Mariya Hake and Philipp Poyntner

Keeping up with the Novaks? Income distribution as a determinant of household debt in CESEE

BOFIT Discussion Papers
Editor-in-Chief Zuzana Fungáčová

Keeping up with the Novaks? Income distribution as a determinant of household debt in CESEE

Mariya Hake [*‡]   Philipp Poyntner [‡]

February 12, 2020

Abstract

This paper constitutes an initial attempt to shed light on the role of income distribution in household debt and financial market access in Central, Eastern and Southeastern Europe (CESEE). Using household-level data from the OeNB's Euro Survey for the period 2009-2018, we address the question whether interpersonal comparisons ("keeping up with the CESEE Joneses" i.e. "the Novaks") affect the probability of having and planning a loan. Applying multilevel probit modeling to take into account the hierarchical structure of the data, our results support the notion that higher income inequality is negatively correlated with the probability of having a loan at the bottom of the distribution, and positively at the top. We show this impact for almost all components of household debt, but evidence is strongest for mortgage, car and foreign currency loans. Interpersonal comparisons turn out to drive loan intentions, however, mainly on the very top of the income distribution.

Keywords: household loans, relative income, income distribution, multilevel models, CESEE

JEL Classification: G0, D1, D3

[1] OeNB, Foreign Research Division, Otto-Wagner-Platz 3, POB 61, A-1011 Vienna, Tel. +43140420-5212. `mariya.hake@oenb.at`, orcid.org/0000-0002-6106-6789.

[2] Vienna University of Economics and Business, Institute for International Economics, Welthandelsplatz 1, A-1020 Vienna, `philipp.poyntner@wu.ac.at`, orcid.org/0000-0001-5346-2100.

1 Introduction

The global financial crisis that started in 2008 has increasingly drawn attention to the importance of, and the threats arising from, household sector debt for macroeconomic stability and GDP growth. At the same time, available literature has shown that also lower levels of household debt might be detrimental for economic growth and a cause for banking crisis. Therefore, policymakers and researchers alike have increasingly turned attention to the factors driving household indebtedness and whether the role of any of these have changed in conditions of high liquidity in the past years. Our analysis of this question for households in the countries of Central, Eastern and Southeastern Europe (CESEE) for the period 2009-2018 constitutes a comprehensive endeavour for this region to relate debt and income distribution (Hake and Poyntner 2019).

While income has been recognized as a major factor for having a loan, only few studies have explored the link between income distribution and household debt but for other regions (e.g. Kumhof et al. 2015, Iacoviello 2008, Loschiavo 2016). Against that background, our paper uses data from a household survey performed in ten CESEE countries[1]. The set of countries is an interesting case as levels of household debt there have remained below the levels experienced in other parts of the world (e.g USA, OECD, euro area countries), while our data show that income has been strongly unequally distributed in some of them with Gini coefficients close to 0.5.

Our paper tests, the "Veblen"-effect - or the "keeping up with the Novaks"- effect[2] as we name it for the individuals in the CESEE countries- the idea that the own consumption is driven by the consumption of a more affluent reference group thus ultimately driving also indebtedness (e.g Carr and Jayadev 2015). For that purpose we apply a less known measure of income inequality i.e the relative income ratio, which gives the average income of other households in similar groups as compared to the own income (in line with papers such as

---

[1]The CESEE country aggregate in this paper includes the EU Member States (Bulgaria, Croatia, the Czech Republic, Hungary, Poland, Romania) and (potential) EU candidate countries (Albania, Bosnia and Herzegovina, North Macedonia and Serbia).

[2]The "Keeping up with the Joneses"- effect goes back to Duesenberry et al. (1949) and Veblen (2017) and stresses the social nature of portfolio decisions. Accordingly, due to externalities (i.e the decisions of the others), the own consumption depends also on the average consumption in the economy. As incomes at the top of the income distribution grow, individuals at the bottom see their relative consumption levels decrease and increase their dissaving and/or borrowing respectively.

Drechsel-Grau and Schmid 2014). Furthermore, income inequality could be a "signaling"-factor indicating the creditworthiness of borrowers to lenders (i.e banks). Accordingly, *ceteris paribus*, with a more unequal income distribution, lenders would become more risk averse and tend to lend to more affluent households rather than households on the lower end of the distribution (Coibion et al. 2014).

Our paper is a pioneer analysis in a number of aspects. It is one of the first to relate household indebtedness and relative income as a measure of income distribution for a set of CESEE countries. Thereby, the survey data at hand traces income and the probability of being indebted in a repeated cross sectional setting for a period of ten years. We are also one of the first to test whether "conspicuous" consumption might play a role as a borrowing motive in CESEE by looking into different loans according to their purpose. In addition, the analysis explores loan intentions, thus, tries to disentangle, to the extent possible, demand and supply factors for household indebtedness. We apply multilevel methodology to account for the correlation of responses from individuals from the same region and country. Finally, we provide a concise test on which are the most suitable reference groups including also a spatial aspect.

To foreground our findings, we show that the relative income position has an impact on households' likelihood of having a loan, but this is valid mainly for households above the median of the both regional and country income distribution. We show this impact for almost all components of household debt, but evidence is strongest for mortgage and car loans. In the preferred specifications, we find that for instance an increase in reference group income by one increases the likelihood of having a loan by about 2% in the 7. decile, controlling for household income. The magnitude of this effect does not vary with comparator groups and is positive, relatively large, and significant across most specifications. Interpersonal comparisons turn out to be a weaker predictor of a household's propensity to have a loan in CESEE countries with a more equal income distribution i.e hinting to a threshold effect. The support we find for loan intentions is valid only in the very top of the income distribution. However, the magnitude and significance of the coefficient of the intentions to demand a loan, is overall lower but also non-negligible.

The rest of the paper is structured as follows. In the next section, we start with a brief discussion on the channels of influence of income distribution on the probability of having a loan. The description of the data and the methodological set-up follow in sections 3 and 4. In section 5 we turn attention to the results as regards the likelihood of having already

a loan digging more into details with respect to its purpose and currency denomination. Section 6 zooms in on the results for the probability of households' loan intentions in CESEE. Going further, sections 7 and 8 focus on alternative measures of income inequality and test the suitability of reference groups on the regional level, respectively. The last section concludes.

2 Literature review: Relative income and household indebtedness

The relevance of income inequality for the occurrence of household debt can be analyzed from both a supply- and a demand-side perspective. There are at least two channels through which income distribution might impact household indebtedness from a demand-side perspective. First, according to the relative income theory of consumption (Veblen and Galbraith 1973; Duesenberry et al. 1949) an individual's utility function depends on the ratio of his or her consumption/income to a weighted average of the consumption/incomes of other persons (i.e. a reference group). Therefore, a more frequent interaction with more affluent individuals would drive up a person's spending when income inequality increases and drive the more frequent occurrence of household indebtedness ("keeping up with the Joneses" or in the CESEE context - "keeping up with the Novaks"-effect). Second, a habit formation effect would prompt an increase in consumer spending, and thus borrowing, when individuals try to avoid cutting down on the level of consumption already attained (e.g. Fasianos et al. 2017, Iacoviello 2008).

From a supply-side perspective, banks use income distribution data next to information on a household's income to assess borrowers' creditworthiness, especially in countries with low credit register coverage (as is the case in some of the CESEE countries in our sample). Coibion et al. (2014) refer to this as the "signaling channel" and show that banks cannot observe to a sufficient extent borrowers' ability to meet debt obligations, so they consider the observed respondents' income together with income inequality in the region or the country. Accordingly, they tend to restrict lending funds when income inequality increases. In addition, higher income inequality might lead to an elevated loanable funds as the people on the higher end of the income distribution generally have a higher propensity to save (Kumhof et al. 2015). In addition, loans in foreign currency in CESEE, although related primarily to mortgage lending, might also be influenced by income inequality as, for instance, banks would consider the exchange risk when assessing the creditworthiness

of the borrower.

With the survey data at hand, we can first focus on an "equilibrium" transaction, i.e. on whether a respondent has a loan or not, and on the loan's purpose and currency denomination. In a second step, we are able to dig deeper into the demand-side aspect of the nexus between relative income and the probability to demand a loan. However, as shown by Bazillier and Hericourt (2017), the two channels are usually activated simultaneously, and the prevailing net effect could go either way. In this paper, we believe to show that that the respondents' position in the country's income distribution would allow for a disaggregated view and hint at a prevalence of either supply or demand factors without, however, excluding the impact of one or the other.

Going further, an extensive body of literature has turned attention to exploring the hypothesis that individuals derive utility from status, which in turn depends on what others believe about people's income (Ireland 1994, Charles et al. 2009). Although income is not observable, consumption is very frequently visible. Therefore, the level of individuals' conspicuous consumption (i.e. consumption that displays social status) can be expected to depend on the income distribution of the entire sample of individuals under observation. Against this background, some goods and loans, respectively, would be driven by conspicuous motives. Therefore, the impact of the income distribution on the likelihood of having a loan and on the loan's purpose (e.g. consumption, car or mortgage) will enable us to make inferences about the existence and magnitude of such motives.

The literature on the relative income hypothesis remains silent on the questions where the "Novaks" are. Are households perceptive of the income of their direct neighbors, households in the same city, region or country? Often, choices defining the "Novaks" are practically bound by data constraints. This is also the case in our study, as data limitations prevent us from performing an analysis at a level more disaggregated than the regional level. However, it is still possible to test differences in outcomes for more coarse neighborhood definitions. In particular, we investigate if neighboring regions affect household debt. To do so, we separately estimate the effects of reference income on debt for the home region and neighboring regions. We also investigate if defining the whole country as reference group instead of the own region changes the results.

On the empirical front, papers only recently started to account for the distribution of income as a driver of household debt, focusing mainly on industrialized countries and the

USA, in particular. For instance, Kumhof et al. (2015) show that for the USA that surge in the income share of the top deciles could largely explain the buildup of leverage among households at the bottom of the income distribution. Building on the central assumption that income shocks are of a permanent nature (and that a a change in income distribution is therefore understood as a permanent shock), the authors present a model that shows that higher leverage and financial crises are the endogenous result of a growing income share of high-income households.

Also for the USA, Coibion et al. (2014) show that in the period from 2001 to 2012, low-income households in high-inequality regions accumulated less debt relative to income than their counterparts in regions with lower income inequality. For Italy, Loschiavo (2016) shows that richer households living in regions with higher income inequality are more likely to be indebted than similarly rich households residing in regions with low income inequality (and vice versa for poorer households). Carr and Jayadev (2015) is a study closely related to ours as they relate relative income of the household and debt while focusing again on US household data. Analyzing the period 1999-2009, they find a confirmation for the "Novaks"-effect as the growth of household debt increases with the share of families with income higher than the own family. In addition, their findings show that low income households are more likely to leverage up than higher income households during the increase in household leverage in the early 2000s in USA.

The only paper intimately related to this study and covering the same set of CESEE countries is Hake and Poyntner (2019). The authors perform an initial analysis about the correlation between the probability of having a loan and income inequality for the period 2009-2017. However, in contrast to the present study, they neither explore in detail the impact on different types of loan, various income inequality measures, loan intentions nor look into the spatial aspect of the definition of the reference groups.

3 Data and descriptive statistics

3.1 Data

This paper is based on data from the OeNB Euro Survey, which is a household survey performed in ten CESEE countries and was commissioned by the Austrian central bank (OeNB). The survey was performed bi-annually between 2007 and 2014 and annually since

2015. It includes six EU member countries (Bulgaria-BG, Croatia-HR, Czech Republic-CZ, Hungary-HU, Poland-PL and Romania-RO) and four EU (potential) candidates (Albania-AL, Bosnia and Herzegovina-BA, Serbia-RS and North Macedonia- MK). In each country and per wave, the target population comprises around 1000 interviewees representative of the country's population, 14 years or older, selected via a multi-stage stratified random sampling procedure. For the period of analysis of this paper (2009 to 2018), this corresponds to a total number of observations of close to 120,000[3]. The survey includes questions on the use of the euro in household's portfolio, both for deposits and loans and the purpose of the latter, but also on different sentiments about the future and experiences from the past as well as questions on wealth and income. It also collects information on socio-demographic characteristics, including age, education and employment status. The survey delivers information on a more disaggregated level (i.e regional level roughly corresponding to the NUTS 2 and on the level of primary sample units - PSUs)[4].

This paper applies newly computed annual income inequality (i.e distribution) measures that are comparable over time (i.e. for 2009 to 2018) and across the ten CESEE countries following Belabed and Hake (2018). The income distributional measures are based on the OeNB Euro Survey question: "What is the total monthly income of the household after taxes?"[5]. Between 2009 and 2016, survey respondents were asked to place their income in one of 20 categories, which were defined such that a maximum of 10% of respondents fall into each category. The ranges of categories were harmonized across the different countries and over the years, amounts were transformed into euro and into purchasing power units (to capture exchange rate and inflation differences) to ensure cross-country comparability. We then took the average of each income category to compute the equalized household income. In the 2017 & 2018 survey waves, respondents were asked to report (at least the approximate) amount of their household income and not in categories.

---

[3]The number of observations corresponds to roughly 2,000 per country and year in the period 2009-2014 and roughly 1,000 observations per country and year in the period 2015-2018. Data for the Czech Republic for 2010 was incomplete and is therefore not included.

