

# Do talk money – Reducing income nonresponse in surveys<sup>\*</sup>

DRAFT: DO NOT CITE AND CIRCULATE

Katharina Allinger<sup>†</sup>      Melanie Koch<sup>‡</sup>

February 2024

## 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 from a list with only three brackets. In all four countries, providing fewer brackets leads to a significant decrease in nonresponse. Thus, when condensed income data are sufficient, fewer brackets are a cost-effective way to reduce nonresponse.

**Keywords:** survey experiment; item nonresponse; household income

**JEL:** C83; D91; G50

---

<sup>\*</sup>We thank Elisabeth Beckmann and Joachim Winter for very helpful comments. The opinions expressed do not necessarily reflect the official viewpoint of the Oesterreichische Nationalbank or the Eurosystem. The study is registered in the AEA RCT Registry, [ID: 0012215](#). Declarations of interest: none.

<sup>†</sup>Oesterreichische Nationalbank, 1090 Vienna, Austria;  
Email: [katharina.allinger@oenb.at](mailto:katharina.allinger@oenb.at)

<sup>‡</sup>Oesterreichische Nationalbank, 1090 Vienna, Austria;  
Email: [melanie.koch@oenb.at](mailto:melanie.koch@oenb.at)

# 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 constant cause of sorrow. One aspect of nonresponse is item nonresponse which means that persons in general are willing to participate in the survey but refuse answering certain questions. The reasons for answer refusal are manifold, ranging from deep mistrust in authority to genuinely not knowing the answer to a question. Unsurprisingly, questions that are especially prone to nonresponse are questions about monetary values, like income.

Income nonresponse is indeed pervasive across surveys. For instance, [Yan et al. \(2010\)](#) collects income nonresponse figures for a selection of U.S. surveys. In these surveys, income nonresponse to open-ended questions ranges from 14% to 35%. For follow-up bracketed questions in some of the surveys, nonresponse ranges from 14% to 19%. [Frick and Grabka \(2010\)](#) note in their paper 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%.

Refusing income questions has also been an issue in the OeNB Euro Survey, which has been conducted in ten countries for more than 15 years. In the 2022 survey wave of the OeNB Euro Survey, nonresponse for 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.

Being members of the research team that is running the OeNB Euro Survey, we conducted an experiment in the 2023 survey wave, testing a new approach to reduce income nonresponse. The main reason to test new approaches is that the current approach, which will be our control treatment, works well in some country samples but barely in others. Thus, we restrict the experiment to four countries, in which nonresponse is particularly high and current approaches to reduce it have not been fruitful.

In the survey, respondents are initially asked to provide exact amounts on, for example, monthly household income. Until wave 2022, if they were not willing or not able to give

an exact amount, i.e. nonresponse occurs, all respondents were offered to choose their income from a list of very granular income brackets. In wave 2023, we use these granular brackets as our control group. In contrast, we ask the other group of respondents, our treatment group, simply if their income falls into the first, second or third income tercile, with boundaries predefined by us. In each of the countries, respondents are randomly assigned to either control or treatment. We have two pre-registered hypotheses, which will be tested independently for each country. First, we expect that income nonresponse is significantly reduced in the group that receives the tercile question, as respondents are more likely to answer questions that are cognitively less demanding and less privacy-invasive. Second, we expect spillover effects from our treatment to the answer behavior for the following question on personal income, asked to all respondents who live in multi-person households.

Our results indeed 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 20 percentage points in Bosnia and Herzegovina, 14 percentage points in Bulgaria, 18 percentage points in Poland and 25 percentage points in Hungary. Thus, we cannot reject our first hypothesis. However, we reject the second hypothesis, as we find that 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. In contrast, in Hungary, the treatment seems to increase subsequent nonresponse to exact amounts. This could be driven by anticipation of the easier tercile question.

Our paper relates to the existing literature on item nonresponse in surveys and, specifically, income nonresponse. This literature has focused on different strategies. For instance, [Moore and Loomis \(2001\)](#) conduct an experiment, where they test the effects of a brief statement before the income questions, designed to reassure the respondent. They find that this reduces nonresponse. 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 bracket 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. Using brackets clearly also has its challenges. Besides generally 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).

Our paper contributes to the aforementioned literature by quantifying the effects of drastically reducing the number of brackets on income nonresponse and potential spillover effects. To our knowledge, this is the first paper that provides experimental evidence on this topic. We think our results can be a helpful input when researchers and data providers have to decide how to elicit income in surveys.

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 and section 5 concludes.

## 2 Data and implementation

### 2.1 The OeNB Euro Survey

We test our treatment in four different country samples of an international household survey, the OeNB Euro Survey. This survey is conducted annually by the Austrian Central Bank (OeNB) in ten countries in Central, Eastern and Southeastern Europe. The main topics covered revolve around household finance issues and respondents are asked to report their household and personal net income. Questionnaires are ex ante harmonized such that income is elicited in a comparable manner. The survey is meant to be representative of each country’s individual population and around 1,000 individuals

per country are interviewed every year by local survey institutes. The experiment was conducted in the 2023 wave of the survey.

## 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. If respondents refused to answer the exact income questions, they got a follow-up question with brackets. Traditionally, income questions are always asked towards the end of the survey. In the 2023 wave, the same questions for the exact amount (see appendix A.1) and the same ordering as for the 2022 wave were used. This means that income questions were the last questions respondents had to answer. Moreover, household income was asked first and personal income was asked only in case there were additional household members besides the respondent.

In all countries, questions on exact income amounts suffer from nonresponse and are often the questions with the highest nonresponse share in the survey. However, shares 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 (see also figure 1 in section 4.1). The numbers are comparable or even higher for the case of personal income, in particular for the countries with the highest nonresponse shares (see figures A1, A2 and A3 in annex A.3).

