Author(s): Hyytinen, Ari; Putkuri, Hanna

Title: Household optimism and overborrowing

Year: 2018

Version:

Please cite the original version:

Hyytinen, A., & Putkuri, H. (2018). Household optimism and overborrowing. Journal of Money, Credit and Banking, 50(1), 55-76. https://doi.org/10.1111/jmcb.12453

To appear:

Journal of Money, Credit and Banking

Household optimism and overborrowing*

Ari Hyytinen
University of Jyväskylä, School of Business and Economics

Hanna Putkuri
Bank of Finland

February 21, 2017

Abstract

We use Finnish household-level data from 1994 to 2013 to measure how often and what kind of forecast errors households make and how the errors are linked to the households' borrowing behaviour and overindebtedness. We find that those households that make the largest optimistic forecast errors have greater debt-to-income ratios. They also are more likely to report that they suffer from excessive debt loads and have problems in coping with their bills. There are no such systematic effects for the households that make pessimistic forecast errors.

JEL: D21, L20

Key words: forecast errors, borrowing, overindebtedness

*We would like to thank two anonymous referees, the editor (Pok-sang Lam), seminar participants at The Annual Meeting of Finnish Economic Association (2012), Finnish HECER (2012), Bank of Finland (2012) and ESEM-EEA (Oslo 2011), as well as Iftekhar Hasan, Panu Kalmi, Markku Kaustia, Matti Keloharju, William Kerr, Karlo Kauko, Steven Ongena, Jouko Vilmunen, and George Waters for helpful comments and discussions. This work was initiated when Hyytinen was visiting Bank of Finland, whose hospitality is greatefully acknowledged. Corresponding author: Ari Hyytinen, School of Business and Economics, University of Jyväskylä, P.O. Box 35, 40014 University of Jyväskylä, Finland. Email: ari.t.hyytinen@jyu.fi.

1 Introduction

Conventional view in economics is that households use unsecured (Sullivan 2008) and secured (Hurst and Stafford 2004) debt to optimally smooth their consumption over time (see, e.g., Alessie and Lusardi 1997).[1] On the other hand, there is evidence that households may borrow imprudently and accumulate too much debt. These behaviours appear to be related to behavioral biases and poor debt literacy (see, e.g., Stango and Zinman 2009, Campbell, Jackson, Madrian and Tufano 2011, and Lusardi and Tufano 2015). This paper shows that households' indebtedness and unsustainable levels of debt are strongly linked to optimistic forecast errors that the households make.

Our data are a rotating, nationally representative panel that covers a large number of Finnish households over a nineteen-year-period from 1994 to 2013. These data are particularly suitable for an analysis of the relation between households' forecast errors and their indebtedness. First, the data allow us to measure at the level of a household how its financial expectations are related to the subsequent realizations. We use the difference between the subjective expectations and the realizations as a direct indicator of the size and nature of households' forecast errors. Second, our forecast error data are comparable to those used elsewhere in the literature. The survey questions on which our measures rely are nearly identical to those used e.g. by Souleles (2004), underlying the Index of Consumer Sentiment in the US. Third, unlike the other data sets available to date, the Finnish data allow us to directly link households' forecast errors to the same households' borrowing behaviour and perceived overindebtedness.[2]

[1] It also is possible that households are unable to borrow as much as they need (see, e.g., Zeldes 1989). For recent reviews of the relevant literature, see Jappelli and Pistaferri (2010), and Attanasio and Weber (2010).

[2] E.g., Jappelli and Pistaferri (2010) note the lack of data that allow implementing such an

We say that an optimistic forecast error is observed when the expected outcome is better than the realized and that a pessimistic forecast error occurs when the opposite happens. We find that on average, households have neither made optimistic nor pessimistic forecast errors over the sample period. The households making the largest optimistic forecast errors have higher debt-to-income ratios. This result is robust to controlling flexibly for age, cohort and time effects, demographics, disposable income as well as for the stock of debt cumulated in the past. We show, moreover, that the households that make the largest optimistic forecast errors also are more likely than others to report that they are overindebted and have problems in coping with their debts and bills.

There is a clear asymmetry in how households' indebtedness and unsustainable levels of debt are linked to the forecast errors. The households making pessimistic forecast errors have lower debt-to-income ratios than those making optimistic forecast errors. They do not suffer from overindebtedness like those who make optimistic forecast errors do.