[4]For more information on the survey, see https://www.oenb.at/en/Monetary-Policy/Surveys/OeNB-Euro-Survey.html

[5]Henceforth,the income data derived from the OeNB Euro Survey refer to net household income and not to households' disposable income.

3.2 Relative income ratio

From demand-side perspective, empirical research has emphasized that interpersonal comparisons are upward-looking: households compare their consumption to richer households and adjust their consumption preferences accordingly (Ferrer-i Carbonell 2005, Carr and Jayadev 2015). Similar to Hake and Poyntner (2019), we follow Drechsel-Grau and Schmid (2014) and define first the households' reference income to account for upward-looking comparisons. In particular, Drechsel-Grau and Schmid (2014) focus on consumption and define reference consumption as the consumption level of all households who are perceived to be richer and the reference consumption ratio as the mean reference consumption to the own consumption. Consequently, the reference income ratio *relinc* as used in our paper follow this approach and is defined as the ratio of the mean reference income to the household's own income.

The reason to opt for this income inequality-measure is threefold. First, as we consider different countries with different currencies, the relative income measure enables the comparison across countries. Second, per definition, this measure incorporates household's own income, allowing thus for a relative reference income measure for each household. Finally, it combines both the impact of income and its distribution, thus alleviating concerns of omitted variables bias.

The relative reference income is therefore defined by:

$$relinc_i = \frac{1}{K} \sum_{\substack{j=1 \\ D_j > D_i}}^{K} (Y_j) \frac{1}{Y_i} \tag{1}$$

where $i = 1, ..., N$ and $j = 1, ..., K$ are households, D_i is the regional income decile of household i, and Y_i is the income of household i. *relinc* is bounded by 1 from below per definition, while the highest value is 82 in the sample (see descriptive statistics in the Appendix).

We decided to define the reference groups as households in the same region. We consider that the "keeping up with the Novaks"-argument refers to individuals being influenced by people they frequently interact with, such as neighbours, family or colleagues. In addition, the regional level is the most granular level of data, which allows us to have a sufficient number of observations as the number of observations at the PSU-level is too small[6].

---

[6]In Section 8, we show that defining the reference group as richer households in the same country (as opposed to region) does not alter the results.

Figure 1: Relative reference income *relinc* by decile and country
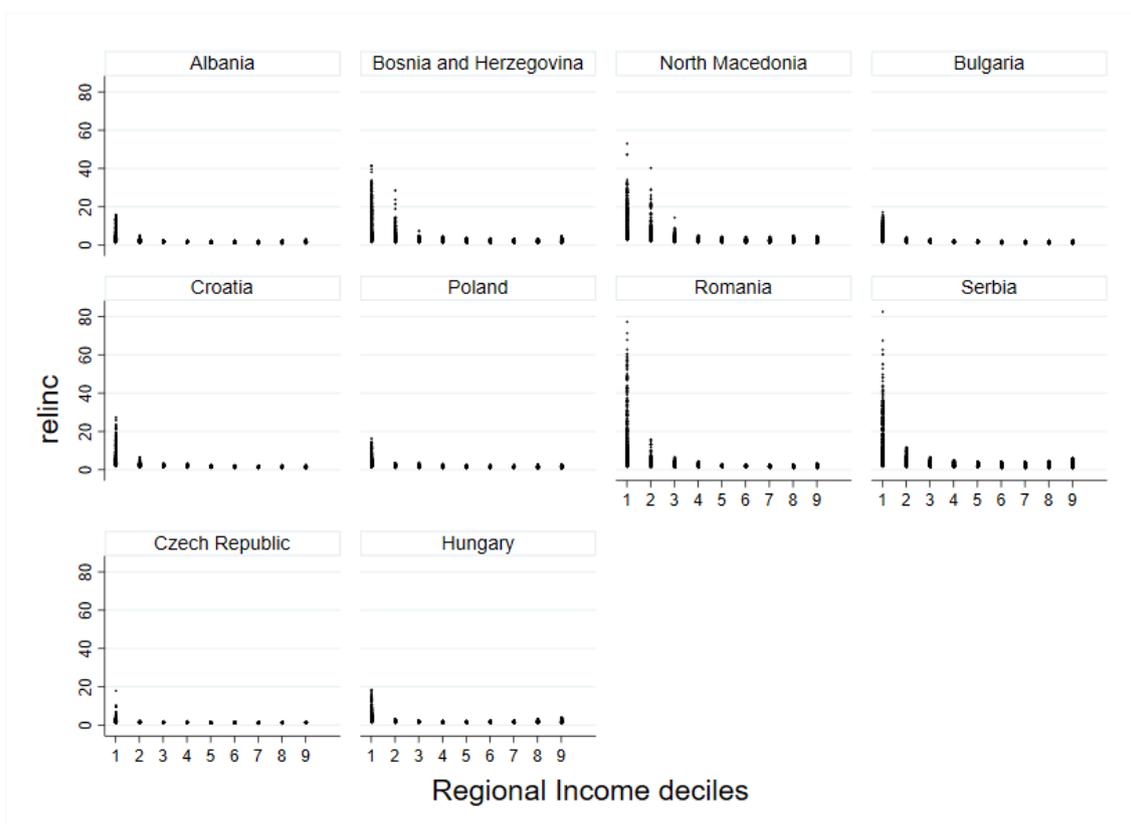
Note: The relative reference income *relinc* is the mean income of households in higher deciles divided by the households' own income. Income deciles are formed at regional level. Source: OeNB EuroSurvey, own calculations.

Figure 1 shows the distribution of households' reference income per country, while the income deciles and the respective comparisons are made and constructed on the regional level. Intuitively, respondents in the lowest income deciles (1st and 2nd) have the highest ratio of up to 82 (in Serbia), implying that the income of a household in the 1st decile could be up to 82 times lower than the average income in all deciles above the 1st decile. For the whole sample, the median of the reference income ratio for all deciles is 2.5, while 99% of all observations are below 15. Judging by the descriptive evidence, there is some country variation as well. Accordingly, the households in the lower end of the income distribution in Bosnia and Herzegovina, North Macedonia, Romania and Serbia tend to be relatively more disadvantaged than households in the same deciles in the rest of the countries.

## 3.3 Descriptive evidence on existing and planned household loans

The OeNB Euro Survey provides information on whether households in the CESEE sample have or plan a loan. Moreover, it also contains information on the loan's purpose (i.e. consumer loan, mortgage loan, car loan or loan for other purposes) and its currency denomination[7]. Noteworthy, the survey data compiled do not give information on the amount of the credit but on its availability, which we consider to somewhat alleviate endogeneity concerns. Figure 2 shows the development of the share of households with loans over time. It also shows the minimum and maximum shares per region in the respective country. Accordingly, the variance of the regional share of respondents with loans is on average the highest in Bosnia and Herzegovina, Romania and the Czech Republic. At the same time, regional variation of indebted households is the lowest in Albania, North Macedonia and Serbia. Overall, the average shares of indebted households are the highest in Czech Republic, Hungary and Croatia reaching close to 40% of total respondents. Figure 3 shows the share of households planning a loan. On average, 8.4% of all respondents plan a loan. The share is the highest in the Czech republic, Hungary and Serbia (see also Table A3 in the Appendix).

3.4 Income distribution and household debt - a first glance at correlations

Figure 4 plots *relinc* and the share of households with a loan (each measure: mean over region and year). We observe a slight negative correlation between relative reference income ratio and loan shares, as depicted by the fitted values indicated with the red line. In Figures 5-7, we plot loan shares and three income inequality measures: the Gini coefficient, the share of the top 10% and top 1% households in the regional income distribution, respectively. The slight negative correlation remains irrespective of the income inequality measure used.

---

[7]The loan question is asked in the following way, *"Do you, either personally or together with your partner, currently have any loans that you are still paying off?"* and *"Do you plan to take out a loan within the next year and if so, in what currency?"*, respectively. If respondents have or plan a loan, they are asked to specify the purpose of the loan: *"to finance a house or apartment," "for consumption goods (furniture, travelling, household appliances, etc.)," "to finance a car" or "for other purposes".*

Figure 2: Current loans
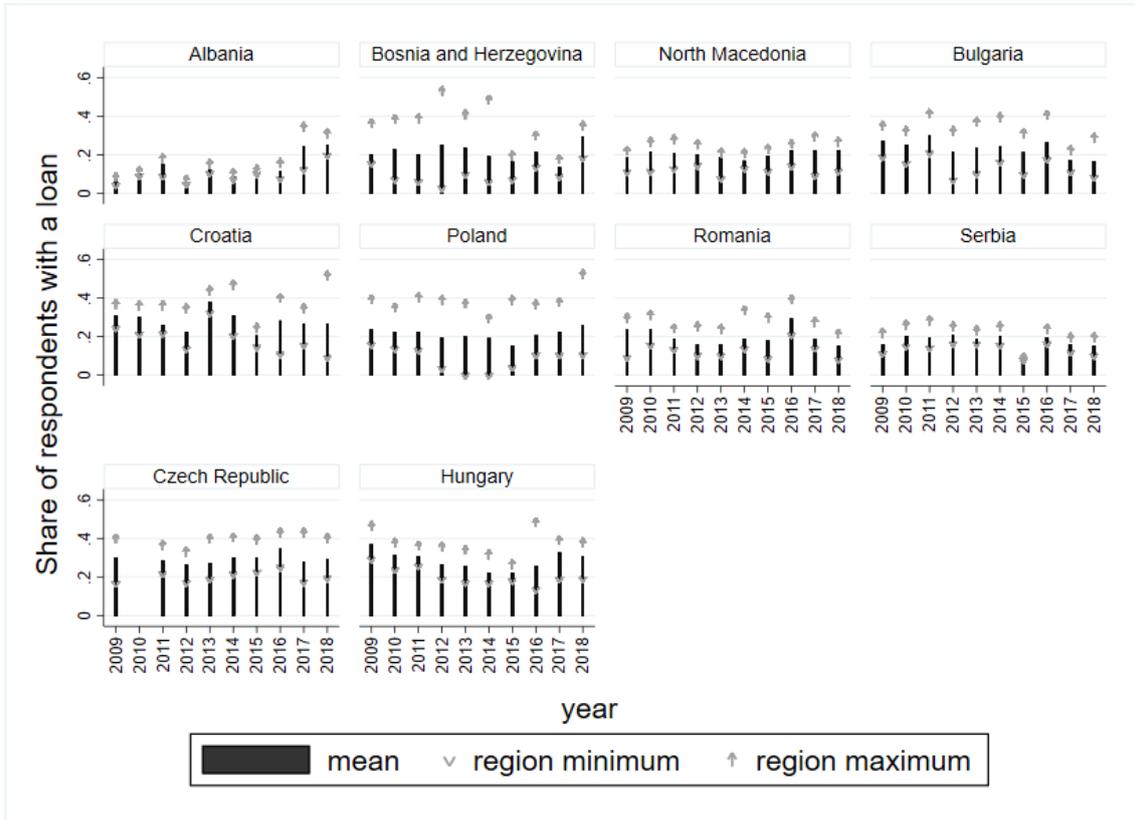
Note: The black bars denote the mean share of households with a loan of regions in a country, the upper arrow shows the maximum share, the v the minimum share. Source: OeNB EuroSurvey, own calculations.