Nonresponse has been partially tackled by asking follow-up questions, in which instead of exact amounts, respondents are offered to choose their income within a list of income brackets. The number of brackets and the range of each bracket is different for each country (see annex A.1). Amounts are guided by administrative income data from the countries and survey responses from previous years. Since income ranges were chosen to be relatively granular, in each country, at least 20 different brackets are given. This number of brackets is definitely at the upper end when comparing it with other surveys.<sup>1</sup> Since the 2022 wave, the follow-up income brackets are asked with an introductory statement,

---

<sup>1</sup> To give just two examples, the Core Questionnaire of the Fed's Survey of Consumer Expectations includes 11 income brackets, while the latest waves of the European Social Survey contains 10 brackets.

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 again in the other countries that are not part of our experiment.

In some countries, the bracket questions decrease income nonresponse tremendously. For instance, Czechia had a refusal rate of around 40% for exact income between 2021 and 2023. However, more than 70% of these respondents were at least willing to answer the brackets question. This reduces the overall share of income nonresponse to about 10% in this country. However, in some countries, the majority of the respondents also refused to answer the income brackets, for instance in Bosnia and Herzegovina (see also 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 in income brackets is only a minor or no remedy at all. We chose these countries with high shares of nonresponse to household income to ensure that we have enough power to test our treatment.<sup>2</sup> The nonresponse shares of exact, categorical and both income questions for these 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 nonresponse to both questions in the full sample.

<sup>2</sup> 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 annex A.3).

### 2.3 The survey experiment

We want to test if nonresponse to income questions can be reduced by lowering the number of brackets that are shown to respondents – in our case, lowering the number from around 20-30 brackets to just three. Therefore, our control and treatment groups consist of respondents who once already refused to answer the household income question which asks for exact amounts. As aforementioned, refusal in previous waves was tackled by asking a follow-up question that is meant to be less obtruding but still relatively accurate. Our control group receives this same question as asked in previous waves (see box 1). Thus, as can be seen in appendix table A1, control respondents are presented lists with at least 24 income brackets from which to choose from. They are very granular with the smallest brackets in each country capturing 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’s income after taxes?

BRACKET LIST	[See appendix table A1 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: if their household income lies in the pre-defined, first, second or third income tercile of a country’s overall income distribution (see box 2).

## 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 2. tercile]	1
Between [Lower bound of 2. tercile] and [Lower bound of 3. tercile]	2
Above [Lower bound of 3. tercile]	3
No income	-77777
Don't know	-88888
No answer	-99999

Hence, the treatment differs in one key aspect from the control. Instead of locating the income in more than 24, granular brackets, respondents are offered only three brackets. We want these three brackets to at least roughly represent terciles so that we get equally sized income groups. As there is no way to know in advance the exact future income distribution within a country, 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.

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%. Thus, 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 treatment. This is helpful to later construct three groups from the control treatment as well. It makes direct comparisons more feasible and reliable as bounds are exactly the same. The final tercile bounds are presented in appendix tables [A2](#) and [A5](#). Moreover, appendix tables [A3](#) and [A6](#) show that the actual tercile bounds in 2023



calculated from the exact income question and the control treatment bracket question are not too far away from the pre-defined values. Tables [A4](#) and [A7](#) show the growth rates of household and personal income from 2022 to 2023.

## 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 Euro Survey are interviewer-led and conducted face-to-face. Moreover, almost all interviews are computer assisted using tablet. Still, in each country, local polling institutes are responsible for conducting the survey, which leads to some differences in technical procedures across countries. 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 differ as well. This is described in more detail in appendix subsection [A.2.1](#). However, in each country, randomization is stratified by interviewer.

## 3 Hypotheses and empirical approach

We have two main hypotheses regarding the effect of our treatment. In summary, 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. Thus, the first hypothesis H1-0 and its alternative, onesided hypothesis H1-a are:

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

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

The intuition is that we expect nonresponse on exact amounts to be caused by two main reasons: not knowing the exact amount or not willing to share the exact amount because of privacy concerns. For both reasons, asking in broader brackets should be a remedy. Respondents do not have to know the exact amount, an approximation is enough, and the interviewer does not learn the exact amount, which preserves at least some privacy.

Moreover, we hypothesize that there might be spillover effects to subsequent income questions. Because the tercile question is easier to digest and invades the privacy of the respondent less, it 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. In general, questions about personal income should be easier to answer in terms of knowledge as knowing the own income is more likely than knowing that of every other household member. Therefore, for the majority of the respondents that will be asked about their personal income as well, we expect the treatment to decrease the refusal to report exact amounts on personal income:

*H2-0: The share of refused answers to the exact amount of personal income is equal in the treatment and in the control group.*

*(H2-a: The share of refused answers 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, this gives us eight, onesided hypotheses to be tested. To test these, we use 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 many brackets questions (control group). Thus,  $treat$  is missing if a respondent gives an exact amount on household income. For individual  $i = 1, \dots, N_c^{treat}$  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  $treat$ . These bivariate tests serve as first indication if the treatment has worked. Given that interviewer and regional effects can be strong (see [Olbrich et al., 2024](#)), we run logistic regressions for each country  $c$ , in which we control for region ( $Z_i$ ) and interviewer ( $J_i$ ) effects using dummy variables:

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

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

Moreover, we adjust for potentially occurring imbalances between treatment and control in the regressions. While we conduct onesided, bivariate tests, for the regressions, we report twosided significance levels.<sup>3</sup> If we then detect significant differences for the proportions within a country and if the  $treat$  coefficient is significant in regressions (1) or

---

<sup>3</sup> In the pre-analysis plan, we stated to report onesided significance tests also for the regressions. However, we have decided against this given the more common convention to report twosided significance in regressions and because for all our results, the significance level is the very similar under one and twosided tests.

