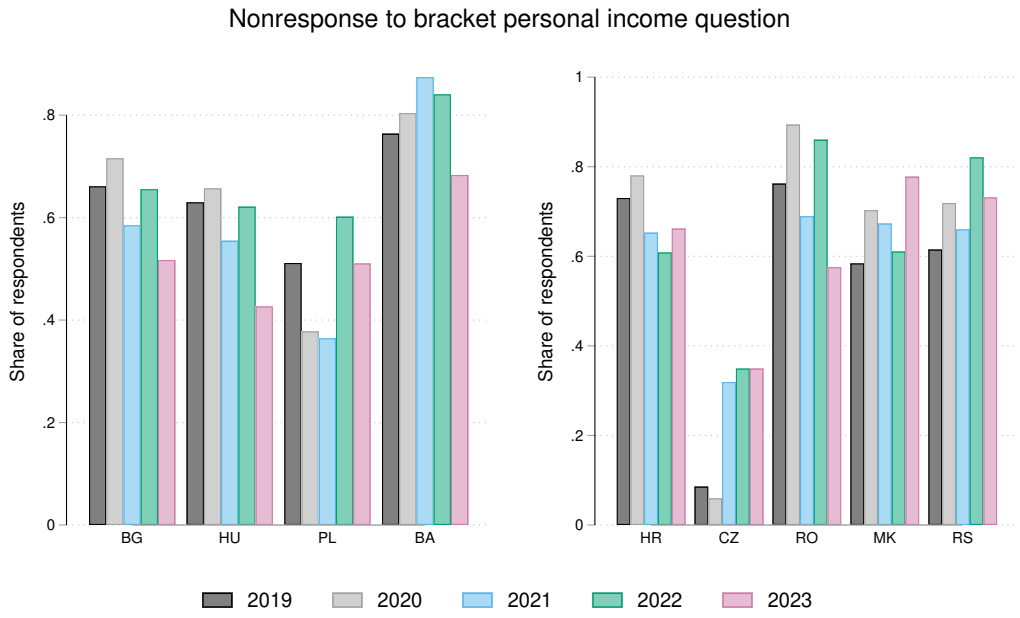
Figures 1 to 3 show that the income refusal for personal income question looks very similar to household income over the last few years.

Figure A.1: Respondents who refused to report the exact amount of personal income



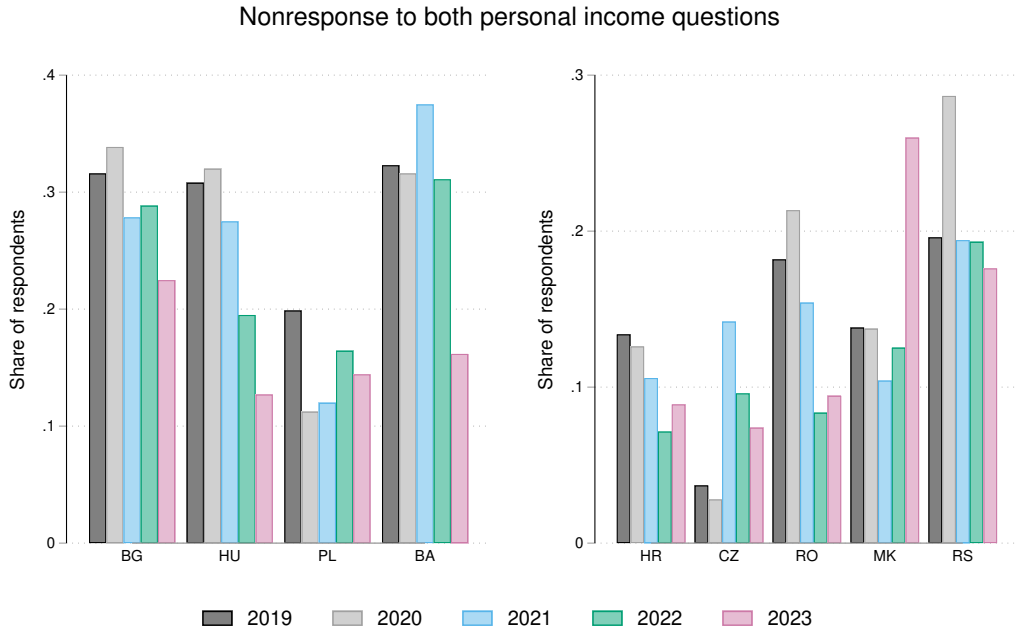
Source: OeNB Euro Survey, waves 2019–2023.
Note: 2021 might not be comparable as income was asked differently.

