NBS Working paper
09/2026

Reach Me if You Can: How to Boost Participation of Rich Households in Wealth Surveys?

Sofie R. Waltl, Judita Jurašeková Kucserová, Zuzana Brokešová, Boris Frankovič, Andrej Cupak

NATIONAL BANK OF SLOVAKIA
EUROSYSTEM

www.nbs.sk

Publisher
© National Bank of Slovakia 2026
research@nbs.sk

Contact
National Bank of Slovakia
Imricha Karvaša 1
813 25 Bratislava

Electronic version
https://nbs.sk/en/publications/research-papers-working-and-occasional-papers-wp-op/

ISSN 2585-9269 (electronic version)

Reach Me if You Can: How to Boost Participation of Rich Households in Wealth Surveys?*

Sofie R. Waltl† ⓡ Judita Jurašeková Kucserová‡ ⓡ Zuzana Brokešová§ ⓡ Boris Frankovič¶ ⓡ Andrej Cupak||

September 17, 2026

Abstract

Wealth surveys systematically under-represent affluent households because wealthy respondents are less likely to participate. We study whether targeted incentives can improve top-tail coverage in the Slovak *Household Finance and Consumption Survey* (HFCS). Using census data, we pre-identify affluent households and randomly assign them to treatment and control groups. Treated households are offered a limited-edition non-cash participation incentive: a proof-quality silver collector coin. The incentive increases participation by 11–14 percentage points, equivalent to a roughly 40% increase relative to the participation rate among non-incentivized wealthy households, with the largest effects occurring when combined with personal interviewer contact. The treatment disproportionately attracts financially sophisticated households, improving coverage of the upper tail of the wealth distribution without reducing response quality. Incorporating these additional households increases measured mean wealth, raises the wealth Gini coefficient by up to 6.5%, and improves alignment between survey aggregates and National Accounts. Our findings show that fieldwork design and participation incentives materially affect the measurement of wealth inequality and may contribute to cross-country differences in HFCS-based statistics.

JEL code: C83, C93, D31.
Keywords: wealth survey, missing wealthy households, unit non-response, over-sampling, survey experiment, incentives.

*We thank Jesse Bricker, Pirmin Fessler, Nishant Yonzan, Jochen Mankart, Jan Klacso, and Pavel Gertler for their helpful comments and suggestions. We also thank the participants at the *Slovak Economic Association Meeting*, the *2nd III/LIS Comparative Economic Inequality Conference* in Luxembourg, the *Inequalities and Opportunities: Insights into Wealth, Income, and Education Disparities* at the *University of Bari*, the *HFCN Meeting* at the *Bank of Greece*, the *International Stone Center Conference on Wealth Inequality* at *Ludwig Maximilian University of Munich*, the *2025 NOeG Winter Workshop* at the *Vienna University of Economics and Business*, the *2026 ESCoE Conference on Economic Measurement* at *King's College London*, the *2026 IARIW General Conference* at the *National Bank of Belgium* as well as seminar participants at the *National Bank of Slovakia*, the *Bratislava University of Economics and Business*, and the *Czech National Bank* for their comments. The views and results presented in this paper are those of the authors and do not necessarily reflect the official views, policies, or positions of the National Bank of Slovakia, the Eurosystem, the Statistical Office of the Slovak Republic, or Eurostat. All remaining errors are our own. ⓡ AEA Random Order Confirmation Code: `a5-Gev0srcnR`.

†Corresponding author. University of Cambridge; Vienna University of Economics and Business, e-mail: `sw2123@cam.ac.uk`.

‡National Bank of Slovakia, e-mail: `judita.jurasekova@nbs.sk`.

§Bratislava University of Economics and Business, e-mail: `zuzana.brokesova@euba.sk`.

¶Statistical Office of the Slovak Republic, e-mail: `boris.frankovic@statistics.sk`.

||National Bank of Slovakia; Bratislava University of Economics and Business, e-mail: `andrej.cupak@nbs.sk`.

Non-technical Summary

Accurate measurement of the wealth distribution is central to the analysis of inequality, the calibration of macroeconomic models, the design of fiscal and monetary policy, and the study of household portfolio choice. However, wealth surveys face a persistent challenge: affluent households are less likely to participate than the rest of the population. As a result, surveys often under-represent the wealthiest households, which may lead to an understatement of average wealth and wealth inequality.

This paper studies whether survey design can help address this problem. The analysis is based on a field experiment conducted within the 2023 Slovak Household Finance and Consumption Survey (HFCS). Using information from the national population census, households with characteristics associated with high economic status were identified in advance and oversampled. Additionally, as an experimental incentive, a limited-edition silver collector coin – issued by the National Bank of Slovakia – was offered to a randomly selected subsample of these households.

The results show that the incentive substantially increased participation among affluent households. Response rates in the incentivized group were approximately 11 to 14 percentage points higher than among comparable households that did not receive the offer. The effect was particularly strong when accompanied by personal contact from an interviewer. Importantly, the additionally incentivized respondents did not provide lower-quality information. Instead, the incentive attracted those wealthy households that were more likely to hold financial assets and display characteristics associated with greater financial sophistication.

Improved participation of wealthy households has important consequences for the measurement of wealth. When the additional respondents recruited through oversampling and offered incentives are incorporated into the sample, estimated average household wealth increases and measured wealth inequality becomes higher. The wealth Gini coefficient rises noticeably, indicating that standard survey procedures may miss a meaningful share of wealth held by households at the top of the distribution. The improved sample also brings survey-based estimates of household wealth closer to corresponding aggregates reported in the National Accounts.

The findings highlight that survey outcomes depend not only on statistical methods used after data collection, but also on the design of the data collection process itself. Measures such as targeted oversampling, appropriate incentives, and effective fieldwork strategies can materially improve the representation of wealthy households and lead to more accurate estimates of wealth and inequality. More broadly, the results suggest that differences in survey practices across countries may contribute to observed differences in measured wealth inequality, underscoring the importance of further harmonization of survey methodologies within international wealth data collection initiatives.

1. Introduction

Accurate measurement of the wealth distribution is central to the analysis of inequality, the calibration of macroeconomic models, the design of fiscal and monetary policy, and the study of household portfolio choice.[1] Across Europe, these analyses increasingly rely on the *Household Finance and Consumption Survey (HFCS)* and related official statistics, including the ECB's *Distributional Wealth Accounts (DWA)* (EG-LMM, 2020; European Central Bank, 2024). Measuring wealth, however, is particularly challenging because wealth is highly concentrated and the households holding the largest fortunes are also among the least likely to participate in surveys (Vermeulen, 2018; Kennickell, 2019; Waltl and Chakraborty, 2022). Privacy concerns, high opportunity costs of time, and limited perceived benefits of participation make affluent households especially prone to non-response (Tourangeau, 2000; Groves and Couper, 2012), and population estimates have been shown to be sensitive to assumptions regarding non-response (Heffetz and Reeves, 2019; Johnson et al., 2026).

Most existing approaches address this through *ex post* corrections, either reweighting (Cantarella et al., 2024), integrating survey data with rich lists and tax records (Vermeulen, 2018; Waltl, 2022), or reconciling survey micro-data with National Accounts aggregates (Saez and Zucman, 2016). Yet much less is known about whether the problem can be addressed *ex ante*, during the data collection process itself.

In this paper, we propose and evaluate such an *ex ante* strategy. We implement a randomized field experiment within the 2023 Slovak chapter of the HFCS (SK-HFCS). Using the Slovak *2021 Census of Population and Housing*, we pre-identify nearly 600 households with high economic status and randomly assign them to treatment and control groups. Building on standard oversampling techniques, we experimentally evaluate the additional effects of targeted incentives and interviewer contact. All selected households receive an invitation letter and a EUR 15 voucher, while the treatment group is additionally offered a proof-quality silver collector coin as a participation incentive, pre-announced in the invitation letter. The design allows us to causally identify the effects of targeted incentives on participation, respondent composition, and response quality.

We find that the incentive increases participation among treated households by 11–14 percent-

[1] See, among others, Guiso et al. (2003), Calvet and Sodini (2014), Saez and Zucman (2016), Landier and Plantin (2017), Kennickell (2019), and Waltl (2022).

age points. The effect is strongest when the coin is combined with personal interviewer contact, suggesting that incentives and interviewer effort are complementary. Treated respondents are also more financially sophisticated, while response quality remains statistically indistinguishable between the treatment and control groups.

More importantly, improved top-tail coverage materially changes measured wealth statistics. The combined oversampling and incentivization strategy increases the effective oversampling rate of the top 10% from -18 in 2021 to 41 in 2023, with oversampling and incentives contributing roughly equally to the overall improvement. Mean and aggregate wealth increase, the gap to National Accounts aggregates narrows, and measured wealth inequality rises. The Gini coefficient increases from 44.4 when refraining from any special treatment of the top tail to 45.6 with oversampling and to 47.3 when incentives are additionally employed.

These findings have implications beyond Slovakia. The HFCS harmonizes questionnaires across participating countries, but sampling and fieldwork protocols differ substantially. Our results suggest that differences in survey design are an overlooked source of variation in internationally harmonized wealth statistics. Consequently, part of the observed cross-country variation in measured wealth inequality may reflect differences in data collection protocols rather than differences in underlying economic reality.

This paper makes three contributions. First, it provides experimental evidence on targeted participation incentives for affluent households in a nationally representative wealth survey, extending the US-based evidence from Hsu et al. (2017) to a European context and to non-monetary collectible incentives. Like Hsu et al. (2017), we embed a randomized incentive experiment within an official wealth survey targeting high-income households and pre-announce the incentive in the invitation letter. We depart from their design in two key respects: we use a non-monetary collectible rather than cash, and we show that a comparably sized participation gain can be achieved at substantially lower cost in purchasing-power terms.[2] Second, it contributes to the literature on survey participation by using randomized incentives to study

[2] The acquisition cost of the silver coin was EUR 30. Slovak GDP per capita (PPP) in 2023 stood at approximately \$39,500 international dollars, compared to approximately \$80,000 for the United States (World Bank, 2024), implying a ratio of roughly 0.5. Using GDP per capita (PPP) as a proxy for average household purchasing power, an incentive of equivalent relative value to EUR 30 in Slovakia would correspond to approximately \$60 in the United States in 2023. Since Hsu et al. (2017) study the 2014 SCF, this amount corresponds to approximately \$47 in 2014 dollars when adjusted for inflation using the CPI. Their most effective treatment combination (\$5 prepaid plus \$150 postpaid, totaling \$155) yielded a broadly comparable participation gain (13.1 percentage points versus our 11–14 percentage points), but at more than three times the purchasing-power equivalent cost of our intervention. The coin's prestige and collector appeal may contribute to this cost-effectiveness, as it likely resonates more strongly with affluent households than an equivalent cash amount.

respondent selection and response quality, in the spirit of Dutz et al. (2026), who show that randomized incentive assignment can reveal and correct for non-response bias that standard adjustments fail to eliminate. We apply this logic in the specific and methodologically challenging context of a wealth survey targeting hard-to-reach affluent households, a setting where selective non-response is particularly consequential. Third, it contributes to the literature on the measurement consequences of selective non-response by showing that an *ex ante* data collection strategy can materially affect measured wealth inequality and the alignment between survey and National Accounts aggregates, complementing existing evidence on the consequences of missing wealthy households for measured wealth concentration and broader macro aggregates (Saez and Zucman, 2016; Vermeulen, 2018).

The remainder of the article proceeds as follows: Section 2 reviews survey modalities and derives hypotheses; Section 3 presents the design; Section 4 evaluates the intervention; Section 5 assesses wider implications for inequality measures and official statistics in general. Finally, Section 6 concludes.

2. The Role of Survey Modalities in Participation Decisions and Measured Inequality

Previous research identifies three key factors that shape the quality of (wealth) surveys. In what follows, we review these findings and draw on them to inform the design of our experiment as well as our assessment dimensions.

2.1. The Role of (Over-)Sampling Strategies

Prior research highlights the importance of pre-identifying and oversampling wealthy households to improve their representation. The *Survey of Consumer Finances (SCF)* and the *Spanish Survey of Household Finances (EFF)* are the leading examples: both use administrative data on income or wealth to pre-identify and oversample affluent households, and both have been shown to capture the top of the wealth distribution substantially better than surveys relying on standard sampling frames (Kennickell, 2008; Bover, 2008; Vermeulen, 2018).

More systematically, Waltl (2022) benchmarks HFCS aggregates against National Accounts across five countries and finds that the gap between survey and National Accounts wealth is smallest in countries that oversample using household-level register data (Spain, France) and largest where no oversampling is applied (Austria). These results demonstrate that, despite harmonized questionnaires, heterogeneous data-collection protocols lead to substantial differences in data quality, adversely affecting the validity of cross-country comparisons.

Our study adopts a similar household-level pre-identification approach using census data, and we assess its effectiveness in improving top-tail coverage relative to previous SK-HFCS waves and to other HFCS countries.

2.2. The Role of Survey Mode

Previous research shows that survey mode significantly affects both response behavior and data quality (Essig and Winter, 2009; Lohmann, 2011; Fessler et al., 2018). In-person interviews (CAPI) typically yield higher response rates but may increase social desirability bias, whereas non–in-person modes (e.g., self-administered questionnaires or telephone interviews) can reduce sensitivity-related non-response but may introduce greater measurement error. For example, moving financial questions away from in-person settings reduces item non-response and improves precision for such questions (Essig and Winter, 2009), while switching entirely from in-person to telephone interviews increases both the total non-response rate and measurement error, thereby biasing derived inequality measures (Fessler et al., 2018). Overall, in-person and non–in-person modes involve a trade-off between response rates and data accuracy.

While our study relies exclusively on in-person interviews, we examine how response behavior is influenced by the successful establishment of personal contacts.

2.3. The Role of Incentives

Previous research shows that incentives can effectively increase participation in household finance and wealth surveys and reduce non-response bias (Singer and Ye, 2013; Schröder et al., 2013). However, their impact varies depending on the type, value, and timing of the incentive, as well as the survey mode (Groves and Couper, 2012; Singer and Ye, 2013).

The literature has focused predominantly on monetary incentives (Schröder et al., 2013; Stecklov et al., 2018; McGonagle, 2020; Chang et al., 2026), finding that cash, larger amounts, and prepaid structures are generally most effective (Pforr et al., 2015).[3] In the context of the SCF specifically,

[3] A meta-analysis of monetary incentive effects in household surveys can be found in Mercer et al. (2015).