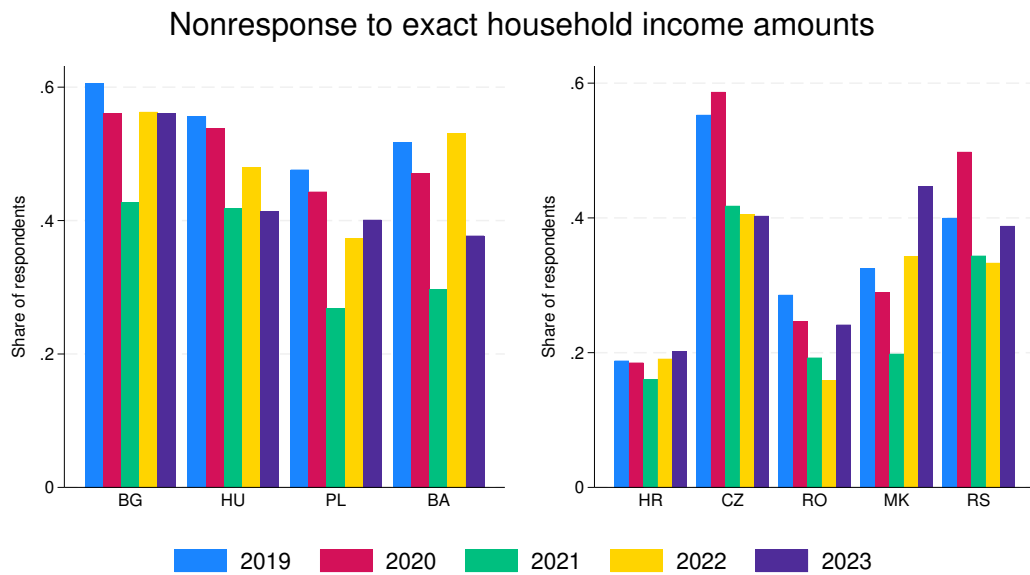
(2) respectively, we interpret this as sign to not reject the alternative hypothesis (H1-a or H2-a) for the specific country. However, as robustness, given that we have in total eight hypotheses, we will adjust our significance values for multiple hypothesis testing using sharpened q-values from [Anderson \(2008\)](#).

## 4 Results

### 4.1 General descriptive statistics

We first take a look at the evolution of nonresponse to exact, categorical and both household income questions across all countries and between 2019 and 2023. The first shows us, whether there were any surprising developments in nonresponse to exact household income. The latter two can already serve as first indication of whether our treatments likely had an effect on nonresponse to income questions or not. As a point of comparison, we also include five other countries covered in the OeNB Euro Survey, which are not part of the experiment.

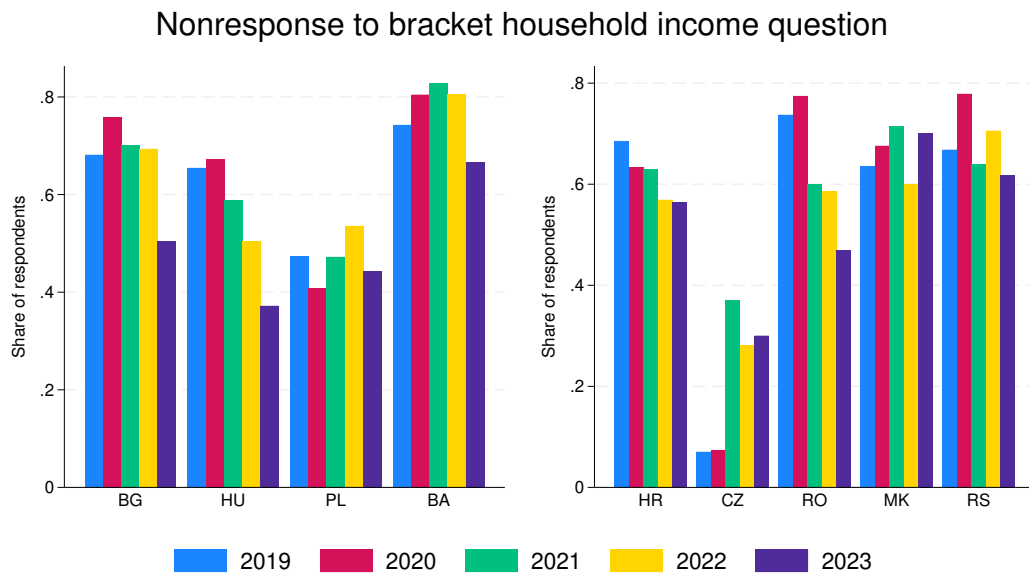
Compared to previous waves, nonresponse shares for exact household income in 2023 do not look out of the ordinary – see [figure 1](#). In the countries included in the experiment (the left panel), they are roughly the same as the 2019-2022 average in Bulgaria and Poland and a bit lower in Hungary and Bosnia and Herzegovina. In Bosnia and Herzegovina, we see the most dramatic drop compared to a year ago.



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

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

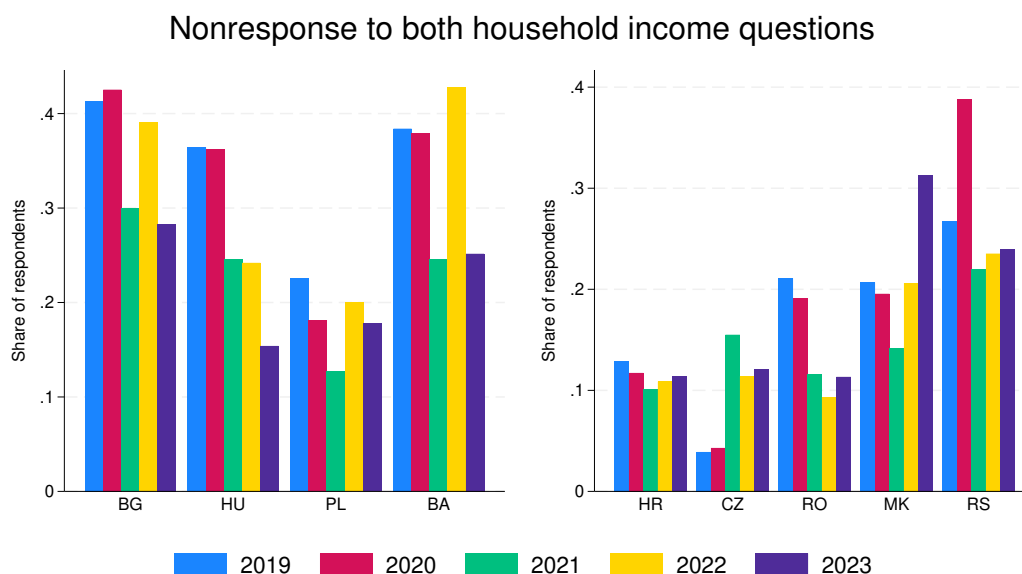
For the bracketed income question in figure 2, the aggregate shares of nonresponse have declined in all countries included in the experiment (see again left panel). The sharpest drop can be observed in Bosnia and Herzegovina, while in Poland the shares have declined compared to 2022, but not so much compared to previous years. In the countries not included in the experiment (right panel), we see mixed developments. In Romania, which was not included in the experiment, we also see a marked drop in nonresponse shares to the categorical question – however, this was accompanied by increases in nonresponse for the exact question.



