the review of income and wealth

Review of Income and Wealth
Series 67, Number 1, March 2021
DOI: 10.1111/roiw.12454

# HOUSEHOLD DEBT AND INCOME INEQUALITY: EVIDENCE FROM ITALIAN SURVEY DATA

by David Loschiavo*

*Bank of Italy*

Does regional income inequality affect a household's likelihood of being indebted? This question is addressed by using survey data on Italian households. The analysis shows that inequality in the regional income distribution has a negative effect on the probability of being indebted. However, this effect is asymmetric across income groups: greater inequality increases (decreases) the share of indebted households that are at the top (bottom) of the income ladder. The study suggests that supply factors are more important than demand ones in explaining these results.

**JEL Codes**: D14, D63, G21

**Keywords**: credit rationing, household debt, income inequality, Great Recession, regional data

## 1. Introduction

In most OECD countries, household debt reached exceptional levels over the decade before the Great Recession. In the aftermath of the crisis, the role of household leverage as a potential source of financial instability has become a central question in policy and academic debates.[1] Differences in household credit market participation are anyway very stark across countries. Much of the literature on household finance attributes cross-country differences to household characteristics (age, education, income, etc.) and their interaction with different economic environments (e.g. Christelis *et al.*, 2015; Coletta *et al.*, 2014; Porta *et al.*, 1998). Among the latter factors the role of income inequality has not been adequately explored.[2]

Income inequality may affect borrowing from both the supply and demand side. Rajan (2010) argues that the upsurge of credit supply to lower income U.S. households was due to a political push to support their consumption in response

*Note*: I would like to thank Laura Bartiloro, Raffaello Bronzini, Silvia Magri, Sauro Mocetti, Salvatore Morelli, Paolo Sestito for useful suggestions and comments. Helpful comments from the seminar participants of the Trends in Inequality Conference of 2–4 November 2017 held in Bologna and of the Banking Research Network Workshop of 24–25 September 2015 held at the Bank of Italy are gratefully acknowledged. The views expressed here are those of the author and should not be attributed to the Bank of Italy. All remaining errors are mine.

*Correspondence to: David Loschiavo, Bank of Italy, Via Nazionale 91, 00184 Rome (david.loschiavo@bancaditalia.it).

[1]According to many scholars household indebtedness induced macroeconomic instability in many countries and played an important role in the Great Recession (Mian and Sufi, 2015).

[2]Nevertheless, in recent years, interest in the impact of income inequality on potential growth and its role in causing the crisis and the weak recovery has risen (e.g. Fitoussi and Saraceno, 2010; Piketty 2014; Summers, 2014; Ostry *et al.* 2014).

to rising income inequality and stagnant incomes.[3] Regarding the demand side, some studies show that widening economic inequality led households to borrow more to allow smoothing consumption despite a volatile income (Krueger and Perri, 2006; Iacoviello, 2008).

This paper aims to empirically explore the relationship between the likelihood of being indebted and local income inequality and to understand the role that supply and demand factors play in mediating this relationship. I use micro-data on Italian households from *Survey on Household Income and Wealth* combined with information on local inequality from *EU-Statistics on Income and Living Conditions* survey. I first explore if the extent of a region's income inequality affects the resident households' credit market participation. Then, following Coibion *et al*. (2014) I compare the probability of debt across income groups located in regions with different degrees of income inequality.[4] Finally, I test alternative hypotheses about the prevalence of demand or supply factors in shaping the relationship between local inequality and household probability of holding a debt, using information on loan demand and credit rationing.

I use Italian regions as the unit of analysis, following a long tradition in economics that uses regions' or cities' heterogeneity as laboratories to understand sources of differences between countries (e.g. Barro, 1991; Gennaioli *et al*., 2014). The heterogeneity across Italian regions is particularly suited to simulating differences across countries. In fact, Italian regions are very heterogeneous in terms of economic structure, stage of economic development, household characteristics, percentage of indebted households and income inequality (see Section 3). In addition to this, comparison across heterogeneous regions of the same country helps isolate the effect of income inequality on household debt from other institutional features affecting the propensity for private debt. These features may substantially differ across countries but do not have a differentiated effect across regions within the same country. In this way it is possible to set aside the confounding effect of some important but difficult to measure determinants of credit supply such as the strength of legal rights (i.e. the degree to which laws protect borrowers and lenders) and the quality of credit information available through credit registers.

To the best of my knowledge, this is the first paper studying the impact of local inequality on debt outcomes for a representative sample of Italian households. The evidence available at the micro-data level is scant and focused on only two other countries. Using Dutch household survey data, Georgarakos *et al*. (2014) provide evidence of the existence of the demand-driven effect of income inequality on household debt. They find that a higher average income in the social circle, as perceived by a household, increases the probability that this household will have outstanding and sizable loans. In contrast, Coibion *et al*. (2014) find evidence of a systematic, supply-driven relationship between local inequality and differential borrowing patterns across richer and poorer U.S. households. However, in both countries (the Netherlands and U.S.A.) the incidence of households with debts is

[3]Regarding the supply side, Kumhof *et al*. (2013) also suggest that permanent positive shocks to the income share of high-income households led to increased supply of loanable funds to poor and middle-income households, allowing the latter to sustain higher consumption levels.

[4]Despite this, some important differences with my paper remain. See Section 2.

relatively high. Thus, the Italian case is also interesting for all the countries with a relatively lower diffusion of debt but a high degree of income inequality and a distribution of debt highly skewed towards high-income households. Furthermore, unlike these previous studies, in this work it is possible to conduct a panel analysis, to observe the effect of local inequality on loan demand and credit rationing, and to address sample selection, simultaneity, and omitted variable bias concerns.

The analysis provides evidence that income inequality negatively affects households' credit market participation and that the more unequal a region is the more the indebted households are concentrated among the richer ones. The findings persist after controlling for socio-demographic differences, sample selection, endogeneity issues, and according to several other robustness checks. Moreover, local inequality does not seem to affect the likelihood of applying for a loan but decreases the probability of loan application refusal for top income households (vice versa for poorer ones). Such results are consistent with Coibion *et al.* (2014) model in which "banks use an applicant's position in the local income distribution, along with the dispersion of that distribution, to make inferences about default risk". Although in line with the model's predictions, my results empirically depart from theirs, mainly due to the differences between Italian and U.S. household debt markets.[5] As highlighted in the next section, this paper additionally contributes to understanding of income inequality's understudied role in household debt diffusion.

The remainder of the paper is organized as follows: Section 2 reviews related literature and highlights the contributions of this analysis; Section 3 describes the dataset and documents some stylised facts motivating this work; Section 4 discusses possible channels through which income inequality might influence borrowing and lending behaviours and lays out the empirical strategy; Section 5 presents the main results; Section 6 extends the analysis to different types of debt and compares pre-crisis and crisis periods; Section 7 discusses endogeneity issues and presents additional robustness checks; and Section 8 concludes. Appendix A in Supporting Information contains a comparison of descriptive statistics from the two surveys used in the paper.

## 2. Related Literature

This paper relates to different strands of literature. First, it contributes to the literature studying the determinants of households' participation in the debt market and, more specifically, to the literature on the importance of relative income for household consumption, debt, and portfolio decisions. Households with incomes below average in their social circle tend to consume a larger share of their income to keep up with peers.[6] The problem of "keeping up with the Joneses" has been proposed to explain U.S. households' overspending and excess of labour supply (Stiglitz, 2012). Georgarakos et al. (2014) provided evidence that the signaling

[5]See Section 5.3 for more details.
[6]The idea that concerns about social status may shape consumption decisions can be traced back to the works of Veblen (1899) and Duesenberry (1949).

motive and concerns about social influence might feature in borrowing decisions.[7] The empirical setting adopted in this work allows to test the influence of "keeping up with the Joneses" on loan demand by comparing differential borrowing patterns across richer and poorer households located in high or low inequality regions.

The paper also relates to the literature on adverse selection in imperfect credit markets and its effects on banks' lending policies. Starting with the classic work of Stiglitz and Weiss (1981), a large body of research shows how imperfect information can lead to credit rationing (e.g. Bester, 1985; Besanko and Thakor, 1987). Since creditworthiness is private information, banks may use local income inequality, together with an applicant's position in the local income distribution, to screen borrowers as in standard models of financial contracting under adverse selection. Coibion *et al*. (2014) present a model that provides one potential supply-side explanation for why differential borrowing behaviours could be related to regional inequality. Each region is composed of two types of households, such that "high-type" households have higher income on average than "low-type" households and are also less likely to default on debt. Banks in each region lend to these households, but do not observe households' types, only their income and another signal (not observed by the econometrician) correlated with the underlying type. The key mechanism in the model is that as local income inequality rises, banks treat an applicant's income as an increasingly precise signal about their type, and therefore will make credit more readily accessible (or cheaper) to high-income households.

Directly inspired by the model of Coibion *et al.* (2014), this paper furthers understanding of income inequality's understudied role in borrowing patterns and lending decisions. To this end, the analysis focuses on the Italian economy, where substantial differences in household debt market participation and income inequality exist across regions. Unlike Coibion *et al*. (2014),[8] the present analysis is on the credit market participation and the influence that inequality may exert on the extension of Italy's household debt market. This work takes advantage of the richness of data allowing loan demand and credit rationing to be observed separately. This helps to test alternative theoretical predictions regarding the role of income inequality in household debt market participation, shedding more light on their relative importance. Moreover, by exploiting the longitudinal feature of SHIW survey, this paper takes also into account unobserved household heterogeneity.[9] Finally, the work suggests a potential explanation of why information contained in the income signal gets stronger when inequality is higher, highlighting how persistent income inequality affects credit rationing (see Section 4).

[7]Other contributions point to the relevance of the "keeping up with the Joneses" phenomena: Bertrand and Morse (2013) have documented the importance of trickle-down consumerism, showing that not only does spending increases if one lives in a community with higher income inequality, but so do bankruptcy and self-reported financial distress; Frank *et al*. (2014) have put forward a similar hypothesis called "expenditure cascades". They provide empirical evidence that increased income inequality is associated with overspending, reflected in higher bankruptcy rates for example. Using U.S. survey data, Bricker *et al*. (2014) find that a household's income rank is positively associated with its expenditures on high status cars, its level of indebtedness, and the riskiness of its portfolio.

[8]In their analysis they are interested in explaining the exceptional rise in debt accumulation in the U.S. across different segments of the population over the course of the 2000s.

[9]Furthermore, the dataset used includes information on households' income so that, unlike the work of Coibion *et al*. (2014), no income imputation is needed.

## 3. Data, Stylised Facts and Motivation

### 3.1. *Data Sources*

The data used in the analysis are obtained from different sources. The main one is the *Banca d'Italia's Survey on Household Income and Wealth* (SHIW), which has been carried out biennially since 1987.[10] The sample used for this analysis consists of the five waves 2004–2012, covering approximately 40,000 observations (nearly 10,000 in the panel dataset).

Since regional variations in income inequality are higher than time variations at the national level, to identify any potential effect of inequality on household behaviour the existing empirical literature focus on the level of *local* inequality. The central idea is that the social interactions and relative positioning among neighbours shape the relationship between borrowing patterns and inequality, making geographic proximity a key dimension for the definition of both the relevant reference group in the case of demand-driven patterns and for the relevance of the signaling channel in the supply-side case.[11] However, narrowing down to a very fine level of spatial aggregation might raise endogenous sorting problems that can bias results. On the other side, households might have stronger incentives to signal their relative income rank at more local levels. Bearing this potential trade-off in mind, the focus of this work will be at the level of Italian regions (NUTS 2). In fact, this is a level of spatial aggregation that still matters for the economic mechanism I want to test (signaling), since it is very close to what is considered a relevant local credit market for households in Italy (Gobbi and Lotti, 2004). At the same time this level of aggregation is wide enough to minimize the risk of sampling error and of unobserved factors that determine both household selection into an area and its borrowing behaviour.[12]

However, the SHIW survey sample is not designed to be *representative* at the *regional* level. To overcome this limitation, I construct regional income inequality measures from the Italian leg of the Eurostat's EU-SILC survey (*EU-Statistics on Income and Living Conditions*), which is designed to be representative at regional level, counting approximately 19,000 households (41,000 individuals).[13]

Given that EU-SILC data are available since 2003, I pool data from the 2004 to 2012 SHIW waves, a time span which allows comparisons between the periods before and during the crisis.

[10]The sampling is in two stages: first municipalities are chosen from different strata from throughout Italy and then households are randomly chosen from registry office records within each chosen municipality. Up to 1987 the survey was conducted with time-independent samples (cross-sections) of households, since 1989 part of the sample has comprised households interviewed in previous surveys (panel households). Comprehensive descriptions of the survey are given by Brandolini and Cannari (1994) Guiso and Jappelli (2002).

[11]The existing literature ranges from as aggregated a geographic level as the state to as fine a level as the Metropolitan Statistical Area. Other works, thanks to a rare availability of data, narrow the context where inequality impacts on household debt at the finer level of zip codes.

[12]Moreover, the analysis is focussed on local income inequality for other two reasons: first, this is likely to be the most relevant metric when households compare themselves to others; second, it avoids measurement issues associated with comparing incomes across very different regions.

[13]In Supporting Information Appendix A, it is shown a general good fit between SHIW and EU-SILC estimates of household incomes in terms of quartile distribution, inequality indices and their time trends.

I use households' residence and merge social and economic data with information drawn from other sources. Data on credit quality are from the *Italian Credit Register* (henceforth, CR) owned by Banca d'Italia. Data on house prices are obtained from *Observatory of the real estate market* managed by Agenzia delle Entrate (AdE - the Italian Revenue Agency).