Bricker (2014) shows that higher monetary incentives increase not only respondent motivation but also the effort interviewers invest in recruiting reluctant households, suggesting that incentive effects operate through both respondent motivation and interviewer behavior. Evidence on the ability of incentives to reduce non-response *bias* is more mixed: while incentives increase participation, they do not consistently attract a more representative set of respondents. Socio-economically disadvantaged groups tend to be most responsive to monetary incentives (Pforr et al., 2015), suggesting that standard cash incentives may be poorly suited to recruiting wealthy households, for whom the marginal utility of a modest payment is low. Beyond participation rates, Dutz et al. (2026) show that randomized incentive assignment can be used to study *who* participates: exploiting exogenous variation in incentives in a Norwegian labor-market survey linked to administrative records, they find that substantial selection into surveys remains even after standard adjustments.

Our design is most directly motivated by Hsu et al. (2017), who conduct the closest existing study to ours. They embed a randomized incentive experiment in an experimental version of the SCF, targeting 900 households drawn from census tracts in the top income quintile in New York City, Los Angeles, and Miami. Like our study, the incentive is pre-announced in the invitation letter and the sample is restricted to high-income areas. Their design varies both the presence of a small prepaid cash incentive ($5) and the size of a promised postpaid incentive ($50, $100, or $150). The $5 prepaid incentive raises participation by 6.5 percentage points; the $150 postpaid incentive adds a further 7 percentage points relative to the $50 baseline; and combining both yields the strongest effect of 13.1 percentage points.

We extend this evidence in two directions: we test a non-monetary collectible incentive rather than cash, and we do so in a European context.

3. The Survey Experiment

3.1. Survey Setting and Sample Construction

We implement the study in the fifth wave of the Slovak Household Finance and Consumption Survey (SK-HFCS), a population-representative household survey. The primary objective of the SK-HFCS is to collect detailed micro-data on households' finances and consumption, assets, liabilities, and income. As part of the broader Household Finance and Consumption Network, the SK-HFCS contributes to cross-country analyses within Europe and the construction of euro area aggregates. In Slovakia, the survey is conducted by the *National Bank of Slovakia*. The fifth wave was fielded between October 2023 and January 2024. Data are collected via computer-assisted, face-to-face interviews (CAPI), and the survey meets quality standards as discussed in Section 2.2.

Our design combines census-based oversampling of affluent households with a randomized incentive treatment. The oversampling strategy is implemented through a dual-frame sampling design, described in Appendix A. This approach yields two pools from which participants are recruited. Recruiting from the unrestricted frame via standard stratified random sampling yields the *General Population Pool (GP)*. [4] By that, each household had a positive probability of being contacted.

To improve coverage of the upper tail of the wealth distribution, we additionally construct the *Rich Population Pool* (RP) using a restricted sample from the Slovak 2021 Census of Population and Housing[5] as a sampling frame. The restrictions are designed to identify households that are likely to be wealthy based on individual and housing characteristics reported in the census that serve as proxies for wealth. Households feeding the (RP) are then sampled from this restricted frame. As discussed in Section 2.1, this census-based oversampling strategy is expected to improve the precision of wealth estimates relative both to previous SK-HFCS waves and to surveys in countries that rely on less precise household-level information for oversampling, or do not oversample affluent households at all.

Assignment to the (RP) is subject to the following set of binding side constraints, all of equal importance, designed to identify households with high economic status:

(1) being an owner-occupier,

(2) residing in a single-family house with a minimum of 200 square meters living area, and built in or after 2001,[6]

(3) being at least 40 years old,[7] and

[4]Precisely, this step consists of a stratified random draw from the list of all districts in Slovakia with the binding side-constraint to sample from each NUTS3 region.

[5]See `https://www.scitanie.sk/en/explanatory-notes` (last accessed in September 2024) for methodological details and general results.

[6]In the Slovak housing market, newer buildings are on average much more valuable than very scarce expensive old ones that would benefit from the well-documented vintage effect in other European cities (Rolheiser et al., 2020). The scarcity of such buildings is a consequence of large-scale destruction in the 20th century.

[7]This restriction applies to the household reference person as defined in the census. The census defines the reference person as the individual with the highest level of educational attainment; in the case of ties, the individual with the higher labor market status is selected.

(4) holding a high occupational status.[8]

The gross sample drawn from $(GP) \cup (RP)$ aimed for 3,500 observations.

The SK-HFCS includes a panel component, meaning that a subset of households participating in one wave is re-invited in subsequent waves. As a precaution, we exclude all these panel households from the (RP) sample pool. This ensures that the estimated effects of incentives are not confounded by households' prior survey experience. Moreover, panel households have already demonstrated a willingness to participate and are therefore not representative of the broader population, which comprises both inherently willing households and those that are reluctant or require incentives to participate.

The resulting (RP) sample forms a targeted pool of wealthy households within which the randomized incentive treatment described below is implemented.

3.2. Participation Incentive

To test whether additional participation incentives would increase participation among over-sampled wealthy households, we implement a randomized experiment. Following the discussion in Section 2.3, we expect higher participation rates among incentivized households.

The experiment is based on a between-subjects design, using households sampled from (RP).[9] For operational reasons, the experiment was conducted in four of Slovakia's eight NUTS 3 regions, covering the western, central, and eastern parts of the country. The targeted household sample size for (RP) was 598[10]. Within (RP), we randomly established a control group (299 households) and a treatment group (299 households).[11] Randomization was carried out by the *Slovak Statistical Office* at the household level using their standard random sampling algorithms: within (RP), each of the 598 households was independently assigned to the treatment or control

[8]'High occupational status' is defined on the basis of average annual earnings by occupation, using data from the EU-wide *Structure of Earnings Survey (SES)*. The classification includes senior managerial positions, a broad range of professional and technical occupations – particularly in science, engineering, ICT, finance, law, and medicine – as well as selected supervisory and associate professional roles. In addition, self-employed individuals operating as legal entities (entrepreneurs) are included in the high occupational status group, following Schröder et al. (2020).

[9]The experimental design was approved by the Ethics Committee of the Bratislava University of Economics and Business on June 23, 2023 (approval number EKEUBA-IV1-8/2023) and pre-registered in the American Economic Association's AEA RCT Registry on October 17, 2023 under the title 'Increasing respondents' willingness to participate in questionnaire surveys" with the RCT ID AEARCTR-0011804.

[10]The original sample size was set at 600. During the data editing process, however, two households were found to be incorrectly assigned to control and treatment groups and were therefore excluded from the analysis.

[11]Based on power analysis calculations, our sample size allows us to detect a medium effect of the intervention relative to the control group, with more than 91% statistical power.

arm. The resulting treatment and control households are geographically dispersed across Slovak districts (see Figure E.1). In larger, urban districts both treated and control households are present, while several smaller districts happen, by chance, to contain sampled households from only one arm. The resulting geographic distribution is depicted in Figure E.1 in Appendix E.

All selected households received an official letter (see Figure E.2) inviting them to participate in a face-to-face survey interview. The letter was presented in an official format, with its credibility reinforced by the signatures and logos of the National Bank of Slovakia and the Slovak Statistical Office. The latter ensured authority as an important element in supporting participation in surveys (see Groves and Couper, 2012; Osier, 2016).[12] The letter served to pre-notify households of their selection for the HFCS. It informed recipients that an interviewer would contact them shortly and provided contact details for any further questions. In addition, the letter pre-announced a EUR 15 participation voucher for all participating households and explained the purpose and importance of the survey.

The control group received the standard invitation letter. The treatment-group letter was nearly identical in content and appearance, but additionally informed recipients that they would receive a proof-quality silver euro collector coin as a reward for participation. This information was prominently highlighted to attract attention and increase salience (see Figure E.2).

We selected a collector coin as the high-value, non-cash incentive for wealthy households. Given that the opportunity cost of survey participation is likely to be substantially higher for this group than for the general population, wealthy households may be less responsive to small cash incentives (see, e.g., Hsu et al., 2017). In contrast, collector coins may have a higher perceived value owing to their artistic design, rarity, and potential collectability, with their market value determined in part by demand among collectors. At the same time, their acquisition cost shortly after minting largely reflects their intrinsic material value, making them a cost-effective option for the experiment.

Specifically, we chose a silver collector coin issued in May 2023 by the National Bank of Slovakia (see Figure 1), with a face value of EUR 10.[13] The total issuance of this coin was limited to 7,200 pieces. The purchase price of the coin for the experiment was EUR 30, and after two years, its

[12] This is the standard form of communication used to request participation in the HFCS and has also been used in all previous SK-HFCS waves.

[13] See `https://nbs.sk/en/banknotes-and-coins/euro-coins/collector-coins/100th-anniversary-of-the-start-of-regular-broadcasting-by-czechoslovak-radio/`, last accessed in September 2025.

Figure 1: Coin Used as an Incentive

Notes: Collector silver coin '100th anniversary of the start of regular broadcasting by Czechoslovak Radio.'

market value had already tripled, likely due to scarcity. Consistent with the argument that the market value of the coin is not fixed, its value was intentionally omitted from the letter.

As noted above, all participants who completed the survey received a EUR 15 voucher as compensation for their time, regardless of group assignment. For participants in the treatment group, the total incentive therefore consisted of both the voucher *and* the collector coin.

3.3. Validation of Targeting and Balance

We assess two conditions required for the design. First, the census-based targeting strategy should identify households that are on average more affluent than households selected through the general population frame. Second, within the sample drawn from (RP), random assignment should generate comparable treatment and control groups in the gross sample, before participation decisions are realized.

To validate the targeting strategy, we compare households entering the survey through the (RP) with those entering through (GP). We expect those entering through the (RP) to be, on average, more affluent and to have a higher economic status as a consequence of the pre-identification strategy. Appendix A supports the validity of the pre-identification: In the gross sample, households sampled from the (RP) pool are rated by interviewers as living in higher-quality dwellings and neighborhoods than households sampled from the (GP). In the net sample, households entering the survey through the (RP) report higher net wealth, financial assets, housing wealth, income, and consumption. This confirms that the census-based restrictions successfully distinguish *ex ante* between more and less affluent households.

We next assess whether randomization generated comparable treatment and control groups within the gross (RP) sample before participation decisions are realized. We expect no *ex ante* differences between the two groups (part of (RP)) in the gross sample – both in subjective and objective wealth measures, which are indeed confirmed in Appendix B.

4. Impact Assessment of the Intervention

This section evaluates the success of the intervention by examining realized and perceived increases in participation rates between the treatment and control groups, comparing top-tail coverage across survey waves and relative to other HFCS countries, identifying the socio-economic characteristics associated with responsiveness to the incentive, and assessing the implications for survey data quality.

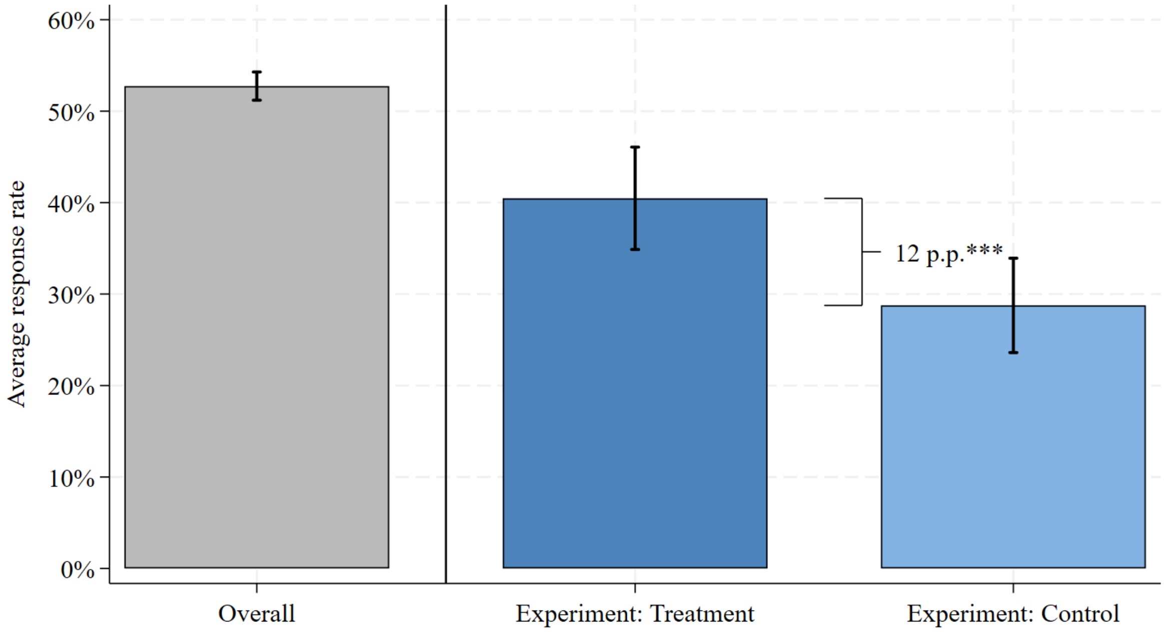
4.1. Likelihood of Participation

We next evaluate the main component of our intervention by comparing participation rates within the (RP) sample between households assigned to the treatment and control groups. We proceed in three steps. First, we estimate the treatment effect by comparing participation rates between the treatment and control groups. Second, we use a regression framework to examine the role of additional survey design features and interview modalities that may influence participation decisions. Finally, we evaluate the effectiveness of the incentivized oversampling strategy by comparing its outcomes with those of previous survey waves and with those observed in other HFCS countries in the current wave.

In a first step, we compute the raw response rates shown in Figure 2. The bar on the left reports the response rate for the full sample, i.e., $(GP) \cup (RP)$, while the two bars on the right distinguish between the treatment and control groups within (RP). The well-documented pattern of lower participation rates among wealthy households is evident once again. At the same time, the use of incentives appears to recover a substantial share of otherwise lost observations. Specifically, we find a statistically significant difference of 12 percentage points in the raw participation rate between the treatment and control groups, which can be attributed solely to the coin incentive.

Next, we analyze the impact of incentives on participation rates within a regression framework to understand several factors potentially increasing or decreasing the likelihood of participating. Therefore, we set up a probit model with the standard normal CDF $\Phi(\cdot)$ as the link

Figure 2: Average Unit Response Rates in the Slovak HFCS



Notes: This figure presents the average unit response rates in the HFCS along with the 95% confidence intervals.
Source: HFCS 2023/24, National Bank of Slovakia.

function,

$$\Pr(Y_{irk} = 1|X_i, \lambda_r, Z_k) = \Phi(\beta_1 COIN_i + \beta_2 VISIT_i + \beta_3 COIN_i \times VISIT_i + \gamma X_i + \lambda_r + \xi Z_k), \quad (1)$$

where Y_{ikr} is an indicator that takes the value 1 if household i chooses to participate in the survey, and 0 otherwise. $COIN_i$ is an indicator distinguishing treated ($COIN_i = 1$) and control ($COIN_i = 0$) households, and $VISIT$ is an indicator taking the value 1 if an interviewer has – after having sent an invitation letter but receiving no response – personally visited the household to pursue them to participate in the survey. The matrix X_i collects survey administrative modalities affecting participant i's recruitment with associated parameters γ. Region fixed effects are denoted by λ_r.