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

Figure 2: Respondents who refused to report household income in brackets

Figure 3 then shows overall income nonresponse to both questions in the full sample. In most countries that were not part of the experiment, joint nonresponse stayed roughly unchanged in 2023 compared to previous waves – with the exception of a marked increase in North Macedonia. Of the countries, where the experiment was conducted, nonresponse has not changed much in Poland, but has decreased markedly in Bulgaria, Hungary and Bosnia and Herzegovina. Nonetheless, Bosnia and Herzegovina and Bulgaria remain the countries with the highest nonresponse shares (above 25%) in 2023 – jointly with North Macedonia and Serbia.



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

Figure 3: Respondents who did not answer any income questions

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 three countries (BA, BG and PL), all respondents were assigned to either control or treatment group – even if they answered the question on household income amounts – because treatment assignment was determined in the beginning of the survey. Therefore, we calculate some descriptive statistics for two different samples in these countries. Here, we always present statistics on the final sample that refused to give an amount for household income and then received either the control or the treatment question. In subsection [A.2.2](#) in the appendix, we show descriptives for all respondents interviewed in these countries separately for control and treatment group. This is to compare the balance of the sample we are interested in (those who refused to answer) to the sample on which the treatment was actually randomized.

Table 2 depicts the share of respondents in each country that was assigned to control and 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, in which the shares are 55% and 45% respectively. [Beckmann and Koch \(2024\)](#) already noted that probably the less advanced randomization procedure in Bosnia and Herzegovina leads to these imbalances. Imbalances were also found in their analysis of an experiment conducted in the 2022 Euro Survey wave.

Table 2: Distribution of treatment groups by country

	BG	HU	PL	BA
Control: many brackets	0.50	0.48	0.52	0.55
Treatment: terciles	0.50	0.52	0.48	0.45
Observations	561	414	406	357

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. We believe these to be the characteristics to be important in predicting nonresponse to income questions. Mostly, socio-demographic variables like gender, age, education, whether the respondent is the main earner of the household and whether she is involved in managing household finances are well balanced between the two groups. However, in Hungary and Poland, there is an imbalance with respect to the number of household members. In addition, an F-test on joint orthogonality shows that in Hungary, all the characteristics together predict treatment assignment (p-value=0.0624). However, we do not have any reason to believe that the fully computerized randomization in Hungary did not work properly. The F-test for Poland is not significant (p-value=0.7276).



Table 3: Differences by treatment across countries

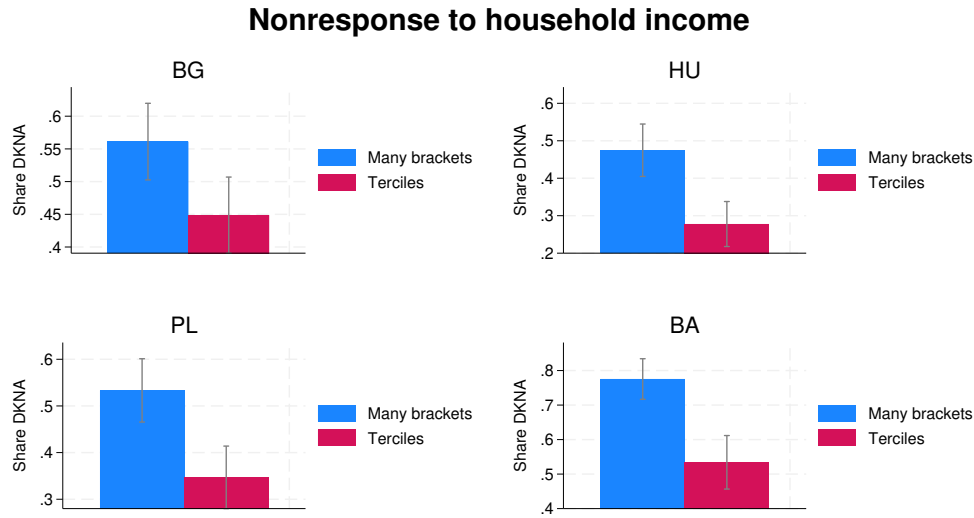
	BG	HU	PL	BA
Gender (f/m)	0.02	−0.07	−0.03	0.04
Age (in years)	0.23	−2.10	−0.59	−2.29
Education (in categories)	0.13	−0.20*	−0.01	0.01
Main earner in household (0/1)	0.02	−0.10**	−0.05	0.03
Manages household finances (0/1)	0.05*	−0.06**	−0.02	0.04
Respondent apprehensive (0/1)	0.03	0.00	−0.03	0.03
Size of Household	−0.01	0.25**	0.22**	−0.03
Observations	561	414	406	357

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.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 4.3 Treatment effects

First, we look at hypothesis one and the effect the tercile treatment has on nonresponse in household income. Figure 4 shows the nonresponse shares for the many brackets and tercile question. As specified in the pre-analysis plan, we use three different tests to see if the shares are significantly different from each other (Fisher’s exact test, chi-square tests and two proportions tests). Table 4 presents the p-values for the one-sided 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.