Sample Restrictions

I apply various selections to the dataset. First, only households with positive income are considered. Second, to minimize potential age related selection effects, I restrict the sample to households whose head is older than 20 or younger than 70. Finally, in order to avoid outliers influencing the estimation's results, I also exclude the observations in the 1st and the 99th percentile of the distribution of disposable income. After having applied the mentioned selections, the sample used in the baseline estimations consists of 29,282 observations (17,038 households) in the cross-section dataset and of 7,762 observations (1,816 households) in the panel dataset.

### 3.2. *Motivating Facts*

Table 1 provides summary statistics on debt holders by SHIW survey year. Data show that the percentage of households with debt increases until 2008 and decreases after the eruption of the crisis. In each wave the frequency of debt monotonically increases with households' income quartiles and is higher among middle age, more educated and larger families.

Figures 1 and 2 show how the percentage of households indebted in Italy is the lowest among Euro area countries, albeit there are sizable differences across Italian regions: credit market participation is higher in Central and in some Northern regions (Lombardy and Veneto), whilst far below the average in Southern regions.

Figure 3 plots Gini coefficients of equivalised[14] income across regions. Inequality is higher in the Southern regions, as well as in some regions of Centre and Northwest Italy (Lazio and Liguria), lower in Northeast regions. Overall, regional differences in income inequality are substantial: in 2012 the average Gini coefficient was of 0.30 with a standard deviation of (0.03) and, therefore, a coefficient of variation of 9.2 percent.[15] These stylised facts support the first research

[14]In the following analysis I will always refer to equivalised household (monetary disposable) income which is normally considered the most appropriate indicator of the standard of living of a family. Equivalised household income is the total household income adjusted by the application of an equivalence scale to facilitate comparison of income levels between households of differing size and composition. In this work I apply the modified OECD scale of equivalence, which assigns a coefficient of 1 to the head of household, 0.5 to other household members aged 14 or more, and 0.3 to those younger than 14. For each household the number of "equivalent adults" is calculated by summing the coefficients assigned to the various members. Household income is then divided by that coefficient and allocated to each household member. For more details see the Canberra Group 2011, 2011, pp. 68–72). Finally, by monetary disposable income it is meant disposable household income net of imputed rents and gross of negative interests.

[15]Income, consumption and wealth inequalities in Italy have been thoroughly documented by Jappelli and Pistaferri (2010). They find that, between 1980 and 2006, income inequality was higher and has grown faster than consumption inequality. Recently, Acciari and Mocetti (2013) using Italian administrative fiscal data show that there is strong heterogeneity of inequality in income distribution among Italian regions, with Southern ones on average more unequal than Northern regions.

TABLE 1
Percentage of households with debt for personal needs by social and economic characteristics[A]

| | 2004 | 2006 | 2008 | 2010 | 2012 | Number of households[b] |
|---|---|---|---|---|---|---|
| **Sex of the head** | | | | | | |
| Male | 25.7 | 24.9 | 26.0 | 23.4 | 23.0 | 5,398 |
| Woman | 13.2 | 16.4 | 18.7 | 15.9 | 15.2 | 2,654 |
| **Age of the head** | | | | | | |
| Less than 35 years | 30.8 | 30.1 | 31.5 | 29.3 | 22.5 | 503 |
| 35 to 44 years | 34.2 | 36.5 | 37.8 | 33.0 | 32.8 | 1,137 |
| 45 to 54 years | 31.4 | 26.2 | 34.1 | 32.0 | 30.6 | 1,664 |
| 55 to 64 years | 19.6 | 22.8 | 22.3 | 19.4 | 18.2 | 1,555 |
| 65 years or older | 4.9 | 6.3 | 6.3 | 5.2 | 6.5 | 3,193 |
| **Education of the head** | | | | | | |
| Primary school or without education | 9.0 | 9.0 | 9.0 | 7.6 | 6.3 | 2,071 |
| Junior high school | 23.0 | 23.4 | 27.1 | 23.6 | 22.0 | 2,804 |
| High school | 31.2 | 30.3 | 31.3 | 26.7 | 27.1 | 2,146 |
| Degree or more | 28.4 | 27.1 | 29.3 | 27.6 | 25.9 | 1,031 |
| **Job status of the head** | | | | | | |
| Employee | 31.3 | 31.5 | 34.0 | 33.5 | 30.3 | 3,240 |
| Self-employed | 31.6 | 29.4 | 31.4 | 21.5 | 24.3 | 824 |
| Non worker | 7.8 | 9.4 | 9.7 | 7.9 | 7.8 | 3,988 |
| **Quartiles of equivalised income** | | | | | | |
| First quartile | 13.5 | 14.8 | 17.7 | 15.8 | 13.7 | 1,699 |
| Second quartile | 19.9 | 17.5 | 19.6 | 18.2 | 17.8 | 2,067 |
| Third quartile | 23.3 | 24.4 | 27.4 | 22.6 | 23.0 | 2,146 |
| Fourth quartile | 29.5 | 31.4 | 29.6 | 26.7 | 25.5 | 2,140 |
| **Quartiles of equivalised net wealth** | | | | | | |
| First quartile | 18.9 | 18.6 | 20.7 | 19.3 | 16.8 | 1,795 |
| Second quartile | 26.1 | 24.2 | 30.8 | 28.4 | 27.5 | 1,702 |
| Third quartile | 21.2 | 23.8 | 23.7 | 18.5 | 21.7 | 2,150 |
| Fourth quartile | 21.7 | 23.0 | 21.2 | 19.0 | 16.5 | 2,405 |
| **Size of the municipality** | | | | | | |
| Up to 20,000 inhabitants | 22.3 | 23.4 | 25.0 | 20.4 | 19.4 | 2,005 |
| More than 20,000 inhabitants | 21.5 | 21.4 | 22.7 | 21.6 | 21.2 | 6,047 |

TABLE 1 (CONTINUED)

| | 2004 | 2006 | 2008 | 2010 | 2012 | Number of households[b] |
|---|---|---|---|---|---|---|
| **Geographical areas** | | | | | | |
| North West | 24.4 | 25.0 | 26.8 | 20.6 | 19.3 | 1,913 |
| North East | 24.5 | 24.5 | 27.2 | 21.7 | 22.6 | 1,586 |
| Centre | 20.8 | 24.7 | 19.7 | 25.1 | 25.6 | 1,714 |
| South and Islands (Mezzogiorno) | 18.6 | 16.8 | 21.5 | 18.4 | 16.7 | 2,839 |
| **Number of households' members** | | | | | | |
| 1 member | 11.1 | 14.9 | 12.7 | 10.3 | 10.4 | 2,160 |
| 2 members | 16.4 | 15.8 | 18.9 | 14.6 | 15.0 | 2,526 |
| 3 members | 28.1 | 27.7 | 31.4 | 27.6 | 28.0 | 1,550 |
| 4 members | 35.7 | 32.4 | 35.8 | 35.9 | 32.8 | 1,330 |
| 5 and more members | 27.6 | 34.0 | 38.0 | 33.6 | 32.2 | 486 |
| **Number of income earners** | | | | | | |
| 1 earner | 15.5 | 18.3 | 18.2 | 15.3 | 15.5 | 4,119 |
| 2 earners | 27.7 | 25.0 | 28.8 | 26.3 | 25.8 | 3,200 |
| 3 earners | 27.7 | 29.7 | 28.7 | 25.7 | 25.1 | 597 |
| 4 earners and more | 37.8 | 29.3 | 33.3 | 27.5 | 22.7 | 136 |
| **Total** | 21.9 | 22.3 | 23.7 | 21.0 | 20.3 | 8,052 |

*Source*: Survey of Household Income and Wealth.

[a]The frequencies are weighted and refer to the whole sample; the 5 categories included in the debt for personal needs are: Buildings, Other real assets, Vehicles, Durable goods, Non-durable goods.

[b]Number of households in 2012 wave.

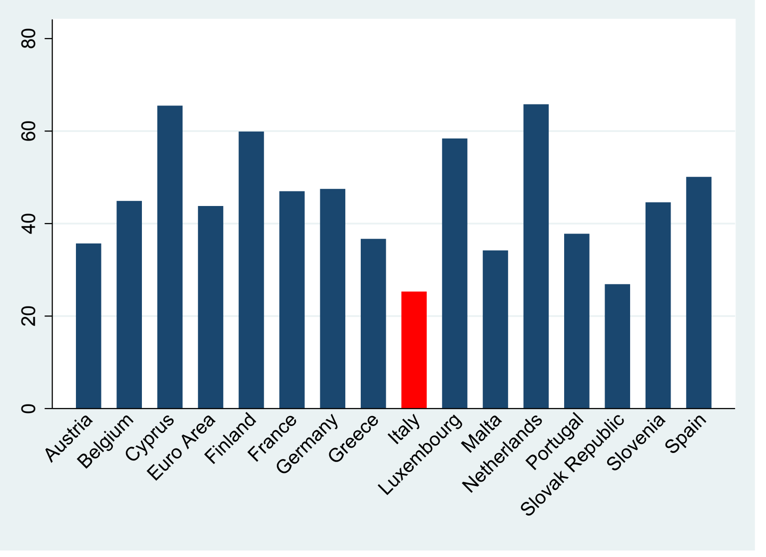


Figure 1. Households with Debt in Euro Area Countries (Percentages)
*Source*: ECB Household Finance and Consumption Survey (2013), Gambacorta *et al.* (2013). [Colour figure can be viewed at wileyonlinelibrary.com]

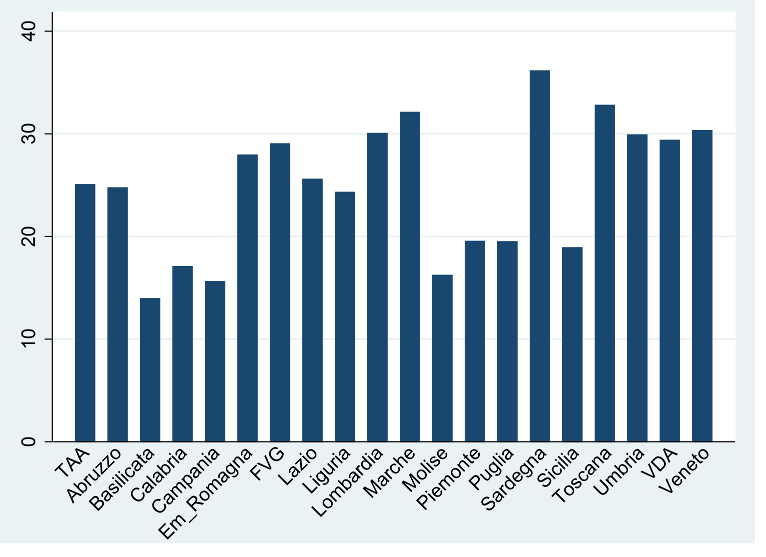


Figure 2. Households with Debt Across Italian Regions (Percentages)
*Source*: Eu-Silc 2012. [Colour figure can be viewed at wileyonlinelibrary.com]

question of this work: is there a link between the degree of inequality in the regional income distribution and the frequency of households with debt?

A *prima facie* evidence is provided by Figure 4 plotting the share of households with debt against the Gini coefficient per annum for all Italian regions. The

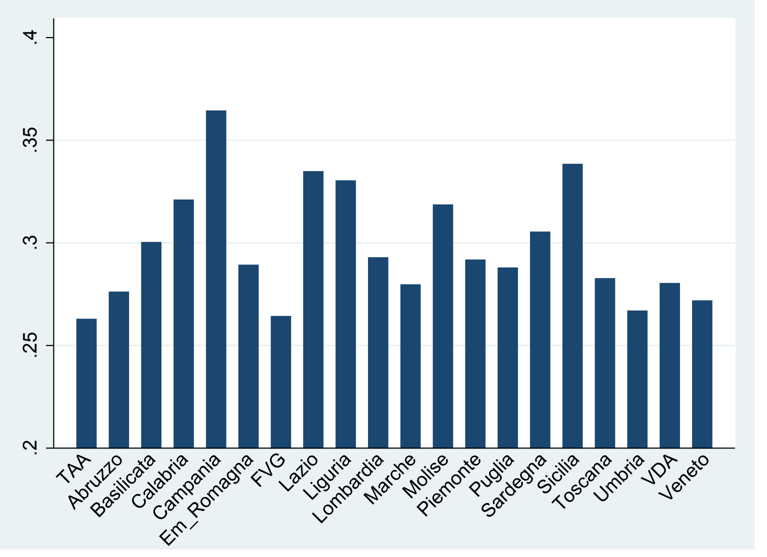


Figure 3. Income Inequality Across Italian Regions.
*Note*: Gini coefficients of equivalised income
*Source*: Eu-Silc 2012. [Colour figure can be viewed at wileyonlinelibrary.com]

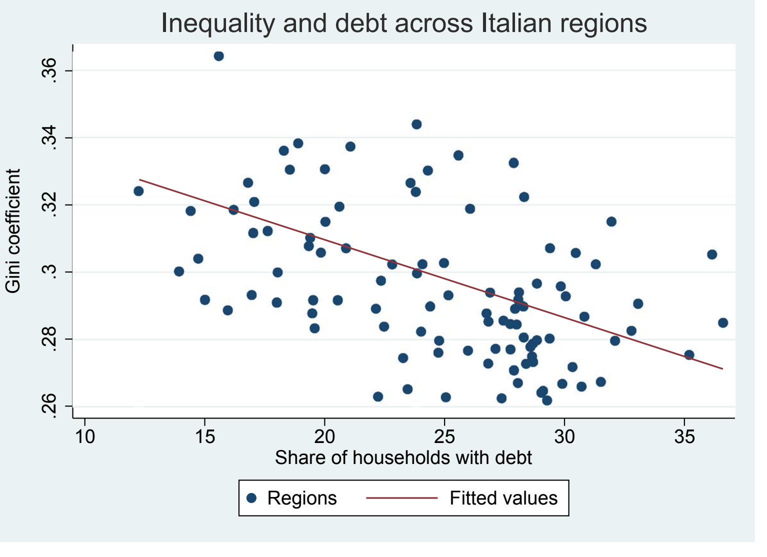


Figure 4. Frequency of Debt Versus Inequality
*Note*: The frequencies are weighted and refer to the whole sample of households.
*Source*: Eu-Silc (waves 2004, 2006, 2008, 2010, 2012). [Colour figure can be viewed at wileyonlinelibrary.com]

negative correlation between the two measures is quite apparent, suggesting that the incidence of household with debt is lower in regions with a higher degree of income inequality and vice versa. This, in turn, leads to the following analysis of how could differences in inequality relate to household debt diffusion.