Figure 3: Planned loans

11
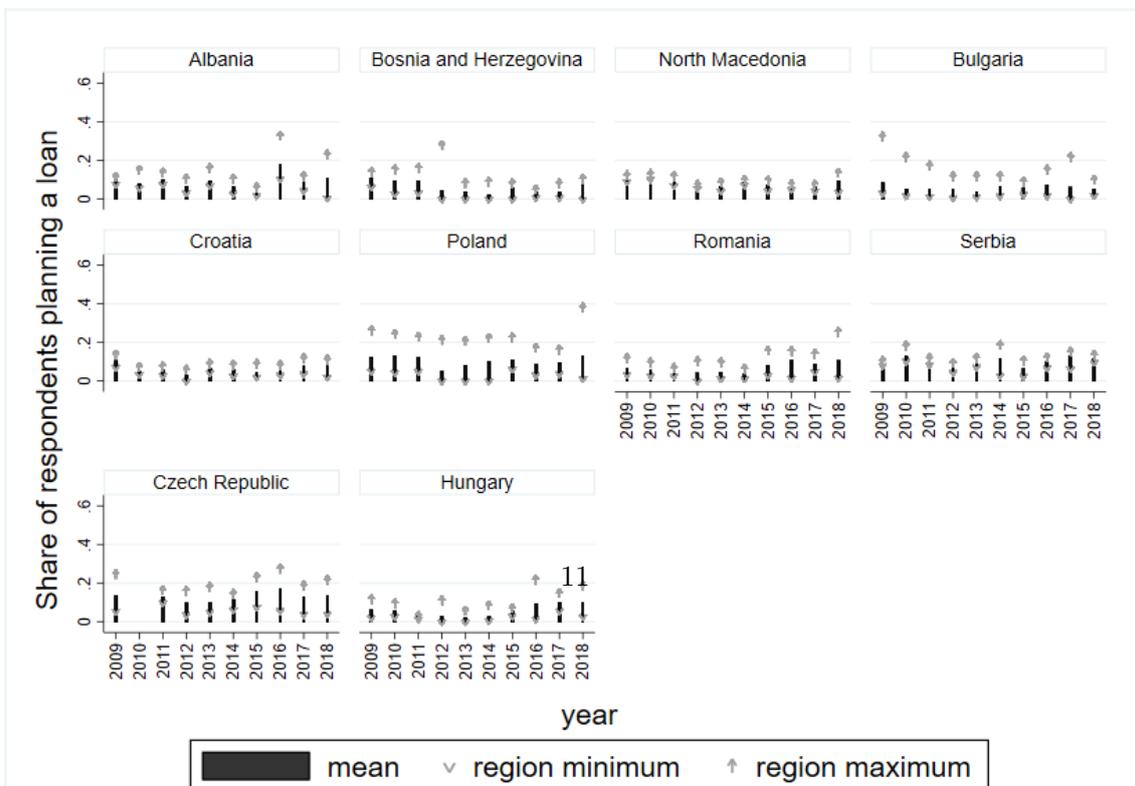
Figure 4: Correlation between *relinc* and
current loan share
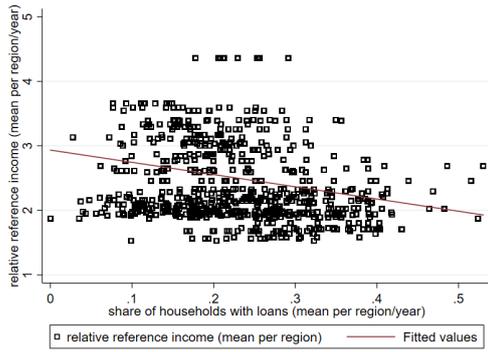
Figure 5: Correlation between Gini
coefficient and current loan share
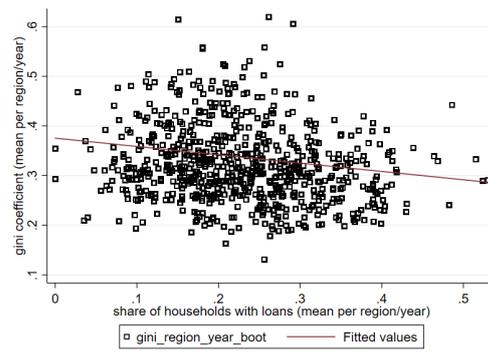

Figure 6: Correlation between Top 10%
share and current loan share
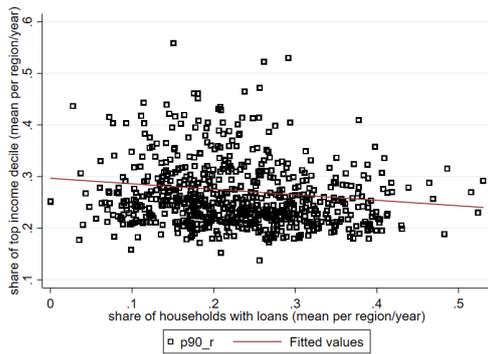
Figure 7: Correlation between Top 1% share
and current loan share
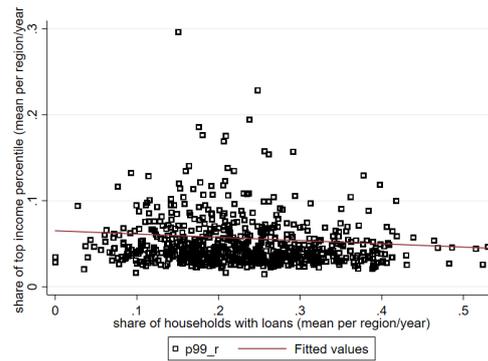
Note: The dots are region-year means for the share of households with loans and *relinc*, the Gini coefficient, the regional share of the Top 10% households and the regional share of the Top 1% households. Source: OeNB EuroSurvey, own calculations.

4 Methodology

Our chosen methodology accounts for the fact that both regional and individual factors might correlate with the household likelihood to have or to intend to take up a loan. In particular, we assume that the "Keeping up with the Novaks"- channel is most likely located at the regional level as the interaction in space and time tend to be the highest among households that are spatially close to each other. Accordingly, the correlation between inequality and household debt is likely to differ between households in different regions and countries, as households' characteristics interact with institutional characteristics. Hence, the assumption that observations are conditionally independent given the covariates is likely to be violated because of interdependence of households in a region. Disregarding this interdependence can lead to spurious "significant" coefficients of the included variables.

Against that background and to account for the nested structure of the data (i.e "clusters"), we apply multilevel models (Rabe-Hesketh and Skrondal 2012), widely used otherwise, for instance, in health (e.g. patients nested in hospitals) or social sciences (e.g. students nested in schools). In our paper, as our outcome of interest is a binary variable (the household has or plans to take on debt), we apply a multilevel probit model. It contains random effects to account for the interdependence of observations at cluster level.

We consider the following two-level models:

$$Pr(loan_{ir} = 1|X_{ir}, U_r) = H(X_{ij}\beta + Z_{ir}U_r) \tag{2}$$

$$Pr(planloan_{ir} = 1|X_{ir}, U_r) = H(X_{ij}\beta + Z_{ir}U_r) \tag{3}$$

with $r = 1, ...76$ clusters, in our case regions, consisting of $i = 1, ..., i_n$ households. X_{ir} is a $1 \times p$ vector of covariates, and β is a vector of regression coefficients. The $1 \times p$ vector Z_{ir} denotes the random effects both in intercepts and coefficients. U_r denotes the random effects. $H(.)$ is the standard normal cumulative distribution function. The same approach has been applied for the different types of loans as well.

The choice of the level of the analysis (i.e regional, country) is key for the model as it determines the level of the random effects. The dependent variable has to show some unexplained variance related to the cluster. Therefore, would hardly any variance be left

after including all the variables on the individual/household level, would an inclusion of the regional level be superfluous.

The intraclass correlation (ICC) gives an indication for the correct choice of the level. In the two levels specification (i.e individual and region effects), the overall error term w_{jt} is decomposed into e_{ijt} and u_{ijt}, where e_{ijt} is the random error term for the $i-th$ respondent within the $j-th$ region and is assumed to have zero mean and constant variance σ_e^2. The regional effects are estimated through u_{jt} which is assumed random and again has a zero mean and a constant variance σ_u^2. The partitioning of the variance in this manner defines a measure to test the suitability of the multilevel modeling i.e the intraclass variance coefficient (ICC or ρ). It measures the strength of 'nesting' with the data hierarchy and is given by:

$$ICC = \rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \tag{4}$$

Hence, ICC tends to be used as a "goodness of fit" for the use of multilevel models i.e the higher this share (i.e in any case significantly different than zero), the more suitable is the application of multilevel modeling.

Table 1 suggests that introducing both regions and countries as levels into our multilevel model improves our estimations.[8] However, small number of clusters at the second level (in our case, countries), can lead to a severe bias. While the literature cannot offer a clear-cut indication on how many level two clusters are sufficient (e.g. Maas and Hox 2005, Schunck 2016), with only 10 countries, the sample size at the top level of our hierarchy is very low by any measure and the probability of downward bias would be present if we included both regions and countries as levels. Against that background i.e due to the cluster size and the lower ICCs at country level, we opted for including only the regional level.[9] In addition, the country variation has been captured by country dummies i.e allowing for different intercepts but not allowing for different coefficient estimates at country level.

---

[8]Table A5 in the Appendix shows the results of a non-multilevel probit estimation with region, country and time fixed effects. The main results hold qualitatively and quantitatively.

[9]In Table A4, we demonstrate that including countries as levels does not alter our main findings for the baseline estimates for current loans.

Table 1: ICCs

# of levels	Level	ICC	Std. Err.	95% Conf. Int.	
1	country	0.035	0.035	0.014	0.084
1	region	0.042	0.006	0.031	0.056
2	country	0.031	0.015	0.012	0.078
2	region	0.049	0.019	0.029	0.084

Note: ICCs calculated after estimation of multilevel probit models without covariates, dependent variable: current loan.

5 Household debt at the "equilibrium"

5.1 Baseline estimations

Table 2 presents the results of the headline estimations for existing "equilibrium" household debt. We include major sociodemographic variables such as age, education and employment status in all estimations. The results in the column 1 with the constant only included show the explanatory power of the included variable on the regional level. Accordingly, nearly 4% of the variation could be explained by the regional level-variables. Going forward, the level effect (i.e without differentiating among the income deciles) indicates that a higher average regional relative income ratio of all households correlates negatively with the likelihood to have a loan (column 2). Therefore, a more unequal regional distribution reduces the likelihood to be indebted. The results in columns (3)-(7) show a positive correlation between the reference income ratio above the 5th income decile of the regional distribution. Interestingly, for the lowest income households (the bottom 20%), higher reference income is correlated with a lower probability of having a loan. We understand these results to be consistent with outcomes where the signaling channel (i.e supply) is larger than the "Keeping up with the Novaks"-channel (i.e demand) as with more unequal income distribution access to finance of less affluent households would worsen. The results for the top of the distribution are in line with both the signaling and the Novaks channel: The signalling channel suggests that higher inequality leads to more and cheaper credit supply for the top of the distribution, and the "Novaks"-channel suggests that higher reference income leads to higher consumption.

Splitting the regions along the level of the average reference income ratio (i.e 2.5) would not change the results significantly but hints that the negative effects in the lower end of

the regional income distribution are driven by low income inequality regions. In a next step, we add proxies for wealth, namely if the households hold savings, cars, or property. The latter two questions were only included from 2011 on, therefore column 6 repeats the estimation in column 2 but for this sub-sample. While we see a smaller correlation for the bottom of the distribution, most likely due to the smaller sample time period, the effects are in size similar, suggesting that our results are not driven by confounding wealth effects affecting debt intake.

The included sociodemographics have largely the expected signs. Older respondents are more likely to have a loan but this effect is non-linear starting to decline above the age of 53 and in line with the negative sign of the "retired"-dummy. Higher education (i.e university degree) is also credit inducing. In addition, female respondents and respondents with children are more likely to have a loan. "The goodness of fit" i.e ICC varies between 1.4% and 3.8% showing that up to 3.8% of the variance is due to variation between levels (i.e regions) thus supporting the application of multilevel models.

As the coefficients so far show the qualitative stance, we estimate the average marginal coefficients to get a better idea on the size of the effect i.e on the economic implications of our findings (Figure 8). For instance, the average marginal effect of respondents' income ratio in the 7th income decile is 0.02, which means that an increase of the relative income ratio in this decile by 1 unit (i.e. for instance, from 2 to 3) would increase the likelihood of a household in this decile having a loan by 2 percentage points. Considering that the share of indebted households in the upper deciles is higher than in the rest of the income distribution, the overall effect of the relative income ratio on household indebtedness would be non-negligible.

16

Table 2: Household loans in CESEE

	(1) const	(2) no interaction	(3) baseline	(4) low ineq reg	(5) high ineq reg	(6) 2011-2018	(7) wealth
Income Distribution							
1th decile#relincome			-0.014***	-0.027***	-0.006*	-0.011**	-0.014***
			(0.004)	(0.009)	(0.003)	(0.005)	(0.005)
2nd decile#relincome			-0.023***	-0.047***	-0.003	-0.019*	-0.024
			(0.008)	(0.016)	(0.009)	(0.010)	(0.010)
3rd decile#relincome			-0.008	-0.018	0.004	-0.006	-0.010
			(0.010)	(0.015)	(0.012)	(0.012)	(0.013)
4th decile#relincome			0.015	-0.004	0.029*	0.027	0.017
			(0.012)	(0.017)	(0.017)	(0.018)	(0.019)
5th decile#relincome			0.031**	0.017	0.034**	0.050**	0.033
			(0.012)	(0.020)	(0.017)	(0.020)	(0.024)
6th decile#relincome			0.050***	0.048**	0.036*	0.056***	0.040*
			(0.014)	(0.020)	(0.020)	(0.020)	(0.021)
7th decile#relincome			0.068***	0.055**	0.060***	0.077***	0.065**
			(0.015)	(0.027)	(0.018)	(0.023)	(0.028)
8th decile#relincome			0.090**	0.067**	0.076***	0.102***	0.097***
			(0.017)	(0.027)	(0.020)	(0.025)	(0.026)
thdecile#relincome			0.086***	0.080**	0.040**	0.105***	0.095***
			(0.019)	(0.034)	(0.020)	(0.030)	(0.034)
Reference income		-0.015**					
		(0.004)					
Income		0.0001***	-0.0001**	-0.0002***	0.0002**	-0.0001*	-0.0002**
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wealth proxies							
Savings							0.010
							(0.016)
House							0.040***
							(0.028)
Car							0.150***
							(0.022)
Sociodemographics							
Female		0.027***	0.026***	0.020	0.042***	0.029**	0.048***
		(0.010)	(0.010)	(0.014)	(0.014)	(0.014)	(0.014)
Age		0.105***	0.105***	0.103***	0.112***	0.108***	0.105***
		(0.004)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)
Age squared		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Children		0.134***	0.134***	0.152***	0.106***	0.140***	0.138***
		(0.013)	(0.012)	(0.015)	(0.018)	(0.014)	(0.014)
Education		0.141***	0.134***	0.102***	0.191***	0.127***	0.110***
		(0.017)	(0.017)	(0.021)	(0.027)	(0.022)	(0.022)
Unemployed		-0.327***	-0.305***	-0.193***	-0.422***	-0.352***	-0.340***
		(0.029)	(0.029)	(0.040)	(0.024)	(0.035)	(0.036)
Self-employed		-0.007	-0.006	0.031	-0.101	-0.022	-0.048
		(0.031)	(0.031)	(0.034)	(0.064)	(0.041)	(0.046)
Student		-0.694***	-0.692***	-0.753***	-0.579***	-0.671***	-0.675***
		(0.059)	(0.059)	9(0.098)	(0.062)	(0.071)	(0.079)
Retired		-0.120***	-0.117***	-0.154***	-0.107***	-0.133***	-0.108**
		(0.025)	(0.025)	(0.035)	(0.033)	(0.040)	(0.042)
cons	-3.443***	0.016***	0.016***	0.015***	0.014***	0.019***	0.020***
	(0.0850)	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
ICC (regional)	0.038	0.015	0.016	0.015	0.014	0.019	0.019
N	119085	119085	105301	65341	39960	58135	49164

