

Reach me if you can: How to Boost Participation of Rich Households in Wealth Surveys?

Sofie R. Waltl¹ (r) Judita Jurašeková Kucserová² (r) Zuzana Brokešová³ (r) Boris Frankovič⁴ (r) Andrej Cupak⁵

Economic Statistics Centre of Excellence (ESCOE)

Disclaimer: The views and results presented in this paper are those of the authors and do not necessarily represent the official opinions of the National Bank of Slovakia or the Eurosystem.

¹University of Cambridge; Vienna University of Economics and Business

²National Bank of Slovakia

³Bratislava University of Economics and Business

⁴Statistical Office of the Slovak Republic

⁵National Bank of Slovakia; Bratislava University of Economics and Business

Motivation



Motivation

- Essential to have a relatively **high proportion of wealthy households** in the sample
- Wealthy households are systematically **under-represented** in surveys ([Vermeulen, 2018](#); [Kennickell, 2019](#))
- **Reasons:** concerns about privacy and sensitivity of financial questions, high time opportunity costs, or a perceived lack of benefit from participating in a survey, ...
- Under-representation affects the **validity** and **reliability** of data and can lead to significant biases
 - E.g., surveys tend to underestimate the top end of the wealth distribution leading to inaccurate conclusions about the extent of wealth and income inequality ([Vermeulen, 2018](#); [Waltl, 2022](#))

Motivation

- Various **ex-post methods** developed to address this issue
 - Re-weighting of survey data to better reflect the population distributions (e.g., [Korinek, Mistiaen, and Ravallion, 2006](#); [Eckerstorfer et al., 2016](#))
 - Combining survey microdata with external sources to adjust collected data ex post (macro or micro) – such as using rich lists or (wealth) tax data (e.g., [Vermeulen, 2018](#); [Waltl and Chakraborty, 2022](#))
 - Combinations of these approaches ([Cantarella, Neri, and Ranalli, 2024](#))
- We take a different approach with an **ex ante perspective** (similar to [Schröder et al., 2013](#))
 - Identification of (likely) wealthy households using administrative Census records (**oversampling**)
 - Randomly assigning them into a treatment and control group to test the effectiveness of a **non-monetary incentive** on the survey participation decision

This paper...

- 1 Improving the over-sampling strategy in the HFCS to better capture high income / wealthy households
- 2 Ex ante intervention: Nudging rich household to take part in the survey by non-monetary (high-value) incentives (**first paper to the best of our knowledge**)
- 3 Assess effectiveness by **comparing incentives to classical oversampling**, and assess effects on other dimensions of **data quality**

Ex ante identification of the rich

The overall SK-HFCS (5th wave; October 2023 to January 2024) sampling strategy consists of two different sample pools from which participants are drawn:

- **General Population Pool (GP)** – *obtained from the database of occupied dwellings from the 2021 Census*
 - Each (private) household residing in Slovakia has a positive probability to be selected
- **Rich Population Pool (RP)** – *obtained from 2021 Census*
 - Consists of **600 wealthy households** (**ex-ante identification** based on 2021 Census making use of information on demographics, wealth, and occupational status [Details](#))
 - Also cross-checked with the list of streets with a high incidence of high income individuals (top 10% in the population) → **94% overlap!**
 - Expect (and test) that households entering the surveyed population from the (RP) sample are on average richer than those entering from the (GP) sample
 - **Target for our experimental study**

Experimental design

- **Withing the (RP) sample:** perform the experiment using a **between-subjects design**
- All SK-HFCS households receive an official **invitation letter** to participate in a **face-to-face wealth survey interview**
- **Official design** with logos of the *National Bank of Slovakia* and the *Statistical Office of the Slovak Republic* and signatures from high-profile representatives of the institutions
- Sample split into a **control group** ($N = 300$) and a **treatment group** ($N = 300$)
- Treatment group receives an **extra incentive** for participating: **pre-announced in invitation letter – a silver proof euro coin** with collector's value.
- Sampled households were randomly assigned to the control or treatment group:
 - 50% **Control group:** standard letter
 - 50% **Treatment group:** information about a reward in the form of a silver proof euro coin was added, **this information was highlighted to attract the reader's attention**

Experimental design



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

... 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...

