

Pre-Analysis Plan:

Do talk money – Reducing income nonresponse in
surveys

Katharina Allinger* Melanie Koch†

September 2023

*Oesterreichische Nationalbank, 1090 Vienna, Austria;
Email: katharina.allinger@oenb.at

†Oesterreichische Nationalbank, 1090 Vienna, Austria;
Email: melanie.koch@oenb.at

1 Introduction

Surveys can be a convenient and cost-effective way to elicit micro-economic 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.

Over the last decades, several methods to reduce nonresponse were developed and tested. Evidently, there is not one solution that fits all. This is also the case in the OeNB Euro Survey. The Euro Survey has been conducted in ten countries for more than fifteen years and, over time, several attempts have been made to reduce income nonresponse in the survey. Being members of the research team that is running this survey, we will use an experiment, 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. Currently, if they are then not willing or not able to give an exact amount, i.e. nonresponse occurs, they are offered to choose their income from a list of very granular income brackets. These brackets can amount to over 30 in some countries. We want to test a treatment in which – instead of presenting respondents these granular brackets after their first income refusal – we simply ask them if their income falls into the first, second or third income tercile, which is predefined by us. The idea is that (i) for many research projects pursued with the OeNB Euro Survey, being able to categorize people in low, medium or high income is sufficient and (ii) respondents are more likely to answer questions that are cognitively less demanding and less privacy-invading.

In this pre-analysis plan, we explain the treatment and implementation of the survey experiment. We describe our hypotheses, outline our empirical strategy and present power calculations.

2 Research question

We want to test if nonresponse to income questions in an international survey 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 an income question which asks for exact amounts. In the survey we use, the refusal is at the moment tackled by asking a follow-up question that is meant to be less obtruding but still relatively accurate. Our control group will receive this same question as asked in previous waves:

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

Thus, control respondents will be presented lists with different income ranges, i.e. brackets. As can be seen in appendix table [A1](#), these lists have at least 24 brackets from which to choose from. They are very granular with the smallest brackets in each country capturing a range as small as 50€.

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.

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. In the following section, we explain how the tercile bounds are calculated. The final bounds are presented in appendix tables [A2](#) and [A3](#).

3 Implementation

3.1 OeNB Euro Survey

We test our treatment in four different country samples of an international 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 population and at least 1,000 individuals per country are interviewed every year by local survey institutes. The field phase of the 2023 wave is aligned across countries and starts the earliest at September 28. The planned duration for field work is between three and six weeks. We will not receive any data before early December.

3.2 Income elicitation

The questions regarding the exact amount of household and personal income have been asked in several previous OeNB Euro Survey waves. Traditionally, income questions are asked towards the end of the survey. In the 2023 wave, the same questions for the exact amount (see appendix subsection A.1) and the same ordering as for the 2022 wave will be used. This means that income questions will be the last questions respondents have to answer, that household income will be asked first and personal income will be only asked in case the respondent does not live alone but has additional household members.

In all countries, questions on exact income amounts suffer from nonresponse and are often the questions with the highest nonresponse share. However, shares differ considerably across countries. Figure 1 depicts the nonresponse share to reporting the exact amount of household income across countries and years. Refusal is either captured as “don’t know” or “no answer.” As can be seen, refusal in 2022 was especially high in Bosnia and Herzegovina and Bulgaria, where more than 50% of respondents did not give an amount for household income. The numbers are comparable or even higher for the case of personal income, in particular for the countries with the highest nonresponse shares.