Dependent variable: binary response indicating if respondents have a loan. Estimation method: multi-level modeling. Country and time fixed effects are included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of regional (second) level covariates. Robust standard errors in parentheses. Variables are defined in the Appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 8: Marginal effects of relative reference income at income deciles for the estimations in Table 2, columns 3,4,5 and 7
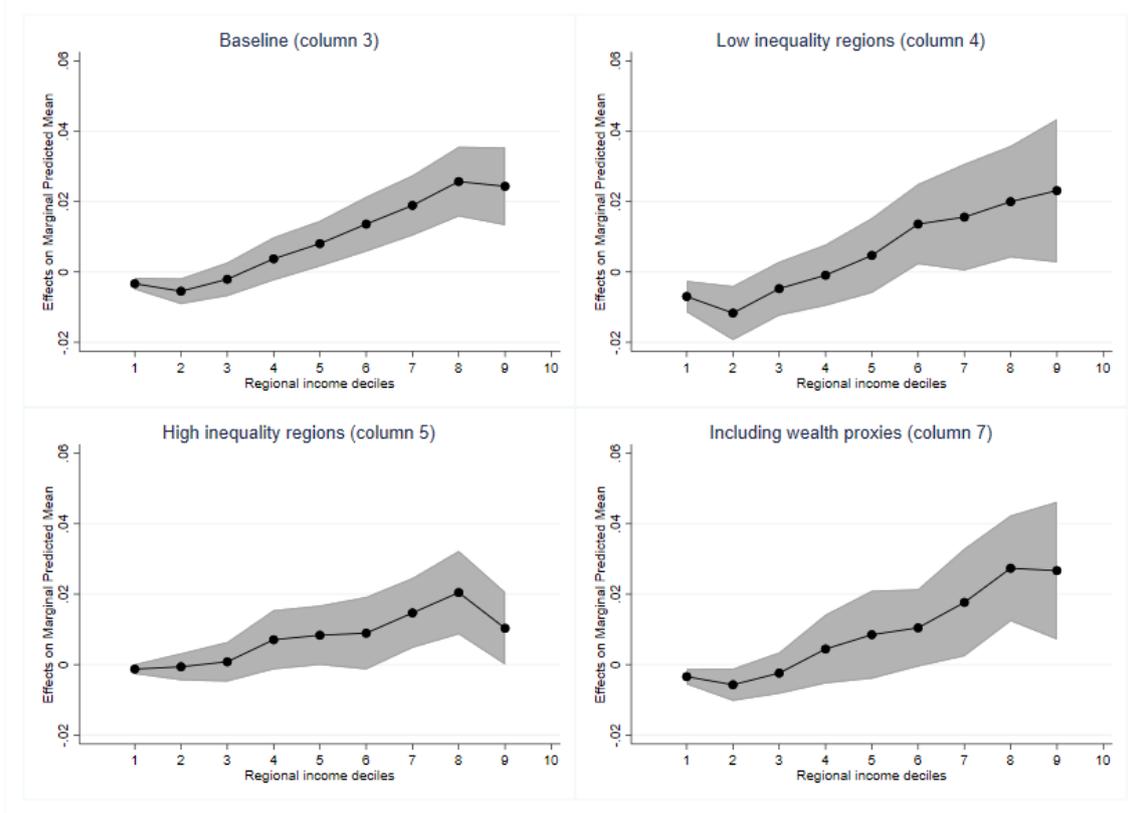
Note: Average marginal effects of relative reference income for selected specifications of estimations presented in Table 2. The grey area depicts the 95% confidence interval. Source: OeNB EuroSurvey, own calculations.

5.2 Current household indebtedness in CESEE - different reference groups

In this section we test the assumption that households are upward-looking in their consumption/income and subsequently in their debt patterns. Our aim is to explore if this measure is driven specifically by certain subgroups. Is the upward-looking component driven by households of the same age, education, or only by those close the households' own income rank?

The estimation results are presented in Table 3. First, for the sake of comparison, we include in column 1 the baseline estimation from Table 2, where the reference group are all households at higher regional income deciles. In column 2, the reference group are all

households only in the next higher decile at regional level. We consider the income of this reference group also very telling as households might focus above all on households that are richer but close to their income rank (Drechsel-Grau and Schmid 2014). Our expectation for larger coefficients in this case are confirmed (for the marginal coefficients see Figure 9). In columns 3 and 4 we define the reference groups to be at the same educational level[10] and in the same age cohort at the regional level, respectively. Our baseline pattern remains unchanged. The magnitude coefficient sizes as in the baseline specification in the former case and somewhat smaller in the latter case. Finally, the reference group in column 5 includes all households, which are more affluent (i.e including the ones in the own regional income decile). An advantage of this approach is that we obtain an estimate also of the top income decile. Also including the 10th decile, the baseline estimation results are confirmed. In a nutshell, there is a robust pattern emerging that does not seem to be entirely driven by specific sub-groups. The largest impact seems to derive from reference income of the neighbouring richer decile.

_____

[10]Education has been defined in categories low, medium and high.

Table 3: Household loans in CESEE - alternative relative reference income concepts

relative reference income group:	(1) baseline	(2) next decile	(3) same education	(4) same age	(5) all richer
Income Distribution					
1th decile#relincome	-0.014***	-0.046***	-0.011***	-0.012***	-0.019***
	(0.0035)	(0.010)	(0.0039)	(0.0033)	(0.0062)
2nd decile#relincome	-0.023***	-0.062***	-0.022***	-0.027***	-0.038***
	(0.0080)	(0.023)	(0.0078)	(0.0076)	(0.012)
3rd decile#relincome	-0.0084	-0.019	-0.0082	-0.014*	-0.019
	(0.0099)	(0.023)	(0.0095)	(0.0082)	(0.015)
4th decile#relincome	0.015	0.016	0.012	-0.00010	-0.016
	(0.012)	(0.024)	(0.011)	(0.011)	(0.017)
5th decile#relincome	0.031**	0.040*	0.027***	0.011	-0.0074
	(0.012)	(0.021)	(0.0099)	(0.0081)	(0.018)
6th decile#relincome	0.050***	0.074***	0.045***	0.023**	0.020
	(0.014)	(0.024)	(0.013)	(0.011)	(0.016)
7th decile#relincome	0.068***	0.092***	0.063***	0.029***	0.025
	(0.015)	(0.025)	(0.013)	(0.011)	(0.018)
8th decilerelincome	0.090***	0.12***	0.079***	0.046***	0.071***
	(0.017)	(0.025)	(0.015)	(0.012)	(0.020)
9th decile#relincome	0.086***	0.081***	0.074***	0.031**	0.055***
	(0.019)	(0.019)	(0.017)	(0.013)	(0.018)
10th decile#relincome					0.094***
					(0.024)
Income	-0.00012**	-0.00013***	-0.000080*	-0.000014	-0.000030
	(0.000047)	(0.000048)	(0.000047)	(0.000046)	(0.000020)
Sociodemographics					
Female	0.027***	0.028***	0.028***	0.027***	0.034***
	(0.0099)	(0.0098)	(0.0098)	(0.0099)	(0.013)
Age	0.11***	0.11***	0.11***	0.10***	0.10***
	(0.0042)	(0.0042)	(0.0042)	(0.0043)	(0.0045)
Age aquared h_age2	-0.0012***	-0.0012***	-0.0012***	-0.0012***	-0.0012***
	(0.000046)	(0.000046)	(0.000047)	(0.000047)	(0.000050)
Children	0.13***	0.13***	0.13***	0.13***	0.15***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.014)
Education	0.13***	0.14***	0.13***	0.14***	0.13***
	(0.017)	(0.017)	(0.018)	(0.017)	(0.018)
Unemployed	-0.31***	-0.31***	-0.31***	-0.30***	-0.29***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)
Self-employed	-0.0075	-0.0057	-0.0073	-0.0025	-0.0072
	(0.031)	(0.031)	(0.031)	(0.032)	(0.029)
Student	-0.69***	-0.69***	-0.68***	-0.69***	-0.69***
	(0.059)	(0.059)	(0.059)	(0.060)	(0.058)
Retired	-0.12***	-0.12***	-0.12***	-0.12***	-0.11***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.027)
	-3.52***	-3.50***	-3.52***	-3.50***	-3.39***
	(0.093)	(0.095)	(0.090)	(0.092)	(0.11)
ICC (regional)	0.016	0.016	0.016	0.016	0.015
N		109780	97060	61096	

Dependent variable: binary response indicating if respondents have a loan. Estimation method: multi-level modeling. Country and time fixed effects for 2009-2018 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Robust standard errors in parentheses. Variables are defined in appendix.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 9: Marginal effects - various reference groups as included in the estimations in Table 3
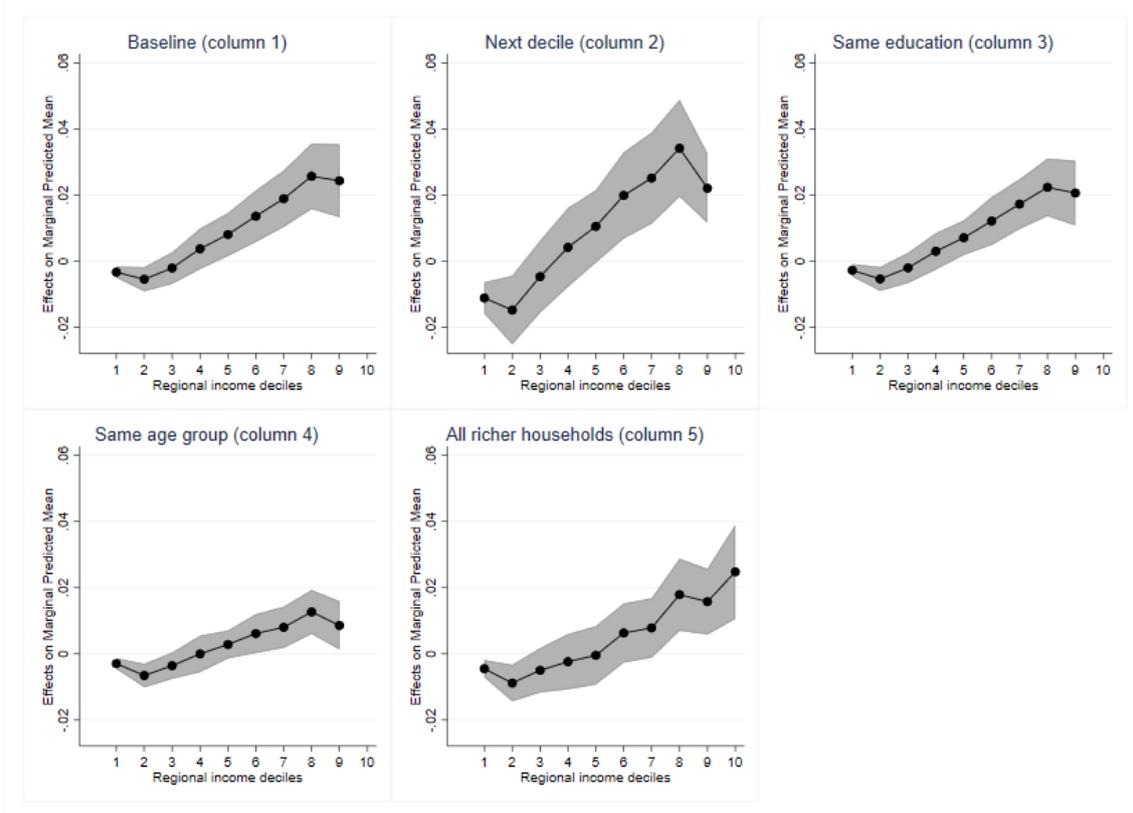
5.3 Exploring heterogeneity - loans by purpose

In a next step, we look into whether conspicuous consumption would matter for household loans in CESEE (Ireland (1994)). For that purpose, we investigate if the nexus between reference income and debt differs depending on the purpose of the "equilibrium" loans. We use information that was gathered by the OeNB Euro Survey from 2010 to 2014, where respondents were asked about their current loans according to four categories: mortgage, car, consumption and other loans.