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

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

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

	BG	HU	PL	BA
One-sided Fisher's exact	0.005	0.000	0.000	0.000
Chi-squared	0.008	0.000	0.000	0.000
One-sided two proportions	0.004	0.000	0.000	0.000

Significant treatment effects are confirmed by the multivariate regressions in table 5. Columns (1), (3), (5) and (7) show the unconditional treatment effects, while we control for the region and the interviewer in columns (2), (4), (6) and (8). In all estimations, treatment effects are highly significant and on average, reduce the share of nonresponse between 11 to almost 28 percentage points.<sup>4</sup> Extrapolating from the collected data, in our sample of 1,000 respondents per country, applying the treatment to all respondents in future waves could generate income information 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. Thus, we cannot reject hypothesis 1a and find that the treatment has its intended effect of lowering refusal rates.

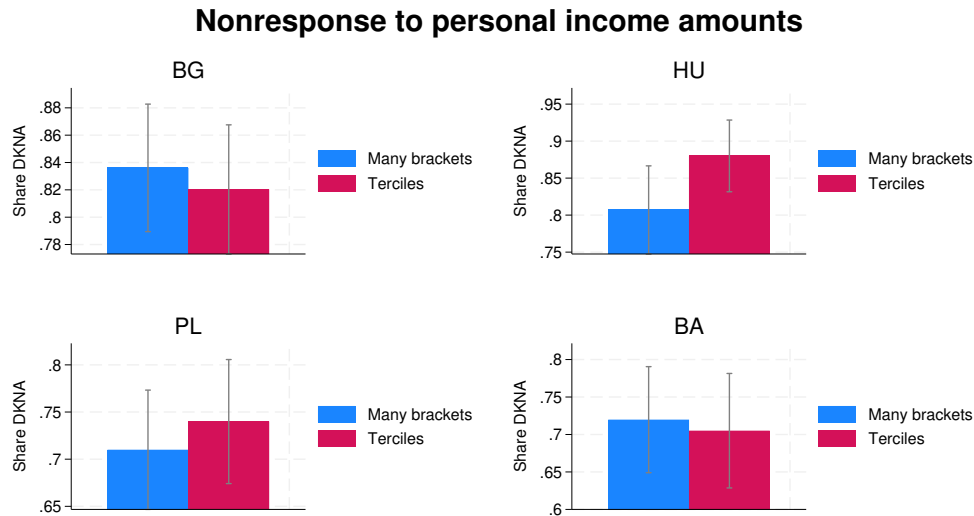
<sup>4</sup> It should also be noted that we have sufficient power and all the estimated effects exceed the minimum detectable changes we calculated in our pre-analysis plan.

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*** (0.039)	-0.148*** (0.046)	-0.191*** (0.040)	-0.278*** (0.052)	-0.182*** (0.042)	-0.196*** (0.054)	-0.229*** (0.047)	-0.204*** (0.049)
DK/NA control	0.56	0.56	0.47	0.47	0.53	0.53	0.78	0.78
Region dummies	No	Yes	No	Yes	No	No	No	Yes
Interviewer d.	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-385.3	-261.2	-264.6	-157.4	-271.6	-198.8	-215.6	-145.1
McFaddens R <sup>2</sup>	0.01	0.16	0.03	0.22	0.03	0.14	0.05	0.28
Observations	561	448	414	293	406	333	357	311

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

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. Therefore, we again run the three bivariate tests first and plot the nonresponse shares in figure 5.



Source: OeNB Euro Survey 2023.  
Note: Confidence intervals are obtained using two-sided  $t$ -tests.

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

As suggested by the figure already, the treatment does not have the intended spillover effects on exact personal income. The share of nonresponse is not significantly different in all countries (see table 6), except for Hungary, in which the effect runs in the opposite

direction. There, the share of nonresponse after respondents answered the tercile question is higher than the share of nonresponse after the many brackets question. However, effects are only borderline significant.

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

	BG	HU	PL	BA
One-sided Fisher’s exact	0.364	0.042	0.299	0.440
Chi-squared	0.641	0.061	0.520	0.780
One-sided two proportions	0.320	0.969	0.740	0.390

Regression results in table 7 confirm the null-effect and small effect sizes for all countries, except for Hungary, in which the treatment effect increases again after controlling for the region and interviewers. Thus, we reject hypothesis 2a that there is a 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	-0.016 (0.033)	-0.008 (0.049)	0.073* (0.039)	0.143** (0.060)	0.030 (0.049)	0.042 (0.070)	-0.015 (0.044)	0.007 (0.051)
DK/NA control	0.84	0.84	0.81	0.81	0.71	0.71	0.72	0.72
Region dummies	No	Yes	No	Yes	No	No	No	Yes
Interviewer d.	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-229.4	-146.7	-148.1	-87.6	-219.6	-164.1	-177.4	-141.7
McFaddens R <sup>2</sup>	0.00	0.13	0.01	0.15	0.00	0.08	0.00	0.13
Observations	500	295	346	183	373	281	296	273

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Summarized, we find that asking income in tercile bounds instead of brackets reduces overall nonresponse tremendously. It does not have any positive spillover effects on subsequent questions on exact income amounts however. It has a negative effect in Hungary, which could be driven by the fact that respondents anticipate that the tercile question is asked again if they refuse the exact amount.

## 5 Conclusion

Item-nonresponse in surveys is common and poses challenges on 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 years.

The former approach to not giving exact income amounts was to ask income from a list of very granular numerical brackets. While this approach decreased refusal rates in some of the country samples, it barely worked in the countries we consider (Bosnia and Herzegovina, Czechia, Hungary and Poland). This former approach 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.

We cannot reject our first hypothesis that 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 income. In contrast, at least in one country, refusal to give an exact income significantly increases after the treatment.

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 whether income terciles are sufficient for the research purposes the data are used for. Still, even if exact income information is necessary to obtain, 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.

Overall, there certainly is no one-fits-all solution to reducing income nonresponse. The outcome of a particular set of income questions may vary depending on a large set of factors, inter alia, where the survey is conducted, its target population, the mode of the survey or the remaining content of the survey ([Essig and Winter, 2005](#); [Moore et al., 2000](#); [Schräpler et al., 2010](#)). Still, the fact that we find similar results for four countries, lends credibility to our results in terms of external validity.