Figure A.2: Respondents who refused to report personal income in brackets



Source: OeNB Euro Survey, waves 2019–2023.
Note: 2021 might not be comparable as income was asked differently. In BG, HU, PL and BA, respondents are randomized into receiving two different kinds of bracket questions (many vs few) in 2023.

Figure A.3: Respondents who did not answer any personal income questions



Source: OeNB Euro Survey, waves 2019–2023.
Note: 2021 might not be comparable as income was asked differently.

The tables below show the descriptive statistics for treatment and control groups for personal income. As in subsection 4.2 of the paper, we show descriptive statistics for the

final sample that refuse to give an amount for personal income. Given that for single-person households, the household income question actually asks about their personal income, we include these respondents in the sample. This means that the respondents we pool do not receive the same order of questions, as some receive only one question (household income) and others two (personal and household income). Nonetheless, we think that these two groups need to be pooled to understand if the treatment works better or worse for household or personal income. Table 15 depicts the share of respondents in each country that is assigned to control and treatment for the income questions. The shares are well balanced in Bulgaria and Poland, but less so in Bosnia and Herzegovina, and Hungary, where the treatment groups make up 44% and 54%, respectively. In Bosnia and Herzegovina, this is also the case in the analysis on household income. In Hungary, the imbalances are mildly larger for personal income.

Table A.15: Distribution of treatment groups for personal income by country

	BG	HU	PL	BA
Control: many brackets	0.50	0.46	0.50	0.56
Treatment: terciles	0.50	0.54	0.50	0.44
Observations	496	360	319	285

Source: OeNB Euro Survey, wave 2023.

Note: Share of respondents who are assigned to control and treatment group respectively. Control is the group in which respondents are asked about income in many brackets and treatment is the group in which they are asked in terciles.

Table A.16: Differences by treatment across countries

	BG	HU	PL	BA
Gender (f/m)	−0.01	−0.06	−0.03	0.07
Age (in years)	−0.91	−1.42	−0.89	−4.32*
Education (in categories)	0.15	−0.22	0.13	−0.01
Main earner in household (0/1)	−0.02	−0.08	−0.02	0.04
Manages household finances (0/1)	0.04	−0.02	−0.01	0.02
Size of Household	−0.01	0.22	0.25*	0.03
Apprehensive before survey (0/1)	0.04	0.08	−0.01	0.02
Observations	496	360	319	285

Source: OeNB Euro Survey, wave 2023.

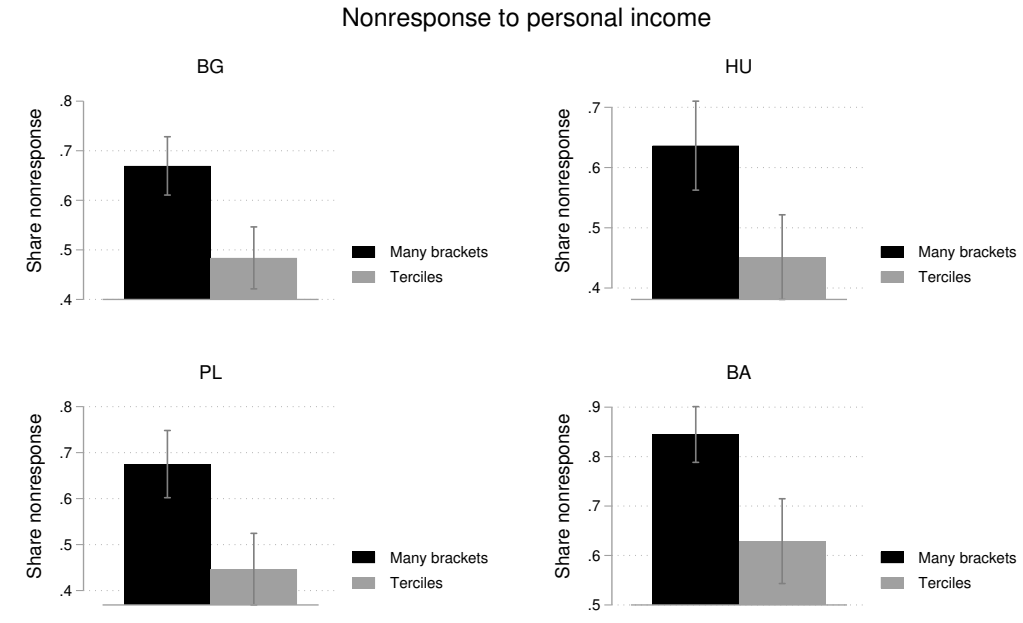
Note: Difference between control and treatment group for each variable. Positive (negative) numbers indicate that the value is larger (smaller) for the control group than for treatment group. Control is the group in which respondents are asked about income in many brackets and treatment is the group in which they are asked in terciles.