## 4. Effects of the Local Income Inequality on Households' Credit Market Participation

### 4.1. *Possible Channels*

Income inequality may influence households' debt from both the demand and supply side. From the former, households below the top percentile in the income distribution might aspire to imitate the consumption patterns of richer ones. Thus, a potential effect of income inequality is that households tend to take on debt to keep up with peers. I call this hypothesis the "keeping up with the Joneses" influence on loan demand.[16]

From the supply side, according to the channel highlighted in Coibion *et al.* (2014) which I would refer to as the "signaling channel" influence on loan supply, top-income households located in highly unequal regions are deemed safer borrowers than corresponding ones residing in low-inequality regions. The rationale of their model is that banks cannot observe borrower's ability to meet debt obligations so that they take observed income, together with its rank in the local income distribution, as a signal: higher income rank means higher ability to pay but the signal is stronger when inequality is higher. So, banks give more importance to it and enable relatively greater access to credit to high-income households either by charging lower interest rates or by denying loans less often.[17] A possible explanation for why increased inequality enhances the signal embedded in the relative income is that local inequality is negatively correlated with income mobility, therefore banks are likely to restrict access to credit for high-income applicants less often than they would if inequality were lower.

Indeed, during the underwriting process banks are interested in a borrower's willingness and capacity to repay obligations. The borrower's willingness to repay is assessed largely by subjective factors, such as the applicant's past credit history, and institutional ones (e.g. the strength of lenders' legal protection and the average time to resolve insolvencies for the judicial system). Capacity is determined by the borrower's ability to generate cash flow to service the interest and principal on the loan in the future. However, since creditworthiness is private information, banks observe neither the ability nor the willingness. It follows that they need to infer the probability of prospective borrowers defaulting based on presently available information that best predicts his/her future ability to meet obligations. In their screening technology banks usually use both hard and soft information: via two surveys carried out in 2006 and in 2009 by the Bank of Italy, Del Prete *et al.* (2013) show

[16]There are several similar, albeit different, channels that may influence loan demand. For example, the so-called "getting ahead of the Joneses" effect may occur when wealthier households care more about their social position than poorer ones. Near-to-top-rank households might use consumption of status goods to signal information about their wealth to others in their reference group. The desire to signal their status as richer than actually it is may push some households to access credit to buy luxury goods. Another potential channel is when increased inequality has a welfare-enhancing "anticipatory feelings" effect which Hirschman and Rothschild (1973) named the "Tunnel Effect". The idea is that individuals observing other people's faster income growth interpret this movement as a sign that their own future income is likely to move in the same direction as their social circle (Senik, 2008). Those at the bottom of the income ladder may therefore be encouraged to enter the debt market as they decide to anticipate future expected income through credit to smooth consumption.

[17]See Section 2 for more details.

that, despite the wide usage of rating and scoring methodologies, Italian banks adopted these devices in a flexible way, giving importance to both hard and soft types of information in their lending decisions.[18]

Current income is obviously the main factor that demonstrates capacity to repay debt in the present. Nevertheless, to mitigate the probability of future default, banks are also interested in income stability over time. Thus, additional information available at the moment of the lending decision needs to be considered along with the level of current income as a basis to develop reasonable expectations about the future borrower's ability to meet obligations.[19] Examples are the job status of the borrowers (employee versus self-employed) or the type of job contract (permanent versus temporary contract). Yet an emerging body of evidence has highlighted that the higher the income inequality the more strongly income position persists over time (i.e. those at the bottom of the income distribution have a good chance of remaining there, as do those at the top (Kopczuk *et al.*, 2010; Stiglitz, 2012; Galor and Zeira, 1993; Piketty, 1997).[20] In other words, a household's rank in the local income distribution at time $t$ is a better predictor of its rank at time $t+x$ if it is located in a highly unequal region rather than in a low inequality one. Preliminary evidence suggests that the same also holds true for Italian households.[21] It follows that income inequality strengthens the signal embedded in the current income, enhancing the screening capability of lenders. This is relevant for banks' screening policy and lending decisions in a context of financial contracting under adverse selection.[22]

[18]Moreover, the relevance of ratings in deciding whether to grant a loan declined during the period between the two surveys: in 2009 the percentage of banks considering rating and scoring methodologies not important for their decision to grant a household loan ranged from 25 percent of the large and medium size banks up to nearly two thirds for smaller banks.

[19]There is wide evidence that banks' credit supply decisions are led by uncertainty about future income more than households' current income level (Magri, 2007; Michelangeli and Sette, 2016).

[20]In other words, the mechanism (or the set of mechanisms) causing the persistence of inequality seems to be more "effective" the higher the current level of inequality is (Corak, 2013); the reasons being greater opportunities for top-income households and the concentration of power so that some groups are in a position to structure policies in their own favour (Acemoglu *et al.*, 2005; Claessens and Perotti, 2007; Dabla-Norris *et al.*, 2015) or to get higher returns on wealth if this is more concentrated (Fagereng *et al.*, 2016; Saez and Zucman, 2016).

[21]SHIW data confirms this prediction. Focusing only on the panel households in years 2004 and 2012 (the beginning and end of the analysis timespan), several measures that capture the correlation in household's income rank between the two years indicate that there is a stronger persistence of the households' position (Income Decile) in the local income distribution when inequality is high. For instance, the Pearson correlation coefficient between the initial (at year 2004) and the final (at year 2012) households' rank is greater for households living in high inequality regions than for ones located in low inequality areas.

[22]Thus, if two borrowers have the same level of current income and belong to the same income position in the local income distribution but are living in areas with different degrees of income inequality, then the income of the borrower residing in the high-inequality region is, ceteris *paribus*, likely more persistent over time than the one of the corresponding borrowers living in a low-inequality area. In this way, the income of the former is a stronger signal of its future income because its income position is more persistent when income inequality is higher. It should be stressed that the relevant dimension of persistence is a positional one: i.e. separately from any changes in the shapes of the marginal distributions that may occur. In fact, the way in which shocks affecting the whole economy impact individual income movement over time are not easily predictable. For example, equiproportionate income growth does not alter each person's position relative to the position of others. On the other hand, the intensity of an income shock to the economy may not be the same along the income distribution (arguably higher for poorer and lower for richer in the case of negative shocks (Denk and Cazenave-Lacroutz, 2015).

All in all, local income inequality may be a useful signal for screening borrowers in the sense that current income position is a stronger signal of creditworthiness when inequality is higher.

### 4.2. *Empirical Strategy*

The aim of this analysis is to examine the effect of local income inequality on households' credit market participation. To this end, I first estimate the probability that a household has a loan, as a function of its position in the local income distribution, conditional on local income inequality. Following Georgarakos *et al.* (2014) and Coibion *et al.* (2014), in the benchmark specification equations of this type are estimated:

$$\begin{aligned} Debt_{irt} &= \alpha IncDecile_{irt} + \beta Gini_{rt} + \gamma IncDecile_{irt} * Gini_{rt} \\ &+ \delta \mathbf{X}_{irt} + \phi \mathbf{Z}_{rt-1} + d_t + m_a + \epsilon_{irt} \end{aligned} \tag{1}$$

where $Debt_{irt}$ denotes a binary ownership indicator of debt of household $i$ that resides in region $r$ and where $2004 \le t \le 2012$; $IncDecile_{irt}$ is the household's equivalised income decile in the local income distribution which expresses its relative rank; and $Gini_{rt}$ is the reion $r$ Gini coefficient,[23] the adopted inequality measure. $\mathbf{X}_{irt}$ and $\mathbf{Z}_{rt-1}$ represent, respectively, vectors of household-specific and location-specific controls. $\epsilon_{irt}$ denotes the error term. Household's own characteristics include the level of equivalised income, the age, age squared, educational attainment and marital status of the head of the household, household size, whether household dissaved. I also use dummy variables for households living in municipalities with fewer than 20,000 inhabitants and for self-employed workers. Controls at regional level include the ratio of new bad debts, the growth rate of loans to household sector and the growth rate of housing prices. In addition, I control for time $d_t$ and macro-area fixed effects $m_a$.[24] Furthermore, in some specifications also regional or province fixed effects are included. Table 2 shows correlation coefficients among the exogenous variables. Concerns of endogeneity bias potentially affecting the estimation of equation (1) are addressed in Section 7.

In the presence of the interaction effect $IncDecile_{irt} * Gini_{rt}$, $\alpha$ ($\beta$) would represent the effect of Income Decile (Gini) on the probability of debt when Gini coefficient (Income Decile) is zero. Both cases do not exist in practice and, consequently, they are not particularly interesting. To enhance the interpretability of coefficients I centre the two control variables first (by subtracting the relevant median value from each case), and then compute the interaction term and estimate the model. After having centred the two variables, $\alpha$ ($\beta$) represents the effect of Income Decile

[23]The Gini concentration index measures the degree of inequality in the distribution of a given variable such as income or wealth; expressed in percentages, it is equal to zero if all households have the same amount of the variable and to one in the case of total inequality, i.e. where a single household possesses the total amount of the variable. I derive the Gini index from monetary disposable household incomes so that I am dealing with net (after taxes and transfers) inequality rather than market (before taxes and transfers) inequality; the former being the more appropriate definition of inequality for the phenomena under scrutiny in this paper.

[24]Standard errors of all pooled estimates are clustered at household level in order to deal with the panel component of the data and correct for serial correlation.

TABLE 2
Correlations among regressors

| | Gini | Income DEc. | Inc. Dec.*Gini | Age | Age squared | Educ. | Household size | Married | Self empl. | Small city | Dissaving | Bad debts | Housing prices | Loans to hh | Equiv. Income |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gini | 1.00 | | | | | | | | | | | | | | |
| Income Decile | −0.04 | 1.00 | | | | | | | | | | | | | |
| Income Decile*Gini | 0.04 | 0.03 | 1.00 | | | | | | | | | | | | |
| Age | 0.02 | 0.05 | 0.02 | 1.00 | | | | | | | | | | | |
| Age squared | 0.02 | 0.04 | 0.02 | 0.99 | 1.00 | | | | | | | | | | |
| Education | −0.07 | 0.39 | 0.04 | −0.28 | −0.29 | 1.00 | | | | | | | | | |
| Households size | 0.12 | −0.10 | −0.06 | −0.22 | −0.25 | 0.07 | 1.00 | | | | | | | | |
| Married | 0.06 | 0.00 | −0.02 | 0.15 | 0.12 | −0.03 | 0.52 | 1.00 | | | | | | | |
| Self-employed | −0.01 | 0.12 | −0.02 | −0.10 | −0.12 | 0.08 | 0.10 | 0.05 | 1.00 | | | | | | |
| Small city | −0.17 | −0.04 | 0.01 | −0.02 | −0.02 | −0.08 | 0.03 | 0.01 | 0.03 | 1.00 | | | | | |
| Dissaving | 0.11 | −0.37 | −0.04 | −0.10 | −0.10 | −0.09 | −0.01 | −0.03 | 0.02 | −0.04 | 1.00 | | | | |
| Ratio of new bad debts | 0.35 | −0.00 | 0.00 | 0.04 | 0.03 | 0.00 | 0.06 | 0.04 | −0.01 | −0.10 | 0.09 | 1.00 | | | |
| Housing prices | −0.04 | −0.01 | −0.00 | −0.04 | −0.04 | −0.05 | 0.01 | −0.01 | −0.00 | −0.00 | −0.04 | −0.52 | 1.00 | | |
| Loans to households | 0.31 | 0.00 | −0.00 | 0.03 | 0.03 | 0.01 | 0.05 | 0.04 | −0.01 | −0.09 | 0.07 | 0.91 | −0.50 | 1.00 | |
| Equivalised Income | −0.22 | 0.83 | −0.03 | 0.06 | 0.05 | 0.41 | −0.15 | −0.01 | 0.15 | −0.01 | −0.33 | −0.10 | −0.05 | −0.08 | 1.00 |

*Note*: Statistics are computed on the SHIW sample selection used for the cross-sections estimates reported in Table 3. See Section 3.1 for more details.

(Gini) on the probability of debt when Gini coefficient (Income Decile) is equal to its median value.[25]

By estimating equation (1) I test different hypotheses of how borrowing and inequality interact. If both $\beta$ and $\gamma = 0$, then local inequality is irrelevant for household borrowing decisions. Differently, if either $\beta$ or $\gamma$ significantly differ from zero then local inequality plays a role in determining the probability of being indebted. In the latter case, the observed equilibrium relationship between the local inequality and debt market participation may be driven by a combination of demand and supply factors.