## 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.
- Beckmann, Elisabeth and Melanie Koch**, 2024, “Perceived sustainable minimum wages – a test of randomization and question scales across countries.” *Mimeo*.
- Essig, Lothar and Joachim Winter**, 2005, “Item nonresponse to financial questions in household surveys: An experimental study of interviewer and mode effects.” *Rationalitätskonzepte, Entscheidungsverhalten und ökonomische Modellierung*, 5.
- 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.
- Heeringa, Steven G. and Richard Suzman**, “Unfolding brackets for reducing item nonresponse in economic surveys.” Technical Report, National Institute on Aging, National Institutes of Health 1995.
- 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.
- Moore, Jeffrey C. and Laura Loomis**, 2001, “Reducing income nonresponse in a topic-based interview.” *Research Report Series, 2002* (06).
- Moore, Jeffrey C., Linda L. Stinson, Edward J. Welniak et al.**, 2000, “Income measurement error in surveys: a review.” *Journal of Official Statistics-Stockholm-*, 16 (4), 331–362.
- Olbrich, Lukas, Elisabeth Beckmann, and Joseph W. Sakshaug**, 2024, “Multivariate assessment of interviewer-related errors in a cross-national economic survey.” *OeNB Working Paper, 253 forthcoming*.
- 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.
- 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.

**Winter, Joachim**, “Bracketing effects in categorized survey questions and the measurement of economic quantities.” Technical Report, Universität Mannheim, Schriftenreihe Rationalitätskonzepte, Entscheidungsverhalten und ökonomische Modellierung 2002.

**Yan, Ting, Richard Curtin, and Matthew Jans**, 2010, “Trends in income nonresponse over two decades.” *Journal of Official Statistics*, 26 (1), 145–164.



# Appendix

## A.1 Additional descriptions and data on income questions

Box A.1: 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.2: Personal income question on exact amounts (only if more than one 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 A1: Income brackets respondents can choose from in control group

	BG	HU	PL	BA
1	1-400 BGN	1-80.000 HUF	1-1000 PLN	1-100 KM
2	401-600 BGN	80.001-100.000 HUF	1000-1499 PLN	101-200 KM
3	601-800 BGN	100.001-120.000 HUF	1500-1999 PLN	201-250 KM
4	801-1000 BGN	120.001-140.000 HUF	2000-2249 PLN	251-300 KM
5	1001-1200 BGN	140.001-160.000 HUF	2250-2499 PLN	301-350 KM
6	1201-1400 BGN	160.001-180.000 HUF	2500-2749 PLN	351-400 KM
7	1401-1600 BGN	180.001-200.000 HUF	2750-2999 PLN	401-500 KM
8	1601-1800 BGN	200.001-220.000 HUF	3000-3249 PLN	501-600 KM
9	1801-2000 BGN	220.001-240.000 HUF	3250-3499 PLN	601-700 KM
10	2001-2250 BGN	240.001-260.000 HUF	3500-3749 PLN	701-800 KM
11	2251-2500 BGN	260.001-280.000 HUF	3750-3999 PLN	801-900 KM
12	2501-2750 BGN	280.001-300.000 HUF	4000-4399 PLN	901-1000 KM
13	2751-3000 BGN	300.001-320.000 HUF	4400-4799 PLN	1001-1100 KM
14	3001-3250 BGN	320.001-340.000 HUF	4800-5199 PLN	1101-1200 KM
15	3251-3500 BGN	340.001-360.000 HUF	5200-5699 PLN	1201-1300 KM
16	3501-3750 BGN	360.001-380.000 HUF	5700-6199 PLN	1301-1500 KM
17	3751-4000 BGN	380.001-400.000 HUF	6200-6999 PLN	1501-1750 KM
18	4001-4250 BGN	400.001-425.000 HUF	7000-7999 PLN	1751-2000 KM
19	4251-4500 BGN	425.001-450.000 HUF	8000-8999 PLN	2001-2250 KM
20	4501-4750 BGN	450.001-475.000 HUF	9000-10999 PLN	2251-2500 KM
21	4751-5000 BGN	475.001-500.000 HUF	11000-12999 PLN	2501-2750 KM
22	5001-5500 BGN	500.001-550.000 HUF	13000-14999 PLN	2751-3000 KM
23	5501-6000 BGN	550.001-600.000 HUF	15000-17500 PLN	3001-3250 KM
24	Over 6000 BGN	600.001-650.000 HUF	Over 17500 PLN	3251-3500 KM
25		650.001-700.000 HUF		3501-3750 KM
26		700.001-750.000 HUF		3751-4000 KM
27		750.001-800.000 HUF		Over 4001 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 A2: 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 A3: 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	5000 PLN	850 KM
3. Tercile, lower bound	2,500 BGN	560,000 HUF	7,500 PLN	1,600 KM

Notes: 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 A4: Household income in local currency, 2022 versus 2023

	BG		HU		PL		BA	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Income 2022	1946	1900	404669	390001	5686	5200	1122	1000
Income 2023	2061	1901	485301	450000	6669	6250	1396	1151
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. Notes: Values are shown in local currency, as reported by the respondents. Includes answers to exact income question and control group bracketed question.

Table A5: Personal 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	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 A6: 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	2700 PLN	500 KM
3. Tercile, lower bound	1250 BGN	290,000 HUF	4,000 PLN	900 KM

Notes: 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 A7: Personal income in local currency, 2022 versus 2023

	BG		HU		PL		> BA	
	Mean	Median	Mean	Median	Mean	Median	Mean	> Median
Income, 2022	1043	900	217823	195000	2991	2900	642	> 530
Income, 2023	1178	1000	264903	250000	3475	3200	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. Notes: Values are shown in local currency, as reported by the respondents. Includes answers to exact income question and control group bracketed question.

## A.2 Additional methodological notes

### A.2.1 Randomization procedure

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

It can happen everywhere that interviewers conduct only a few interviews. Therefore, in each randomization approach, a strict alternating order between control and treatment group is enforced. This ensures that each interviewer will have a balanced mix between the two groups. The computerized approach automatically switches between groups. In Bosnia and Herzegovina and Bulgaria, group assignment will depend on whether the

interviewer types in an odd or even number. This number is determined by how many interviews the interviewer has already conducted in the specific PSU. This means that, in contrast to the computer, interviewers will restart counting after each PSU they have finished and interrupted interviews will be counted in. In the PAPI-case, questionnaires receive running numbers and are stacked in alternating order.