Alongside with a high preference of buying a home, favorable credit supply conditions and increased availability of housing has led to a higher demand in mortgage-financed

housing. Column 1 of Table 4 shows a high positive correlation between reference income and mortgage loans. Interestingly, the coefficients' sign is positive even at the third and fourth decile. This might suggest a stronger demand-driven link. For consumption and other loans, we see a similar pattern as in our baseline specification. For car loans, only the second decile shows a significant (negative) coefficient.

Table 4: Household loans in CESEE by purpose

	(1) house	(2) car	(3) consumption	(4) other
Income Distribution				
1th decile#relinc	0.008	0.000	-0.012*	-0.011**
	(0.006)	(0.005)	(0.006)	(0.005)
2rd decile#relinc	0.032**	-0.042**	-0.036***	-0.029
	(0.014)	(0.018)	(0.011)	(0.021)
3rd decile#relinc	0.033**	-0.021	-0.013	-0.004
	(0.014)	(0.020)	(0.015)	(0.026)
4th decile#relinc	0.057***	-0.013	0.013	0.006
	(0.020)	(0.024)	(0.019)	(0.030)
5th decile#relinc	0.053**	-0.030	0.014	0.012
	(0.023)	(0.022)	(0.019)	(0.027)
6th decile#relinc	0.106***	0.006	0.019	0.032
	(0.022)	(0.023)	(0.020)	(0.028)
7th decile#relinc	0.095***	-0.001	0.052**	0.023
	(0.025)	(0.023)	(0.022)	(0.028)
8th decile#relinc	0.114***	0.013	0.063***	0.064**
	(0.023)	(0.029)	(0.020)	(0.032)
9th decile#relinc	0.107***	0.005	0.079***	0.070**
	(0.024)	(0.027)	(0.021)	(0.030)
Income	-0.00005***	0.00007**	-0.0002***	-0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)
cons	-3.694***	-3.723***	-3.748***	-3.416***
	(0.222)	(0.167)	(0.147)	(0.195)
ICC (regional)	0.014	0.011	0.022	0.017
N	56841	56841	56841	56841

Dependent variable: binary response indicating if respondents have a loan, by loan purpose. Estimation method: multi-level modeling. Sample where loan purpose question was included: 2010-2014. Country and time fixed effects for 2010-2014 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Robust standard errors in parentheses. Coefficient estimates for socioeconomic variables excluded in the table. Variables are defined in appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 10: Marginal effects of relative reference income on loans by purpose at income deciles for the estimations in Table 4
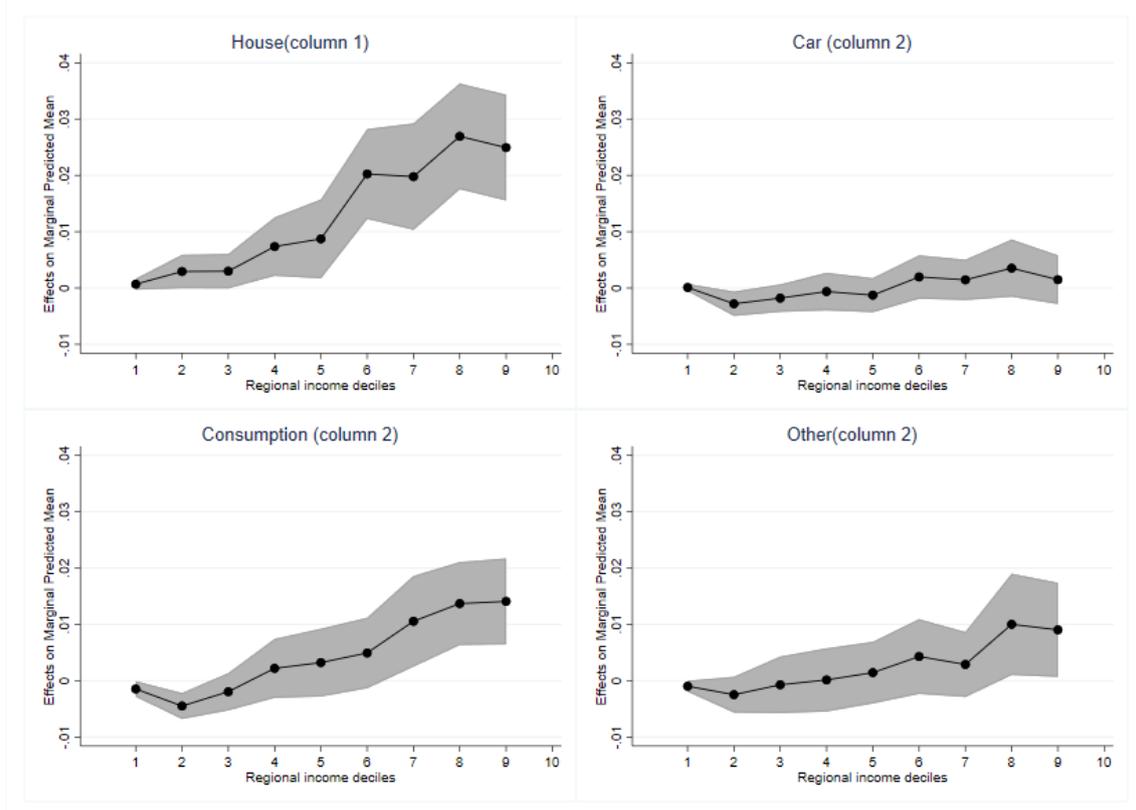
Note: Average marginal effects of relative reference income for loan purposes as in Table 4. The grey area depicts the 95% confidence interval. Source: OeNB EuroSurvey, own calculations.

5.4 The currency composition of household loans

Despite the approaches adopted to alleviate endogeneity concerns, it might be that if a factor influences both the income distribution and the probability of having a loan at the same time, this could lead to a bias in the estimated coefficients. Therefore, we opted for exploring the currency denomination of the household loans to explore possible endogeneity concerns, as discussed below.

Respondents have been asked to state the currency denomination of their loans and place their replies into five categories: solely denominated in: (1) the whole amount is in foreign currency, (2) the whole amount is in local currency, (3) the loan is predominantly

denominated in foreign or (4) in local currency, (5) the loan is equally split. Thus, we perform several estimations according to whether an household has a loan with one of the five categories. The estimation results are presented in Table 5.

The headline results show that the strongest link between reference income and indebtedness can be found for loans, which are fully denominated in foreign currency (i.e about 20% of households with a loan) or fully denominated in local currency (i.e about 58% households with a loan). There is some smaller and statistically weak positive correlation for predominantly local currency loans. For the rest of the categories in-between (mostly foreign, equal split), no significant correlation can be found (however, the sample size of these categories are also significantly lower). A new and an interesting result in Table 5 is that higher reference income ratio correlates with a higher probability of holding foreign currency debt already in the third decile (as compared to the baseline estimation in column (1), where it is in the fifth decile and above).

These results strengthen our assumption that omitted variable bias/endogeneity might be less of a concern as shocks that affect income inequality and loans at the same time are more likely to affect loans in local currencies. An example would be favourable economic conditions leading to more local currency credit supply and higher reference income at the same time. As the survey data also show, most of households with loans fully denominated in foreign currency have those loans in a foreign-owned bank (56%). Accordingly, if a shock causes the bias, the effect on foreign currency loans should be smaller or nil, which is, overall, not the case as our results show. The only notable differences can be observed in the the lower part of the distribution. This might be related to a potentially stronger signalling effect for domestically-owned banks. Alternatively, domestically-owned banks would acquire more information about the local income distribution as a signal.

Table 5: Household loans in CESEE by currency

	(1) baseline	(2) only fx	(3) mostly fx	(4) equal	(5) mostly local	(6) only local
Income Distribution						
1th decile# relincome	-0.011***	-0.005	-0.002	-0.009	-0.003	-0.019***
	(0.003)	(0.004)	(0.011)	(0.007)	(0.004)	(0.005)
2nd decile#relincome	-0.020***	0.001	-0.004	-0.046*	0.002	-0.039***
	(0.007)	(0.010)	(0.018)	(0.023)	(0.010)	(0.010)
3rd decile#relincome	-0.007	0.044***	-0.049*	-0.024	-0.011	-0.038***
	(0.010)	(0.012)	(0.027)	(0.026)	(0.017)	(0.011)
4th decile#relincome	0.013	0.023	-0.010	-0.000	-0.000	-0.007
	(0.012)	(0.018)	(0.026)	(0.029)	(0.019)	(0.013)
5th decile#relincome	0.025**	0.037**	-0.021	0.016	0.032	-0.006
	(0.012)	(0.017)	(0.030)	(0.038)	(0.020)	(0.015)
6th decile#relincome	0.040***	0.056**	0.000	0.004	0.035	0.007
	(0.014)	(0.023)	(0.031)	(0.038)	(0.022)	(0.016)
7th decile#relincome	0.053***	0.072***	-0.011	0.021	0.044**	0.021
	(0.014)	(0.023)	(0.029)	(0.036)	(0.022)	(0.018)
8th decile#relincome	0.069***	0.058**	0.021	0.064	0.046*	0.042**
	(0.015)	(0.024)	(0.034)	(0.041)	(0.025)	(0.019)
9^{th}decile#relincome	0.054***	0.066**	0.015	0.012	0.042*	0.035*
	(0.014)	(0.026)	(0.029)	(0.043)	(0.023)	(0.021)
Income	-0.0001	0.00007	0.00005	-0.00005	-0.0000009	-0.0001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
cons	-3.530***	-4.148***	-4.325***	-4.134***	-3.704***	-3.358***
	(0.092)	(0.199)	(0.222)	(0.290)	(0.158)	(0.110)
ICC (regional)	0.016	0.024	0.012	0.018	0.018	0.020
N	105301	105301	105301	105301	105301	105301

Dependent variable: binary response indicating if respondents have a loan. Estimation method: multi-level modeling. Country and time fixed effects for 2009-2018 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Robust standard errors in parentheses. Coefficient estimates for socioeconomic variables excluded in the table. Variables are defined in the Appendix.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 11: Marginal effects at income deciles for the estimations in Table 5
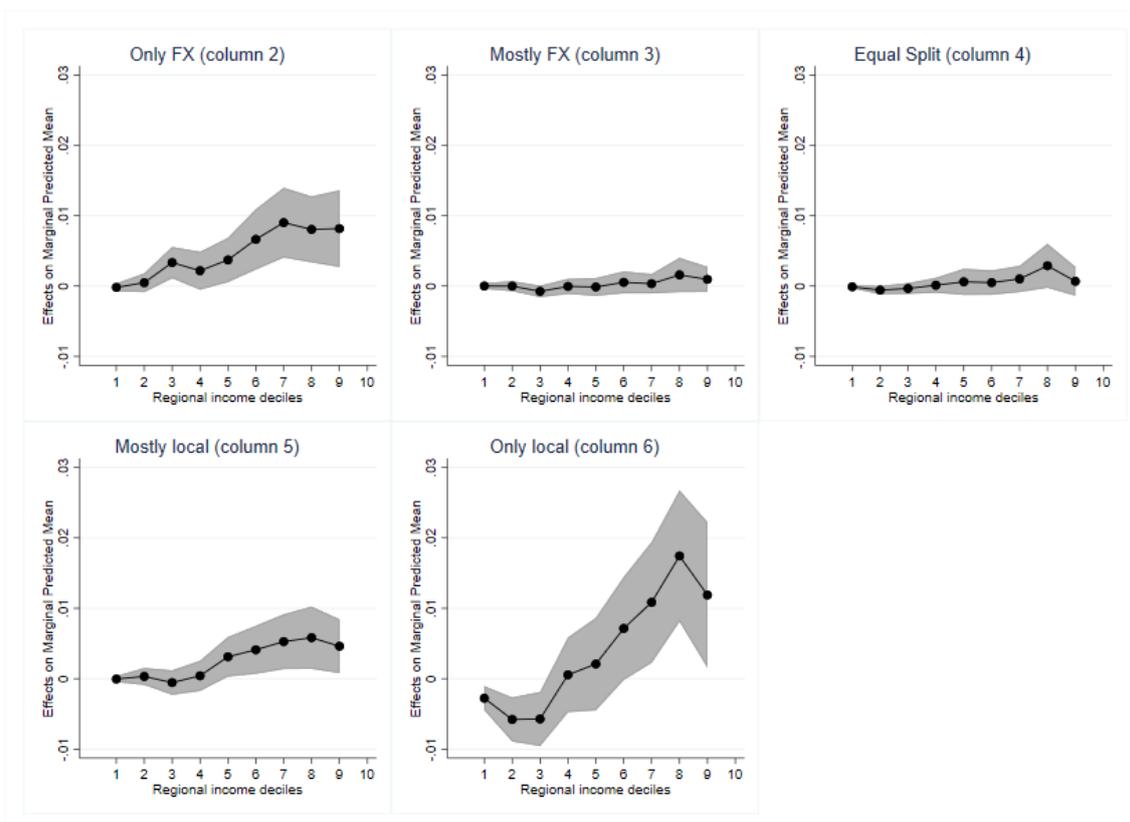
Note: Average marginal effects of relative reference income for loans by currency composition presented in Table 5. The grey area depicts the 95% confidence interval. Source: OeNB EuroSurvey, own calculations.