The adopted specification allows to investigate this further since the sign of the interaction coefficient ($\gamma$) helps to determine whether local inequality affects households' borrowing patterns differently across income groups and to shed more light on the prevailing channel of effects described in Section 4.1. This, in turn, will help to assess if credit demand or supply factors are more important in practice. Table 1 shows that, at national level, the percentage of households with debt increases with (quartiles of equivalised) income. I expect that this holds true even within each region so that $\alpha > 0$, i.e. low-income households within an area are relatively less indebted than the high-income ones. Moreover, Figure 4 shows that income inequality is, on average, negatively correlated with frequency of debt. It follows that I expect that $\beta < 0$, i.e. a region with higher inequality is associated with a lower frequency of debt. Given the expected signs of $\alpha$ and $\beta$, the interaction ($\gamma$) is then the key parameter to determine whether local inequality affects households' borrowing patterns across income groups.

On one hand, if $\gamma < 0$, the negative impact of local inequality on the frequency of debt strengthens as household income decile increases or, vice versa, low and middle income households living in high inequality regions are *more* likely indebted than similar households in low-inequality regions. Such a scenario suggests that *credit demand* is the driving force behind household borrowing and, in particular, is consistent with the "keeping up with the Joneses" hypothesis: i.e. with the idea that rising inequality leads to increased loan demand through social preferences.

On the other hand, if $\gamma > 0$ the negative impact of inequality on the frequency of debt weakens as household income decile increases or, the other way round, low income households living in high inequality regions are *less* likely indebted than similar households in low-inequality regions. This scenario points to the relevance of *supply factors* in determining the likelihood of having a loan. In particular, according to the finding of Coibion *et al*. (2014), increased inequality allows high-income households to borrow relatively more frequently since "higher levels of inequality imply that applicant incomes are stronger signals of creditworthiness".

Of course, the two scenarios are not mutually exclusive, since, actually, it is possible that both relative demand and supply factors are at work at the same time. In fact, by only observing a household with zero debt it is not possible to distinguish if this is either the exclusive outcome of the demand process or it reflects rejected loan applications. In other words, there are various ways for a household to have zero observed debt holding: there are those who are not interested in debt

[25]Motivations for employing variable centring include also reducing multicollinearity between "conditional main effects" and interaction effect. See for instance Brambor *et al*. 2006, p. 71).

market participation; those who want a positive amount of debt and actually do not apply for a loan because they are discouraged either directly from the loan officer or, indirectly, by the prospect of possible rejection; and those that apply and may be rejected by the lender.

To better evaluate the role that credit demand and supply factors play in determining the observed outcome, I thereby take a futher step by exploiting a different group of SHIW questions that allows to single out *i*) those households that asked for a loan in the year regardless of whether they got it or not from those that did not apply for a loan, from *ii*) those households whose applications have been rejected from those that have been accepted. In this way it is possible to separately observe the role of local inequality, interacted with household position in the local income distribution, among determinants of loan demand and of bank's lending evaluation process.

Loan Demand

In order to evaluate separately those factors acting on the *demand-side*, I estimate the probability for a household of demanding a loan, using as explanatory variables only a subset of the ones in the baseline specification (1) that literature suggests being relevant for household borrowing decisions. I exclude thereby both the ratio of new bad debts and the growth rate of loans to household sector from vectors of location-specific controls $\mathbf{Z}_{rt-1}$ and run the following regression:

$$\begin{aligned} Demand_{irt} = {} & \alpha IncDecile_{irt} + \beta Gini_{rt} + \gamma IncDecile_{irt} * Gini_{rt} \\ & + \theta \mathbf{X}_{irt} + \tau \mathbf{Z}^{D}_{rt-1} + d_t + m_a + \epsilon_{irt} \end{aligned} \tag{2}$$

where $Demand_{irt}$ denotes a binary variable indicating whether or not household $i$ that resides in region $r$ asked for a loan in the year $t$. $\mathbf{Z}^{D}_{rt-1}$ is the vector of location-specific controls relevant for household loan demand that now includes only regional growth rate of house prices; while the other controls are unchanged with respect to (1).

Equation (2) allows to test if local inequality influence household loan demand, with the interaction effect being the key variable to determine through which of the channels described this possible influence may be exerted. For the sake of brevity, let's assume that in line with previous results of household finance literature, $\alpha > 0$, and, according to the demand driven hypotheses described in Section 4.1, inequality increases loan demand (expected sign $\beta > 0$). Given the expected signs of $\alpha$ and $\beta$, if $\gamma < 0$, then the positive impact of local inequality on loan demand weakens as household income decile increases, a result that would be consistent with "keeping up with the Joneses" channel; on the contrary, if $\gamma > 0$ the positive impact of local inequality on the loan demand would strengthen as household income decile increases, a result compatible with the "getting ahead of the Joneses" hypothesis (see footnote 16). Clearly, in the case both $\beta$ and $\gamma = 0$, then local inequality would be irrelevant for household loan demand.

Credit Rationing

In order to evaluate separately the *supply side* of the market, I also estimate the probability that a household application is rejected, conditional on having applied for a loan:

$$
\begin{aligned}
(3) \qquad Raz_{irt} = {} & \alpha IncDecile_{irt} + \beta Gini_{rt} + \gamma IncDecile_{irt} * Gini_{rt} \\
& + \mu \mathbf{X}^S_{irt} + \rho \mathbf{Z}^S_{rt-1} + d_t + m_a + \epsilon_{irt}
\end{aligned}
$$

where $Raz_{irt}$ denotes a binary variable indicating whether or not household $i$ that resides in region $r$ has been credit rationed in the year $t$, provided that it applied for a loan.[26] I use as controls only the subset of exogenous variables included in specification (1) that literature has identified as relevant for bank decision to grant a loan.[27] Thus, $\mathbf{X}^S_{irt}$ and $\mathbf{Z}^S_{rt-1}$ are, respectively, the vectors of household-specific and of location-specific controls that affect lender decisions to rationing; while other controls remain unchanged with respect to specification (1).

Equation (3) allows to assess the relative importance of supply factors in determining the outcome observed in (1). If both $\beta$ and $\gamma = 0$, regional inequality doesn't affect the probability of a loan being rejected. Differently, if from the estimation of equation (3) it is observed that $\alpha < 0$, $\beta > 0$ and $\gamma < 0$, then banks reject otherwise similar applications by top rank applicants *less* frequently in high-inequality regions than in low-inequality regions. Such results would be compatible with the signaling channel highlighted by Coibion *et al.* (2014) and would, thereby, suggest that supply factors play an important role in explaining debt market participation. On the other hand, if $\alpha < 0$, $\beta > 0$ and $\gamma > 0$, banks reject otherwise similar applications by low income applicants *less* frequently in high-inequality regions than in low-inequality regions, a result (still supply driven) that would be compatible with contributions suggesting that permanent positive shocks to the income share of high-income households led to increased supply of loanable funds to poor and middle-income households, allowing the latter to sustain higher consumption levels (see Rajan, 2010; Kumhof *et al.*, 2013 and Section 1).

In equation (3), the probability of rationing is observable only for households that have demanded a loan. To control for the possible selectivity implied in excluding those not asking for a loan, I also estimate a standard Heckprobit approach, thus directly assessing the relevance of selection. Following Magri (2007), the chosen exclusion restriction necessary to identify the model is the dummy variable for household living in a small municipality, with fewer than 20,000 inhabitants. The hypothesis is that *Small City* is an important factor in modelling the entry costs in

[26]The equation mirrors a stringent definition of being credit constrained in the sense that one can only be constrained if actually applies for credit and is rejected. Other laxer definitions adopted in literature include as constrained (i) those who did not actually apply but who wished to have credit nevertheless and did not receive it; (ii) those who report that they have been rejected or unable to gain all of the amount they applied for or who report that they were discouraged from applying. Because of the SHIW questionnaire design, it not possible to distinguish households belonging to the last two categories.

[27]$\mathbf{X}^S_{irt}$ include the equivalised income, age, age squared, educational attainment and marital status of the head of the household, household size, a dummy variables for self-employed workers. Controls at regional level $\mathbf{Z}^S_{rt-1}$ include the ratio of new bad debts and the growth rate of housing prices.

the debt market and thus for the decision whether to borrow or not, but that this is not an important factor in the lender's decision to grant a loan.

In the following Section, I first report the results of the econometric estimation of the probability of holding a debt (1) and then I turn to the results for the loan demand (2) and credit rationing equations (3) respectively.

## 5. The Empirical Results

### 5.1. *Probability of Holding a Debt*

Table 3 presents the main estimates for the effect of local inequality on the probability of being indebted. I first discuss the effect of inequality on the probability of holding any type of debt and then turn to the effects of other independent variables.

Column 1 displays results from a simple LPM regression (Linear Probability Model) run on the pooled dataset (years 2004–2012). The LPM estimation has the advantage of making the interpretation of the coefficients easier, in particular in the case of interaction term.[28] The coefficients from the LPM regression are presented with robust standard errors clustered at the household level in order to deal with the panel component of the data and correct for serial correlation.[29] As expected from descriptive statistics reported in Section 3.2, the coefficient associated to $IncDecile_{irt}$ is positive so that households belonging to top-income groups have a significantly higher probability to be indebted than poorer ones. Furthermore, the relevance of regional income inequality for household borrowing and its negative correlation with the frequency of debt, as anticipated in Figure 1, are confirmed by the significantly negative coefficient of $Gini_{irt}$: the probability of being indebted is higher the less unequal is the region where the household resides. Finally, the coefficient of the interaction between local inequality and household income decile is positive and significant, which suggests that the negative effect of $Gini_{irt}$ on the probability of being indebted is weaker (stronger) when household rank is high (low). The result points to the *supply-side* hypothesis, presented in Section 4.1, according to which poorer household loan demands are rejected by banks more frequently when local income inequality is higher.

Figure 5 presents these results graphically. Panel (a) of Figure 5 plots the relationship between regional Gini coefficient and the estimated likelihood of debt for each household income decile. The graph shows that, as income rank increases, the negative effect of local inequality on probability of debt gets smaller and smaller

[28]Interaction effects interpretation is more complicated in nonlinear models because, like the marginal effect of a single variable, the magnitude of the coefficient depends on all the covariates in the model. In addition, it can have different signs for different observations, making simple summary measures of the interaction effect difficult. Finally, it requires computing the cross derivative or cross difference. However, I report pooled probit and panel probit estimates in each specification after having computed a consistent estimator for the interaction effect for nonlinear model according to the methodology presented in Ai and Norton (2003). It is worth noting that testing for the significance of partial effects in this context remains an open question (see Greene, 2010; Hodge and Shankar, 2014).

[29]In the robustness analysis I test the sensitivity of the results when clustering standard errors both at provincial and at regional level to allow for possible correlations of the unobserved features at a local level.

TABLE 3
DIFFERENT ESTIMATES OF PR(LOAN > 0)—COMPARISON BASELINE

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Income Decile | 0.02*** | 0.02*** | 0.02*** | 0.03*** | 0.02*** | 0.02*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Gini | −0.51** | −0.54** | −1.14*** | −1.15*** | −0.60 | −0.62 |
| | (0.22) | (0.23) | (0.39) | (0.39) | (0.60) | (0.59) |
| Income Decile*Gini | 0.18*** | 0.19*** | 0.29*** | 0.26*** | 0.28*** | 0.23** |
| | (0.06) | (0.06) | (0.10) | (0.10) | (0.10) | (0.10) |
| Equivalised Income | 0.00 | 0.00 | 0.00*** | | 0.00*** | 0.00* |
| | (0.00) | (0.00) | (0.00) | | (0.00) | (0.00) |
| Age | 0.01*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Age squared | −0.00*** | −0.00*** | −0.00*** | −0.00*** | −0.00*** | −0.00*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Elementary school | 0.00 | −0.00 | 0.05** | 0.04* | 0.05* | 0.04 |
| | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| Middle school | −0.02** | −0.02** | −0.00 | −0.00 | 0.00 | −0.00 |
| | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) |
| University | −0.04** | −0.04*** | −0.04 | −0.03 | −0.05* | −0.05* |
| | (0.01) | (0.01) | (0.03) | (0.02) | (0.02) | (0.02) |
| Household size | 0.03*** | 0.03*** | 0.04*** | 0.04*** | 0.04*** | 0.04*** |
| | (0.00) | (0.00) | (0.01) | (0.01) | (0.01) | (0.01) |
| Married | 0.08*** | 0.08*** | 0.08*** | 0.08*** | 0.08*** | 0.08*** |
| | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) |
| Self-employed | −0.05*** | −0.05*** | −0.05*** | −0.05** | −0.05*** | −0.05** |
| | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) |
| Small city | 0.00 | 0.00 | −0.03 | −0.03 | −0.03 | −0.04** |
| | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) |
| Dissaving | 0.12*** | 0.12*** | 0.14*** | 0.14*** | 0.14*** | 0.13*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Ratio of new bad debts | 0.01 | 0.02 | −0.00 | −0.01 | 0.08* | 0.08* |
| | (0.03) | (0.03) | (0.04) | (0.04) | (0.05) | (0.05) |

TABLE 3 (CONTINUED)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Housing prices by region | −0.01*** | −0.01*** | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Loans to household | 0.05 | 0.05 | 0.17 | 0.17 | 0.08 | 0.04 |
| | (0.10) | (0.10) | (0.14) | (0.14) | (0.14) | (0.14) |
| Macro-region FE | Yes | Yes | Yes | Yes | No | No |
| Region FE | No | No | No | No | Yes | No |
| Province FE | No | No | No | No | No | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 29282 | 29282 | 7762 | 7762 | 7762 | 7720 |
| $R^2$/Pseudo $R^2$ | 0.08 | 0.07 | | | | |
| Predicted Probability | | 0.253 | 0.153 | 0.154 | 0.153 | 0.155 |
| LR Test (*p*-value) | | | 0.000 | 0.000 | 0.000 | 0.000 |