Given that the HFCS fieldwork is conducted by a large group of interviewers ($n = 150$), concerns may arise regarding potential interviewer effects. Crossley et al. (2021) distinguish three main types of interviewer effects: unit non-response, whereby interviewers differ in their success rates in recruiting respondents; item non-response, whereby interviewers differ in their ability to obtain answers to specific questions; and response effects, whereby interviewers influence the responses provided by participants. To account for such effects, we explicitly control for the characteristics of the interviewer (years of interviewing experience and highest educational attainment) denoted by Z_k and associated parameters ξ.

Overall, our empirical strategy follows a stepwise specification approach to evaluate the effects of both the coin incentive and the personal visit on survey participation. We then assess the robustness of the estimated effects to the inclusion of successive sets of control variables. Finally, we introduce interaction terms between the key treatment variables to identify the recruitment strategy associated with the highest participation rates.

Results are reported in Table 1. Model (1) confirms the previous univariate test: the propensity to participate in the survey increases by roughly 12 percentage points when being promised a coin. This result is highly statistically significant.

Model (2) assesses whether personal contact is key to successful recruitment. In terms of magnitude, this feature is even more important than the coin as evidenced by the large and statistically significant effect size of 44 percentage points This is consistent with Bricker (2014), who shows that interviewer effort plays an important independent role in recruiting reluctant households in the SCF.

Model (3) combines both factors within a single specification to assess whether they offset one another or independently affect the participation decision. The results indicate the latter, as the estimated effect sizes remain largely unchanged relative to models (1) and (2). This further suggests that the coin increases participation above and beyond traditionally used measures aimed at improving response rates.

Model (4) shows that the likelihood of participation declines with the number of visit attempts. This suggests that, for some households, the incentive is insufficient to compensate for the required time and effort, while others may have more general reservations about participating in the survey. Reaching this group may therefore require alternative recruitment strategies.

Models (5) to (7) add sets of controls to assess the general robustness of effects. Model (5) adds some survey administration information (namely the daytime and location of initial contact), model (6) region fixed effects, and model (7) interviewer characteristics. Regarding model (5), we tested a range of alternative functional forms and specifications. None of the characteristics was statistically significant on its own, irrespective of the specification employed.[14] A χ^2-test confirms that all interviewer effects are jointly insignificant (χ^2_4=4.40, p-value= 0.3544). Likewise, the regional fixed effects added to model (6) are jointly statistically insignificant at conventional levels (χ^2_6= 11.43, p-value= 0.0761). Model (7) clusters standard errors at the interviewer-level, yet this also does not change the results, further demonstrating the robustness. Model (8) further reinforces the findings from model (3): the effect on participation is largest when offering a coin *and* establishing a personal contact. The two aspects are not substitutes, but each aspect adds towards an increased participation probability as shown by the strongly significant main and interaction effects. Similarly, model (9) reiterates that increasing the number of contact attempts has a negative effect and confirms that offering a coin does not mitigate this effect.

Table 1: Probit Estimates of the Probability of Survey Participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
COIN	0.117***		0.115***	0.109***	0.111***	0.135***	0.136**	0.118**	0.136**
	(0.039)		(0.039)	(0.040)	(0.040)	(0.047)	(0.057)	(0.058)	(0.056)
VISIT		0.435***	0.433***	0.447***	0.443***	0.450***	0.448***	0.346***	0.448***
		(0.059)	(0.060)	(0.062)	(0.063)	(0.064)	(0.080)	(0.041)	(0.080)
Number of contacts				-0.122***	-0.118***	-0.126***	-0.115***	-0.113***	-0.115***
				(0.032)	(0.033)	(0.034)	(0.041)	(0.041)	(0.041)
COIN x VISIT								0.168***	
								(0.060)	
COIN x Number of contacts									-0.018
									(0.061)
Survey controls	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes	Yes	Yes	Yes
Int. characteristics	No	No	No	No	No	No	Yes	Yes	Yes
Interviewer-level clustered S.E.	No	No	No	No	No	No	Yes	Yes	Yes
Pseudo R^2	0.012	0.080	0.092	0.109	0.111	0.118	0.128	0.131	0.128
N	598	598	598	598	598	598	598	598	598

Notes: The table reports estimation results for the probability of participating in the survey estimated on the gross sample following model (1). Marginal effects at the mean (MEM) are reported. Robust/clustered S.E. are reported in parentheses. Survey controls include information about the daytime of being contacted and its square as well as the rural/urban classification indicator. Regional fixed effects are captured by a set of dummy variables for 8 regions in Slovakia. Interviewer characteristics include age, age squared, level of education, and years of experience.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Source: HFCS 2023/24, National Bank of Slovakia.

Because the survey is conducted exclusively in CAPI mode, we cannot directly estimate the effect of interview mode. However, previous research reviewed in Section 2.2 suggests that face-to-face interviewing is particularly effective in encouraging participation among affluent households. We therefore examine whether pre-announcing the coin in the invitation letter

[14]While, as expected, a positive parameter is estimated for age and a negative one for age squared, both estimated parameters are not statistically significantly different from zero at conventional levels. Also, an alternative specification of the experience relying on a dummy distinguishing between little experience (less than one year) and substantial experience (more than one year) did not alter results.

is especially effective when the interviewer's first contact with the household takes place in person. Model (8) hence augments the baseline specification by an interaction term[15] between the treatment dummy and the way of first-time contact, i.e., personal visit versus the base category of any other mode such as telephone contact. Indeed, a personal contact on top of the incentive boosts the likelihood of participation.

On the other hand, as shown before in model (4), the more attempts are taken to successfully contact a household, the less likely they participate at the end. We also interact this effect with $COIN$ to understand whether there is heterogeneity between the treated and control groups. Model (9) indicates no statistically significant difference in the relationship between the number of contact attempts and participation probabilities across the treatment and control groups. While the estimated coefficient is slightly larger for the control group, the effect size is small and statistically insignificant.

Overall, we document a sizable positive effect of the incentive on the likelihood that affluent households participate in the survey, consistent with the prior literature reviewed in Section 2.3. These results compare favorably with Hsu et al. (2017), the closest existing study. Their most effective treatment – combining a $5 prepaid cash incentive with a $150 postpaid offer – raises participation by 13.1 percentage points among high-income households in the U.S. SCF. Our silver coin, with an acquisition cost of EUR 30, achieves a broadly comparable gain of 11–14 percentage points. Beyond the similar effect size, our incentive operates through a single prepaid component with no postpaid element, and targets a European rather than a US context. Together, these results suggest that non-monetary collectible incentives can be as effective as large cash payments in recruiting hard-to-reach wealthy households, at substantially lower relative cost.

Further, our analyses immediately lead to the following recommendations. First, an in-person survey mode should be preferred over other forms. Second, a personal contact should be established as soon as possible, while excessive contact attempts appear to be redundant. Third, the use of incentives and their prominent pre-announcement can boost participation of traditionally under-represented wealthy households in such surveys. The incentive, in the form of a rare and valuable coin, proves to be a valid choice for attracting wealthy households. Finally,

[15] For our interaction terms, we follow Imbens and Rubin (2015) in de-meaning continuous variables when interacting them with our treatment effect, and the approach of Ai and Norton (2003) to compute the correct marginal effect of a change in two interacted (dummy) variables for logit and probit models, as well as the correct standard errors.

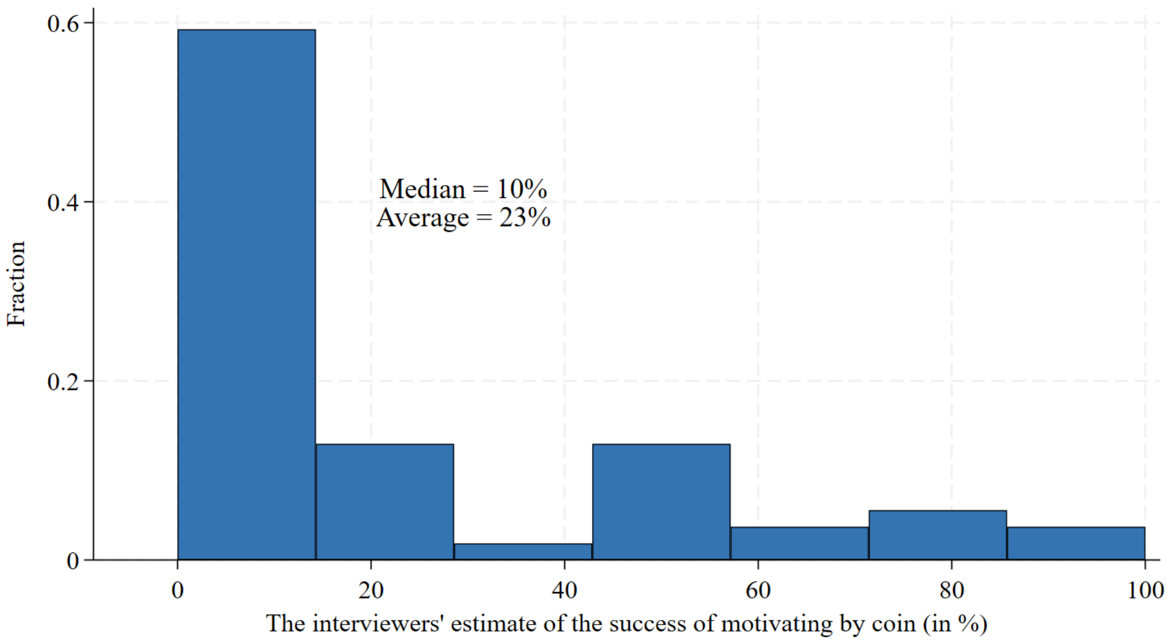
personal contact and incentives are not substitutes, but work best if applied jointly.

4.2. Interviewers’ Perceived Impact of the Coin

So far, we have focused on *revealed outcomes*, specifically participation decisions, to assess the success of our intervention. To complement these findings, we also consider *stated perceptions* regarding the role of the coin as a participation incentive. In particular, we utilize a survey question posed to interviewers after the completion of the main survey, designed to gauge the perceived impact of the coin:

> *Q: Please enter your estimate of the percentage by which the offered coin helped motivate these selected households to participate in the survey [... %]*

Figure 3: Interviewers’ Perceptions of the Coin’s Impact on Household Participation Decisions



Notes: This figure presents the distribution of answers to interviewers’ perceived success of the coin.
Source: HFCS 2023/24, National Bank of Slovakia.

A total of 54 interviewers (who worked with households from the experimental sub-sample) responded to this question, producing the subjective success distribution shown in Figure 3. This yields a median success rate of 10% (and a mean of 23%). These subjective probabilities closely resemble realized effect sizes computed from households’ behavior, strengthening the internal validity of our study design.

4.3. Differential Household Responses to the Incentive

While the coin clearly increases participation rates among wealthy households, as shown in Section 4.1, it is important to assess whether it also affects the composition of respondents. Although balance is mechanically ensured in the gross sample, it is not clear *a priori* whether the incentive is equally effective across different household types.

We therefore conduct two sets of tests. First, we examine selection into the survey by comparing the socio-economic characteristics of the gross and net samples. Second, we assess whether the incentive changes the composition of respondents by testing for differences between the treatment and control groups within the net sample. In addition to standard socio-economic characteristics, we pay particular attention to indicators that proxy for financial sophistication.

By first comparing socio-economic characteristics between the gross and net samples, we assess whether types of households generally deciding to participate differ from those not willing to do so. We therefore compare summary statistics of the gross (Table B.1) and the total net (Table 2) samples. Both the gross and net samples report a stark gender imbalance in line with the known under-representation of women at the top of the wealth distribution (Schneebaum et al., 2018). The imbalance is even more pronounced in the net sample. Further, households with more children are less often represented in the net sample than in the gross sample, while households headed by entrepreneurs (according to our classification) are more often found in the net sample than in the gross sample, suggesting that – regardless of being incentivized or not – among wealthier households, female-led households as well as those with more children are less likely to participate. In contrast, households led by a highly educated person – while mechanically over-represented in the gross sample due to our identification strategy – are generally less likely to participate, yet can be motivated by incentives as discussed below. While average age appears to be somewhat higher in the net sample than in the gross sample, this effect must be attributed to the time lag between the census interview and the SK-HFCS interview and is hence a mechanical feature of our data.

Next, we assess the net sample in more detail and test for differences in socio-economic characteristics between the treatment and control groups. While Table 2 finds no statistically significant differences in *means* for most characteristics – including age, gender, number of children, entrepreneurial status, net wealth, total assets, financial assets, housing wealth, income, and consumption – some differences emerge along other dimensions. In particular, the promised coin appears to attract particularly highly educated households and those more likely to hold shares or investment funds as their share is statistically significantly increased. Consistent with this, we also observe a small but systematic upward shift in medians of net wealth, total assets, and gross household income, as confirmed in Table 2.

Table 2: Balance within the (RP) Net Sample by Treatment

	Control (N = 86 of 299)			**Treated** (N = 121 of 299)			**Differences in Means**	**Differences in Medians**
	Mean/Share	Median	SD	Mean/Share	Median	SD	(p-value)	(p-value)
Net Wealth [in 1000 EUR]	596.6	335.3	(1229.3)	656.6	508.0	(651.9)	60.0 (0.650)	172.7** (0.026)
Total Assets [in 1000 EUR]	635.6	355.0	(1230.5)	698.9	531.0	(651.6)	63.3 (0.632)	176.0*** (0.009)
Total Financial Assets [in 1000 EUR]	33.4	14.7	(56.0)	37.8	20.0	(44.3)	4.3 (0.535)	5.3 (0.140)
Value of Main Residence [in 1000 EUR]	315.1	250.0	(193.3)	333.6	280.0	(178.6)	18.6 (0.480)	30.0 (0.198)
Floor area [sqm]	193.6	180.0	(76.5)	198.0	185.0	(84.0)	4.4 (0.702)	5.0 (0.804)
Has Shares or Funds	0.2		(0.4)	0.3		(0.5)	0.1* (0.095)	
Gross Annual Household Income [in 1000 EUR]	54.9	42.1	(49.8)	59.0	50.0	(36.1)	4.1 (0.491)	7.9** (0.042)
Monthly Consumption [EUR, equiv.]	893.7	805.9	(469.2)	841.9	808.6	(337.9)	-51.8 (0.356)	2.7 (0.875)
Gender = male	0.7		(0.4)	0.7		(0.4)	0.0 (0.976)	
Age [years]	51.5	50	(9.0)	52.5	51	(8.9)	1.0 (0.444)	1 (0.609)
Highest Education Level = yes	0.5		(0.5)	0.7		(0.5)	0.1* (0.087)	
Number of Children	1.1	1	(1.1)	1.0	1	(1.0)	-0.1 (0.450)	0 (0.533)
Entrepreneur = yes	0.3		(0.4)	0.3		(0.5)	0.0 (0.433)	

Notes: Based on households entering the HFCS via the (RP) pool. Reported p-values refer to t-tests for differences in means for continuous variables, Z-tests for shares, and to Wilcoxon signed-rank test for differences in medians. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Census 2021, Statistical Office of the Slovak Republic; HFCS 2023/24, National Bank of Slovakia.