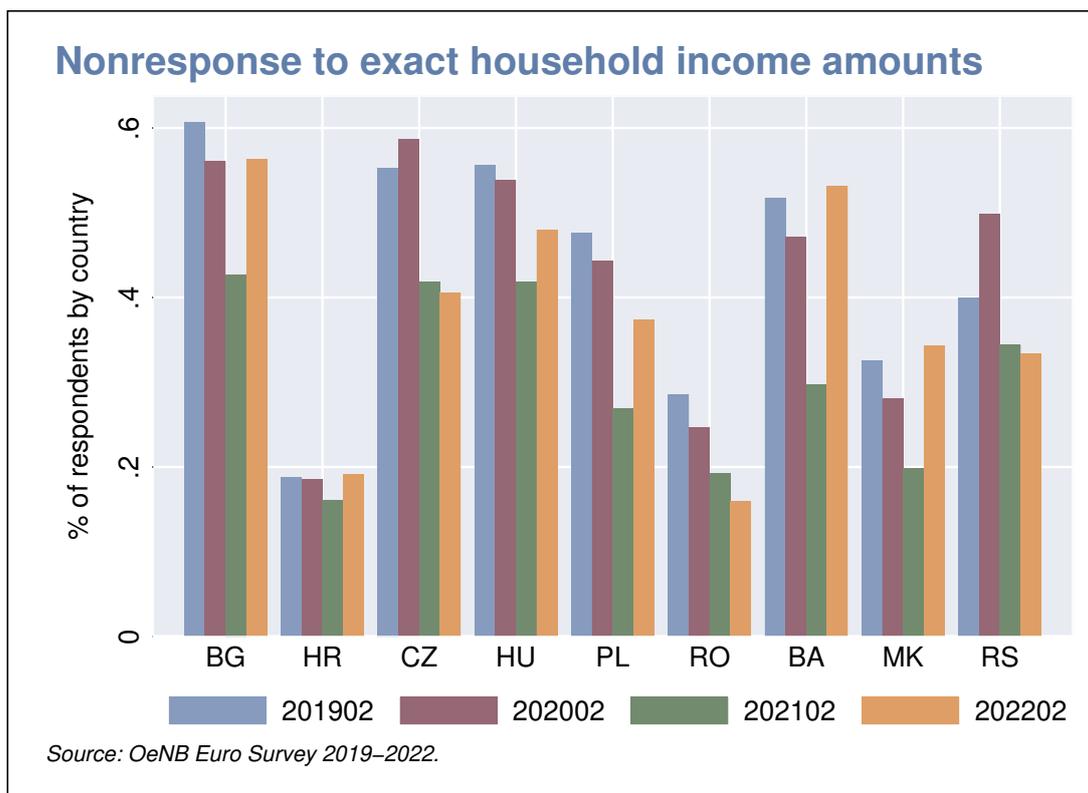


Figure 1: Respondents who refused to report the exact amount of household income

As described in section 2, nonresponse is currently 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. 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. Moreover, since the 2022 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 will keep this statement for our treatment and control groups, as it will be also used again in the other countries that are not part of our study.

In some countries, the bracket questions decrease income nonresponse tremendously. Figure 2 shows the share of the nonresponse in the bracket questions. In countries like Czechia, which in the last survey wave had a refusal rate of around 40% for exact income, 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 example in Bosnia and Herzegovina. Thus, asking respondents to choose their income in very granular brackets does not seem to work in several countries.