* p < .05, ** p < .01, *** p < .001

In table 16, we show mean comparisons for selected individual characteristics across treatment groups. Most selected variables are well-balanced between treatment and control group. As for household income, in Hungary, there is an imbalance with respect to the number of household members. Moreover, in Bosnia and Herzegovina there are imbalances related to age. However, in none of the countries do all criteria together predict treatment assignment. F-tests on joint orthogonality result in insignificant test statistics.

Figure 4 shows the nonresponse shares for the many brackets and tercile question for personal income. Table 17 presents the p-values for the one-sided Fishers exact test, chi-square tests and two proportions tests. In all four countries, the share of nonresponse for the tercile question is significantly lower than for the many brackets question, under all three tests.

Figure A.4: Mean comparison personal income nonresponse between brackets and terciles



Source: OeNB Euro Survey, wave 2023.
Note: Confidence intervals are obtained using two–sided t–tests.

Table A.17: Personal income – p-values for smaller nonresponse shares in treatment group

	BG	HU	PL	BA
One-sided Fisher's exact	<0.001	<0.001	<0.001	<0.001
Chi-squared	<0.001	<0.001	<0.001	<0.001
One-sided two proportions	<0.001	<0.001	<0.001	<0.001

Average treatment effects are also significant in multivariate regressions in table 18. Columns (1), (3), (5) and (7) show the unconditional treatment effects, while we control for the region in columns (2), (4), (6) and (8). In all estimations, treatment effects are highly significant and on average, reduce the share of nonresponse between 18 to 23 percentage points. The effect sizes are larger than those for household income in the case of Bulgaria, while the do not differ too much for the other countries.

Table A.18: Personal income nonresponse share – logit regressions with controls

	BG		HU		PL		BA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment: terciles	−0.181***	−0.189***	−0.181***	−0.214***	−0.221***	−0.227***	−0.205***	−0.228***
	(0.038)	(0.036)	(0.048)	(0.047)	(0.049)	(0.047)	(0.046)	(0.045)
DK/NA control	0.67	0.67	0.64	0.64	0.68	0.68	0.81	0.84
Control variables	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-329.2	-286.4	-242.4	-209.4	-210.2	-166.8	-151.3	-117.3
McFaddens R^2	0.03	0.11	0.02	0.16	0.04	0.16	0.05	0.21
Observations	496	471	360	360	319	289	285	267

* $p < .05$, ** $p < .01$, *** $p < .001$

DK/NA control shows the nonresponse share in the control group. Control variables include age, gender, education, main earner in household, manages household finances, size of household, whether respondent was apprehensive before the survey, region dummies, gender of interviewer, age of interviewer, interviewer education and interviewer experience.

The Working Paper series of the Oesterreichische Nationalbank is designed to disseminate and to provide a platform for discussion of either work of the staff of the OeNB economists or outside contributors on topics which are of special interest to the OeNB. To ensure the high quality of their content, the contributions are subjected to an international refereeing process. The opinions are strictly those of the authors and do in no way commit the OeNB.

The Working Papers are also available on our website (http://www.oenb.at) and they are indexed in RePEc (http://repec.org/).

Publisher and editor Oesterreichische Nationalbank
Otto-Wagner-Platz 3, 1090 Vienna, Austria
PO Box 61, 1011 Vienna, Austria
www.oenb.at
oenb.info@oenb.at
Phone (+43-1) 40420-6666

Editor Martin Summer

Cover Design Information Management and Services Division

DVR 0031577

ISSN 2310-533X (Online)

© Oesterreichische Nationalbank, 2026. All rights reserved.