Taken together, these findings suggest that the coin disproportionately attracts more financially sophisticated households. As these differences are present in medians but not in means, they indicate a shift in the composition of respondents toward greater financial sophistication rather than an increase in aggregate top wealth. Put differently, the results demonstrate that a valuable collector coin is an effective incentive for boosting participation among otherwise hard-to-reach wealthy and financially sophisticated households. This finding is consistent with Dutz et al. (2026), who emphasize that incentives may shift not only the level of participation but also the composition of respondents.

4.4. The Effect on the Quality of Responses

So far, we have focused primarily on unit (non-)response as a measure of survey quality. However, an equally important dimension is the quality and accuracy of the responses provided by participating households. Previous research has shown that response quality is not homogeneous across survey participants but varies systematically with individual characteristics, some of which are correlated with wealth. In particular, older and less formally educated respondents have been found to provide lower-quality responses (Groves and Couper, 2012; Biemer et al., 2013). Furthermore, Angel et al. (2019) document gender differences in response quality. Using market-based valuations to replace owner-reported housing values, Naidin et al. (2025) show that misreporting is associated with residential tenure length and tenure type, dwelling characteristics, and household income and wealth.

Against this background, one may be concerned that the incentive affects not only participation but also response quality. While the coin may encourage households to take part in the survey, it is not clear whether it also promotes careful and accurate reporting. This concern is particularly relevant because the incentive rewards participation alone and does not impose any requirements regarding respondent effort or data quality.

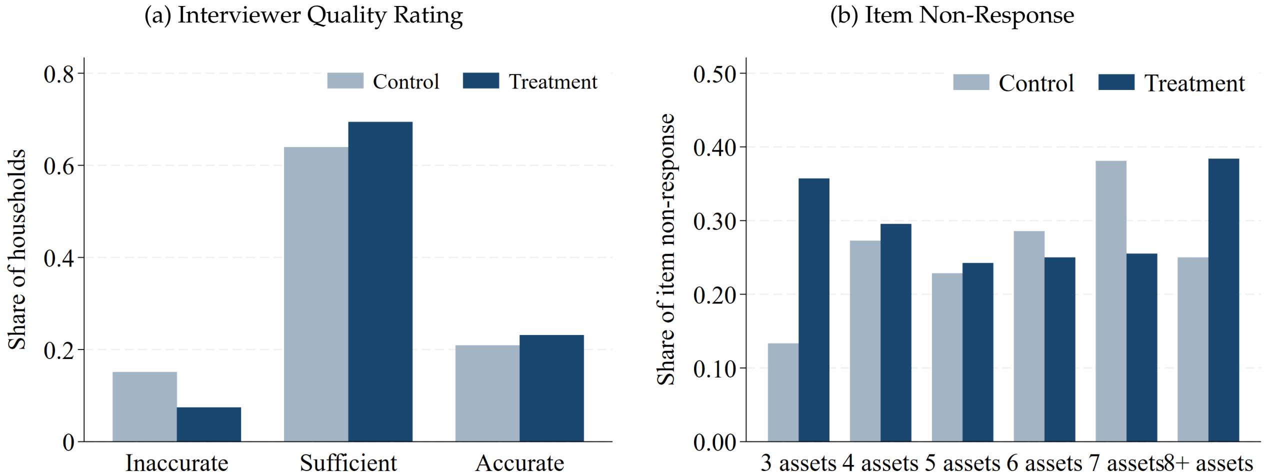
Two competing *a priori* hypotheses can be formulated. On the one hand, recipients of the coin may feel obliged to provide careful and accurate responses, perceiving the incentive as compensation for their participation, which would imply a positive effort effect. On the other hand, the incentive may attract respondents who are primarily motivated by the reward and therefore exert less effort when completing the survey, giving rise to a free-rider effect.

To assess which of these mechanisms dominates, we again draw on the (RP) pool and examine whether response quality differs between the treatment arms and, if so, in which direction.

We conduct two tests. First, we use SK-HFCS para-data containing interviewer evaluations and exploit a question, which assesses the overall reliability of responses to wealth and income questions. Interviewers are asked (HR0600): 'How would you rate the reliability of the information provided by the respondent about his/her income and assets? (1) Inaccurate, (2) Sufficient, (3) Accurate.' The results are reported in Figure 4. Although the treatment group exhibits a slight shift toward more favorable ratings, the differences are not statistically significant.

Second, we consider the fraction of item non-response and 'I don't know' answers among all questions eliciting the value of owned assets. To ensure comparability, we first compute average item non-response rates (INR) among households holding a similar number of different asset classes and then aggregate these using a weighted average. Details of the construction of this measure are provided in Appendix C.

Figure 4: Quality of Responses



Notes: Panel 4a presents the distribution of categories regarding the reliability of answers provided by respondents to questions on assets and incomes. Panel 4b shows average item-non-response by portfolio size per treatment group following formula (4).
Source: HFCS 2023/24, National Bank of Slovakia.

Results shown in Figure 4 indicate that, on average, the treatment group provided slightly more 'I don't know' responses to questions eliciting the value of possessed assets than the control group ($INR(T) = 29.7\%$ versus $INR(C) = 25.9\%$). However, the difference is statistically insignificant (Kolmogorov–Smirnov equality-of-distributions test, $D = 0.333$, p-value $= 0.893$), suggesting no meaningful effect of the incentive on item non-response. Hence, the results indicate that the incentive does not systematically affect response quality, implying neither a deterioration due to free-riding nor an improvement due to increased respondent effort.

5. What Changes when the Top-Tail is Better Captured, and Which Sampling Features Drive these Changes?

This section derives counterfactual population-representative wealth distributions for Slovakia under proportional sampling as a benchmark and then sequentially incorporates oversampling and, finally, incentives to identify the marginal contribution of each sampling innovation. Based on this analysis, we also draw broader conclusions about how differences in sampling strategies across Europe may affect cross-country comparisons of wealth statistics.

5.1. Constructing Counterfactual Distributions

We first compute household-specific net wealth[16] via

$$\text{Net Wealth}_i = \text{Assets}_i - \text{Liabilities}_i.$$

From there, we compile three wealth distributions that differ in the selection criteria applied for observations entering the sample and, accordingly, their respective survey weights. First, when exclusively relying on the (GP) sample, that entirely refrains from oversampling, we mimic the practices of surveys not applying an oversampling strategy (see Figure 6 for an overview of the sampling strategies applied across HFCS countries). Second, we merge the (GP) sample with the control group of the (RP) sample, i.e., $(GP) \cup (RP)_{\text{Control}}$. This mimics the approach of countries that adopt some kind of oversampling strategy relying on household-specific information correlating with wealth, yet provide no targeted participation incentives. Third, we rely on the full $(GP) \cup (RP)$ sample representing also the novel sampling features of combining oversampling with incentives.

As we aim to compute population-representative wealth distributions from different samples, we need to adjust the used survey weights accordingly. We do this by compiling weights *as if* we had used stratified sampling only, i.e., refraining from both oversampling and incentivizing wealthy households, to achieve weights for the (GP) sample denoted by $w_i^{(GP)}$. Similarly, we compute weights considering oversampled households yet no incentivized ones, denoted by $w_i^{(GP)\cup(RP)_{\text{Control}}}$, and weights for the full sample $w_i^{(GP)\cup(RP)}$. See Appendix D for details.

With these adjusted weights, all three samples become population-representative and any differences in the resulting statistics purely stem from qualitative differences in the survey administration.

5.2. Implication for the Measured Net Wealth Distribution

Building on the discussion of oversampling in Section 2.1 and incentives in Section 2.3, we formulate testable hypotheses regarding differences in average, aggregate, and distributional

[16]We follow standard practices in the HFCS and exclude public and occupational pension wealth.

wealth measures across the three samples in a within-country research design. Furthermore, we relate our results to recent initiatives linking HFCS micro-data with macroeconomic aggregates on household wealth from the National Accounts. This comparison is particularly relevant for the construction of Distributional National Accounts (EG-LMM, 2020; Blatnik et al., 2024). Specifically, we compare total financial assets and total net financial wealth (financial assets less liabilities) derived from the three survey samples with the corresponding aggregates reported in the National Accounts balance sheet for the household sector[17] for 2023, which matches the survey fieldwork period. As variables in the SK-HFCS and the National Accounts are not all well align-able, we focus on a truncated wealth measure here that consists of well comparable instruments only as assessed by EG-LMM (2020).[18] Concretely, we use the common concept of *coverage ratios (CR)* – defined as the respective HFCS total divided by the corresponding National Accounts aggregate – to assess the quality of alignment. Usually, survey aggregates undershoot National Accounts aggregates, i.e., $CR < 1$. Our hypotheses are based on Waltl (2022) linking different sampling strategies to CRs and showing that well-designed oversampling approaches together with the use of administrative data in the survey compilation process are associated with higher CRs.

Table 3: Hypothesized and Measured Effects of Oversampling and Incentivization

Sample	Median Net Wealth	Mean Net Wealth	Mean-to-Median Ratio	Gini Coefficient	CR (Fin. Assets)	CR (Liabilities)	EOW (5%)	EOW (10%)
	Hypothesised							
(GP)	lowest	lowest	lowest	lowest	lowest	lowest	lowest	lowest
$(GP) \cup (RP)_{\text{Control}}$	mid	mid	mid	mid	mid	mid	mid	mid
$(GP) \cup (RP)$	highest	highest	highest	highest	highest	highest	highest	highest
	Measured							
(GP)	124867	156922	1.26	44.4	28.69%	57.14%	-3	1
$(GP) \cup (RP)_{\text{Control}}$	124895	161433	1.29	45.6	28.76%	57.52%	23	15
$(GP) \cup (RP)$	125412	168296	1.34	47.3	30.27%	57.54%	48	41

Notes: The top panel summarizes the hypothesized effects of the intervention on measured net wealth. The bottom panel reports results.

Source: HFCS 2023/24, National Bank of Slovakia.

[17] We follow the National Accounts concept that most closely corresponds to the survey definition of a household, namely the household sector excluding Non-Profit Institutions Serving Households (NPISH, sector S.14).

[18] We define total financial assets via the following survey instruments: DA1140 (Value of self-employment businesses), DA1400 (Real estate incl. property used for business activities), DA2101 (Deposits; Value of sight and savings account), DA2102 (Mutual funds), DA2103 (Bonds), DA2104 (Value of non-self-employment businesses), and DA2105 (Shares). For measuring liabilities, we rely on the HFCS question DL1000 (Total outstanding balance of household's liabilities). We compare these survey items to the following instruments in the financial accounts as defined in the European System of National and Regional Accounts (ESA 2010). Assets: F22 (Transferable deposits), F29 (Other deposits), F3 (Debt securities), F51 (Equity), and F52 (Investment fund shares/units). Liabilities: F4 (Loans, long-term and short-term).

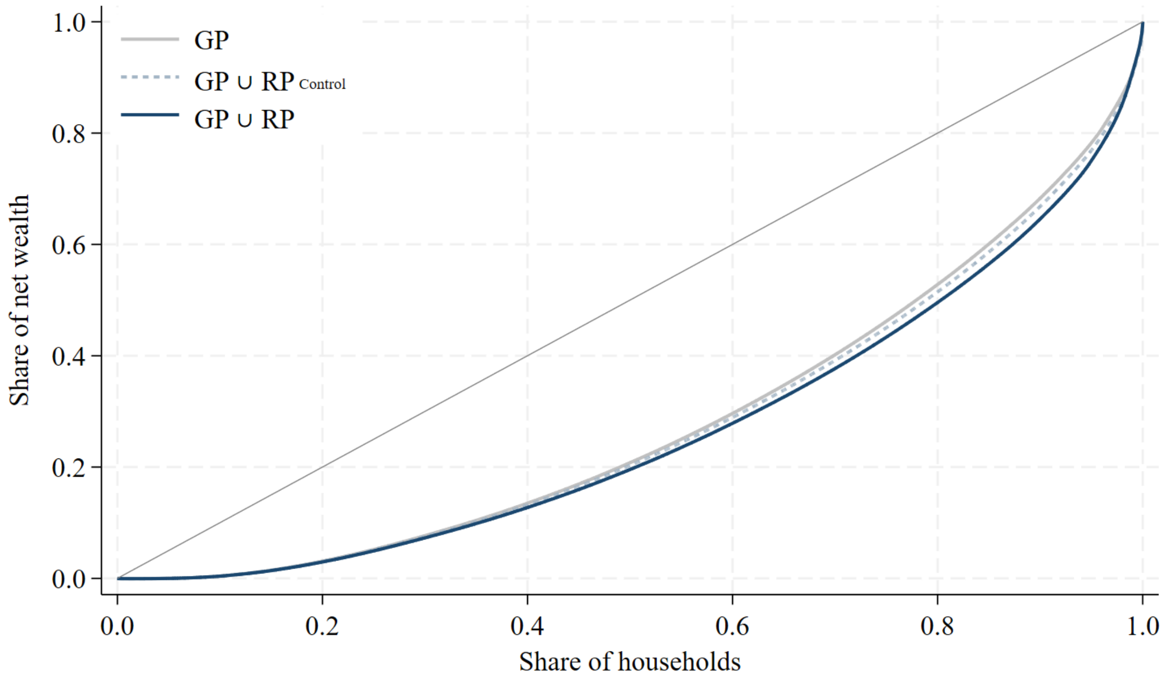
These considerations yield a certain expected hierarchy within the three samples summarized in the top panel of Table 3. Quantitative results are reported in the bottom panel. Indeed, in line with expectations, measured median and mean household wealth increase when adopting oversampling (+2.9% for the mean and +0.02% for the median) and even more when making additionally use of incentives for recruiting wealthy households (+7.2% for the mean and +0.4% for the median). The relatively small changes in the median are unsurprising, as the intervention targets only the upper tail of the distribution, while the median reflects the central household. In contrast, the mean – although a measure of central tendency as well – is highly sensitive to changes in the tails. This measure therefore shows much larger changes in response to incorporating oversampling and incentives, respectively. The measured strong responsiveness of mean wealth, alongside the stability of the median, suggests that the intervention indeed primarily affects the upper tail of the distribution without materially altering the rest. In line with this, inequality indices find substantially higher country-wide degrees of wealth inequality measured by the mean-to-median ratio when enhancing the survey administration (+6.4% overall).