Experimental design

The Coin

- “100th anniversary of the start of regular broadcasting by Czechoslovak Radio”
- Number of coins minted: 7,200 (proof quality)
- Purchase value at time of the experiment: EUR 30
- Face value: EUR 10
- Collectors value: already tripled



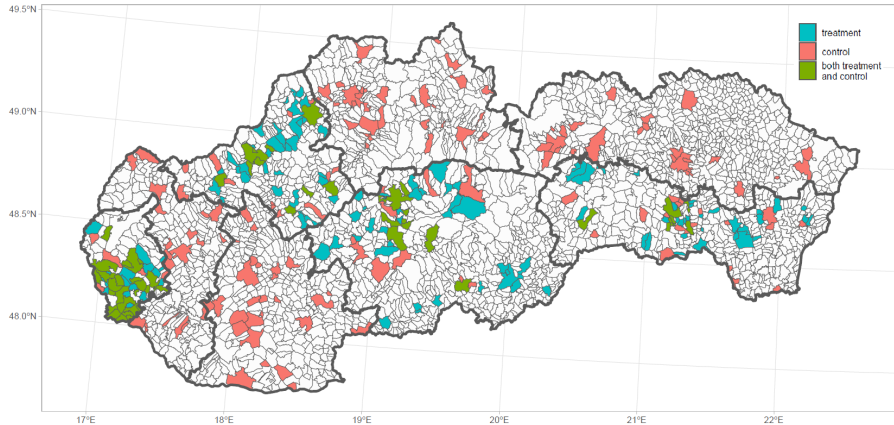
Experimental design

Balance in the gross sample based on the treatment group

	Control (N = 299)			Treated (N = 299)			Differences in Means (p-value)	Differences in Medians (p-value)
	Mean/Share	Median	SD	Mean/Share	Median	SD		
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 (0.263)	0 (0.307)
Daytime of contact	13.000	13	(2.513)	13.137	14	(2.783)	0.137 (0.527)	1 (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 (0.965)	
Household Reference Person								
Gender = male	0.676		(0.469)	0.635		(0.482)	0.041 (0.303)	
Age [years]	49.806	49	(7.644)	50.321	48	(7.785)	-0.515 (0.415)	
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 (0.437)	
Entrepreneur = yes	0.177		(0.383)	0.194		(0.396)	-0.017 (0.600)	

Geographic Balance

Geographical coverage of the experiment



Experimental design

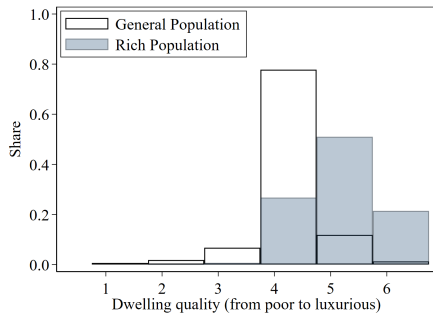
Balance in the net sample based on the treatment group

	Control (N = 86 of 299)			Treated (N = 121 of 299)			Differences in Means (p-value)	Differences in Medians (p-value)
	Mean/Share	Median	SD	Mean/Share	Median	SD		
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)	