6 Zooming on the demand of household loans in CESEE

Disentangling the effects of demand and supply factors for loans has proven to be a daunting task in the literature. As shown in Section 5, current loans express equilibrium transactions without a prior information on whether they could be both supply- and/or demand-driven. In this section, we attempt to take a closer look on the demand side of loans as the data set includes a question on the intentions of respondents to take up a loan. We ground our analysis in this section on a question where respondents have to indicate whether they plan to approach a bank to demand a loan in the following 12 months [11]. In a subsequent

---

[11]The question in the questionnaire reads:"Do you plan to take out a loan within the next year and if so, in what currency?"

step, information has been given also on the purpose of the intended loan (i.e house, car, consumption or for other purposes).[12]

The results in Table 6 test whether the intended demand for loans correlates with the change of the income distribution. Overall, we find some indication for a positive and significant correlation between larger regional relative income ratio and loan plans above the 7th deciles although this correlation tends to be weaker than in the case of current "equilibrium" loans. Interestingly, when testing whether this result is valid for all planned loan categories, it remains only for intended loans for car purchases. This would support a conspicuous motive for demand as cars have been shown to be a highly (consumption)-visibility good (e.g. Heffetz 2011). Loans aimed to acquire consumption goods do not seem to support the conspicuous motive. The results in columns (7) and (8) hint at threshold effects as the coefficients remain significant for the upper quantiles and when relative mean income ratio increases to levels above the mean of 2.5. The ICC in column (1) shows the share of the variance which is explained by variables from the level 2 (regions) i.e 3.9%. In this case, although this would be a smaller portion than in the case of current loans, the application of multilevel estimations is still suitable. Overall, the ICC remains close to 2%.

---

[12]"What is the purpose of your loan or your loans? To finance a house or apartment, for consumption goods (furniture, travelling, household appliances, etc.), for a car or for other purposes." This question has been included in the questionnaire since 2014.

Table 6: Loan intentions of households in CESEE

	(1) const	(2) all	(3) house	(4) car	(5) consumption	(6) other	(7) regions low	(8) regions high
Income Distribution								
1th decile# relincome		0.004	0.006	0.004	0.001	-0.005	0.032	0.005
		(0.006)	(0.004)	(0.006)	(0.009)	(0.008)	(0.043)	(0.004)
2nd decile#relincome		-0.028*	-0.012	-0.028*	-0.001	0.013	-0.009	0.015
		(0.017)	(0.015)	(0.017)	(0.026)	(0.014)	(0.040)	(0.010)
3rd decile#relincome		0.040*	-0.007	0.040*	0.022	-0.028	-0.005	0.013
		(0.023)	(0.026)	(0.023)	(0.029)	(0.026)	(0.048)	(0.015)
4th decile#relincome		-0.008	-0.063**	-0.008	0.034	0.009	-0.011	0.039*
		(0.030)	(0.029)	(0.030)	(0.036)	(0.031)	(0.050)	(0.021)
5th decile#relincome		0.029	-0.004	0.029	0.028	-0.019	0.003	0.079***
		(0.032)	(0.031)	(0.032)	(0.045)	(0.031)	(0.050)	(0.027)
6th decile#relincome		0.015	0.010	0.015	0.038	0.034	-0.022	0.027
		(0.025)	(0.025)	(0.030)	(0.041)	(0.037)	(0.055)	(0.030)
7th decile#relincome		0.044*	0.022	0.044*	0.034	0.014	0.008	0.046*
		(0.038)	(0.038)	(0.025)	(0.041)	(0.034)	(0.057)	(0.025)
8th decile#relincome		0.026*	0.007	0.026	0.023	-0.004	-0.008	0.065**
		(0.025)	(0.025)	(0.025)	(0.045)	(0.035)	(0.055)	(0.032)
9th decile#relincome		0.064**	0.005	0.064**	0.022	-0.006	0.014	0.045**
		(0.027)	(0.33)	(0.027)	(0.038)	(0.035)	(0.054)	(0.022)
Sociodemographics								
Female		-0.145***	0.007	-0.146***	0.024	-0.069***	-0.191	-0.007
		(0.029)	(0.022)	(0.029)	(0.019)	(0.023)	(0.014)	(0.032)
Age		0.021***	0.021**	0.021***	0.028***	0.037***	0.029***	0.038***
		(0.007)	(0.006)	(0.007)	(0.005)	(0.007)	(0.004)	(0.006)
Age squared		-0.000***	-0.000***	0.000***	0.000***	0.000***	-0.000***	-0.00***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Children		0.004	0.029*	-0.004	0.023	0.016	0.010	0.019
		(0.020)	(0.015)	(0.020)	(0.017)	(0.017)	(0.013)	(0.013)
Education		0.056**	0.035	0.056**	-0.027	-0.001	0.061***	0.057**
		(0.027)	(0.028)	(0.027)	(0.029)	(0.030)	(0.019)	(0.027)
Unemployed		-0.183***	-0.189***	-0.183***	-0.181***	-0.060	-0.176***	-0.211***
		(0.045)	(0.026)	(0.045)	(0.047)	(0.041)	(0.019)	(0.033)
Self-employed		-0.015	0.146***	0.015	-0.043***	0.200	0.099***	0.096
		(0.063)	(0.052)	(0.063)	(0.050)	(0.042)	(0.039)	(0.079)
Student		-0.580***	-0.742***	-0.580***	-0.616***	-0.101	-0.564***	-0.455***
		(0.086)	(0.73)	(0.086)	(0.084)	(0.080)	(0.063)	(0.084)
_cons	-1.423***	-1.556***	-1.556***	-2.135***	-0.743***	-0.105	-1.591***	-1.94***
	(0.025)	(0.132)	(0.132)	(0.175)	(0.132)	(0.133)	(0.166)	(0.166)
ICC (regional)	0.039	0.013	0.013	0.013	0.017	0.014	0.015	0.012
N	113722	71322	71322	71322	71322	76297	76297	23390

Dependent variable: the share of respondents, who plan a loan in the following 12 months (dummy variable). Estimation method: multi-level modeling. Country and time fixed effects for 2009-2018 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Robust standard errors in parentheses. Variables are defined in appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Loan intentions could also be understood as the crossing point of sociodemographic characteristics of respondents along with their sentiments about past developments both of the country or of the respondents themselves. Against that background, we include in Table 7 information from the OeNB Euro Survey on (1) developments of the financial situation of the household in the past 12 months and (2) current economic situation of the

country.[13] We opted for focusing on sentiments about past and current developments so as to alleviate causality concerns. Accordingly, a good financial situation of the household and a good economic situation of the country increase the likelihood to take up a loan. In addition, the availability of savings might act as a security for taking up loan. On the contrary, people might finance their purchases when dissaving and thus not affecting their demand for loans. Similar to the results of currently having a loan (Table 2), we could not detect any significant effect of savings.

Notably, we could trace some discrepancies between the direction of the effect of some individual sociodemographic characteristics in the case of "equilibrium" and planned household loans. While female respondents tend to be more likely to have a loan, they tend to plan loans less often. Also having children in the family reduces the likelihood to be indebted but does not impact the likelihood of demanding a loan. Finally, in line with the threshold effect self-employed respondents are likely to plan loans only in less unequal regions. Moreover, these respondents rather plan foreign currency loans and have intensified their intentions in the outer years of the sample period.

---

[13]The questions in the survey read as follows:"Over the last 12 months, the financial situation of my household has got better" and "Currently, the economic situation of my country is very good". The replies of the respondents range in five categories between "strongly agree" and "strongly disagree".

Table 7: Loan intentions of households in CESEE

	(1) sentiments	(2) savings	(3) prior 2012	(4) post 2012	(5) w/t current loan	(6) fx loan
Income Distribution						
1th decile# relincome	-0.001	-0.001	-0.001	0.005	0.005	-0.010
	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)	(0.008)
2nd decile#relincome	-0.008	-0.001	-0.002	0.015	0.009	-0.031
	(0.009)	(0.009)	(0.017)	(0.012)	(0.010)	(0.022)
3rd decile#relincome	0.000	0.000	-0.000	0.021	0.014	-0.038
	(0.010)	(0.015)	(0.016)	(0.023)	(0.018)	(0.024)
4th decile#relincome	0.011	0.011	0.011	0.034	0.037*	-0.060*
	(0.017)	(0.017)	(0.021)	(0.026)	(0.021)	(0.034)
5th decile#relincome	0.046***	0.030*	0.047**	0.037	0.051**	-0.035
	(0.017)	(0.018)	(0.024)	(0.026)	(0.023)	(0.028)
6th decile#relincome	0.029	0.002	-0.011	0.046*	0.091	0.059*
	(0.020)	(0.021)	(0.028)	(0.028)	(0.023)	(0.033)
7th decile#relincome	0.044*	0.026	0.031	0.059**	0.045*	0.060
	(0.038)	(0.018)	(0.025)	(0.026)	(0.026)	(0.043)
8th decile#relincome	0.068***	0.022	0.034	0.039	0.032	0.020
	(0.022)	(0.022)	(0.027)	(0.026)	(0.022)	(0.034)
9th decile#relincome	0.072***	0.033**	0.048**	0.047**	0.041**	-0.046
	(0.019)	(0.017)	(0.021)	(0.026)	(0.022)	(0.032)
Sentiments						
Past sit HH	0.049**					
	(0.024)					
Econ sit country	0.119***					
	(0.030)					
Savings		0.026				
		(0.020)				
Sociodemographics						
Female	-0.008	-0.011	0.006	-0.030*	0.011	-0.145***
	(0.018)	(0.014)	(0.021)	(0.017)	(0.016)	(0.031)
Age	0.030***	0.030***	0.031**	0.031***	0.027***	-0.002
	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)	(0.007)
Age squared	-0.000***	-0.000***	-0.000***	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Children	0.001	0.011	0.030	0.032**	0.017	-0.012
	(0.013)	(0.010)	(0.014)	(0.013)	(0.011)	(0.019)
Education	0.058***	0.067**	0.069***	0.058**	0.064***	0.072**
	(0.019)	(0.020)	(0.020)	(0.029)	(0.020)	(0.032)
Unemployed	-0.184***	-0.186***	-0.185***	-0.177***	-0.186***	-0.029
	(0.023)	(0.045)	(0.028)	(0.034)	(0.023)	(0.057)
Self-employed	0.090*	0.104***	0.060	0.110***	0.094***	0.153***
	(0.049)	(0.034)	(0.045)	(0.037)	(0.036)	(0.058)
Student	-0.536***	-0.508***	-0.505***	-0.603***	-0.578***	-0.133
	(0.063)	(0.056)	(0.081)	(0.061)	(0.064)	(0.083)
_cons	-1.677***	-1.556***	-1.679***	-2.135***	-1.631***	-1.047***
	(0.153)	(0.132)	(0.143)	(0.175)	(0.116)	(0.231)
ICC (regional)	0.018	0.013	0.021	0.013	0.019	0.024
N	59018	84995	51775	71322	76769	13985

Dependent variable: the share of respondents, who plan a loan in the following 12 months (dummy variable).
Estimation method: multi-level modeling. Country and time fixed effects for 2009-2018 included in all estimations.
Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second)
level covariates. Robust standard errors in parentheses. Variables are defined in appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

7   Alternative measures of income inequality

In a next step of our analysis, we test the robustness of our estimations and we explore the effects by other inequality measures. For instance, the Gini coefficient (ranging between 0 and 1), although widely used in the literature, is an inequality measure that incorporates movements in the whole income distribution, thus being "noisier" than the stricter upward comparisons measure, which we use in our analysis. Therefore, we could expect that the effect on household indebtedness to be smaller than by the relative income ratio. For instance, if a household is in the middle of the distribution, Gini changes attributed to income losses in the bottom of the distribution might not affect its behaviour because the household compares in an upward-looking way rather than downward-looking way. Indeed, the results in column 1 of Table 8, could be largely confirmed, however, only the effects for the upper half are significant. Going further, the ratio between income in the 90th and 10th percentile as depicted in column 2 yields results qualitatively similar to the results using relative reference income, albeit quantitatively the effect is smaller. In column 3, the results using the ratio between the 75th and the 25th percentile is used, with the bottom and top deciles having a weakly significant coefficient.

Column 4 to 7 show the results for the share of the top 1%, top 5% and top 10% on total income, respectively. Using the share of the Top 1% yields very similar results as reference income. For the share of the Top 5% and Top 10%, only the coefficients in the upper part of the distribution yield significant estimates. These results could hint to a signalling effect resulting in lower loan supply for the bottom of the distribution driven by the share of the top income households. No effect can be found for the shares of the bottom 10% and bottom 20%, respectively. The last column finally calculates the difference between the log income of percentiles 90 and 10, as used by Coibion et al. (2014).

Overall, these estimations suggest that: (i) our main result that higher income inequality is correlated with a lower probability of having a loan for the bottom third of the distribution and a higher probability of having a loan for the top of the distribution, is confirmed by various inequality measures other than the relative reference income used above, (ii) the strongest link between income inequality and debt can be captured by the P90/P10 percentile ratio and the share of the top 1% on total income. Overall, movements at the top of the income distribution seem to play a key role for both the signalling channel and the "Novaks" (i.e demand)-channel, while movements at the bottom of the regional income distribution tend to be less central.