We also compute Lorenz curves (Lorenz, 1905) from the three samples. Adjusting for the missing rich using Pareto-type adjustments (an *ex post* adjustment for the missing wealthy) pushes Lorenz curves further to the right (as shown in Figure 6 in Waltl and Chakraborty, 2022) and, hence, increases Gini indices and by that measured wealth inequality.

We replicate this assessment for the three samples and report the corresponding Lorenz curves in Figure 5 and their summary statistics, the Gini coefficients, in Table 3. Consistent with expectations, improved top-tail coverage through ex-ante interventions causes the Lorenz curves to bend progressively further away from the 45-degree line. Mechanically, the Gini coefficient increases by 2.3% when oversampling is introduced and by a further 6.5% when incentives are added.

These effects are substantial when viewed in the context of the variation in HFCS Gini coefficients across countries and survey waves. Across waves, the median absolute cross-country difference in Gini coefficients ranges from 0.035 to 0.051 (see Table E.2). Likewise, the variation in Slovakia's Gini coefficient across survey years is only modestly larger than the differences induced by the alternative sampling strategies reported in Table 3. This suggests that cross-country differences in HFCS-based Gini coefficients may partly reflect differences in sampling design rather than genuine differences in wealth inequality.

Figure 5: Lorenz Curves for Different Sub-samples



Notes: The figure shows Lorenz curves for Net Wealth separately computed from the three samples (GP), $(GP) \cup (RP)_{\text{Control}}$ and $(GP) \cup (RP)$.
Source: HFCS 2023/24, National Bank of Slovakia.

Finally, we assess total financial wealth and liabilities only, and compare the HFCS totals to aggregates from the National Accounts. In line with expectations, CRs are generally low. However, for financial assets they increase by 0.23% when adopting oversampling and rise by a further 4.98% when also making use of incentives. Overall, coverage ratios are up by 5.2% when adopting our full sampling strategy. The much larger marginal increase stemming from incentives is in line with the interpretation that our incentives are successful to attract more financially sophisticated households to the survey via the collector coin as incentive than with generic low-value participation incentives only (see Section 4.3). This is also visible in the slightly shifted distributions of Financial Asset holdings among incentivized households shown in Figure A.2 (Panel d).

Regarding liabilities, coverage is high even without refinements of the sampling strategy in line with the notion that mortgages are much more evenly distributed across the wealth distribution than financial assets. Adding oversampling and incentives still increases coverage yet changes are small.

Overall, the increase in coverage is small and does not fundamentally address the problem of misalignment of these two data sources, yet both oversampling and incentives contribute to a better match.

5.3. The Impact of Survey Modalities over Time

While the previous subsection focused on how oversampling and incentives alter measured wealth within Slovakia, we next assess whether these survey modalities also improve the survey's ability to capture the upper tail over time. To this end, we rely on a standard outcome measure reflecting the ability of a survey to capture the top of the wealth distribution (see also HFCN, 2023): the *effective oversampling rate of the wealthy (EOW)*. This measure is computed for each country c and survey wave w, and compares the *theoretical share* of households belonging to the wealthiest $x\%$ of the population with the *realized share* $S_{c,w}(1-x)$ of sampled households,

$$EOW(x|c,w) = \frac{S_{c,w}(1-x) - x}{x}. \tag{2}$$

An EOW greater than zero therefore indicates successful oversampling, while higher values imply more effective coverage of the upper tail of the wealth distribution.

We first exploit the repeated nature of the HFCS, which is conducted every two to three years and therefore enables longitudinal within-country comparisons. Historically, the Slovak HFCS achieved particularly low EOW values. In the 2021 wave, the rate was even negative, implying that wealthy households were under-represented relative to their population share and that there was effectively no successful oversampling:

$$EOW(10\%|SK,2021) = -18 \quad \text{and} \quad EOW(5\%|SK,2021) = -17,$$

as reported in HFCN (2023). Only Czechia and Austria – two countries that entirely refrained from oversampling wealthy households – achieved worse outcomes.[19] In the 2023 wave, by contrast, the survey achieved

$$EOW(10\%|SK,2023) = 41 \quad \text{and} \quad EOW(5\%|SK,2023) = 48,$$

representing a substantial improvement over the previous wave. For $EOW(10\%)$, this corre-

[19]In the 2017 wave, Slovakia recorded the worst outcome among all participating countries, while the 2014 wave achieved low but still positive rates: $EOW(10\%|SK,2017) = -17$, $EOW(5\%|SK,2017) = -20$, $EOW(10\%|SK,2014) = 5$, and $EOW(5\%|SK,2014) = 15$.

sponds to an increase of 59 points in absolute terms and 328% in relative terms, highlighting the markedly improved effectiveness of reaching households located at the very top of the wealth distribution.

To disentangle the marginal contributions of oversampling and incentives, we recompute the EOW measures for each of the three samples separately. The results reported in Table 3 indicate that both interventions substantially improve top-tail coverage. Relative to the baseline sample (GP), oversampling alone (the control group) increases $EOW(5\%)$ by 26 percentage points and $EOW(10\%)$ by 14 percentage points. Adding incentives leads to further sizeable gains: compared with the baseline, $EOW(5\%)$ increases by 51 percentage points and $EOW(10\%)$ by 40 percentage points. A direct comparison of the two oversampled groups shows that the incentive itself accounts for an additional improvement of 25 percentage points in $EOW(5\%)$ and 26 percentage points in $EOW(10\%)$.

5.4. Cross-Country Implications within the HFCN

The previous analysis demonstrates that oversampling and incentives materially affect top-tail coverage within Slovakia. Because the SK-HFCS is embedded in the *Household Finance and Consumption Network (HFCN)* – a pan-European framework of coordinated wealth surveys bringing together survey specialists from the ECB, national central banks, and statistical institutes[20] – these findings also raise a broader question: to what extent do cross-country differences in measured wealth inequality reflect heterogeneity in survey design and implementation, rather than genuine differences in underlying economic disparities?

All nationally conducted surveys participating in the HFCN follow an *ex ante* harmonized core questionnaire, including harmonized question wording and variable definitions, while allowing for country-specific language versions and optional supplementary modules. At the same time, sampling strategies – and in particular the application and design of oversampling procedures – are not harmonized and vary substantially across countries and survey waves. Barcaroli et al. (2021) identify cross-country differences in legal restrictions aimed at protecting household privacy as the primary reason why access to and use of administrative data in survey compilation differ markedly across countries, although part of the observed variation also reflects institutional path dependence.

[20] See `https://www.ecb.europa.eu/stats/ecb_surveys/hfcs/html/index.en.html` for more details. Last accessed in September 2024.

To place the Slovak results into this broader international context, we plot $EOW(10\%|c, 2021)$ and $EOW(10\%|c, 2023)$ for all HFCN countries c[21] and group them according to their adopted oversampling strategy. Additionally, for Slovakia, we compute EOW again separately for the three samples (GP), $(GP) \cup (RP)_{\text{Control}}$, and $(GP) \cup (RP)$ (see Table 3).

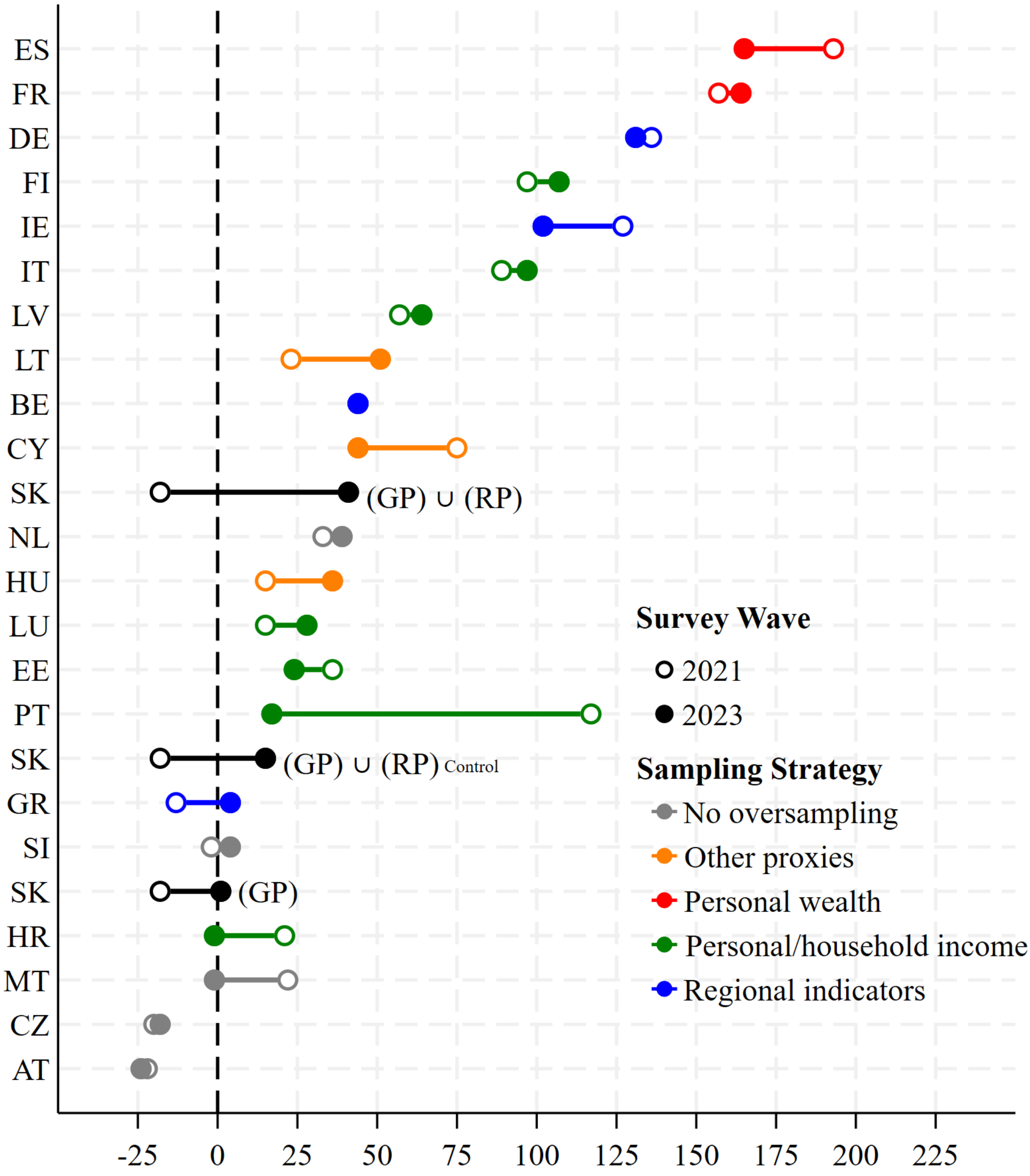
Figure 6 illustrates how Slovakia compares with other countries. Relative to the 2021 results, Slovakia already shows an improvement even without implementing oversampling or incentives, likely reflecting the updated sampling frame based on the 2021 census. Introducing oversampling improves Slovakia's ranking by three positions. However, compared with countries using similar proxy-based oversampling approaches, Slovakia still achieves a relatively modest effective oversampling rate. When incentives are combined with oversampling, Slovakia improves by a further six ranks, ultimately advancing from 20th place (out of 22 countries) to 11th place. In fact, Slovakia recorded the largest increase in EOW across all participating countries between the two survey waves, highlighting the substantial improvement in top-tail coverage achieved through our intervention.

Taken together, the findings from Section 5.2, Section 5.3 and Section 5.4 highlight the strong sensitivity of standard aggregate and distributional wealth statistics to sampling strategies as well as quality measures of the survey. As such indicators are regularly reported as part of official statistics, used for decision-making and international comparisons, their stark sensitivity to choices in survey administration protocols creates major concerns: The pan-European HFCS does not apply harmonized guidelines with respect to sampling frames or sampling strategies. As the three samples constructed here mimic some strategies adopted by different countries, our results suggest that a substantial share of observed cross-country differences in aggregate and distributional household wealth indicators may reflect statistical artifacts arising from methodological differences, rather than genuine differences in inequality.

This raises concerns about the suitability of the HFCS, in its current form, for producing fully comparable wealth statistics across Europe and underscores the need for more harmonized sampling protocols in the future.

[21] We use the following abbreviations: Belgium (BE), Czech Republic (CZ), Germany (DE), Estonia (EE), Ireland (IE), Greece (GR), Spain (ES), France (FR), Croatia (HR), Italy (IT), Cyprus (CY), Latvia (LV), Lithuania (LT), Luxembourg (LU), Hungary (HU), Malta (MT), The Netherlands (NL), Austria (AT), Portugal (PT), Slovenia (SI), Slovakia (SK), and Finland (FI).

Figure 6: Change in the Effective Oversampling Rate of the Top 10%.



Notes: Lines connect EOW values in the 2021 and 2023 waves. Countries are ordered according to their 2023 EOW values. Slovakia is highlighted in black; the three Slovak rates are computed using the (GP), $(GP) \cup (RP)_{\text{Control}}$, and $(GP) \cup (RP)$ samples. EOWs are computed according to formula (2). Other proxy-based oversampling approaches include dwelling size (HR), electricity consumption (CY), and dwelling values (HU).
Source: HFCS 2021 and 2023, European Central Bank.

6. Conclusions

We perform an experiment within the sampling procedure of the Slovak chapter of the Household Finance and Consumption Survey. We test whether proxies of household characteristics recorded in the census can be used to pre-identify wealthy households and whether the promise of a valuable collector coin can incentivize them to participate in the survey.

Both aspects are successful: the pre-identification successfully selects households that *ex post* turn out to be substantially richer in terms of income, consumption, and wealth than households sampled using a population-representative sampling strategy. This ensures the internal validity of our intervention.