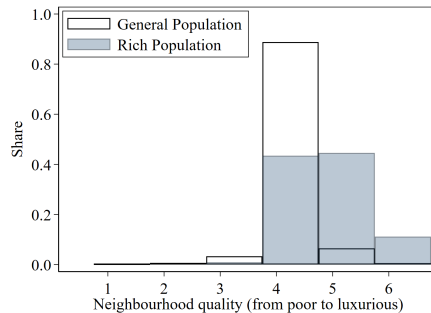
Oversampling validity

I. How do **interviewer ratings** diverge between the (*GP*) and (*RP*) samples?

(a) Dwelling quality



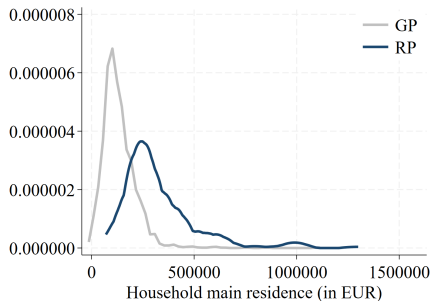
(b) Neighborhood quality



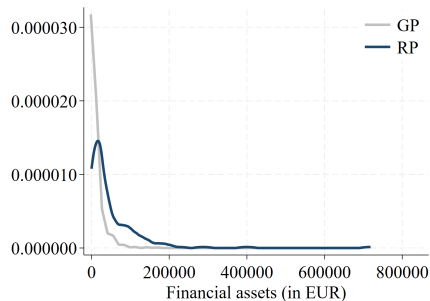
Oversampling validity

II. How does the distribution of holdings in **several asset classes** diverge between the (*GP*) and (*RP*) samples?

(a) Household main residence



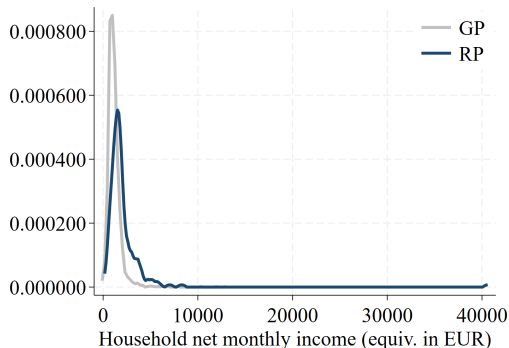
(b) Financial assets



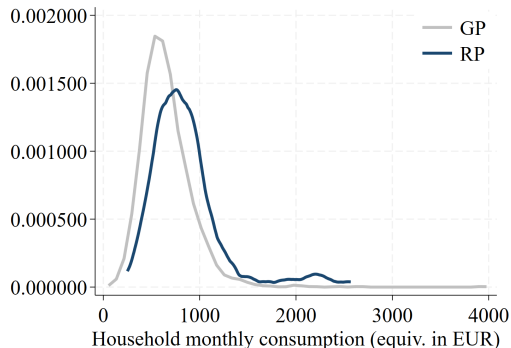
Oversampling validity

III. How does the distribution of **income and consumption** diverge between the (*GP*) and (*RP*) samples?

(a) Household net monthly income

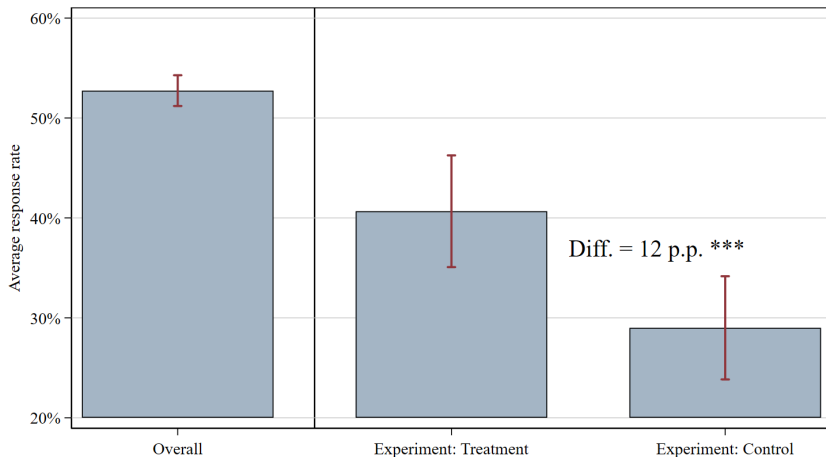


(b) Household monthly consumption



Impact of the incentive

Offering a silver coin to (potentially) rich HHs increases their likelihood to participate in the HFCS by **12 p.p.**: 122 of 300 (treatment) vs 87 of 300 (control)



Impact of the incentive

- We also use a regression framework to **model the probability of participation upon receiving the treatment (=COIN)**
 - Both LPM and Probit (very similar results)
- Controlling for **the way the household was contacted** (in person vs by telephone; =VISIT) and the **number of contacts** (Hsu et al., 2017)
- Accounting for the **interviewer characteristics** ([Crossley et al., 2021](#)) + other controls (urban/rural, regions, time of being contacted...)
- **Clustered S.E.** at the interviewer level
- Allowing for **interaction effects**: COIN $\times$ VISIT, COIN $\times$ Number of contacts, ...
 - Marginal effects of interaction terms in non-linear models computed following ([Ai and Norton, 2003](#))