Randomization was tested in all countries in the 2022 wave, 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, we decided to stick to randomizing all questionnaires instead of only those for which income nonresponse occurred, except for Hungary. Therefore, in the next subsection, we also show some data on randomization for the overall samples in Bosnia and Herzegovina, Bulgaria and Poland.

## A.2.2 Descriptive statistics for all respondents in BA, BG and PL

Table A8: 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

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 refused to give an amount for household income. In Hungary, it was not randomized across all respondents.

Table A9: 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)	0.10	-0.04	0.06
Main earner in household (0/1)	0.04	-0.00	0.03
Manages household finances (0/1)	0.05***	-0.02	0.00
Respondent apprehensive (0/1)	0.02	-0.00	-0.01
Size of Household	-0.09	0.24***	0.01
Observations	1001	1012	946

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.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### A.3 Analyses on personal income questions

Charts [A1](#) to [A3](#) show that the income refusal for personal income question looks very similar to household income over the last few years.

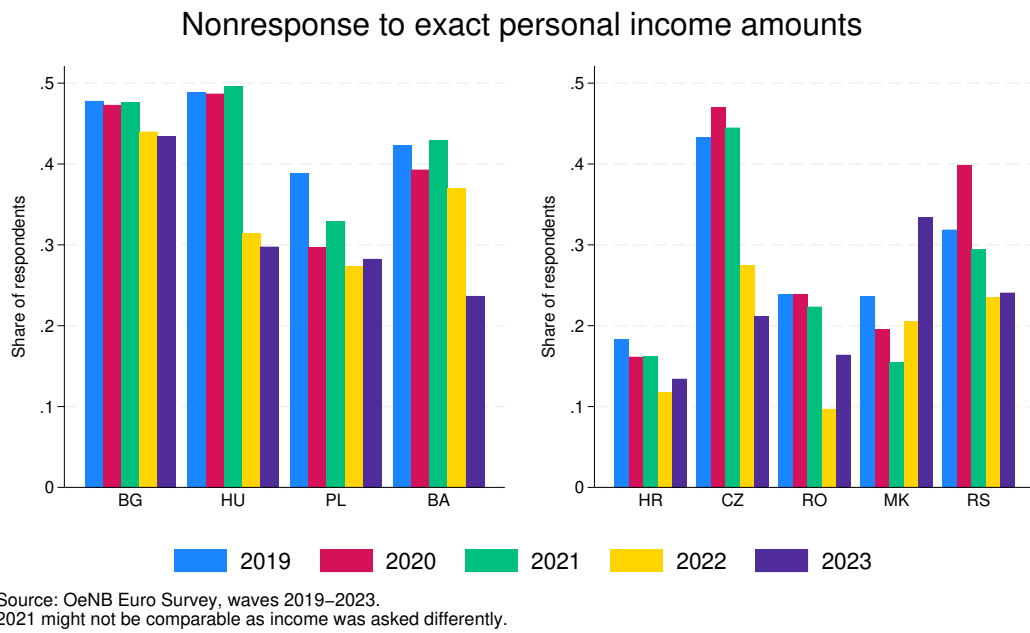


Figure A1: Respondents who refused to report the exact amount of personal income

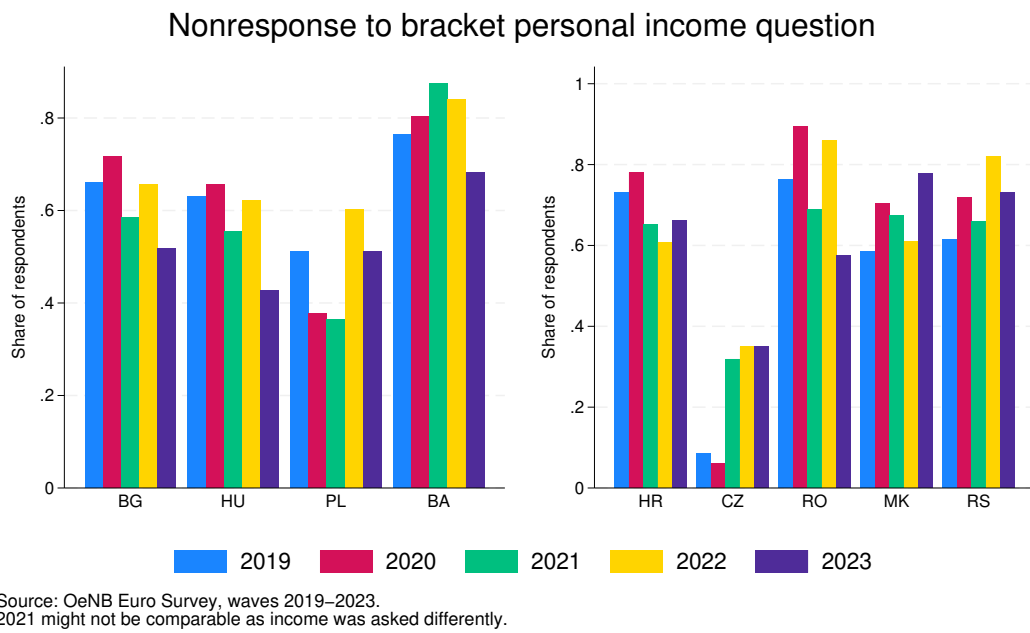
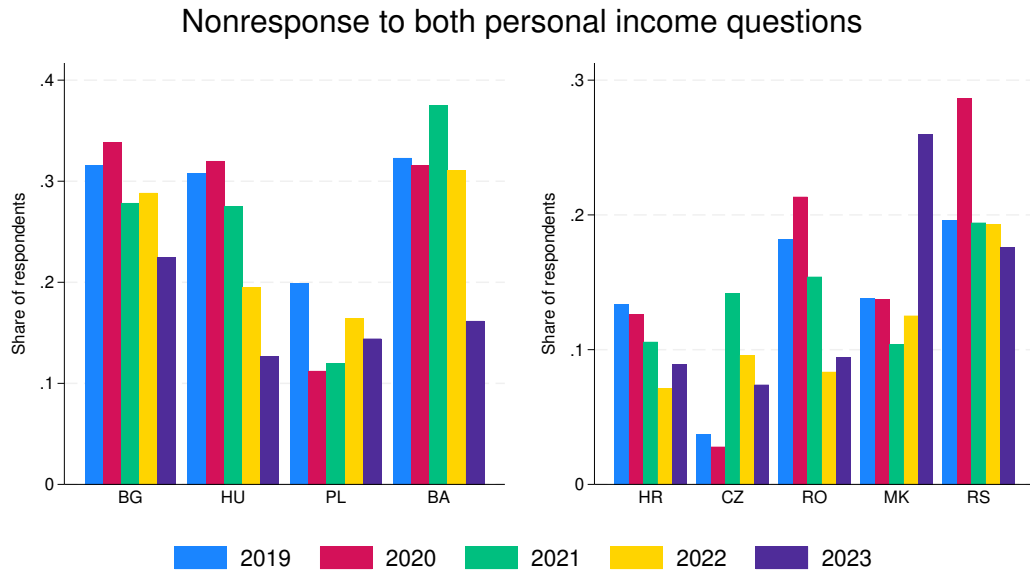