A randomly chosen 50% of all pre-identified rich households receive additional participation incentives in the form of a valuable collector coin. These households indeed participate with an increased probability of 11-14 percentage points. The effect survives an ample set of robustness tests. Next to participation rates, also the share of financially sophisticated households is higher when offering an incentive. We show that establishing a personal interaction between the interviewer and to-be-recruited survey participant, and the incentive are not substitutes but independently contribute to the success of motivating a household to participate in the survey. Next to pure changes in participation rates, the intervention also moves up the Slovak HFCS in a quality ranking across all countries conducting this survey. While Slovakia had one of the lowest effective oversampling rates among all HFCS countries in the previous wave – meaning that wealthy households were in fact *under-represented* relative to their population share – the rate increased from -18 to 41 in the present wave, moving Slovakia from the bottom fifth to the upper third of the country ranking. Additionally, the intervention does not decrease collected survey data quality as confirmed by several measures. These findings imply that targeted incentives are a viable tool for improving the representativeness, and hence external validity and overall quality, of the survey.

To assess the implications for survey results, we conduct a counterfactual analysis and compute standard aggregate and distributional wealth statistics for three samples: a baseline sample without oversampling or incentives, an expanded sample including oversampled households, and a third sample additionally incorporating incentivized households. The results show a systematic increase in measured wealth inequality along this hierarchy: mean and aggregate wealth increase substantially across samples and the Gini index rises by 2.3% when oversampling is introduced and by 6.5% when both oversampling and incentives are applied. Other inequality measures exhibit similar patterns.

These findings suggest that differences in survey design are an overlooked source of variation in internationally harmonized wealth statistics. Given the substantial heterogeneity in sampling strategies across countries participating in the Europe-wide HFCS, the comparability of cross-country results and established country rankings may therefore be compromised. Consequently, part of the observed cross-country variation in measured wealth inequality may reflect differences in data collection protocols rather than differences in underlying economic reality. From this perspective, a greater harmonization of sampling and fieldwork protocols across participating countries may represent a sensible way forward to improve the comparability and reliability of European wealth statistics.

References

Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1):123–129.

Angel, S., Disslbacher, F., Humer, S., and Schnetzer, M. (2019). What did you really earn last year? Explaining measurement error in survey income data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(4):1411–1437.

Barcaroli, G., Ilardi, G., Neri, A., and Tuoto, T. (2021). Optimal sampling design for household finance surveys using administrative income data. *Rivista di statistica ufficiale*, 2:29–70.

Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., and Sudman, S. (2013). *Measurement Errors in Surveys*. John Wiley & Sons.

Blatnik, N., Bobasu, A.-G., Krustev, G., and Tujula, M. (2024). Introducing the Distributional Wealth Accounts for euro area households. *Economic Bulletin Articles*, 5.

Bover, O. (2008). Oversampling of the wealthy in the Spanish Survey of Household Finances (EFF). *Irving Fisher Committee Bulletin*, 28:399–402.

Bricker, J. (2014). Survey incentives, survey effort, and survey costs. *Finance and Economics Discussion Series*, (2014-74). https://ssrn.com/abstract=2503478. Last accessed in September 2024.

Calvet, L. E. and Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *The Journal of Finance*, 69(2):867–906.

Cantarella, M., Neri, A., and Ranalli, M. G. (2024). Estimating the distribution of household wealth: methods for adjusting survey data estimates using national accounts and rich list data. *Review of Income and Wealth*, 70(3):551–580.

Chang, A. C., Hsu, J. W., Ma, E., Bachtell, K., and Sjoblom, M. (2026). Does it pay to send multiple pre-paid incentives? Evidence from a randomized experiment. *Journal of Survey Statistics and Methodology*, page smaf047.

Crossley, T. F., Schmidt, T., Tzamourani, P., and Winter, J. K. (2021). Interviewer effects and the measurement of financial literacy. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 184(1):150–178.

Dutz, D., Huitfeldt, I., Lacouture, S., Mogstad, M., Torgovitsky, A., and van Dijk, W. (2026). Selection in surveys: Using randomized incentives to detect and account for nonresponse bias. *The Review of Economic Studies*, 93(5):3086–3132.

EG-LMM (2020). Understanding household wealth: linking macro and micro data to produce distributional financial accounts. *ECB Statistics Paper Series*, No 37 / July 2020.

Essig, L. and Winter, J. K. (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.

European Central Bank (2023a). Household finance and consumption survey: Statistical tables, wave 2 (2017). https://www.ecb.europa.eu/home/pdf/research/hfcn/HFCS_Statistical_Tables_Wave_2017_July2023.pdf.

European Central Bank (2023b). Household finance and consumption survey: Statistical tables, wave 3 (2021). https://www.ecb.europa.eu/home/pdf/research/hfcn/HFCS_Statistical_Tables_Wave_2021_July2023.pdf.

European Central Bank (2024). Experimental Distributional Wealth Accounts (DWA) for the Household Sector – Methodological Note. https://data.ecb.europa.eu/sites/default/files/2024-01/DWA%20Methodological%20note_0.pdf.

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

Groves, R. M. and Couper, M. P. (2012). *Nonresponse in household interview surveys*. John Wiley & Sons.

Guiso, L., Haliassos, M., and Jappelli, T. (2003). Household stockholding in Europe: where do we stand and where do we go? *Economic Policy*, 18(36):123–170.

Heffetz, O. and Reeves, D. B. (2019). Difficulty of reaching respondents and nonresponse bias: Evidence from large government surveys. *Review of Economics and Statistics*, 101(1):176–191.

HFCN (2023). Household finance and consumption survey: Methodological report for the 2021 wave. *ECB Statistics Paper Series*, No. 45.

Hsu, J. W., Schmeiser, M. D., Haggerty, C., and Nelson, S. (2017). The effect of large monetary incentives on survey completion: Evidence from a randomized experiment with the Survey of Consumer Finances. *Public Opinion Quarterly*, 81(3):736–747.

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Johnson, D. S., Meinhardt, M., and Sabelhaus, J. (2026). Why did people stop responding to federal economic surveys? What can be done? *The Brookings Institution*. https://www.brookings.edu/articles/why-did-people-stop-responding-to-federal-economic-surveys-what-can-be-done/. Accessed: 2026-04-14.

Kennickell, A. B. (2008). The role of over-sampling of the wealthy in the Survey of Consumer Finances. In for International Settlements, B., editor, *The IFC's contribution to the 56th ISI Session*, volume 28, pages 403–408.

Kennickell, A. B. (2019). The tail that wags: differences in effective right tail coverage and estimates of wealth inequality. *The Journal of Economic Inequality*, 17(4):443–459.

Landier, A. and Plantin, G. (2017). Taxing the rich. *The Review of Economic Studies*, 84(3):1186–1209.

Lohmann, H. (2011). Comparability of eu-silc survey and register data: The relationship among employment, earnings and poverty. *Journal of European Social Policy*, 21(1):37–54.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219.

McGonagle, K. A. (2020). The effects of an incentive boost on response rates, fieldwork effort, and costs across two waves of a panel study. *Methoden, Daten, Analysen*, 14(2):241.

Mercer, A., Caporaso, A., Cantor, D., and Townsend, R. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, 79(1):105–129.

Naidin, M. D., Waltl, S. R., and Ziegelmeyer, M. H. (2025). Objective housing sales and rent prices in representative household surveys: Implications for wealth, inequality, housing market, and affordability statistics. *Review of Income and Wealth*, 71(1):e12692.

Osier, G. (2016). *Unit non-response in household wealth surveys: Experience from the Eurosystem's Household Finance and Consumption Survey*. Number 15. ECB Statistics Paper.

Pforr, K., Blohm, M., Blom, A. G., Erdel, B., Felderer, B., Fräßdorf, M., Hajek, K., Helmschrott, S., Kleinert, C., and Koch, A. (2015). Are incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. *Public Opinion Quarterly*, 79(3):740–768.

Rolheiser, L., van Dijk, D., and van de Minne, A. (2020). Housing vintage and price dynamics. *Regional Science and Urban Economics*, 84:103569.

Saez, E. and Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2):519–578.

Schneebaum, A., Rehm, M., Mader, K., and Hollan, K. (2018). The gender wealth gap across European countries. *Review of Income and Wealth*, 64(2):295–331.

Schröder, C., Bartels, C., Grabka, M. M., König, J., Kroh, M., and Siegers, R. (2020). A novel sampling strategy for surveying high net-worth individuals—a pretest application using the Socio-Economic Panel. *Review of Income and Wealth*, 66(4):825–849.

Schröder, M., Saßenroth, D., Körtner, J., Kroh, M., and Schupp, J. (2013). Experimental evidence of the effect of monetary incentives on cross-sectional and longitudinal response: Experiences from the Socio-Economic Panel (SOEP). *SOEPpapers on Multidisciplinary Panel Data Research*, No. 603.

Singer, E. and Ye, C. (2013). The use and effects of incentives in surveys. *The Annals of the American Academy of Political and Social Science*, 645(1):112–141.

Stecklov, G., Weinreb, A., and Carletto, C. (2018). Can incentives improve survey data quality in developing countries? Results from a field experiment in India. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 181(4):1033–1056.

Tourangeau, R. (2000). *The psychology of survey response*. Cambridge University Press.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 64(2):357–387.

Waltl, S. R. (2022). Wealth inequality: A hybrid approach toward multidimensional Distributional National Accounts in Europe. *Review of Income and Wealth*, 68(1):74–108.

Waltl, S. R. and Chakraborty, R. (2022). Missing the wealthy in the HFCS: micro problems with macro implications. *The Journal of Economic Inequality*, 20(1):169–203.

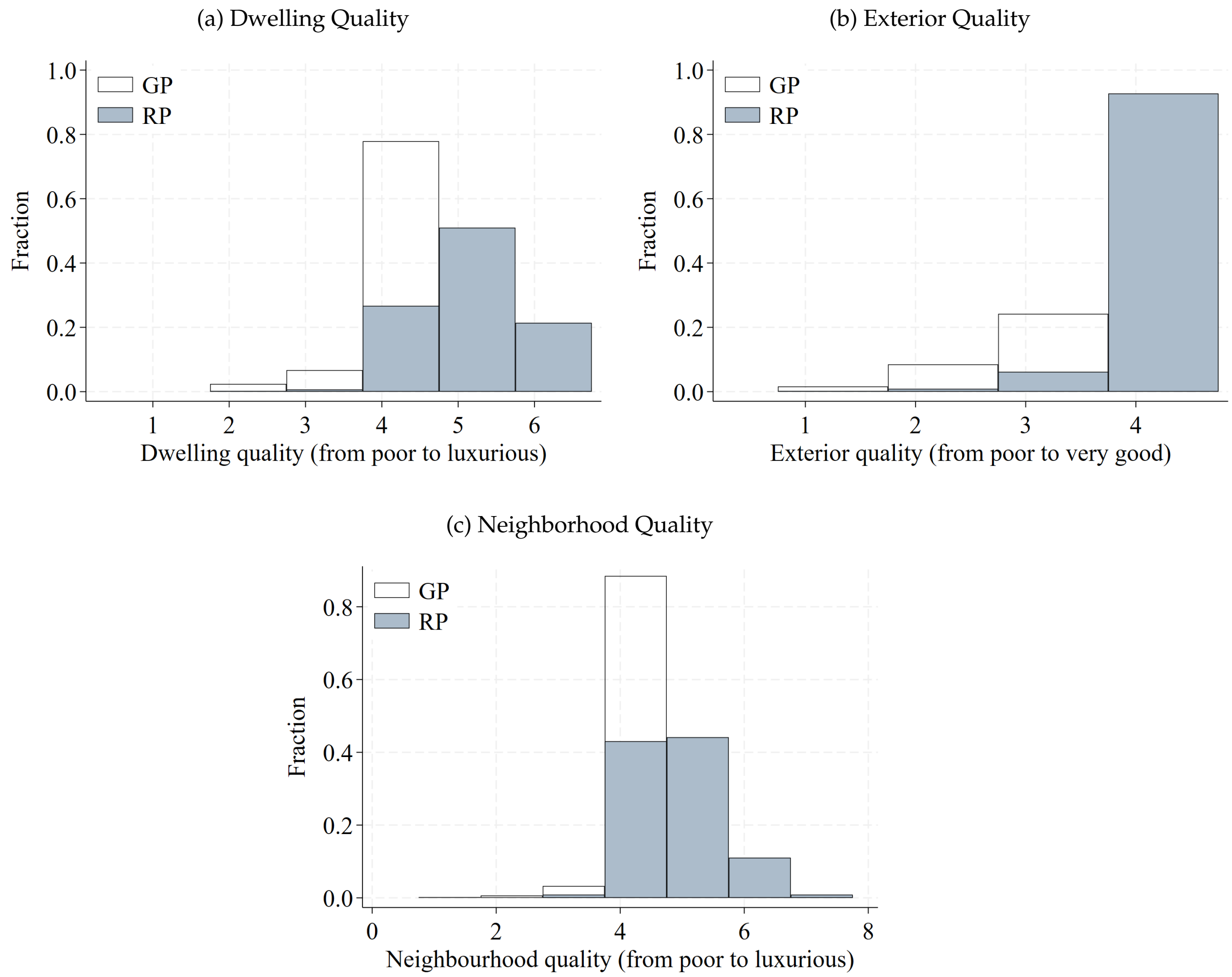
World Bank (2024). World development indicators: GDP per capita, purchasing power parity (current international $). *World Bank*. https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD. Last accessed in September 2024.

Appendix

A. Assessing the Success of the Ex-Ante Identification

In this appendix, we perform a variety of tests to verify the success of the ex-ante identification strategy.

Figure A.1: External Validity: Interviewer Ratings



Notes: This figure presents the distribution of interviewer ratings of the survey participants' main residences. Panel (a) refers to interior dwelling quality (SCO200), panel (b) refers to exterior dwelling quality (SCO400), and panel (c) refers to the overall neighborhood quality (SC0600). Questions are measured on a 4 and 6 steps Likert scale.
Source: HFCS 2023/24, National Bank of Slovakia.

The first set assesses the gross sample, i.e., the characteristics of all households that were sampled through the (RP)-strategy. As not all sampled households agreed to participate, we have to rely on the para-data collected by interviewers. These include ratings of the quality of the main residence of the invited household as well as the neighborhood quality (both measured on Likert scales). As interviewers also visit non-responding households, these ratings are up to date. As it is known that rich households tend to live in more expensive and better-quality housing units and that the dwellings' value (largely determined by location) tends to increase when moving up the net wealth distribution (see, for instance, Figures 5 and 6 in the appendix of Waltl, 2022), we can indeed test whether survey interviewers – who are not aware which household was sampled using the base or oversampling strategy – provide a better rating of dwellings occupied by participants entering the pool from the (RP) sample than those from the (GP) sample.