31

Table 8: Alternative Inequality measures

	(1) Gini	(2) P90/P10	(3) P75/P10	(4) Top1%	(5) Top5%	(6) Top10%	(7) Bottom10%	(8) Bottom20%	(9) logP90-logP10
Income Distribution									
1st decile#ineq	-0.264	-0.022***	-0.069**	-1.983***	-0.604**	-0.213	-1.800	-1.129	-0.078**
	(0.170)	(0.005)	(0.030)	(0.493)	(0.243)	(0.193)	(1.991)	(0.894)	(0.031)
2nd decile#ineq	-0.117	-0.013***	-0.045	-1.347***	-0.344*	-0.084	-1.916	-1.035	-0.052*
	(0.154)	(0.004)	(0.029)	(0.323)	(0.200)	(0.174)	(1.966)	(0.885)	(0.029)
3rd decile#ineq	0.010	-0.007*	-0.024	-0.870***	-0.128	0.107	0.209	-0.201	-0.014
	(0.165)	(0.004)	(0.030)	(0.336)	(0.218)	(0.186)	(1.875)	(0.858)	(0.029)
4th decile#ineq	0.126	0.000	-0.007	-0.316	0.091	0.202	-0.339	-0.329	0.001
	(0.164)	(0.004)	(0.029)	(0.344)	(0.216)	(0.184)	(1.929)	(0.864)	(0.029)
5th decile#ineq	0.190	0.003	0.001	0.082	0.238	0.274	0.244	-0.145	0.013
	(0.183)	(0.004)	(0.030)	(0.469)	(0.262)	(0.214)	(2.021)	(0.885)	(0.032)
6th decile#ineq	0.272	0.008**	0.016	0.256	0.369	0.397**	1.286	0.343	0.034
	(0.178)	(0.004)	(0.030)	(0.398)	(0.240)	(0.195)	(1.752)	(0.801)	(0.030)
7th decile#ineq	0.345*	0.010**	0.024	0.898**	0.540**	0.509**	1.173	0.334	0.050*
	(0.179)	(0.004)	(0.030)	(0.408)	(0.250)	(0.209)	(1.955)	(0.865)	(0.030)
8th decile#ineq	0.456**	0.017***	0.041	1.314***	0.724***	0.655***	1.617	0.630	0.079**
	(0.185)	(0.004)	(0.031)	(0.446)	(0.252)	(0.213)	(1.865)	(0.819)	(0.032)
9th decile#ineq	0.381**	0.013***	0.028	0.991**	0.597**	0.515***	0.691	0.217	0.054*
	(0.176)	(0.003)	(0.029)	(0.407)	(0.245)	(0.196)	(2.061)	(0.897)	(0.028)
10th decile#ineq	0.566***	0.020***	0.053**	1.792***	0.965***	0.777***	1.634	0.706	0.092***
	(0.180)	(0.004)	(0.027)	(0.556)	(0.281)	(0.227)	(2.325)	(0.996)	(0.027)
_cons	-3.483***	-3.438***	-3.426***	-3.431***	-3.461***	-3.497***	-3.442***	-3.425***	-3.443***
	(0.103)	(0.090)	(0.112)	(0.087)	(0.092)	(0.101)	(0.102)	(0.107)	(0.106)
ICC (regional)	0.016***	0.016***	0.016***	0.016***	0.016***	0.015***	0.015***	0.015***	0.015***
N	119085	119085	119085	119085	119085	101049	101049	101049	101049

Dependent variable: binary response indicating if respondents have a loan. Estimation method: multi-level modeling. Country and time fixed effects for 2009-2018 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Socioeconomic variables included in all estimations. omitted in the table. Robust standard errors in parentheses. Variables are defined in appendix.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

8 Where are "the Novaks"? Keeping up with other regions

Until now, we constructed the relative income variable income of richer households in the same region. In addition, we will focus on relaxing the assumption that respondents compare themselves only with "Novaks" from the same region. For that purpose, we include also households in neighbouring regions and test whether this would add to the explanatory value. For demand-side effects, it is possible that households compare themselves not only to their "neighbours" (households in the same regions), but also to other households in their relative vicinity because of various possible social ties. Additionally, from a supply-side perspective, banks that operate trans-regionally could have clients in other regions. Against that background, in an explanatory spatial analysis, we test if including reference income

for neighbouring regions would affect our results[14]. To disentangle "home region" and "neighbour region" effects, we construct the neighbours reference income separately. *Relinc neighbour* denotes the relative reference income of households in higher income deciles in neighbouring regions of the same country.

Table 9 depicts the baseline regression in column one, while column two contains the added neighbour region income effect. There are significantly negative estimates for the correlation between the mean income of richer households in neighbouring regions for decile 5, only significant at the 10% level. This result suggests that (i) the bulk of effects is for households in the same region, not neighboring regions and (ii) it might be that the signalling effect seems to be stronger than the "Keeping up with the Novaks"-effect for neighboring regions effects. However, as some if the results for the influence of "home region"-relative income appear to lose their significance, possible co-movements between relative income at home and in negative regions should be investigated. These results suggest that at least some of the signalling effect is due to effects in neighbouring regions. This could potentially come from banks with possible household customers across different regions.

Finally, we try out also a different definition of a reference group, namely that the reference group is defined as households in higher income deciles in the same country (as opposed to the same region). The results, shown in column 3, demonstrate that the main results are robust, with negative effects of higher reference income at the bottom of the distribution and positive effects on loans at the top. Some minor differences include that if we define the "Novaks" on country level, the second decile result is not significant anymore, while there is some low level of statistical significant for the fourth decile. This result suggests that the "Novaks" seem to be present both at country and regional level, with somewhat negligible changes in quantitaive terms of our estimates.

---

[14]Various other "neighbour" concepts are possible, e.g. households closer than 100km, 200km, etc. The analysis performed here is only explanatory, a full spatial analysis is beyond the scope of this paper.

Table 9: Spatial Dimension: Neighbouring regions

	(1) Baseline	(2) *relinc neighbour*	(3) *relinc country*
Income Distribution			
1st decile#relinc	-0.014***	-0.009	-0.008***
	(0.004)	(0.010)	(0.003)
2nd decile#relinc	-0.023***	-0.033	-0.013
	(0.008)	(0.024)	(0.008)
3rd decile#relinc	-0.008	0.028	-0.003
	(0.010)	(0.030)	(0.001)
4th decile#relinc	0.015	0.017	0.024*
	(0.012)	(0.043)	(0.013)
5th decile#relinc	0.031**	0.086***	0.043***
	(0.012)	(0.032)	(0.015)
6th decile#relinc	0.050***	0.080*	0.072***
	(0.014)	(0.041)	(0.015)
7th decile#relinc	0.068***	0.086*	0.091***
	(0.015)	(0.046)	(0.016)
8th decile#relinc	0.090***	0.120***	0.119***
	(0.017)	(0.036)	(0.018)
9th decile#relinc	0.086***	0.087*	0.109***
	(0.019)	(0.044)	(0.021)
1st decile#relinc neighbour		-0.005	
		(0.010)	
2nd decile#relinc neighbour		0.009	
		(0.024)	
3rd decile#relinc neighbour		-0.038	
		(0.030)	
4th decile#relinc neighbour		-0.004	
		(0.043)	
5th decile#relinc neighbour		-0.058*	
		(0.034)	
6th decile#relinc neighbour		-0.033	
		(0.041)	
7th decile#relinc neighbour		-0.020	
		(0.044)	
8th decile#relinc neighbour		-0.031	
		(0.038)	
9th decile#relinc neighbour		-0.002	
		(0.038)	
cons	-3.515***	-3.504***	-1.280***
	(0.093)	(0.094)	(0.050)
/			
ICC (regional)	0.016***	0.016***	0.015***
	(0.003)	(0.003)	(0.003)
N	105301	104963	106227

Dependent variable: binary response indicating if respondents have a loan. Estimation method: multi-level modeling. *relinc neighbour* denotes the relative reference income of households in higher income deciles in neighbouring regions of the same country. Country and time fixed effects for 2009-2018 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Coefficient estimates for socioeconomic variables omitted in the table. Robust standard errors in parentheses. Variables are defined in appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

9 Conclusions

In this paper, we analyze the nexus between household income and inequality in CESEE, using data from a household survey-the OeNB Euro Survey- for the period 2009 to 2018. We compute and apply various income inequality measures on regional level - a primer for many countries and years. Our empirical analysis aims at pinpointing the association between income inequality and the probability of currently being indebted or planning a loan. As households are nested into regions, we employ multilevel models that account for intra-regional correlation. Our main findings are twofold. First, there is a negative correlation between the relative mean income of richer households and the probability of being indebted for the very bottom of the distribution. This result suggests a strong signaling effect of income inequality, meaning that presumably banks react to rising income inequality by targeting richer households with more, and cheaper, credit while restraining credit supply to poorer households. Second, there is a positive correlation between reference income and the probability of having a loan for the upper half of the regional income distribution. This result is consistent with the aforementioned supply-side channel, but also with a demand-side channel ("Keeping up with the Novaks"). This result is robust to variations in the definitions of reference income, and holds also when using traditional inequality measures such as the Gini coefficent, 90/10 percentile ratios and the shares of the Top 1% and Top 5% of the income earners.

Our paper is one of the first to look more into detail in the demand side of loans and income inequality. Although, the coefficients in this case are quantitatively weaker, we can confirm a positive correlation at the very top of the distribution, supporting thus the "Keeping up with the Novaks"-effect. Moreover, this holds especially for car loans (which the literature has often found to be exemplar for "conspicuous" consumption). In addition, loan plans seem to be affected by higher inequality especially in regions with higher inequality.

Our findings highlight the complex features of household debt determination along the lines of the income distribution. Our results show that policymakers should pay attention to income inequality when focusing on financial stability. Additionally, rising inequality could limit access to finance for low-income households, especially in regions where income inequality is already high, an important finding to better understand financial inclusion.

References

Bazillier, R. and Hericourt, J. (2017), 'The circular relationship between inequality, leverage, and financial crises', *Journal of Economic Surveys* **31**(2), 463–496.

Belabed, C. A. and Hake, M. (2018), *Income inequality and trust in national governments in Central, Eastern and Southeastern Europe*, Oesterreichische Nationalbank Working Paper.

Carr, M. D. and Jayadev, A. (2015), 'Relative income and indebtedness: evidence from panel data', *Review of income and wealth* **61**(4), 759–772.

Charles, K. K., Hurst, E. and Roussanov, N. (2009), 'Conspicuous consumption and race', *The Quarterly Journal of Economics* **124**(2), 425–467.

Coibion, O., Gorodnichenko, Y., Kudlyak, M. and Mondragon, J. (2014), Does greater inequality lead to more household borrowing? new evidence from household data, Technical report, National Bureau of Economic Research.

Drechsel-Grau, M. and Schmid, K. D. (2014), 'Consumption–savings decisions under upward-looking comparisons', *Journal of Economic Behavior & Organization* **106**, 254–268.

Duesenberry, J. S. et al. (1949), 'Income, saving, and the theory of consumer behavior'.

Fasianos, A., Raza, H. and Kinsella, S. (2017), 'Exploring the link between household debt and income inequality: an asymmetric approach', *Applied Economics Letters* **24**(6), 404–409.

Ferrer-i Carbonell, A. (2005), 'Income and well-being: an empirical analysis of the comparison income effect', *Journal of public economics* **89**(5-6), 997–1019.

Hake, M. and Poyntner, P. (2019), 'Household loans in cesee from a new perspective: the role of income distribution', *Focus on European Economic Integration* (Q2/19), 75–93.

Heffetz, O. (2011), 'A test of conspicuous consumption: Visibility and income elasticities', *Review of Economics and Statistics* **93**(4), 1101–1117.

Iacoviello, M. (2008), 'Household debt and income inequality, 1963–2003', *Journal of Money, Credit and Banking* **40**(5), 929–965.

Ireland, N. J. (1994), 'On limiting the market for status signals', *Journal of public Economics* **53**(1), 91–110.

Kumhof, M., Rancière, R. and Winant, P. (2015), 'Inequality, leverage, and crises', *American Economic Review* **105**(3), 1217–45.

Loschiavo, D. (2016), 'Household debt and income inequality: evidence from italian survey data', *Bank of Italy Temi di Discussione (Working Paper) No* **1095**.

Maas, C. J. and Hox, J. J. (2005), 'Sufficient sample sizes for multilevel modeling', *Methodology* **1**(3), 86–92.

Rabe-Hesketh, S. and Skrondal, A. (2012), Multilevel and longitudinal modeling using stata, Technical report, StataCorp LP.

Schunck, R. (2016), 'Cluster size and aggregated level 2 variables in multilevel models. a cautionary note', *Methods, Data, Analyses* **10**(1), 120.

Veblen, T. (2017), *The theory of the leisure class*, Routledge.

Veblen, T. and Galbraith, J. K. (1973), *The theory of the leisure class*, Houghton Mifflin Boston.