*Note*: The Table reports marginal effects (and associated standard errors) from the following three binary regressions used to model the probability of having an outstanding loan: (1) a linear probability model, estimated via OLS, (2) a nonlinear probability model, estimated via pooled probit, and (3) to (6) a nonlinear probability model, estimated via panel probit random effect estimation. Marginal effects are expressed at the mean value of the independent variables and, in the case of interaction terms in nonlinear models, are computed by taking into account cross derivatives (Ai and Norton, 2003). A likelihood-ratio test on the significance of the panel level variance component is included at the bottom of the output under the null that the panel-level variance component is unimportant, and the panel estimator is not different from the pooled estimator. Standard errors are corrected for heteroscedasticity and (in the pooled estimations) clustered at the household level. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

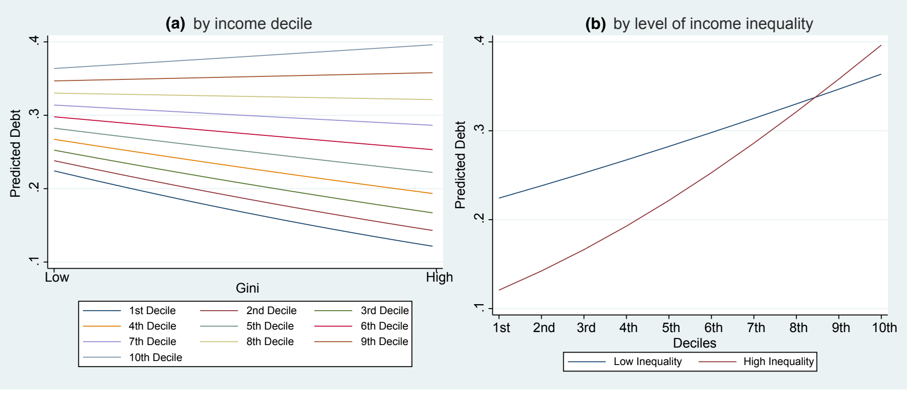


Figure 5. Probability of Debt, Income Rank and Local Inequality

*Note*: Figures report the effect on the probability of debt based on the probit pooled regression in Table 3. [Colour figure can be viewed at wileyonlinelibrary.com]

till reversing for the top income deciles. Panel (b) of Figure 5 plots the relationship between income decile and likelihood of debt for the highest and the lowest level of local inequality. The graph shows that increased inequality allows high-income (low-income) households to borrow more (less) frequently. In fact, households with rank to the right of the crossing are more likely indebted on average as inequality increases; whilst households to the left of the crossing are less likely indebted as inequality increases.

To take into account the well-known limitations of the linear probability model, Column 2 of Table 3 reports the estimated marginal effects[30] from a pooled probit model which shows that the results are substantially equivalent both in terms of the statistical significance, signs and size of the coefficients, as well as in their economic implications.

SHIW data has a household panel component which I exploit to control for unobserved heterogeneity via individual household effects and to track changes in household behaviour over time. All the regressions presented in this work are weighted using SHIW sampling weights, thus reducing the risk of attrition bias.[31]

[30]Marginal effects are calculated at sample means of the independent variables which are usually preferred over average marginal effects when some of the regressors are mathematical transformations of other regressors. In unreported estimations I have checked whether the marginal effects are similar using both approaches across the various estimations of Table 3 and I have found no significant difference in terms of both significance levels and coefficient magnitudes.

[31]Panel households' characteristics may partially differ from those of the whole sample because of panel attrition. In order to account for this potential source of bias and to improve the representativeness of the panel over time, in the SHIW sample design panel households are post-stratified on the basis of the distribution of the whole sample measured in the previous wave and the survey weights are attrition-adjusted (for details see Faiella and Gambacorta, 2007). In addition, and more importantly, it is unlikely that estimates of equation (1) are biased because of selective attrition. For this to be the case, the probability of exit the survey should be, rather implausibly, differentiated according to the degree of local income inequality between same income rank households.

Column 3 of Table 3 reports marginal effects from a random effects panel probit.[32] This model allows to shape in a more precise way differences in the households' behaviour, even though it greatly reduces the number of observations. Although the LPM/pooled probit and random-effects probit panel regression estimates are not directly comparable as regards the size of the coefficients, the results are substantially equivalent both in terms of the statistical significance and signs of the coefficients, as well as in their economic implications. However, at the bottom of column 3 in Table 3, I report a Likelihood Ratio test comparing the pooled probit with the random effects model. The test rejects the null hypothesis under which the random effect makes no contribution to the residual error of each equation and the same parameter values for each equation would result if the observations were all pooled over time. It follows that the panel specification should be considered the preferred one.

It is worth noting that the random effects specification requires the strong assumption that the effects and the regressors are uncorrelated. However, the use of the fixed effects model in nonlinear models is plagued by the incidental parameter problem (Greene, 2012) and, in the context of the present analysis, implies also the loss of information highlighted in footnote 32. Mundlak (1978) approach is frequently used as a compromise between the fixed and random effects models since it preserves the specification of the latter model, but relaxes the assumption that the random effects are distributed independently of regressors. In line with this approach, I have also estimated an augmented specification of equation (1) that contains the panel-level means of the time-varying covariates. By testing if these generated panel-level means are jointly zero, it is possible to determine whether the random effects assumptions are satisfied or not (Greene, 2012, pp. 35–36). I have thereby computed a Wald statistic for the null hypothesis that the coefficients in the augmented model equal zero. The chi-squared statistic equals 18.08 with twelve degrees of freedom. The critical value from the chi-squared table for 95 percent significance is 21.03, so I cannot reject the null that the generated regressors are zero. This suggests that time-invariant unobservables are not related to regressors and that the random-effects model is appropriate.

Column 4 of Table 3 shows that excluding the household's equivalised income, which may induce collinearity with the Income Decile, does not alter results.

In order to distinguish the effect of regional income inequality from other regional characteristics, in the baseline specification I have used area fixed effects at one level of aggregation (macro-region) higher than the geographic area (region) used to construct the income distribution and the income inequality measure. To address the issue of other regional unobservables that may determine both regional inequality and likelihood of debt, I run regressions with regional level fixed effects,

[32]A fixed-effects probit model estimation would drop all households that exhibit no variation in the dependent variable over time. In other words, it would drop all the households which, during the whole period of analysis, are always either indebted or not indebted and would maintain only households switching borrowing status. The latter case would be outside of my modeling strategy of borrowing behaviour focussed, as it is quite standard in household finance literature, on the decision taken in every period on the allocation of resources and the amount of borrowing. More importantly, the adoption of a fixed-effects model would require a drastic drop in the number of observations and, in particular, leave with a modest number of households holding a debt.

which allows averaging out any regional time-invariant effect. Column 5 of Table 3 shows that, despite it is no longer possible to separate the effect of Gini coefficient from other regional characteristics, the coefficient on the interaction term between the household's income decile and regional inequality is still significant and very similar in magnitude to the baseline results.

A relevant local credit market for households in Italy is considered to be a province (Gobbi and Lotti, 2004). Furthermore, provinces with relatively better-developed financial systems are likely to also have higher frequency of debt. Thus, including province fixed effects accounts for any such unobserved time-invariant location attributes, common to the credit market areas, that may affect household likelihood of debt. Results of specifications with province fixed effects reported in Column 6 of Table 3 are almost identical to the ones found with regional fixed effects, thus confirming the baseline findings.

In line with previous empirical evidence, I find that age has a non-linear effect: the probability of debt increases with age and decreases with age in a quadratic term. This result is compatible both with the demand channel (the need for a loan is strongest for youngest households and decreases beyond a certain age threshold) and with the supply channel (banks consider elder households safer borrowers until a certain age threshold when the longer life expectancy and increasing income profile of the younger prevail). The probability of debt is higher among well-educated households (head of household is a high-school), who are more likely to have rising income expectations and to afford lower entry costs due to a better financial education. Estimation's results confirm other evidence from descriptive statistics in Table 1. The frequency of debt increases with household size, is higher among married households and lower among self-employed ones whose income is subject to greater volatility.

### 5.2. *Probability of Demanding a Loan*

In order to single out the role of local inequality among determinants of loan demand, I estimate the probability of a household demanding a loan, using the specification of equation (2). I present the results of the estimations in the left panel (columns 1 to 4) of Table 4. The estimates are practically the same for all the models employed: LPM, pooled probit and panel probit. The key finding is that both the estimated coefficients of $Gini_{irt}$ and $IncDecile_{irt} * Gini_{rt}$ are not significantly different from zero so that there is no evidence that demand-side factors related to local inequality levels matter for the borrowing decisions of households. In other words, both the "keeping up with the Joneses" and "the getting ahead of the Joneses" hypotheses do not find empirical evidence in data on loan applications. Hence, results point mainly toward channels operating through credit supply—namely through the banks' use of local income inequality as additional signal for identifying credit worthy customers.

### 5.3. *Probability of Being Credit Rationed*

Next, I turn to supply side of the market by estimating the probability that a household's loan request is rejected, conditional on having applied for it, using the specification of Equation (3). Columns 6 to 8 in the right panel of Table 4 show

TABLE 4
THE PROBABILITY OF DEMANDING A LOAN AND OF BEING CREDIT RATIONED

| | Loan Demand | | | | | Credit rationing | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit | LPM | Pooled Probit | Pooled Probit |
| Income Decile | 0.00 | 0.00 | 0.00 | 0.00* | | −0.01 | 0.01 | −0.02*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | | (0.01) | (0.01) | (0.01) |
| Gini | 0.14 | 0.15 | 0.01 | 0.01 | 0.01 | 1.26 | 0.66 | 0.99 |
| | (0.13) | (0.12) | (0.15) | (0.15) | (0.15) | (0.77) | (0.75) | (0.76) |
| Income Decile*Gini | 0.03 | 0.03 | −0.01 | −0.01 | −0.01 | −0.45** | −0.44** | −0.32* |
| | (0.03) | (0.03) | (0.04) | (0.04) | (0.04) | (0.19) | (0.19) | (0.19) |
| Equivalised Income | 0.00 | 0.00 | −0.00 | | 0.00 | −0.01** | −0.02** | |
| | (0.00) | (0.00) | (0.00) | | (0.00) | (0.00) | (0.01) | |
| Age | −0.00 | −0.00 | −0.00 | −0.00 | −0.00 | −0.02* | −0.01* | −0.02* |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.01) | (0.01) | (0.01) |
| Age squared | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00* |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Elementary school | 0.00 | 0.00 | 0.02** | 0.02** | 0.02** | 0.12** | 0.10** | 0.12** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.05) | (0.05) | (0.05) |
| Middle school | −0.00 | −0.00 | −0.00 | −0.00 | −0.00 | 0.08** | 0.07** | 0.08** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.03) | (0.03) | (0.03) |
| University | −0.01* | −0.01* | −0.01 | −0.01 | −0.01 | −0.05* | −0.15*** | −0.16*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.03) | (0.05) | (0.05) |
| Household size | 0.01*** | 0.01*** | 0.01** | 0.01** | 0.01* | −0.03* | −0.02 | −0.02 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.01) | (0.01) | (0.01) |
| Married | 0.00 | 0.01 | 0.01* | 0.01* | 0.01* | −0.04 | −0.04 | −0.04 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.03) | (0.03) | (0.03) |
| Self-employed | 0.00 | 0.01 | −0.00 | −0.00 | −0.00 | 0.08** | 0.08** | 0.07** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.03) | (0.03) | (0.03) |
| Small city | −0.01* | −0.01* | −0.01 | −0.01 | −0.01 | | | |
| | (0.01) | (0.00) | (0.01) | (0.01) | (0.01) | | | |
| Dissaving | 0.05*** | 0.04*** | 0.03*** | 0.03*** | 0.03*** | | | |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | | | |

TABLE 4 (CONTINUED)

| | Loan Demand | | | | | Credit rationing | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit | LPM | Pooled Probit | Pooled Probit |
| Housing prices | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.01) | (0.01) | (0.01) |
| New bad debts | | | | | | 0.14** | 0.14*** | 0.14*** |
| | | | | | | (0.06) | (0.05) | (0.05) |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 20683 | 20683 | 5071 | 5071 | 5071 | 1575 | 1575 | 1575 |
| $R^2$/Pseudo $R^2$ | 0.02 | 0.03 | | | | 0.17 | 0.19 | 0.18 |
| Predicted Probability | | 0.05 | 0.03 | 0.03 | 0.03 | | 0.15 | 0.16 |
| LR Test ($p$-value) | | | 0.00 | 0.00 | 0.00 | | | |

*Note*: The left panel of the Table reports marginal effects (and associated standard errors) from the following three binary regressions used to model the probability of demanding a loan: (1) a linear probability model, estimated via pooled OLS, (2) a nonlinear probability model, estimated via pooled probit, and (3)–(5) a nonlinear probability model, estimated via panel probit random effect estimation. The right panel of the Table reports marginal effects (and associated standard errors) from the following three binary regressions used to model the probability of demanding a loan and being credit rationed: (6) a linear probability model, estimated via pooled OLS, (7)-(8) a nonlinear probability model, estimated via pooled probit. Marginal effects are expressed at the mean value of the independent variables and, in the case of interaction terms in nonlinear models, are computed by taking into account cross derivative (Ai and Norton, 2003). A likelihood-ratio test on the significance of the panel level variance component is included at the bottom of the output under the null that the panel-level variance component is unimportant, and the panel estimator is not different from the pooled estimator. Standard errors are corrected for heteroscedasticity and (in the pooled estimations) clustered at the household level. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

estimates, respectively, of LPM and of probit regressions for the probability of being credit rationed.[33] In line with previous evidence on Italian data,[34] income increases the likelihood of application acceptance. In fact, the estimated coefficient of $IncDecile_{irt}$ is negative indicating that a higher household income rank is associated in the bank's evaluation with a higher ability to repay the debt. The important finding is that the coefficient of the interaction term $IncDecile_{irt} * Gini_{rt}$ is significantly negative, indicating that applications from top (low) income households in more (less) unequal regions are less (more) likely to be rejected than those from top (low) income households in less (more) unequal regions.