Impact of the incentive

Probit regression: Participation y/n

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
COIN	0.117*** (0.039)		0.113*** (0.037)	0.105*** (0.037)	0.106*** (0.037)	0.131*** (0.042)	0.130** (0.052)	0.138*** (0.052)	0.130** (0.053)
VISIT		0.341*** (0.033)	0.339*** (0.033)	0.366*** (0.037)	0.361*** (0.038)	0.369*** (0.039)	0.366*** (0.050)	0.285*** (0.066)	0.366*** (0.050)
Number of contacts				-0.122*** (0.027)	-0.118*** (0.027)	-0.124*** (0.029)	-0.115*** (0.036)	-0.112*** (0.036)	-0.114*** (0.039)
COIN x VISIT								0.163* (0.087)	
COIN x Number of contacts									-0.003 (0.051)
Constant	0.288*** (0.026)	0.078*** (0.024)	0.024 (0.031)	0.161*** (0.039)	-0.080 (0.454)	-0.184 (0.462)	-0.667 (0.650)	-0.527 (0.643)	-0.667 (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

Revealed vs. Stated Success

- So far: assessment based on revealed actions (**participation decision**)
- What about **subjectively stated success** as reported by the **interviewers**?
Please estimate the percentage of cases in which the coin offered helped motivate the selected households to participate in the survey.
- Subjectively perceived **median success rate of 10%** – closely aligns with the effect size observed through revealed behavior

Heterogeneity Across Household Types

- To assess whether our intervention is **differentially successful across household types**, we conduct a series of **heterogeneity tests**.
- The **coin does not selectively attract wealthier households**. We find **no statistically significant differences in net wealth, total assets, the value and size of the main residence, or net monthly income and consumption**. Instead, the coin merely **increases participation likelihood**.
- We observe **no significant differences in demographic characteristics**, including **age, gender, number of children, and entrepreneurial status**. The only exception is **education**: the promised coin **slightly increased participation among less educated households**.
- Finally, we test whether the coin attracts **financially sophisticated households**. Using the **diversity of asset portfolios** as a proxy, we find that treated households hold on average **5.32 asset classes**, compared to **4.64** in the control group, indicating that the coin **disproportionately attracts financially sophisticated households**.

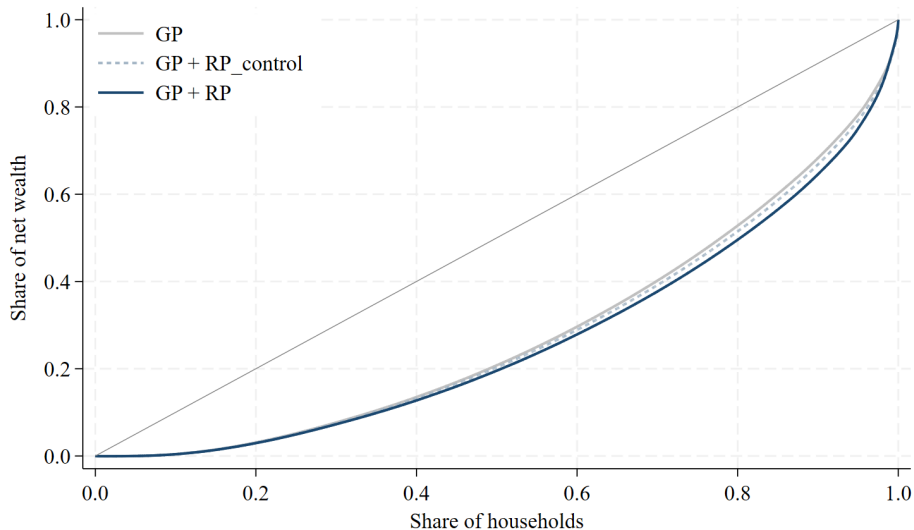
Implications for the measured wealth distribution

Hypothesized and measured effects of oversampling and incentivization

<i>Hypothesised</i>						
Sample	Median Net Wealth	Mean Net Wealth	Mean-to-Median Ratio	Gini Coefficient	Coverage (Fin. Assets)	Coverage (Liabilities)
(GP)	lowest	lowest	lowest	lowest	lowest	lowest
$(GP) \cup (RP)_{\text{Control}}$	mid	mid	mid	mid	mid	mid
$(GP) \cup (RP)$	highest	highest	highest	highest	highest	highest
<i>Measured</i>						
Sample	Median Net Wealth	Mean Net Wealth	Mean-to-Median Ratio	Gini Coefficient	Coverage (Fin. Assets)	Coverage (Liabilities)
(GP)	124867	156922	1.26	44.4	28.69%	57.14%
$(GP) \cup (RP)_{\text{Control}}$	124895	161433	1.29	45.6	28.76%	57.52%
$(GP) \cup (RP)$	125412	168296	1.34	47.3	30.27%	57.54%