Figure A.1 shows results: Indeed, interviewers rate both the interior dwelling and neighborhood quality as much better in the (RP) sample than in the (GP) sample. On a scale ranging between 1 (poor) and 6 (luxurious), the median (mean) rating for the interior dwelling quality equals 4 (4.04) in the (GP) sample and 5 (4.86) in the (RP) sample. Likewise, neighborhood quality is rated on average 4 (4.02) in the (GP) and 5 (4.68) in the (RP) sample.

In addition, we match other socio-economic characteristics reported in the census to survey responses. The survey relied on voluntary interviews, while the census is a complete, legally mandatory data collection combining register data with personal interviews.

To assess the internal validity, we test how closely the census information matches responses provided in the survey interview. As the census took place in 2021 but the survey was fielded in 2023/24, we do expect some deviations yet no fundamental changes. Out of the 207 oversampled HFCS households, 155 (roughly 75%) are the same in terms of household composition. Occupation and employment status naturally vary more between the two samples yet still remain largely aligned.

Next, we test another aspect of internal validity by checking whether the pre-identification indeed led to a systematically different population in terms of economic prosperity than the standard stratified random sampling procedure. This is confirmed for the gross sample in Figure D.1: the distribution of employment income of pre-identified wealthy households has a much thicker right tail than that of the general population. This is also confirmed in the net sample shown in Figure A.2: Participants entering the survey via the (RP) pool have higher net wealth, financial assets, more expensive main residences, hold more valuable vehicles and valuables, have higher incomes and consumption expenses.

It has been shown in the past that rich households also tend to differ in other demographic dimensions.[22] Thus, we test whether we find the patterns when comparing demographics between the (RP) and (GP) samples. Indeed, we confirm that households recruited via the (RP) sample are on average higher educated, more likely to be entrepreneurs (27% vs. 11%), more often headed by a male household member (72% vs. 61%), are more likely to have attained a university degree (61% vs. 24%) and have on average more children (1.07 vs. 0.60 children per household). This further increases confidence in the internal validity of the applied pre-identification strategy.

B. Balance Checks

To correctly estimate a treatment effect associated with the coin, balance in the gross sample is required, i.e., no deviations between household characteristics across the sampled treatment and control households. We assure this by randomly assigning households to the treatment or control groups, respectively.

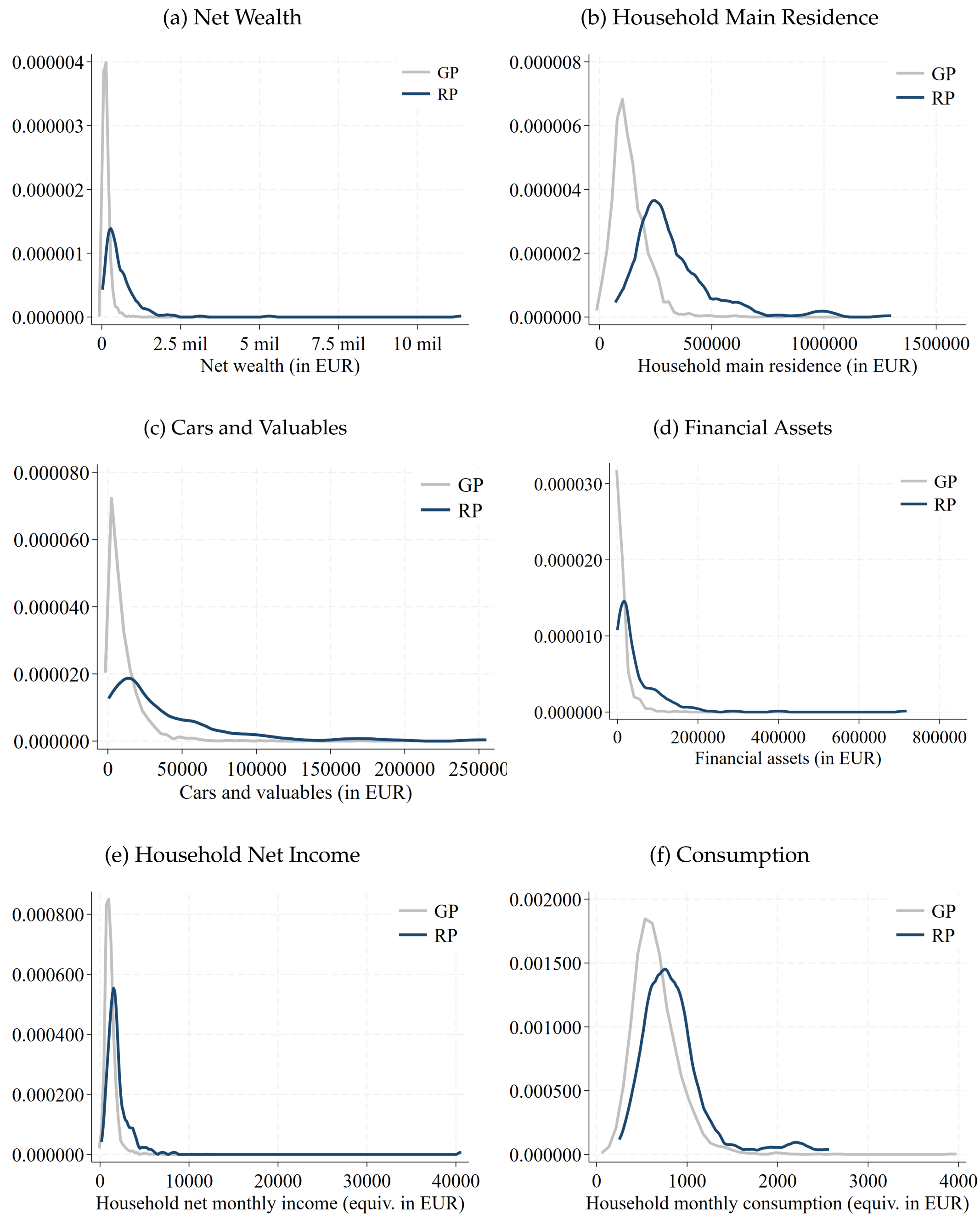
To confirm the success of the random allocation, we compute summary statistics separately for the two groups based on both census data (used for pre-identification) and HFCS para-data. Table B.1 indeed confirms balance: There are no statistically significant differences between the two groups in either data source.

C. Item Non-Response Rates

Let N_a^G be the number of households i in the treatment group $(G = T)$ or control group $(G = C)$ possessing an asset portfolio consisting of $a \in \{1, \dots, A\}$ different classes and $N_G = \sum_{a=1}^{A} N_a^G$ the overall number of households in group G. The HFCS in total elicits holdings in 14 asset classes. Few households possess less than three asset classes or more than 8 and hence we

[22] See https://nbs.sk/dokument/5dbda6d8-8ff1-4767-b52f-fe5c5c9524c8/stiahnut/?force=true, last accessed in August 2025.

Figure A.2: Internal Validity: Distribution of Several Asset Classes in (GP) and (RP)



Notes: In all cases, the results of the Kolmogorov–Smirnov tests for equality of the distribution across the two samples (i.e., GP and RP) reject the null hypotheses, i.e., the two compared distributions are significantly different. *Source*: HFCS 2023/24, National Bank of Slovakia.

merge them into broader categories. This leaves us with $A = 6$. Further, let $m_i \in \{0, 1, \ldots, a\}$ be the number of household i's missing items; we compute the average item-non-response rate among households in G owning wealth in a asset classes by

$$INR(a, G) = \frac{1}{N_a^G} \sum_{i=1}^{N_a^G} \frac{m_i}{a}. \tag{3}$$

For example, suppose a household owns five types of assets (e.g., a main residence, cars, savings accounts, mutual funds, and private pension wealth) and does not report the value of

Table B.1: Balance within the (RP) Gross Sample by Treatment

	Control (N = 299)			Treated (N = 299)			Differences in Means	Differences in Medians
	Mean/Share	Median	SD	Mean/Share	Median	SD	(p-value)	(p-value)
HFCS para-data								
Personal visit =yes	0.779		(0.415)	0.793		(0.405)	0.013	
							(0.690)	
Number of contacts	1.298	1	(0.652)	1.241	1	(0.587)	-0.057	0
							(0.263)	(0.307)
Daytime of contact	13.000	13	(2.513)	13.137	14	(2.783)	0.137	1
							(0.527)	(0.497)
Urban area = yes	0.569		(0.495)	0.555		(0.497)	-0.013	
							(0.742)	
Census variables								
Floor area [sqm]	247.25	223.85	(66.72)	247.49	232.00	(67.86)	-0.24	8.15
							(0.965)	(0.737)
Household Reference Person								
Gender = male	0.676		(0.469)	0.635		(0.482)	0.041	
							(0.303)	
Age [years]	49.806	48	(7.644)	50.321	49	(7.785)	-0.515	1
							(0.415)	(0.484)
Highest Education Level=yes	0.649		(0.478)	0.696		(0.461)	-0.047	
							(0.223)	
Working (excl. pensioners)= yes	0.893		(0.310)	0.890		(0.314)	0.003	
							(0.896)	
Number of Children	3.064	3	(0.882)	3.010	3	(0.796)	0.054	0
							(0.437)	(0.957)
Entrepreneur = yes	0.177		(0.383)	0.194		(0.396)	-0.017	
							(0.600)	

Notes: The table reports summary statistics for census variables and HFCS para-data evaluated for the (RP) sample drawn for the HFCS. Reported p-values refer to t-tests for differences in means for continuous variables, Z-tests for shares, and to Wilcoxon signed-rank test for differences in medians. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Census 2021, Statistical Office of the Slovak Republic; HFCS 2023/24, National Bank of Slovakia.

its cars. This household's item-non-response rate would then be $1/5$, or 20%. The indicator compares this rate to the average item-non-response rate among all households in group G that also hold five types of assets. These rates are aggregated to overall group-specific average item-non-response rates, i.e.,

$$INR(G) = \frac{1}{A} \sum_{a=1}^{A} INR(a, G). \tag{4}$$

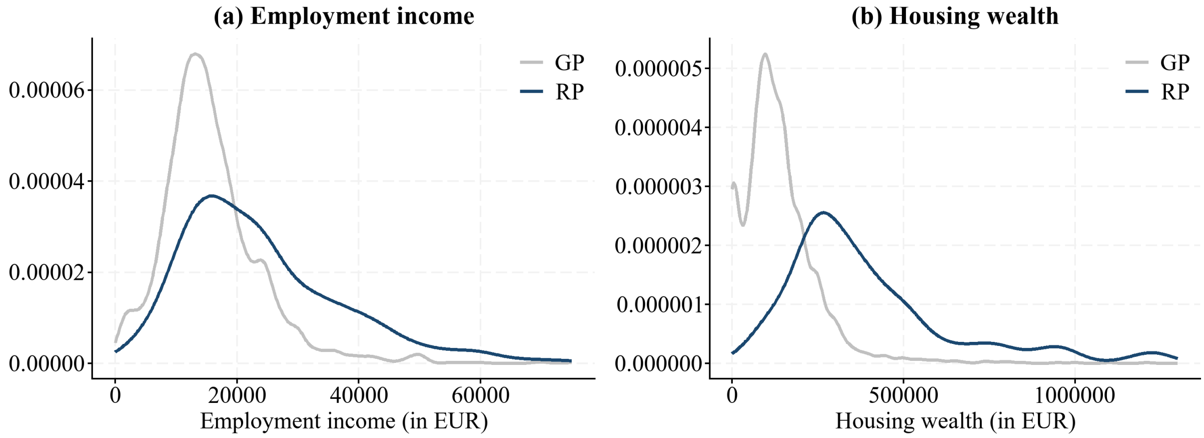
This allows us to adequately compare item-non-response rates between the treatment ($G = T$) and control ($G = C$) groups.

D. Adjusting Weights for Counterfactual Analyses

A common way to mitigate non-response bias in surveys is to calibrate design weights to known population totals. If the calibration variables are correlated with the outcomes of interest, this improves the precision and representativeness of survey estimates. In Slovak social surveys, calibration typically relies on the gender-age distribution, the number of pensioners, self-employed and unemployed persons, and the total number of private households. As these characteristics are related to both the survey objectives and the determinants of non-response, calibration helps restore representativeness with respect to the general population.

In the HFCS, however, the oversampling of wealthy households requires additional calibration because the standard variables alone cannot fully offset the intentionally enlarged upper tail of the wealth distribution. We therefore also calibrate to the number of individuals with tertiary education and the distribution of employment income, for which external benchmark totals are available.

Figure D.1: Distribution of Income from Employment and Housing Wealth



Notes: The figure illustrates unweighted distributions of income from employment and housing wealth separately for the (GP) and the (RP) subsample. A heavy tail can be observed for the oversampled group, supporting an effective strategy.
Source: HFCS 2023/24, National Bank of Slovakia.

The (RP) sample is not calibrated separately yet jointly with the (GP) sample to ensure that the *overall* HFCS sample represents the reference population. For the counterfactual analysis, the (GP) and the $(GP) \cup (RP)_{\text{Control}}$ samples are also weighted separately using the same procedure.

Table D.1: Representativeness of the Unweighted (RP) Sample

Deciles	Employment Income	Self-employment Income	Housing Wealth
d1	0.67%	5.26%	0.48%
d2	2.68%	5.26%	0.00%
d3	2.34%	4.21%	1.93%
d4	5.69%	4.21%	2.90%
d5	8.70%	5.26%	0.48%
d6	7.02%	9.47%	0.97%
d7	9.70%	13.68%	2.42%
d8	13.38%	9.47%	5.31%
d9	20.74%	16.84%	19.32%
d10	29.10%	26.32%	66.18%

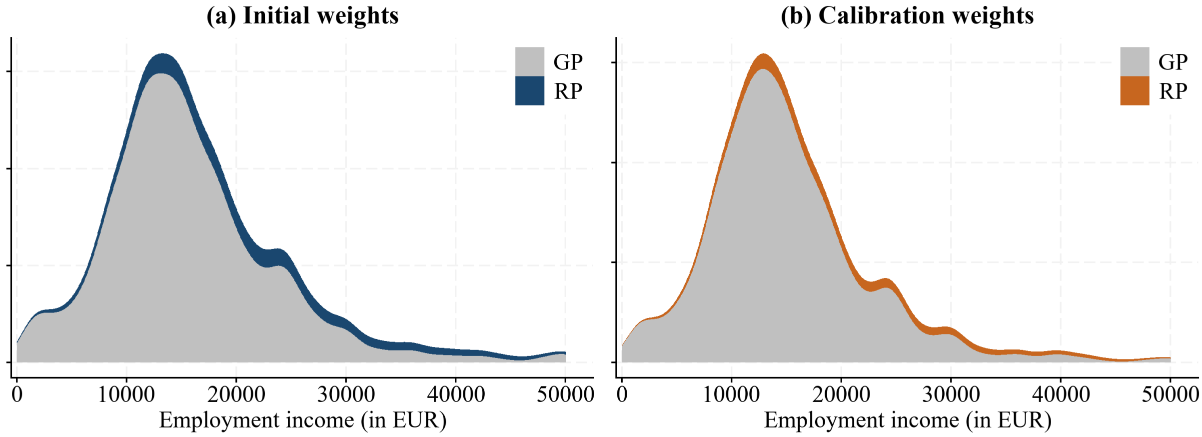
Notes: The table reports the representativeness of the unweighted (RP) subsample relative to benchmark distributions derived from (i) administrative data on employment income and (ii) the weighted representative (GP) sample for self-employment income and housing wealth. It shows the share of each decile covered by the (RP) sample, where values above 10% indicate over-representation.
Source: HFCS 2023/24, National Bank of Slovakia.