To control for the possible selectivity implied in excluding those households not asking for a loan, I have also estimated a Heckprobit model identified through the *Small City* variable. Results show that, in the case under scrutiny, the null hypothesis of no correlation between error terms of the two equations (demand and rationing probit models) is not rejected. In other words, coefficients of determinants of credit rationing presented in Table 4 are not biased because of sample selection.

Overall, the result for credit rationing is in line with the supply side interpretation of the results obtained in Section 4.1. It is also consistent with Coibion *et al.* (2014) model, and their empirical findings on US households debt accumulation, in which banks use an applicant's position in the local income distribution, along with the dispersion of that distribution, to make inferences about default risk. Despite this, it is worth stressing that my results empirically depart from theirs, mainly due to the differences between Italian and US household debt markets. First of all, I find that a higher income is associated with a higher likelihood of being indebted while they find that debt accumulation over the course of the early to mid-2000s in the US was, on average, greater for lower income households. More importantly, in my results the signaling channel seems to strengthen in the crisis period when credit supply conditions were tighter. Thus, the effect of local inequality on household debt appears to be larger the stronger is credit rationing in the system. Vice versa, Coibion *et al.* (2014) evidence supports the notion that the growth in household borrowing during the mid-2000s was driven in large part by credit supply expansions targeted at lower-income households. However, as said in Section 4.1, there could be manifold causes that explain how local income inequality is related to a signal of households' creditworthiness.

All in all, results suggest that once other household and location characteristics are controlled for, top-income households located in highly unequal regions are deemed safer borrowers than the corresponding ones residing in low-inequality regions, likely because of the higher persistence of their income profile, which may suggest a greater capability of meeting debt repayments in the future.

[33]Since restricting the analysis to the households that have applied for a loan would require a drastic drop in the number of observations when using the panel dimension, unlike the case of loan demand, for credit rationing I do not estimate the panel probit model.

[34]See Crook and Hochguertel (2006), Magri (2007) and, more recently on mortgages market, Michelangeli and Sette (2016).

## 6. Extensions

### 6.1. *Decomposition by Type of Debt*

I now consider whether the results obtained are sensitive to the difference between types of debt: mortgages and consumer loans. I focus only on the probability of being indebted because the SHIW survey questions on loan demand and credit rationing do not allow for distinction by types of debt. Then, I reproduce the previous regressions and report results in Table 5 for each type of loan. Columns 1 to 3 document that the results for mortgages are almost identical to those found for total debt, albeit the parameter $\gamma$ is estimated to be statistically significant only at a 10 percent level for the random effects panel probit. Although the main effect of local inequality is not significant in Columns 4 and 5 , I find still a strong, statistically significant relationship between local inequality and the probability of holding consumer loans across different income groups as expressed by the interaction effect $\boldsymbol{IncDecile_{irt} * Gini_{rt}}$ and as shown in Column 6. All in all, results from the decomposition by type of debt indicate that both contribute to the total debt patterns described above.

### 6.2. *Pre-Crisis Vs Crisis Subsamples*

The dataset used covers a time span which allows comparisons between the periods before and during the crisis. In the years prior to the crisis, Italy experienced a fast growth in household credit market participation. In the part of the latter period included in the data, significant growth occurred in the housing market, driven by an increase in house prices and consumer credit lending, alongside relaxed credit conditions. These factors led towards a convergence of Italian household credit market participation with higher rates of participation in other countries. However, the financial crisis caused a drop in the demand for credit and a higher selectivity of banks in lending. This has affected the share of households with debt and the composition of borrowers with a rebalance towards high-income ones (see Table 1).[35] Considering these stylised facts, re-estimating the model (1) separately for the two sub-periods may help to better scrutinise the results presented on the effects of inequality on the probability of holding a debt. Tables 6 and 7 show that evidence is stronger during the crisis period. In particular, Columns (1) to (3) of Table 6 show that during the pre-crisis period inequality still negatively affected the probability of holding a debt, even though no significant asymmetric effect is found across different income groups. On the other hand, Columns (1) to (3) of Table 7 document that the results for the crisis period are almost identical to those found for the whole timespan analysed (cf. Table 3).

It is easier to interpret the different patterns observed between the two periods by considering the behaviour for the different types of debt. Columns (4) and (5) of Table 6 show that before the crisis the differentiated effect of local inequality was fully in place for consumer loans while mortgages were completely unaffected by it. In contrast, the last two columns of Table 7 document how the asymmetric effect

[35]For a more in depth analysis of the recent evolution of credit market participation in Italy see Magri and Pico (2014).

TABLE 5
DIFFERENT ESTIMATES OF PR(LOAN > 0)—DECOMPOSITION BY TYPE OF DEBT

| | Mortgages | | | Consumer Loans | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | LPM | Pooled Probit | RE Panel Probit | LPM | Pooled Probit | RE Panel Probit |
| Income Decile | 0.01*** | 0.01*** | 0.00** | 0.01*** | 0.01*** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Gini | −0.49*** | −0.54*** | −0.11 | −0.19 | −0.18 | −0.82*** |
| | (0.18) | (0.18) | (0.08) | (0.18) | (0.18) | (0.25) |
| Income Decile*Gini | 0.12** | 0.16*** | 0.03* | 0.16*** | 0.15*** | 0.18*** |
| | (0.05) | (0.05) | (0.02) | (0.05) | (0.05) | (0.07) |
| Equivalised Income | 0.00** | 0.00 | 0.00** | −0.00 | −0.00 | 0.00 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Household controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 29282 | 29282 | 7762 | 29282 | 29282 | 7762 |
| $R^2$/Pseudo $R^2$ | 0.06 | 0.09 | | 0.04 | 0.05 | |
| Predicted Probability | | 0.13 | 0.01 | | 0.14 | 0.09 |
| LR Test (*p*-value) | | | 0.00 | | | 0.00 |

*Note*: The Table reports marginal effects (and associated standard errors) from the following three binary regressions used to model the probability of having, respectively, outstanding mortgage or consumer loans: (1) and (4) a linear probability model, estimated via OLS, (2) and (5) a nonlinear probability model, estimated via pooled probit, and (3) and (6) a nonlinear probability model, estimated via panel probit random effect estimation. Marginal effects are expressed at the mean value of the independent variables and, in the case of interaction terms in nonlinear models, are computed by taking into account cross derivatives (Ai and Norton, 2003). A likelihood-ratio test on the significance of the panel level variance component is included at the bottom of the output under the null that the panel-level variance component is unimportant, and the panel estimator is not different from the pooled estimator. Standard errors are corrected for heteroscedasticity and (in the pooled estimations) clustered at the household level. ***, **, and * denote significance at 1%, 5%, and 10%, respectively.

TABLE 6
Different estimates of Pr(loan > 0)—sample split: pre-crisis period (2004–2008)

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Total Debt | | | Mortgages | Consumer Loans |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Income Decile | 0.02*** | 0.02*** | 0.02*** | 0.01** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Gini | −0.76*** | −0.85*** | −1.45*** | −0.14* | −0.72*** |
| | (0.30) | (0.31) | (0.41) | (0.08) | (0.27) |
| Income Decile*Gini | 0.12 | 0.15 | 0.17 | 0.01 | 0.22*** |
| | (0.09) | (0.09) | (0.11) | (0.02) | (0.07) |
| Household controls | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 17794 | 17794 | 7113 | 7113 | 7113 |
| $R^2$/Pseudo $R^2$ | 0.08 | 0.08 | | | |
| Predicted Probability | | 0.26 | 0.16 | 0.01 | 0.09 |
| LR Test ($p$-value) | | | 0.00 | 0.00 | 0.00 |

*Note*: See Table 3.

TABLE 7
DIFFERENT ESTIMATES OF PR(LOAN > 0)—SAMPLE SPLIT: CRISIS PERIOD (2010–2012)

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Total Debt | | | Mortgages | Consumer Loans |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Income Decile | 0.02*** | 0.02*** | 0.01*** | 0.00** | 0.01** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Gini | −0.53 | −0.51 | −0.76** | −0.07* | 0.41** |
| | (0.33) | (0.34) | (0.31) | (0.03) | (0.17) |
| Income Decile*Gini | 0.25*** | 0.25*** | 0.25*** | 0.02** | 0.09** |
| | (0.08) | (0.08) | (0.08) | (0.07) | (0.04) |
| Household controls | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 11488 | 11488 | 6496 | 6496 | 6496 |
| $R^2$/Pseudo $R^2$ | 0.07 | 0.07 | | | |
| Predicted Probability | | 0.242 | 0.098 | 0.002 | 0.049 |
| LR Test (*p*-value) | | | 0.000 | 0.000 | 0.000 |

*Note*: See Table 3.

of inequality was fully common to both types of debt during the crisis period. In summary, it can be said that evidence found in Section 5.1 is driven by both forms of debt during the crisis period and only by the consumer credit market before the crisis. Overall, these results are in line with the idea that local inequality affects the likelihood of debt through the credit supply channel. In fact, before the crisis, when credit conditions were laxer, local inequality may not have affected the mortgage market because of the collateralised nature of this form of debt. Despite mortgages representing much larger loan amounts than other forms of debt, in case of default banks may take possession and sell the secured property. Before the crisis, the housing market was in an expanding phase with rising prices and it was easier for lenders to sell property (in case of foreclosure) at a value sufficient to cover the remaining principal of a loan. Consumer loans, on the other hand, are mostly uncollateralised (except for auto loans) and so may be considered riskier by banks, even though are much smaller in size. In the wake of the crisis, the tensions in the banks' funding availability and cost, the adverse shocks to household income, and a depressed housing market led first to a tightening of credit standards, then to a reduction in lending to Italian households. Hence, banks' incentive to devote resources toward identifying applicants underlying creditworthiness should have strengthened, leading to a wider utilization of the information provided by local income inequality as found in Table 7.

## 7. Identification Issues and Sensitivity Analysis

Is it possible that the evidence shown so far is affected by some source of bias? There are at least two reasons why this may be the case. One potential source of bias may arise from simultaneity if the key variable (Income Decile X Gini) is jointly determined with the dependent variable. In addition to this, the association between the probability of being indebted and local inequality is subject to omitted variables concerns. In this section these sources of bias potentially affecting results will be addressed. Furthermore, the last subsection is devoted to examining the sensitivity of findings to a variety of different issues.

### 7.1. *Endogeneity Issues*

In the specification adopted the key variable is the interaction between Income decile and Gini coefficient. Endogeneity is a concern only if it affects one of these two variables. The other variables are controls allowing reduction of omitted variable bias on the interaction term.

Having said that, one may be still worried about the endogeneity of the Income Decile and Gini coefficient (and consequently of the interaction term). Regarding the Income Decile, the equivalised income variable (the basis upon which Income Deciles are computed) is built to avoid any endogeneity problem. In fact, I have adopted a definition of monetary disposable income that is net of imputed rents and gross of negative interests. Hence households' equivalised incomes are neither increased because they own a house (bought through a mortgage) nor reduced because of mortgage payments or other forms of debts. It follows that holding a debt does not affect the household equivalised income and, consequently, its

Income Decile. To reduce any residual concern of simultaneity bias due to contemporaneous household-level controls, Table 8 shows that results are qualitatively unaffected when two-year[36] lagged household-level controls are replaced into the baseline specification.

The Gini coefficient may be affected by endogeneity for two reasons: simultaneity and omitted variable bias. To alleviate concerns of simultaneity, similarly to Coibion *et al.* (2014), I consider a specification in which the measure of local income inequality as well as all other controls are fixed at the values of the beginning period (2004). This removes potential bias due to contemporaneous controls that may be influenced by the treatment effect. Moreover, such a specification can be interpreted as a "difference-in-differences" approach across income groups and regional inequality levels with the coefficient of the interaction between local income inequality and household Income Decile being the key parameter that determines whether such differences have been important. This change does not have a significant impact on the results (see Table 9).

The omitted variable problem in equation (1) is that there might be other location-specific characteristics that are related to local income inequality and, at the same time, with the probability of holding a debt. As it is uncertain that all the location-specific covariates included in the estimations so far are enough to capture all relevant heterogeneity, I assess the stability of the coefficients of interest to the inclusion of these covariates. As explained by Altonji *et al.* (2005), if selection on observables is related to selection on unobservables, then the changes in the coefficient are informative about the bias induced by excluding relevant characteristics. However, to understand whether these changes are large, one needs to look at the explanatory power of the additional covariates, as captured by the $R^2$. To provide a quantitative assessment, I exploit Oster (2017) suggestion to calculate a bound for the coefficient of interest, by assuming that selection on observables is directly proportional to selection on unobservables and that the maximum $R^2$ that could be explained including also all other (unobservable) relevant features would be 1.3 times the one we observe.[37] To be cautious I additionally adopt a more conservative approach to $R_{\max}$ used in González and Miguel (2015).