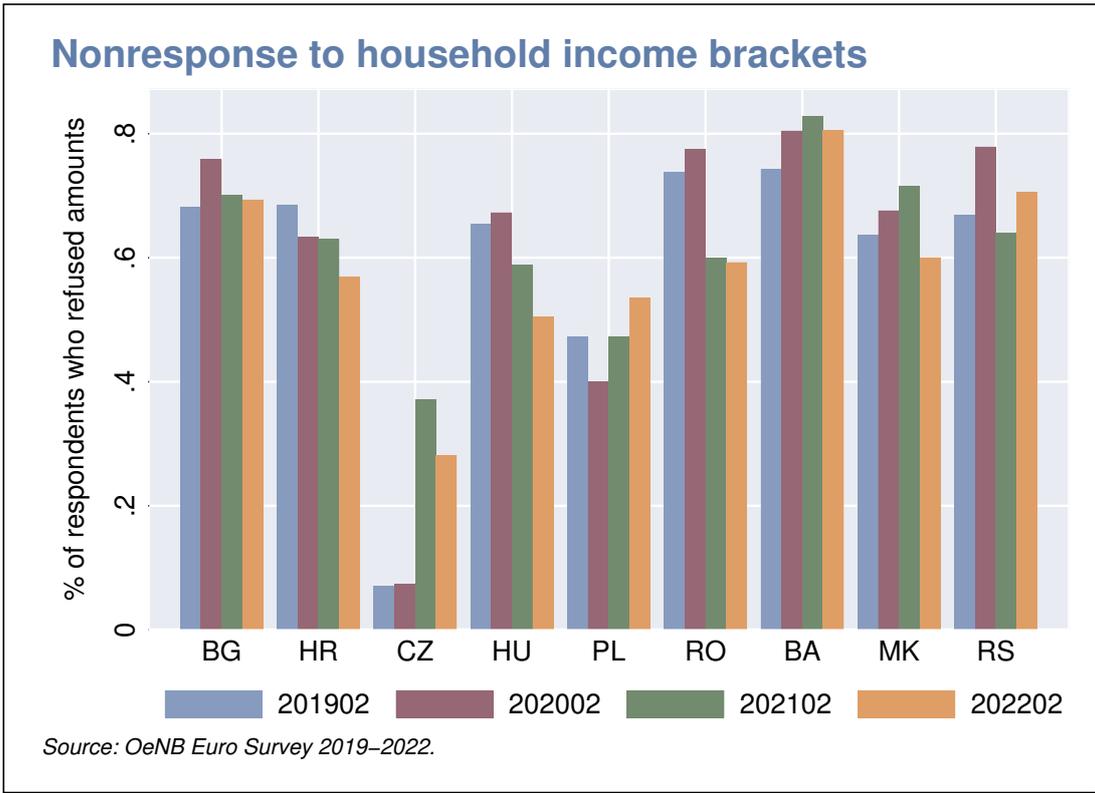


Figure 2: Respondents who refused to report household income in brackets

Figure 3 shows the combination of the two questions, namely the share of respondents in each country that did not answer any income question. Based on these findings, we have

decided to conduct the experiment in four countries: Bosnia and Herzegovina, Bulgaria, Hungary and Poland. In all these countries, nonresponse to exact income amounts is very high and asking in income brackets is only a minor or no remedy at all. The high share of nonresponse also ensures that we have enough power to test our treatment (see section 6 for power calculations). Although we will also randomize groups for the personal income elicitation, our focus is on the randomization of the household income questions, which will be asked first. Moreover, survey respondents without any additional household members do not get the personal income question at all.