Not every individual in the (RP) subsample is necessarily a high-income earner, as many are included because they belong to a wealthy household. Consequently, the distribution of employment income within the (RP) spans the full range of possible values, although it exhibits a pronounced upper tail (Figure D.1). To avoid overstating aggregate employment income in the HFCS, the over-representation of high-income households must be corrected through calibration. This requires identifying the point in the income distribution at which oversampling becomes effective. To this end, we benchmark the (RP) income deciles against the corresponding deciles derived from administrative data. As shown in Table D.1, substantial over-representation emerges only among the top 20% of the employment income distribution. We therefore calibrate this segment separately to account for the intentional oversampling of affluent households.

To jointly account for multiple wealth-related dimensions, we additionally incorporate self-employment income and housing wealth into the calibration procedure. As no external benchmarks are available for these variables, we use the (GP) sample as a proxy for the population distribution. To do so, the (GP) sample is first calibrated separately and used to estimate the deciles of self-employment income and housing wealth. These benchmarks are then incorporated into the weighting of the full $(GP) \cup (RP)$ dataset. As illustrated in Figure D.1 and confirmed by Table D.1, substantial over-representation is concentrated in the upper tail, implying that the top two deciles should be calibrated separately to preserve the representativeness of the HFCS sample.

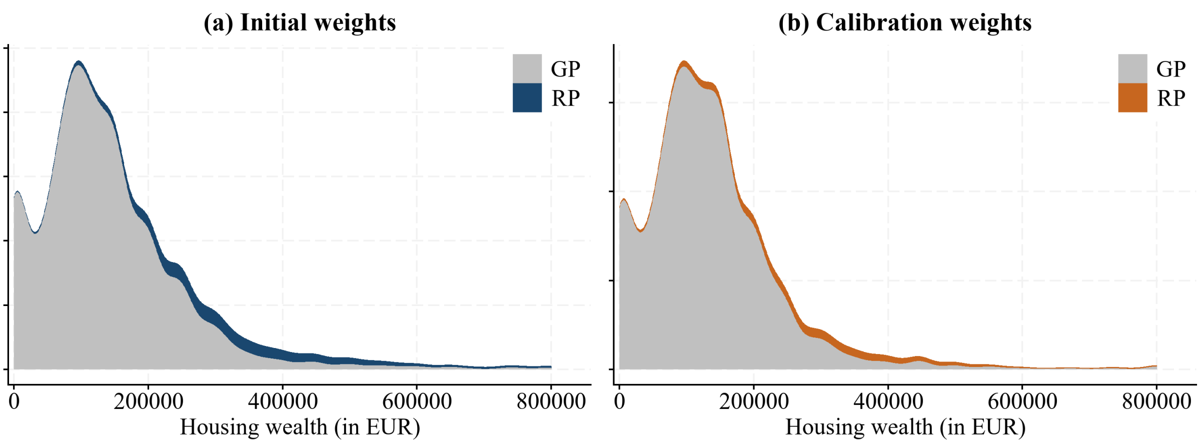
Figure D.2 and Figure D.3 illustrate the results of the calibration process. The (RP) distribution remains relatively stable across employment income values, which, compared to the (GP) sample, exhibits over-representation in the upper tail. To address the issue of over-representation, the share of the (RP) sample in the *overall* HFCS sample is slightly reduced after calibration.

Figure D.2: Distribution of Income from Employment – Initial vs. Calibration



Notes: The figure shows stacked distributions of employment income for the (GP) and (RP) subsamples, comparing initial and final calibration weights. The (RP) group is over-represented at higher income levels but is reduced after calibration, yielding joint representativeness with the (GP) sample.
Source: HFCS 2023/24, National Bank of Slovakia.

Figure D.3: Distribution of Housing Wealth – Initial vs. Calibration



Notes: The figure shows stacked housing wealth distributions for the (GP) and (RP) subsamples under initial and final calibration weights. The (RP) group exhibits a pronounced upper tail.
Source: HFCS 2023/24, National Bank of Slovakia.

E. Supplemental Tables and Figures

Table E.1: LPM Estimates of the Probability of Survey Participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
COIN	0.117***		0.113***	0.105***	0.106***	0.131***	0.130**	0.138***	0.130**
	(0.039)		(0.037)	(0.037)	(0.037)	(0.042)	(0.052)	(0.052)	(0.053)
VISIT		0.341***	0.339***	0.366***	0.361***	0.369***	0.366***	0.285***	0.366***
		(0.033)	(0.033)	(0.037)	(0.038)	(0.039)	(0.050)	(0.066)	(0.050)
Number of contacts				-0.122***	-0.118***	-0.124***	-0.115***	-0.112***	-0.114***
				(0.027)	(0.027)	(0.029)	(0.036)	(0.036)	(0.039)
COIN x VISIT								0.163*	
								(0.087)	
COIN x Number of contacts									-0.003
									(0.051)
Constant	0.288***	0.078***	0.024	0.161***	-0.080	-0.184	-0.667	-0.527	-0.667
	(0.026)	(0.024)	(0.031)	(0.039)	(0.454)	(0.462)	(0.650)	(0.643)	(0.651)
Survey controls	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes	Yes	Yes	Yes
Int. characteristics	No	No	No	No	No	No	Yes	Yes	Yes
Interviewer-level clustered S.E.	No	No	No	No	No	No	Yes	Yes	Yes
Adj. R^2	0.013	0.085	0.097	0.121	0.118	0.118	0.123	0.126	0.121
N	598	598	598	598	598	598	598	598	598

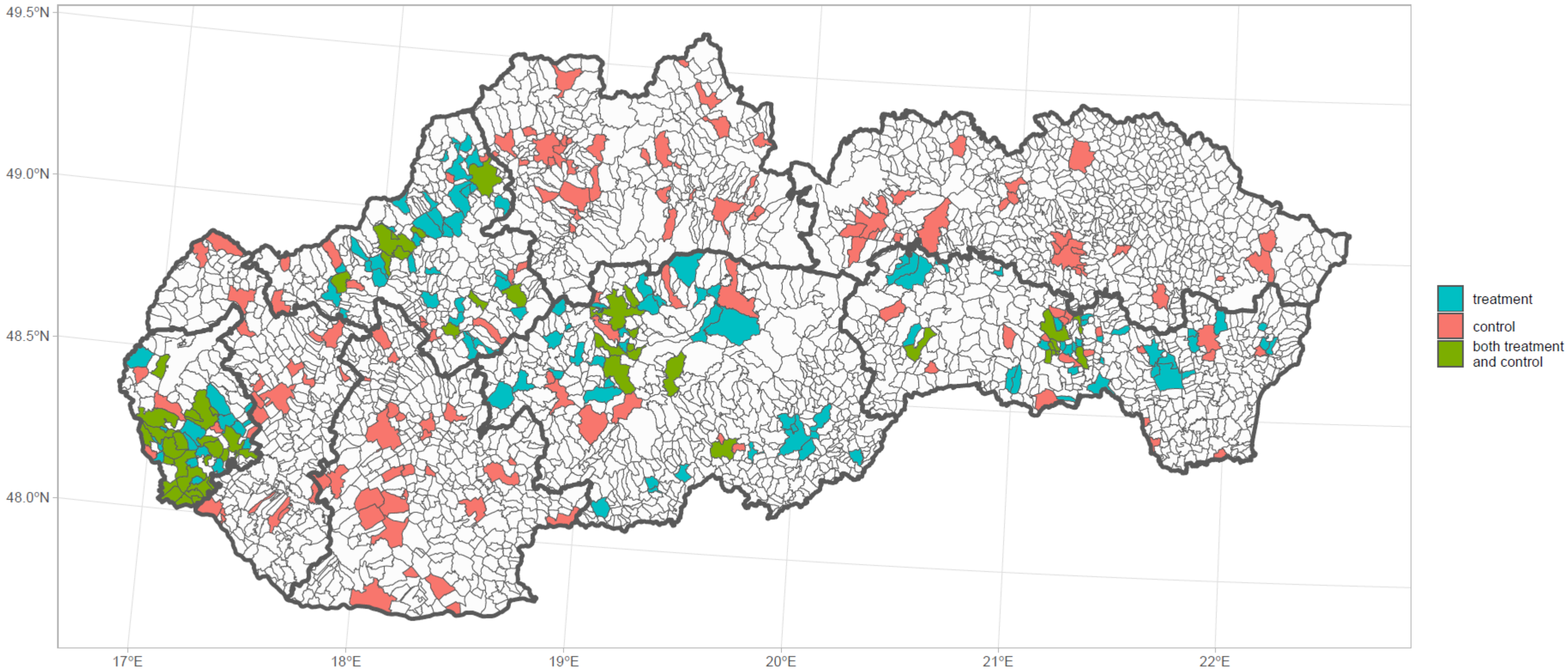
Notes: The table reports estimation results for the probability of participating in the survey estimated on the gross sample following equation (1) yet estimated as linear probability model. Marginal effects at the mean (MEM) are reported. Robust/clustered S.E. are reported in parentheses. Survey controls include information about the daytime of being contacted and its square as well as the rural/urban classification indicator. Regional fixed effects are captured by a set of dummy variables for 8 regions in Slovakia. Interviewer characteristics include age, age squared, level of education, and years of experience.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Source: HFCS 2023/24, National Bank of Slovakia.

Table E.2: HFCS Net Wealth Gini Coefficients

Country	2017	2021	2023
Belgium (BE)	0.632	0.619	0.598
Czech Republic (CZ)	–	0.563	0.529
Germany (DE)	0.739	0.727	0.725
Estonia (EE)	0.709	0.709	0.655
Ireland (IE)	0.703	0.666	0.660
Greece (GR)	0.602	0.580	0.587
Spain (ES)	0.677	0.680	0.674
France (FR)	0.675	0.676	0.677
Croatia (HR)	0.606	0.623	0.539
Italy (IT)	0.606	0.671	0.646
Cyprus (CY)	0.707	0.633	0.610
Latvia (LV)	0.679	0.685	0.690
Lithuania (LT)	0.589	0.580	0.648
Luxembourg (LU)	0.652	0.638	0.631
Hungary (HU)	0.649	0.614	0.598
Malta (MT)	0.602	0.547	0.517
Netherlands (NL)	0.782	0.679	0.652
Austria (AT)	0.730	0.693	0.701
Poland (PL)	0.567	–	–
Portugal (PT)	0.679	0.655	0.648
Slovenia (SI)	0.594	0.593	0.566
Slovakia (SK)	0.540	0.456	0.473
Finland (FI)	0.662	0.683	0.707
Euro area	0.695	0.694	0.686
Median	0.657	0.647	0.647
MAD	0.051	0.035	0.046

Notes: Gini indices are computed for New Wealth according to *Table J4 Net wealth inequality indicators* for the 2nd and 3rd HFCS waves European Central Bank (2023a,b) and first results for the present wave. Median absolute deviation (MAD) for wave t is computed via $\mathrm{MAD}(t) = \mathrm{median}_c \, |G_{c,t} - \mathrm{median}_{c'}(G_{c',t})|$.
Source: HFCS 2017, 2021 and 2023, European Central Bank.

Figure E.1: Geographical Distribution of Sampled Households by Pool



Notes: This figure presents the geographical distribution of oversampled households by treatment status. Colors denote household-level treatment status; apparent geographic clustering reflects the small number of sampled (RP) households per district, not district-level assignment.
Source: Statistical Office of the Slovak Republic.

Figure E.2: Invitation Letter

NÁRODNÁ BANKA SLOVENSKA EUROSYSTÉM

ŠTATISTICKÝ ÚRAD SLOVENSKEJ REPUBLIKY

Bratislava September 19, 2023

Dear Household,

We would like to invite you to participate in the European Household Finance and Consumption Survey (HFCS). In Slovakia, this survey is conducted by the National Bank of Slovakia in cooperation with the Statistical Office of the Slovak Republic. The HFCS is conducted regularly every three years and this year it will be taking place for the 5th time all over Europe.

Your household has been randomly selected from all households in Slovakia and you will be visited by an interviewer from the Statistical Office of the Slovak Republic in the following period. Your participation is extremely important for the quality of the survey, so as a token of our gratitude for your cooperation, we will offer you a reward in the form of shopping vouchers worth €15 and a **silver euro coin** (proof quality with collector's value), which will be given to you by the interviewer from the Statistical Office of the Slovak Republic after you have answered the questionnaire.

The questions in the survey are mainly financial in nature. They cover ownership of real and financial assets, financial liabilities (mortgages and other loans), household income and consumption. We are also interested in the composition of the household, the employment of the adult members of the household and the pension provision. These data provide an important insight into the economic situation of households, which helps us to propose policies to support households in Slovakia and to set up effective support for financial stability.

All information you provide to us will be processed anonymously and treated as sensitive data in accordance with Act No. 18/2018 Coll. on the Protection of Personal Data, as well as with the European legislation on the confidentiality of personal data (GDPR).

If you have any questions, please do not hesitate to contact the interviewer from the Statistical Office of the Slovak Republic, who will visit you at home in the following period, or other employees of the Statistical Office of the Slovak Republic, who can be contacted at the phone number 02/69 250 410 (on weekdays from 9 a.m. to 3 p.m.) and at the e-mail address roman.schieber@statistics.sk. Information on the survey can be found on the website www.nbs.sk/hfcs and on the website of the Statistical Office of the Slovak Republic (www.statistics.sk).

Thank you in advance for your cooperation.

Yours sincerely

Martin Reiner, PhD.
Executive Director in charge of managing the currency, statistics, research and banking operations section of the NBS

Ing. Peter Peťko, MBA
Chairman
Statistical Office of the Slovak Republic

Notes: This figure shows the invitation letter to participate in the HFCS translated to English. The original letter was sent in Slovak. The bold text was added to the version sent to the treatment group, while the letter sent to the control group is identical except that no silver coin is mentioned.