Figure A2: Respondents who refused to report personal income in brackets



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

Figure A3: Respondents who did not answer any personal income questions

The tables below show the descriptive statistics for treatment and control groups for personal income. As in section 4.2, we show descriptive statistics for the final sample that refused 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 did not receive the same order of questions, as some received 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 A10 depicts the share of respondents in each country that was 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 was already the case in the analysis on household income. In Hungary, the imbalances are mildly larger for personal income.



Table A10: 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	497	360	319	285

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 [A11](#), we show mean comparisons for selected individual characteristics across treatment groups. Most selected variables are well-balanced between treatment and control group. As in section [4.2](#), in Hungary and Poland, there is an imbalance with respect to the number of household members. Moreover, in Hungary there are imbalances related to education and in Bosnia and Herzegovina 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.

Table A11: Differences by treatment across countries

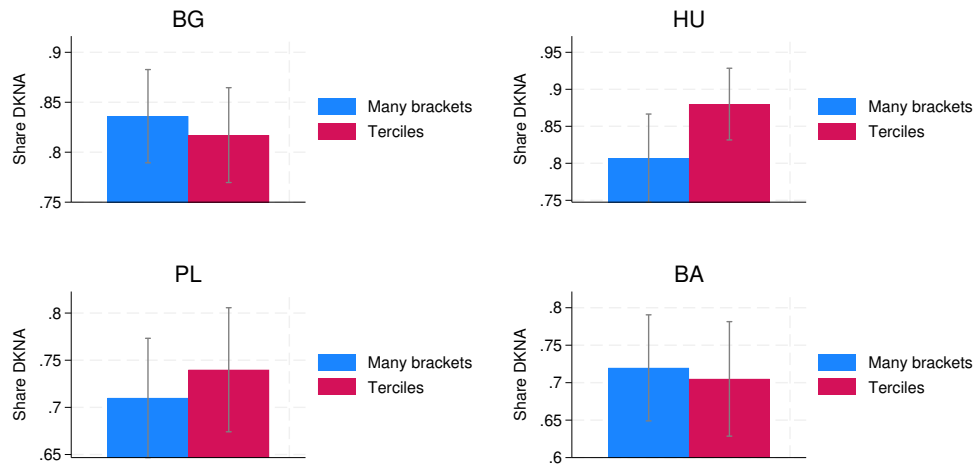
	BG	HU	PL	BA
Gender (f/m)	-0.01	-0.06	-0.03	0.07
Age (in years)	-0.78	-1.42	-0.89	-4.32**
Education (in categories)	0.12	-0.21*	0.12	-0.00
Main earner in household (0/1)	-0.02	-0.08	-0.02	0.04
Manages household finances (0/1)	0.04*	-0.02	-0.01	0.02
Respondent apprehensive (0/1)	0.04	0.08	-0.01	0.02
Size of Household	-0.00	0.22*	0.25**	0.03
Observations	497	360	319	285

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.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Figure A4 shows the nonresponse shares for the many brackets and tercile question for personal income. Table A12 presents the p-values for the one-sided Fisher's 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.

### Nonresponse to personal income



Source: OeNB Euro Survey 2023.  
 Note: Confidence intervals are obtained using two-sided *t*-tests.

Figure A4: Mean comparison personal income nonresponse between brackets and terciles

Table A12: Personal income – p-values for differences in nonresponse shares

	BG	HU	PL	BA
One-sided Fisher's exact	0.000	0.000	0.000	0.000
Chi-squared	0.000	0.000	0.000	0.000
One-sided two proportions	0.000	0.000	0.000	0.000

Treatment effects are also significant in multivariate regressions in table [A13](#). Columns (1), (3), (5) and (7) show the unconditional treatment effects, while we control for the region and the interviewer 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 26 percentage points. The effect sizes are larger than those for household income in the case of Bulgaria, while they do not differ too much for the other countries.

Table A13: Personal income nonresponse share – logit regressions with controls

	BG		HU		PL		BA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment: terciles	-0.183*** (0.038)	-0.219*** (0.043)	-0.181*** (0.048)	-0.192*** (0.066)	-0.221*** (0.049)	-0.262*** (0.068)	-0.205*** (0.046)	-0.219*** (0.051)
DK/NA control	0.67	0.67	0.64	0.64	0.68	0.68	0.81	0.84
Region dummies	No	Yes	No	Yes	No	Yes	No	Yes
Interviewer d.	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-329.8	-235.9	-242.4	-151.9	-210.2	-140.9	-151.3	-108.3
McFaddens R <sup>2</sup>	0.03	0.19	0.02	0.17	0.04	0.13	0.05	0.26
Observations	497	421	360	263	319	235	285	258

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$