Implications for the measured wealth distribution

Shift in the Lorenz curves for different subsamples



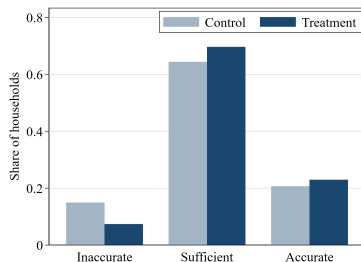
Quality of Responses

- Do participants entering the sample after having been **incentivised** provide **better or worse responses**? Two potential hypotheses:
 - (+) Feel obliged to properly engage as they were paid
 - (-) Only participate for the sake of the reward but don't care about the survey
- Make use of **interviewer rating**:
 - "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*
- Assess **item non-response**

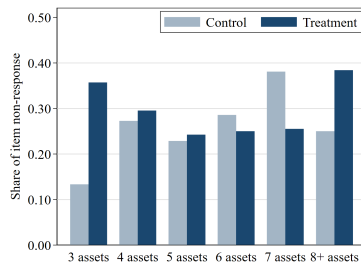
Quality of Responses

The treatment had little yet no statistically significant effect on the quality of the provided data. (Tendency to improve data quality.)

(a) Interviewer rating on response quality



(b) Share of item non-response



Summary

- Experiment within the SK-HFCS aimed to test an **incentivised probabilistic strategy** to recruit more **rich households**
- Incentives were **low-cost** yet still **significantly pushed up response rates**
- **Treatment effect** of roughly 12 p.p. – very robust
- Best results in terms of participation achieved when **aiming for a personal visit in combination with an incentive**
- **Financially sophisticated** households react slightly more strongly to the incentive.
- Measured wealth inequality and concentration increases when better capturing the top tail – as sampling strategies vary strongly across countries participating in the HFCS, this result questions the **reliability of comparative statistics based on the HFCS**.
- The **quality of responses** remained **unaffected**

Thank you for your attention!

Questions?

References I

- Ai, Chunrong and Edward C Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters* 80 (1):123–129.
- Cantarella, Michele, Andrea Neri, and Maria Giovanna Ranalli. 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.
- Crossley, Thomas F, Tobias Schmidt, Panagiota Tzamourani, and Joachim K Winter. 2021. "Interviewer effects and the measurement of financial literacy." *Journal of the Royal Statistical Society Series A: Statistics in Society* 184 (1):150–178.
- Eckerstorfer, Paul, Johannes Halak, Jakob Kapeller, Bernhard Schütz, Florian Springholz, and Rafael Wildauer. 2016. "Correcting for the missing rich: An application to wealth survey data." *Review of Income and Wealth* 62 (4):605–627.
- Kennickell, Arthur 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.

References II

- Korinek, Anton, Johan A Mistiaen, and Martin Ravallion. 2006. "Survey nonresponse and the distribution of income." *The Journal of Economic Inequality* 4:33–55.
- Schröder, Mathis, Denise Saßenroth, John Körtner, Martin Kroh, and Jürgen Schupp. 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.
- Vermeulen, Philip. 2018. "How fat is the top tail of the wealth distribution?" *Review of Income and Wealth* 64 (2):357–387.
- Waltl, Sofie R. 2022. "Wealth inequality: A hybrid approach toward multidimensional distributional National Accounts in Europe." *Review of Income and Wealth* 68 (1):74–108.
- Waltl, Sofie R and Robin Chakraborty. 2022. "Missing the wealthy in the HFCS: micro problems with macro implications." *The Journal of Economic Inequality* 20 (1):169–203.

- (1) being an owner-occupier,
- (2) residing in a single-family house with a minimum of 200 sqm living space and built in or after 2001,⁶
- (3) being at least 40 years old,⁷
- (4) have a high occupational status.⁸

⁶In 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. The scarcity is a consequence of large-scale destruction in the 20th century.

⁷This restriction applies to the reference person of the household used in the census. The census selects the person with the highest educational level attained as the reference person. In case of ties, the person with the higher labor status is selected.

⁸High occupational status is defined based on occupations with high average annual earnings, using data from the EU-wide *Structure of Earnings Survey (SES)*. The classification comprises senior managerial positions, a wide range of professional and technical occupations – particularly in science, engineering, information and communication technologies, finance, law, and medicine – as well as selected supervisory and associate professional roles. In addition, self-employed legal entities (entrepreneurs) are included in the high occupational status group.