Accordingly, in Table 10 I adjust coefficients reported in Column 1 of Table 3 using these two different approaches: (1) the Gonzalez and Miguel approach ($R_{\max} = 2R* - R$), (2) the Oster approach ($R_{max} = \min\{1.3R*, 1\}$). Each of these approaches gives different estimates of $R_{\max}$ and, therefore, a different estimate of the bounds. I do this for each possible combination of the key variables potentially

[36]It should be recalled that SHIW survey is carried out biannually so that lagged household controls referring to previous SHIW wave are two-year lagged.

[37]In a recent paper, Oster (2017) shows that the following is a consistent estimator of the effect of $\gamma$ on the outcome variable:

$$\hat{\hat{\gamma}} = \hat{\gamma}* - (\hat{\gamma} - \hat{\gamma}*) \times \frac{R_{max} - R^*}{R^* - R}, \tag{4}$$

where $\hat{\gamma}*$ and $R^*$ are the coefficient estimate and $R^2$ from the regression including observable covariates, and $\hat{\gamma}$ and R are the coefficient and $R^2$ from the uncontrolled regression. In addition, $R_{\max}$ is the $R^2$ in a regression of the outcome variable on all observable and unobservable controls, which is clearly unknowable (given its reliance on unobservables).

TABLE 8
Different estimates of Pr(loan > 0)—specification with fully lagged controls

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Total Debt | | | Mortgages | Consumer Loans | Mortgages | Consumer Loans |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit | Pooled Probit | Pooled Probit |
| Income Decile | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.01* |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Gini | −0.73* | −0.76* | −0.83* | −0.09 | −0.40 | −0.56 | −0.41 |
| | (0.44) | (0.45) | (0.45) | (0.07) | (0.28) | (0.37) | (0.31) |
| Income Decile*Gini | 0.25** | 0.27** | 0.37*** | 0.03* | 0.20** | 0.19* | 0.11 |
| | (0.12) | (0.12) | (0.12) | (0.02) | (0.08) | (0.10) | (0.08) |
| Equivalised Income | 0.00* | 0.00* | 0.00 | 0.00 | −0.00 | 0.00*** | −0.00 |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Household controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 12244 | 12244 | 5946 | 5946 | 5946 | 12244 | 12244 |
| $R^2$/Pseudo $R^2$ | 0.06 | 0.05 | | | | 0.09 | 0.04 |
| Predicted Probability | | 0.260 | 0.145 | 0.007 | 0.082 | 0.125 | 0.144 |
| LR Test ($p$-value) | | | 0.000 | 0.000 | 0.000 | | |

*Note*: See Table 3. To avoid potential simultaneity bias due to the contemporaneous household-level controls, in this specification household-level controls are replaced with their value at previous SHIW wave (i.e. they are two-year lagged).

TABLE 9

Different estimates of Pr(loan > 0)—all controls from year = 2004

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Total Debt | | | Mortgages | | Consumer Loans | |
| | LPM | Pooled Probit | RE Panel Probit | Pooled Probit | RE Panel Probit | Pooled Probit | RE Panel Probit |
| Income Decile | 0.01*** | 0.01*** | 0.02** | 0.01*** | 0.01*** | 0.01*** | 0.01 |
| | (0.00) | (0.00) | (0.01) | (0.00) | (0.00) | (0.00) | (0.01) |
| Gini | −0.41 | −0.48 | −0.58 | −0.57* | −0.07 | −0.04 | −0.36 |
| | (0.37) | (0.39) | (0.47) | (0.33) | (0.10) | (0.27) | (0.27) |
| Income Decile*Gini | 0.25** | 0.27*** | 0.52* | 0.15* | 0.10 | 0.21*** | 0.35* |
| | (0.10) | (0.11) | (0.32) | (0.09) | (0.07) | (0.07) | (0.19) |
| Household controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Location controls | No | No | No | No | No | No | No |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 14427 | 14427 | 7755 | 14427 | 7755 | 14427 | 7755 |
| $R^2$/Pseudo $R^2$ | 0.07 | 0.07 | | 0.07 | | 0.05 | |
| Predicted Probability | | 0.252 | 0.158 | 0.123 | 0.126 | 0.143 | 0.089 |
| LR Test (*p*-value) | | | 0.000 | | 0.000 | | 0.000 |

*Note*: See Table 3. In order to avoid treatment effect in this specification the Gini coefficients and all household-specific controls are from the beginning of the period of analysis (2004).

TABLE 10
DIFFERENT ESTIMATES OF THE BOUNDS ON THE POTENTIAL BIAS DUE TO UNOBSERVABLES

| | (1) | (2) | (3) González and Miguel (2015) | (4) Oster (2017) approach |
|---|---|---|---|---|
| *Panel A* | | | | |
| Income Decile*Gini | 0.15** | 0.18*** | [0.18, 0.21] | [0.18, 0.19] |
| | (0.06) | (0.06) | | |
| R2 | 0.001 | 0.075 | | |
| Rmax | | | 0.149 | 0.098 |
| *Panel B* | | | | |
| Income Decile | 0.01*** | 0.02*** | | |
| | (0.00) | (0.00) | | |
| Gini | −0.63*** | −0.51** | | |
| | (0.19) | (0.22) | | |
| Income Decile*Gini | 0.10 | 0.18*** | [0.18, 0.26] | [0.18, 0.21] |
| | (0.06) | (0.06) | | |
| $R^2$ | 0.008 | 0.075 | | |
| Rmax | | | 0.143 | 0.098 |
| *Panel C* | | | | |
| Gini | −0.66*** | −0.51** | [−0.51, −0.36] | [−0.51, −0.46] |
| | (0.19) | (0.22) | | |
| $R^2$ | 0.002 | 0.075 | | |
| *R*max | | | 0.149 | 0.098 |
| Controls | No | Yes | | |
| Area and Year FE | Yes | Yes | | |
| N. of observations | 29282 | 29282 | | |

*Note*: The dependent variable is a binary ownership indicator of household debt. Set intervals estimation using an equal proportional selection assumption. González and Miguel's approach uses $R\text{max} = 2R^*-R$; Oster's approach uses $R\text{max} = 1.3R^*$, where $R$ and $R^*$ are the $R^2$ from column (1) and (2) respectively. Controls include all the variables reported in Table 3 Column (1). Heteroskedasticity robust standard errors clustered at household level in parenthesis with ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

affected by unobserved heterogeneity, namely, the interaction term (Income Decile X Gini) alone (the top panel of Table 10); the interaction term together with main effects (the middle panel); and the Gini coefficient alone (the bottom panel). Building on Oster (2017), the result of each extension is an interval of values for the Income Decile X Gini (or Gini in isolation) coefficient that are consistent with the degree of omitted variable bias accommodated by the $R_{max}$ value, the difference between the $R^2$ in the uncontrolled and controlled regression, and the change in the estimated coefficient across those specifications. Specifically, Columns 1 and 2 in each panel present results from the uncontrolled and the controlled regressions, respectively, and the corresponding $R^2$ values. The increase in the $R^2$ across these two columns captures the amount of variation in the dependent variable that is explained by the observed covariates. This, together with the maximum amount of variation that can be potentially explained by $R_{max}$, gives rise to an estimate of the upper bound for the interval of coefficients. If this upper bound is greater than zero, it would provide further evidence that the true causal effect of the interaction term on the debt outcome is indeed likely to be positive. It goes without saying that, to confirm that the true effect of inequality on likelihood of debt is negative, the upper bound must be negative when Gini coefficient is taken in isolation.

I report these bounds in the last two columns of Table 10. They confirm that the true effect of the interaction term is positive, and possibly even larger than estimates in Table 3. The intervals of coefficient values in the case of Gini coefficient are also consistent with previous estimates, even though the negative effect of local inequality appears slightly smaller. In my view, these results make it less likely that the estimated effects are fully driven by omitted variable bias.

Overall the analysis shows that, by exploiting lagged controls specifications and the á la Altonji approach, it is possible to provide robust results with respect to endogeneity concerns.

### 7.2. *Additional Robustness Checks*

In this Section, several additional robustness checks are considered. First of all, in order to account for any macro region-specific time trends that may influence household borrowing decisions, a specification with a full set of area-year dummies is tested. Results reported in Table 11 confirm the robustness of baseline findings.

As is quite standard in the household finance literature, in my modeling strategy, households decide every period on the allocation of their resources and on borrowing. However, one may argue that for many households with mortgages outstanding in a given period, the decision to take up such loans was made many years prior to the interview. To examine the sensitivity of results to this issue, I have re-estimated the panel model for mortgages, focusing only on households that take up such loans (i.e. with switch borrowing status) during the period covered by the data. Specifically, I use the sample of households without outstanding mortgages in 2004 (i.e. the initial observation period in the sample) and estimate the probability of taking up such a loan in any of the subsequent four waves. This panel model conditions on the same set of covariates as the ones used in the baseline specification (presented in Table 3). The estimated effects of local inequality on the likelihood of taking up a mortgage from this "inflow" sample are still significant, albeit at 10 percent confidence level, and have the same economic interpretation of the baseline specification (see Table 12).

Standard errors of all the pooled estimates are clustered at household level in order to deal with the panel component of the data and correct for serial correlation. As said in Section 3, the SHIW sampling is in two stages: first municipalities are chosen from different strata and then households are selected at random. It follows that observations are independent across municipalities but there could be still within-municipality error correlation (e.g. because of neighbourhood effects) affecting standard errors and, on top of that, unobserved provincial or regional effects. However, including fixed effects at different geographical level of aggregation generally does not control for all the within-cluster correlation of the error and one should still use the cluster-robust estimate of the variance matrix . Correcting standard errors for clustering on municipality, province, and region, results do not change (for the sake of brevity, I do not report relevant Tables).

The Gini coefficient is sometimes criticised as being too sensitive to relative changes around the middle of the income distribution. Tables 13 and 14 show that the choice of Gini coefficient as income inequality indicator is unlikely to influence

TABLE 11
Different estimates of Pr(loan > 0)—specification with interaction Macro-region × Year FE

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | | Total Debt | | Mortgages | Consumer Loans |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Income Decile | 0.02*** | 0.02*** | 0.02*** | 0.00** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Gini | −0.60*** | −0.64*** | −1.04*** | −0.11 | −0.70*** |
| | (0.23) | (0.24) | (0.39) | (0.07) | (0.24) |
| Income Decile*Gini | 0.18*** | 0.19*** | 0.28*** | 0.03* | 0.17*** |
| | (0.06) | (0.06) | (0.10) | (0.02) | (0.07) |
| Household controls | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| Macro-region × Year FE | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 29282 | 29282 | 7762 | 7762 | 7762 |
| $R^2$/Pseudo $R^2$ | 0.08 | 0.07 | | | |
| Predicted Probability | | 0.253 | 0.152 | 0.011 | 0.083 |
| LR Test (*p*-value) | | | 0.000 | 0.000 | 0.000 |

*Note*: See Table 3. With respect to the baseline scenario, this specification includes the interaction between macro-region and time fixed effects to account for any location-specific time trends that may influence individual borrowing decisions.

TABLE 12
Estimates of Pr(Mortgage > 0)—only households not indebted in 2004

| | (1) |
|---|---|
| | RE Panel Probit |
| **Mortgage** | |
| Income Decile | 0.08** |
| | (0.04) |
| Gini | −4.90 |
| | (3.31) |
| Income Decile*Gini | 1.66* |
| | (0.86) |
| Equivalised Income | 0.02* |
| | (0.01) |
| Household controls | Yes |
| Location controls | Yes |
| Macro-region FE | Yes |
| Year FE | Yes |
| N. of observations | 5034 |
| LR Test (*p*-value) | 0.000 |

*Note:* See Table 3. In this specification only households that were not indebted at the beginning of the period of analysis (2004) are considered. In this way the risk of considering decisions taken many years prior the interview is reduced.

the results. Table 13 uses as income inequality measure the Theil Index, which is more sensitive than Gini coefficient to changes that affect the upper tail of the distribution, while Table 14 uses the mean logarithmic deviation which is more sensitive to changes in the lower tail. In all cases results are very similar to the ones of the baseline scenario, supporting the robustness of the findings under varying inequality measures. Furthermore, I have tested a functional form that employs quintiles, instead of deciles, to model the households' income ranks and the results are insensitive even to such a transformation.

## 8. Conclusion

The literature on household finance has only recently paid attention to the distribution of access to finance. Using Italian household-level survey data, this paper empirically explored the relationship between household debt and local income inequality, highlighting the role that belonging to different income groups plays in mediating this relationship. The analysis provides evidence that income inequality negatively affects the probability of being indebted. Moreover, richer households living in highly unequal regions are relatively more likely to be indebted than richer households situated in low-inequality regions (and vice versa for poorer ones). The work also tested alternative views about the prevalence of demand or supply factors in shaping the interaction between inequality and household debt (namely the "keeping up with the Joneses" hypothesis versus the "signaling channel" one). In fact, local inequality does not seem to affect the likelihood to apply for a loan while greater inequality decreases the probability of loan application refusal for top income households (and increases for the poorer ones). Such results

TABLE 13
DIFFERENT ESTIMATES OF PR(LOAN > 0)—DIFFERENT MEASURES OF INEQUALITY: THEIL INDEX

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | | Total Debt | | Mortgages | Consumer Loans |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Income Decile | 0.02*** | 0.02*** | 0.01*** | 0.01** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Theil Index | −0.43*** | −0.47*** | −0.59*** | −0.11** | −0.36** |
| | (0.13) | (0.14) | (0.23) | (0.05) | (0.11) |
| Income Decile *Theil Index | 0.10*** | 0.12*** | 0.13* | 0.01 | 0.09* |
| | (0.04) | (0.04) | (0.07) | (0.01) | (0.05) |
| Household controls | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 29282 | 29282 | 7762 | 7762 | 7762 |
| $R^2$/Pseudo $R^2$ | 0.08 | 0.07 | | | |
| Predicted Probability | | 0.254 | 0.154 | 0.012 | 0.086 |
| LR Test (*p*-value) | | | 0.000 | 0.000 | 0.000 |

*Note*: The Table reproduces the results in Table 3 using the Theil Index measure of inequality rather than the Gini coefficient.