Appendix

Table A1: Variables description

Income distribution variables. Source: OeNB Euro Survey	
Relinc ratio	The ratio of the average income of all individuals in a region who are above individual's income decile to the specific individual's income.
Position in income decile	Variable ranging between 1 and 10 and expressing in which decile of the regional income distribution the respondent is positioned.
Household debt variables. Source: OeNB Euro Survey	
Loan	Dummy variable that takes a value of 1 if a respondent has a loan; respondents answering "don't know" or "no answer" are excluded from the analysis
Mortgage/consumer/car	Dummy variables that take a value of 1 if a respondent has one of these loan categories : respondents answering "don't know" or "no answer" are excluded from the analysis
Planned loan	Dummy variable that takes a value of 1 if a respondent plans a loan in the next 12 months; respondents answering "don't know" or "no answer" are excluded from the analysis
Sociodemographic variables. Source: OeNB Euro Survey	
Age	Age of the respondents (the age is in the range 14+)
Female	Dummy variable that takes the value one if the respondent is female.
Children	Number of children aged 6 years and younger.
Education (low, medium, high)	Dummy variables, degree of education (university level, medium level and basic education), omitted category: education low.
Employment status	Dummy variable coded as one if respondent belongs to selected occupational category (student, unemployed/other, working, self-employed). Omitted category: retired
Regional level, other income inequality measures. Source: OeNB Euro Survey	
Regional Gini coef	Variable measuring income inequality constructed per region and year.
Top 1% income share	Variable measuring income share of top one percent constructed per region and year.
Top 10% income share	Variable measuring income share of top decile constructed per region and year.
Top 20% income share	Variable measuring income share of top quintile constructed per region and year.
Bottom 10% income share	Variable measuring income share of bottom decile constructed per region and year.
Bottom 20% income share	Variable measuring income share of Bottom quintile constructed per region and year.

Continued on next page

Sentiments. Source: OeNB Euro Survey	
Memories of restr deposits	Dummy variable taking the value of one if the respondent remember times where the access to deposits was restricted.
Fin sit of household	Dummy variable taking value of one if respondent expects that the financial situation of his/her household to improve in the coming 12 months.
Econ sit my country improve	Dummy variable taking value of one if respondent expects that the economic situation of the country will improve in the following five years.

Table A2: Summary statistics of main variables

	N	mean	median	min	max
Income variables					
Relinc ratio	106,462	2.47	1.92	1.08	82.70
Income	120,400	612.80	521.80	0	7594.44
Household debt					
Loan	120,400	0.23	0	0	1
Consumer loan	56,841	0.05	0	0	1
Mortgage loan	56,841	0.04	0	0	1
Car loan	56,841	0.02	0	0	1
Fx loan[15]	120,400	0.06	0	0	1
Plan a loan	113,728	0.08	0	0	1
Plan-consumer	113,728	0.02	0	0	1
Plan-mortgage	113,728	0.02	0	0	1
Plan-car	113,728	0.01	0	0	1
Plan-fx-loan	51353	0.03	0	0	1
Sociodemographics					
Female	120,400	0.53	1	0	1
Age	120,393	44.39	44	14	98
Have children	120,400	0.37	0	0	1
Education	120,088	2	2	1	3
Unemployed	119,638	0.18	0	0	1
Self-employed	119,400	0.07	0	0	1
Student	119,638	0.08	0	0	1
Retired	119,638	0.20	0	0	1

---

[15]Current loans solely or predominantly denominated in foreign currency.

Table A3: Descriptive statistics (average over 2009 to 2018, weighted)

	AL	BA	MK	BG	HR	PL	RO	RS	CZ	HU
Loan (%)	12.5	20.9	20.1	24.2	28.2	21.5	20.2	17.6	29.1	29.4
Plan a loan (%)	8.9	6.0	9.0	5.9	6.4	10.6	6.4	10.1	12.8	5.4
Age, conditional on...										
... having a loan	41.2	44.9	43.8	42.8	44.7	43.3	43.3	45.3	42.7	42.6
... planning a loan	38.5	42.2	40.6	40.2	39.9	40.6	39.1	43.0	39.6	38.9
relinc, **conditional on...**										
... having a loan	2.0	2.9	3.0	1.9	2.1	1.9	2.5	3.1	1.7	2.0
... not having a loan	2.1	3.1	3.6	2.1	2.4	2.0	3.0	3.4	1.7	2.1
... planning a loan	2.0	2.8	3.7	2.0	2.1	2.0	3.4	3.0	1.7	2.0
... not planning a loan	2.1	3.1	3.5	2.1	2.3	2.0	2.9	3.4	1.7	2.0
Loan if... %										
...Retired	2.8	15.5	16.6	9.0	18.5	13.1	12.2	12.0	7.8	13.8
...Student	3.9	2.3	3.6	2.5	1.9	3.1	5.4	2.6	12.7	
...Unemployed/other	7.4	16.5	13.5	24.8	22.7	17.9	16.8	9.0	46.0	36.7
...Employed	18.3	34.8	33.7	33.0	42.2	26.2	28.3	27.0	35.9	37.5
...Self-employed	19.3	37.8	27.8	32.0	36.8	33.0	22.2	32.0	40.8	40.2
Loan, plan if... %										
...Retired	2.6	4.6	5.3	1.7	2.8	4.9	2.8	5.3	2.6	1.6
...Student	6.4	2.7	3.6	2.3	1.1	5.1	5.3	3.6	4.4	3.4
...Unemployed/other	6.6	5.1	7.9	5.2	5.0	12.0	5.6	7.0	18.4	5.3
...Employed	11.7	11.9	13.3	8.2	10.4	12.6	9.1	15.2	15.8	7.6
...Self-employed	13.2	15.2	14.7	9.4	10.7	14.0	10.0	19.0	15.3	5.5

Table A4: Household loans in CESEE - country and region levels

	(1)\nconst	(2)\nno interaction	(3)\nbaseline	(4)\nlow ineq reg	(5)\nhigh ineq reg	(6)\n2011-2018	(7)\nwealth
Income Distribution							
1th decile# relincome			-0.014***	-0.027***	-0.006	-0.010*	-0.014***
			(0.005)	(0.007)	(0.004)	(0.005)	(0.005)
2nd decile#relincome			-0.023***	-0.046***	-0.003	-0.019*	-0.024**
			(0.008)	(0.017)	(0.009)	(0.010)	(0.010)
3rd decile#relincome			-0.009	-0.018	0.004	-0.006	-0.010
			(0.012)	(0.015)	(0.016)	(0.017)	(0.017)
4th decile#relincome			0.014	-0.003	0.029	0.026	0.016
			(0.013)	(0.018)	(0.019)	(0.020)	(0.021)
5th decile#relincome			0.030***	0.017	0.034**	0.049**	0.031
			(0.011)	(0.020)	(0.015)	(0.021)	(0.020)
6th decile#relincome			0.050***	0.048***	0.036**	0.054***	0.038*
			(0.012)	(0.016)	(0.014)	(0.019)	(0.019)
7th decile#relincome			0.067***	0.054**	0.057***	0.074***	0.062**
			(0.015)	(0.025)	(0.016)	(0.027)	(0.031)
8th decile#relincome			0.089***	0.068**	0.078***	0.100***	0.094***
			(0.018)	(0.029)	(0.017)	(0.031)	(0.032)
9th decile#relincome			0.085***	0.078*	0.042**	0.101**	0.090**
			(0.024)	(0.045)	(0.021)	(0.040)	(0.042)
Reference income		-0.015***					
		(0.005)					
Income		0.00009*	-0.0001*	-0.0002*	0.0002	-0.0001	-0.0002*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wealth proxies							
Savings							0.009
							(0.016)
House							0.039
							(0.025)
Car							0.151***
							(0.034)
Sociodemographics							
Female		0.027***	0.027***	0.021	0.042***	0.029*	0.047***
		(0.010)	(0.010)	(0.014)	(0.005)	(0.016)	(0.016)
Age		0.105***	0.105***	0.103***	0.112***	0.108***	0.105***
		(0.006)	(0.006)	(0.008)	(0.009)	(0.006)	(0.005)
Age squared		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Children		0.134***	0.134***	0.152***	0.106***	0.140***	0.138***
		(0.021)	(0.021)	(0.022)	(0.032)	(0.021)	(0.021)
Education		0.141***	0.134***	0.102***	0.191***	0.127***	0.110***
		(0.026)	(0.026)	(0.026)	(0.048)	(0.036)	(0.035)
Unemployed		-0.317***	-0.305***	-0.194**	-0.421***	-0.352***	-0.341***
		(0.063)	(0.062)	(0.089)	(0.047)	(0.065)	(0.065)
Selfemployed		-0.008	-0.008	0.031	-0.101	-0.023	-0.049
		(0.045)	(0.045)	(0.044)	(0.107)	(0.060)	(0.068)
Student		-0.694***	-0.692***	-0.753***	-0.578***	-0.671***	-0.675***
		(0.108)	(0.108)	(0.176)	(0.087)	(0.103)	(0.112)
Retired		-0.119***	-0.117***	-0.154***	-0.106***	-0.133**	-0.108*
		(0.030)	(0.030)	(0.028)	(0.028)	(0.054)	(0.058)
cons	-3.047***	-3.037***	-2.778***	-3.596***	-3.233***	-3.247***	
	(0.159)	(0.172)	(0.160)	(0.184)	(0.172)	(0.155)	
ICC (regional)		0.035	0.048	0.069	0.007	0.035	0.034
N		105301	105301	65341	39960	58135	49164

Dependent variable: binary response indicating if respondents have a loan. Estimation method: multi-level modeling, country and region levels. Time fixed effects for 2009-2018 included in all estimations. Intraclass correlation coefficient denotes the explained portion of the variance by inclusion of the regional (second) level covariates. Robust standard errors in parentheses. Variables are defined in Appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Household loans in CESEE - probit estimation

	(1) const	(2) no interaction	(3) baseline	(4) low ineq reg	(5) high ineq reg	(6) 2011-2018	(7) wealth
Income Distribution							
1th decile# relincome			-0.014***	-0.027***	-0.006**	-0.011***	-0.014***
			(0.003)	(0.008)	(0.003)	(0.004)	(0.004)
2nd decile#relincome			-0.023***	-0.047***	-0.003	-0.019**	-0.025***
			(0.007)	(0.015)	(0.007)	(0.008)	(0.009)
3rd decile#relincome			-0.008	-0.018	0.003	-0.005	-0.010
			(0.009)	(0.018)	(0.012)	(0.013)	(0.014)
4th decile#relincome			0.015	-0.003	0.029**	0.028*	0.018
			(0.010)	(0.019)	(0.013)	(0.015)	(0.016)
5th decile#relincome			0.032***	0.018	0.035**	0.051***	0.033*
			(0.011)	(0.020)	(0.015)	(0.016)	(0.018)
6th decile#relincome			0.051***	0.049**	0.037**	0.057***	0.041**
			(0.012)	(0.020)	(0.015)	(0.017)	(0.019)
7th decile#relincome			0.070***	0.056***	0.059***	0.078***	0.066***
			(0.012)	(0.021)	(0.016)	(0.018)	(0.020)
8th decile#relincome			0.093***	0.071***	0.080***	0.105***	0.100***
			(0.012)	(0.021)	(0.016)	(0.018)	(0.020)
9th decile#relincome			0.089***	0.081***	0.046**	0.108***	0.098***
			(0.013)	(0.023)	(0.018)	(0.020)	(0.023)
Reference income		-0.015***					
		(0.003)					
Income		0.00009*	-0.0001*	-0.0002*	0.0002	-0.0001	-0.0002*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Wealth proxies							
Savings							0.010
							(0.016)
House							0.042*
							(0.023)
Car							0.042***
							(0.023)
Sociodemographics							
Female		0.027***	0.027***	0.020	0.042***	0.029**	0.047***
		(0.010)	(0.010)	(0.012)	(0.016)	(0.013)	(0.015)
Age		0.105***	0.105***	0.103***	0.112***	0.108***	0.105***
		(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Age squared		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Children		0.134***	0.134***	0.152***	0.107***	0.141***	0.139***
		(0.006)	(0.006)	(0.008)	(0.009)	(0.008)	(0.009)
Education		0.141***	0.134***	0.103***	0.191***	0.127***	0.110***
		(0.009)	(0.010)	(0.012)	(0.015)	(0.013)	(0.014)
Unemployed		-0.317***	-0.304***	-0.193***	-0.421***	-0.351***	-0.340***
		(0.015)	(0.015)	(0.020)	(0.022)	(0.021)	(0.022)
Selfemployed		-0.006	-0.007	0.032	-0.100***	-0.021	-0.047*
		(0.020)	(0.020)	(0.024)	(0.038)	(0.026)	(0.028)
Student		-0.694***	-0.692***	-0.752***	-0.578***	-0.672***	-0.676***
		(0.034)	(0.034)	(0.043)	(0.058)	(0.051)	(0.056)
Retired		-0.119***	-0.117***	-0.154***	-0.105***	-0.133***	-0.109***
		(0.021)	(0.021)	(0.027)	(0.032)	(0.028)	(0.031)
cons	-0.766***	-3.442***	-3.471***	-3.292***	-3.589***	-3.461***	-3.483***
	(0.004)	(0.070)	(0.071)	(0.088)	(0.121)	(0.095)	(0.105)
N	120392	105301	105301	65341	39960	58135	49164

Dependent variable: binary response indicating if respondents have a loan. Estimation method: probit. Region, country, time fixed effects for 2009-2018 included in all estimations. Robust standard errors in parentheses. Variables are defined in Appendix.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

BOFIT Discussion Papers

A series devoted to academic studies by BOFIT economists and guest researchers. The focus is on works relevant for economic policy and economic developments in transition / emerging economies.