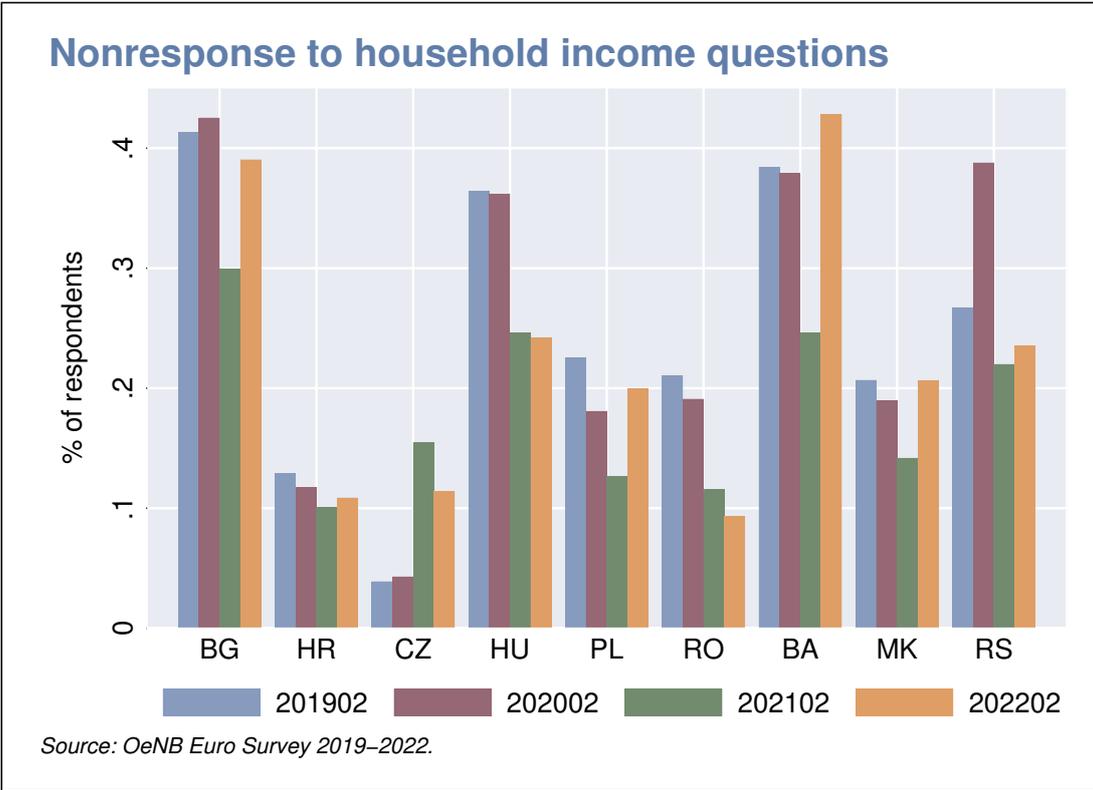


Figure 3: Respondents who did not answer any income questions

3.3 Determining the tercile bounds

The goal of our treatment is to reduce nonresponse and eventually being able to categorize respondents into three income groups. We want the three groups to at least roughly represent terciles, i.e. equally sized groups. Based on this, we calculate the tercile bounds by using 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 wage growth from year to year. For example, growth rates between 2021 and 2022 were between 8-14%. Thus, we increase previous year bounds by around 10-15% to calculate tercile bounds for 2023. 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. This makes direct comparisons more feasible and reliable as bounds are exactly the same.

3.4 Randomization

All interviews in the Euro Survey are interviewer-led and conducted face-to-face. Moreover, almost all interviews are computer assisted using tablet. From the samples in our study, only some minor fraction in Poland will be paper-based (the share of paper-based was 27% in 2022). We will not use a unified randomization mechanism across countries, but still, in each country, randomization is stratified by interviewer. 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 previous wave, using the same methods as planned for this survey wave. 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 wave 2022, all questionnaires 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. Thus, sample sizes between treatment and control could be imbalanced. However, judging by data from the full-questionnaire-randomization from wave 2022, imbalances will likely be small. At least, for our sub-sample of interest – respondents who refused to answer exact household income – the treatments in 2022 showed only minor imbalances.

4 Hypotheses

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.)

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 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. 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.

While we test for the same hypotheses in each country and expect to find a significant treatment effect in all countries, we know that it is likely that there will be differences across countries. At least, this is what we have found in previous attempts to reduce nonresponse (see, for instance, section 3.2). As we do not know why respondents have refused to answer income in the past, we refrain from guessing in which countries our treatment might be more or less effective. In general, response behavior is shaped by cultural, institutional, personal or other reasons. Regarding the personal reasons driving respondent's answer behaviour, we may look into the socioeconomic factors correlated

with different response behaviors in our treatment and control group. However, these are additional analyses that will suffer from a low number of observations and are not part of our main hypotheses and experiment.

5 Methodology

Let *treat* be an indicator variable that equals 1 if a respondent belongs to the treatment group and 0 if the respondent belongs to the control group. Thus, *treat* is missing if a respondent gives an exact amount on household income. Then, for individual $i = 1, \dots, N_c^{treat}$ in country $c \in \{BA, BG, HU, PL\}$, y_{hci} is the binary variable that equals 0 if the respondent gives an answer to either the household income bracket list or the tercile question and 1 if the respondent refuses to answer the question. Analogously, y_{pci} equals 0 if the respondent gives an exact amount on personal income and 1 in case of nonresponse. Please note again that y_{pci} will be missing for single-person households by definition. We expect that the shares of y_{hci} and y_{pci} equal to 1 are different between treatment and control group. Therefore, we conduct the following analysis:

1. For each country c , we perform Fisher's exact test, chi-square tests and two proportions tests for the equality of proportions of y_{hci} by *treat* and y_{pci} by *treat*.
2. For each country c , we run logit regressions, in which we control for region (Z_i) and interviewer (J_i) using dummy variables:

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

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

Given that our alternative hypotheses are onesided, we will look at onesided significance values, also in the regression analysis. If we then 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 sign to reject a hypothesis 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 \(2008\)](#).

6 Power

As previously mentioned, in each country, a sample of around 1,000 people will be interviewed. This total sample size is predetermined by the conditions of the contract between OeNB and its external partners. All of these respondents are asked the question on household income. However, sample sizes for treatment and control groups then depend on how many people refuse to answer the income question with exact amounts. For the power tests, we use nonresponse shares for exact household income from the 2022 wave of the Euro Survey as benchmark (see again figure 2). Since we are interested in how the share of persons who refuse to answer changes with the treatment, we estimate proportion power tests.

Table 6 reports the minimum detectable effect (MDE) sizes 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 onesided tests as our hypotheses clearly state that the shares of nonresponse are expected to decrease with the treatment. In the Table, assumed sample sizes and nonresponse proportions for the control group are given as well.

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 1: 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.

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.

Appendix

A.1 Additional descriptions on income questions

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).]

Personal income question on exact amounts (only if more than one household 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 for 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: Personal income tercile bounds for 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