TABLE 14

Different estimates of Pr(loan > 0)—different measures of inequality: Mean Logarithmic Deviation

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | | Total Debt | | Mortgages | Consumer Loans |
| | LPM | Pooled Probit | RE Panel Probit | RE Panel Probit | RE Panel Probit |
| Income Decile | 0.02*** | 0.02*** | 0.02*** | 0.01** | 0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Mean Logarithmic Deviation | −0.45*** | −0.47*** | −0.65** | −0.10* | −0.45** |
| | (0.17) | (0.18) | (0.29) | (0.06) | (0.19) |
| Income Decile*Mean Log. Dev. | 0.14*** | 0.15*** | 0.27*** | 0.03** | 0.15*** |
| | (0.04) | (0.05) | (0.08) | (0.02) | (0.05) |
| Household controls | Yes | Yes | Yes | Yes | Yes |
| Location controls | Yes | Yes | Yes | Yes | Yes |
| Macro-region FE | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes |
| N. of observations | 29282 | 29282 | 7762 | 7762 | 29282 |
| $R^2$/Pseudo $R^2$ | 0.08 | 0.07 | | | |
| Predicted Probability | | 0.254 | 0.154 | 0.012 | 0.086 |
| LR Test (*p*-value) | | | 0.000 | 0.000 | 0.000 |

*Note*: The Table reproduces the results in Table 3 using the Mean Logarithmic Deviation measure of inequality rather than the Gini coefficient.

are consistent with models in which banks use an applicant's position in the local income distribution, along with the dispersion of that distribution, to make inferences about default risk; and are in line with the most recent survey-based evidence on U.S. data according to which supply factors are more important than demand ones in explaining the mentioned result. These findings persist after controlling for socio-demographic differences, different types of debt, unobserved individual heterogeneity thanks to panel data, sample selection, endogeneity concerns, and several robustness checks. Comparison between pre-crisis and crisis period indicates that, in line with the supply-side interpretation of the results, evidence is stronger during the latter one, when credit supply conditions were particularly tight. This paper suggests that household income may be considered a stronger signal of creditworthiness in highly unequal regions because higher inequality implies less income mobility over time and a skewed access to investment opportunities and/or political influence. It follows that banks, which screen borrowers considering their capacity to meet their obligations in the future, are less prone to grant credit to poor households located in more unequal regions. In conclusion, inequality can become self-sustained as it produces unequal access to finance and ultimately unequal opportunities, which can reinforce any initial economic inequality.

## References


Acciari, P. and S. Mocetti, "The Geography of Income Inequality in Italy," Questioni di Economia e Finanza (Occasional Papers) 208, Bank of Italy, Economic Research and International Relations Area, 2013.

Acemoglu, D., S. Johnson, and J. A. Robinson, "Institutions as a Fundamental Cause of Long-Run Growth," in P. Aghion and S. Durlauf (eds), *Handbook of Economic Growth, volume 1 of Handbook of Economic Growth*, Elsevier, Amsterdam, Netherlands, chapter 6, 385–472, 2005.

Ai, C., and E. C. Norton, "Interaction Terms in Logit and Probit Models," *Economics Letters*, 80, 123–29, 2003.

Altonji, J., T. Elder, and C. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113, 151–84, 2005.

Barro, R. J., "Economic Growth in a Cross Section of Countries," *The Quarterly Journal of Economics*, 106, 407–43, 1991.

Bertrand, M. and A. Morse, "Trickle-down consumption." Working Paper 18883, National Bureau of Economic Research, 2013.

Besanko, D., and A. V. Thakor, "Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets," *International Economic Review*, 28, 671–89, 1987.

Bester, H., "Screening vs. Rationing in Credit Markets with Imperfect Information," *American Economic Review*, 75, 850–55, 1985.

Brambor, T., W. R. Clark, and M. Golder, "Understanding Interaction Models: Improving Empirical Analyses," *Political Analysis*, 14, 63–82, 2006.

Brandolini, A. and L. Cannari, "Methodological Appendix: The Bank of Italy's Survey of Household Income and Wealth," in A. Ando, L. Guiso, and I. Visco (eds), *Saving and the Accumulation of Wealth: Essays on Italian Household and Government Behaviour*. Cambridge University Press, Cambridge, 1994.

Bricker, J., R. Ramcharan, and J. Krimmel, "Signaling Status: The Impact of Relative Income on Household Consumption and Financial Decisions. Finance and Economics Discussion Series 2014–76," Board of Governors of the Federal Reserve System (U.S.), 2014.

Canberra Group, (2011). *The Canberra Group Handbook on Household Income Statistics*. United Nations, Geneva, second edition, 2011.

Christelis, D., M. Ehrmann, and D. Georgarakos, Exploring Differences in Household Debt Across Euro Area Countries and the United States. Working Papers 15-16, Bank of Canada, 2015.

Claessens, S., and E. Perotti, "Finance and Inequality: Channels and Evidence," *Journal of Comparative Economics*, 35, 748–73, 2007.

Coibion, O., Y. Gorodnichenko, M. Kudlyak, and J. Mondragon, Does Greater Inequality Lead to More Household Borrowing? New Evidence from Household Data. NBER Working Papers 19850, National Bureau of Economic Research, Inc, 2014.
Coletta, M., R. De Bonis, and S. Piermattei, The determinants of household debt: a cross-country analysis. Temi di discussione (Economic working papers) 989, Bank of Italy, Economic Research and International Relations Area, 2014.
Corak, M., "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 27, 79–102, 2013.
Crook, J. and S. Hochguertel, Household debt and credit constraints: Comparative micro evidence from four OECD countries. Working Paper 05, University of Edinburgh, 2006.
Dabla-Norris, E., K. Kochhar, N. Suphaphiphat, F. Ricka, and E. Tsounta, Causes and Consequences of Income Inequality: A Global Perspective. IMF Staff Discussion Notes 15/13, International Monetary Fund, 2015.
Del Prete, S., M. Pagnini, P. Rossi, and V. P. Vacca, Getting organized to lend in a period of crisis: findings from a survey of Italian banks. Questioni di Economia e Finanza (Occasional Papers) 154, Bank of Italy, Economic Research and International Relations Area, 2013.
Denk, O. and A. Cazenave-Lacroutz, Household finance and income inequality in the euro area. OECD Economics Department Working Papers 1226, OECD, 2015.
Duesenberry, J. S. *Income, Saving and the Theory of Consumer Behavior*. Harvard University Press, Cambridge, MA, 1949.
Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri, "Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality," *American Economic Review*, 106, 651–55, 2016.
Faiella, I. and R. Gambacorta, The Weighting Process in the SHIW. Temi di discussione (Economic working papers) 636, Bank of Italy, Economic Research and International Relations Area, 2007.
Fitoussi, J. P., F. Saraceno, "Europe: How Deep Is a Crisis? Policy Responses and Structural Factors Behind Diverging Performances," *Journal of Globalization and Development*, 1, 1–19, 2010.
Frank, R. H., A. S. Levine, and O. Dijk, "Expenditure Cascades," *Review of Behavioral Economics*, 1, 55–73, 2014.
Galor, O., and J. Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60, 35–52, 1993.
Gambacorta, R., G. Ilardi, A. Locatelli, C. Rampazzi, and R. Pico, Main Results of the Household Finance and Consumption Survey: Italy in the International Context. Questioni di Economia e Finanza (Occasional Papers) 161, Bank of Italy, Economic Research and International Relations Area, 2013.
Gennaioli, N., R. La Porta, F. Lopez De Silanes, and A. Shleifer, "Growth in Regions," *Journal of Economic Growth*, 19, 259–309, 2014.
Georgarakos, D., M. Haliassos, and G. Pasini, "Household Debt and Social Interactions," *Review of Financial Studies*, 27, 1404–33, 2014.
Gobbi, G., and F. Lotti, "Entry Decisions and Adverse Selection: An Empirical Analysis of Local Credit Markets," *Journal of Financial Services Research*, 26, 225–44, 2004.
González, F. and E. Miguel, "War and Local Collective Action in Sierra Leone: A Comment on the Use of Coefficient Stability Approaches," *Journal of Public Economics*, 128, 30–33, 2015.
Greene, W., "Testing Hypotheses about Interaction Terms in Nonlinear Models," *Economics Letters*, 107, 291–96, 2010.
Greene, W. H., *Econometric Analysis*. Pearson, Boston, 7th edition, 2012.
Guiso, L. and T. Jappelli, Stockholding in Italy. CSEF Working Papers 82, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy, 2002.
Hirschman, A. O., M. Rothschild, "The Changing Tolerance for Income Inequality in the Course of Economic Development: With a Mathematical Appendix," *The Quarterly Journal of Economics*, 87, 544–66, 1973.
Hodge, A., S. Shankar, "Partial Effects in Ordered Response Models with Factor Variables," *Econometric Reviews*, 33, 854–68, 2014.
Iacoviello, M., "Household Debt and Income Inequality, 1963–2003," *Journal of Money, Credit and Banking*, 40, 929–65, 2008.
Jappelli, T., and L. Pistaferri, "Does Consumption Inequality Track Income Inequality in Italy?," *Review of Economic Dynamics*, 13, 133–53, 2010.
Kopczuk, W., E. Saez, and J. Song, "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937," *The Quarterly Journal of Economics*, 125, 91–128, 2010.
Krueger, D., and F. Perri, "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 73, 163–93, 2006.
Kumhof, M., R. Ranciere, and P. Winant, Inequality, Leverage and Crises: The Case of Endogenous Default. IMF Working Papers 13/249, International Monetary Fund, 2013.

Magri, S., "Italian Households' Debt: The Participation to the Debt Market and the Size of the Loan," *Empirical Economics*, 33, 401–26, 2007.
Magri, S. and R. Pico, "The Household Credit Market After Five Years of Crisis: Evidence From the Survey on Income and Wealth," Questioni di Economia e Finanza (Occasional Papers) 241, Bank of Italy, Economic Research and International Relations Area, 2014.
Mian, A., and A. Sufi, *House of Debt Economics Books*. University of Chicago Press,Chicago, IL, 2015.
Michelangeli, V. and E. Sette, How Does Bank Capital Affect the Supply of Mortgages? Evidence from a Randomized Experiment. Temi di discussione (Economic working papers) 1051, Bank of Italy, Economic Research and International Relations Area, 2016.
Mundlak, Y., "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85, 1978.
Oster, E., "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 35, 1–18, 2017.
Ostry, J. D., A. Berg, and C. G. Tsangarides, Redistribution, Inequality, and Growth. IMF Staff Discussion Notes 14/02, International Monetary Fund, 2014.
Piketty, T., "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies*, 64, 173–89, 1997.
______, *Capital in the Twenty-First Century*. Belknap of Harvard University Press, Cambridge, MA, 2014.
Porta, R. L., F. L. de Silanes, A. Shleifer, and R. W. Vishny, "Law and Finance," *Journal of Political Economy*, 106, 1113–55, 1998.
Rajan, R., *Fault Lines: How Hidden Fractures Still Threaten The World Economy*. Princeton University Press, Princeton, NJ, 2010.
Saez, E., and G. Zucman, "Wealth Inequality in the United States Since 1913: Evidence from Capitalized Income Tax Data*," *The Quarterly Journal of Economics*, 2016.
Senik, C., "Ambition and Jealousy: Income Interactions in the 'Old' Europe versus the 'New,'' *Europe and the United States", Economica*, 75, 495–513, 2008.
Sierminska, E. and M. Medgyesi, "The Distribution of Wealth between Households," European Commission Research Note 11, European Commission, 2013.
Stiglitz, J. E. *The Price of Inequality: How Today's Divided Society Endangers Our Future*. W. W. Norton & Company, 2012.
Stiglitz, J. E., and A. Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71, 393–410, 1981.
Summers, L. H., "U.S. Economic Prospects: Secular Stagnation, Hysteresis, and the Zero Lower Bound," *Business Economics*, 49, 65–73, 2014.
Veblen, T., *The Theory of the Leisure Class. An Economic Study of Institutions*. The Macmillian Company, New York, 1899.

## Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

**Figure A1:** Quartiles of equivalised income (1).
*Source*: Eu-Silc and SHIW. (1) Monetary disposable income.
**Figure A2:** Inequality indices (1).
*Source*: Eu-Silc and SHIW. (1) Monetary disposable equivalised income.
**Table A1:** Summary statistics.
*Notes:* The sample is restricted to the households with 20–70 year old head of household. The statistics are calculated using sampling weights. The Table shows the statistics from the sample restricted to observations with positive equivalised income. The sample is further restricted to remove outliers. See text for more details. Total debt is the sum of Mortgages and Consumer credit. The number of observations in Panel A is 29,282 from the waves 2004, 2006, 2008, 2010, 2012 of SHIW. The number of observations in Panel B is 75,177 from the waves 2004, 2006, 2008, 2010, 2012 of EU-SICL.