We also separate households that make "prudent" optimistic forecast errors from those that make "non-prudent" forecast errors. A prudent optimistic forecast error refers to those cases in which a household experiences a negative surprise relative to what it expected but, despite of that happening, it reports that its financial situation does not actually get worse (in absolute terms). A household is said to make a non-prudent optimistic forecast error if it both experiences a negative surprise relative to what it expected and if its financial situation actually gets worse. To sharpen our baseline results, we show that households that make non-prudent optimistic forecast errors have greater debt-to-income ratios than those who do not make forecast errors or who make pessimistic forecast errors. Those that make non-prudent optimistic forecast errors are also more likely to report that they are overindebted and experience

empirical analysis.

problems with their debt than others. We do not find similar results for those who make prudent optimistic forecast errors.

Our findings contribute first and foremost to the scant literature on the economic consequences of households' expectational errors. In an important contribution, Souleles (2004) used the Michigan Survey of Consumer Attitudes and Behaviour together with the Consumer Expenditure Survey to document that in the US, households' forecast errors are correlated with their demographics (and thus ex post biased) and that optimistic forecast errors are negatively related to consumption. Puri and Robinson (2007) used the Survey of Consumer Finance and a difference between self-reported life expectancy and statistical life tables to construct a measure of consumer optimism. Using this measure, they documented that besides heterogeneity in the degree of optimism, there is a negative relation between holding optimistic views and many economic choices, such as saving, balance payment habits of credit card holders, long-term planning and non-smoking.[3]

Our analysis differs from these previous papers in two major ways: First, neither Souleles (2004) nor Puri and Robinson (2007) link households' forecast errors to their total indebtedness or to their problems in coping with payments. Second, we consider different degrees and types of forecast errors (e.g., prudent vs. non-prudent). In addition, we show that the borrowing behaviour of those making pessimistic errors differs from the borrowing behaviour of those making optimistic forecast errors.

Our results also contribute to the literature on the financial fragility of households. Nearly half of the US households report that they are probably not able

[3] See also Hayashi (1985), and Pistaferri (2001), who evaluate the sensitivity of consumption to income shocks using elicited expectations data, as well as Dominitz (1998), Manski (2004) and Jappelli and Pistaferri (2010), who discuss the use of subjective expectations data from various angles.

to come up with $2000 in a month should a need arise (see Lusardi, Schneider and Tufano 2011). This suggests that already small unexpected financial shocks can cause problems to them. In our data, nearly 5% of the households make "large", clearly optimistic forecast errors. The mean drop in annual disposable income among these households is lower than the threshold of 1500 euros used by Lusardi, Schneider and Tufano (2011) to measure which European households are likely to be financially fragile. Besides drawing on savings and relying on intra-family transfers, borrowing is one of the most commonly used means to cope with financial shocks, both in the US and other industrialized countries (Lusardi, Schneider and Tufano 2011). Our findings show that optimistic forecast errors that on average are not extremely large in monetary terms are strongly associated with problematic borrowing behaviour and problems in coping with payments. This suggests a link between the financial fragility of households and the type of forecast errors that they make.

Finally, our analysis is related to analyses of debt illiteracy and behavioral biases in the market for debt. Lusardi and Tufano (2015) show that those who appear to have a limited understanding of debt contracts and interest calculations pay more for their borrowing and are more like to be excessively indebted. Stango and Zinman (2009) focus on the systematic tendency of consumers to underestimate loan interest rates (due to exponential growth bias). Gathergood (2012) and Disney and Gathergood (2013) provide complementary evidence on how low financial literacy is related to borrowing using UK data.

The rest of this paper is as follows: Section 2 describes the data. Section 3 reports how households' forecast errors are related to borrowing behaviour. Section 4 then shows how the errors are related to measures of overindebtedness and problems in coping with bills. Section 5 offers concluding remarks.

2 Data

Our main source of data is Statistics Finland's Income Distribution Statistics (IDS). It is a nationally representative data set, covering private Finnish households and their members. The data set is annual and the sampling is based on a rotating panel: Each year, about half are new households; the rest have been included in the IDS once before. This rotation means that each household shows up in the data for two consecutive years.[4]

The IDS consists of two components: The first is a register-based component for which the data are collected from administrative registers, such as census data, tax registers, and social and pension registers. This register-based component of the data contains, for example, detailed demographic information about the households, as well as data on the sources of their income and borrowing behaviour. The second component of the IDS data comes from an interview-based database of Statistics Finland, called Income and Living Conditions Survey. This survey component includes, for example, questions about the expectations of households about the subsequent development of their financial situation.

As we explain in greater detail below, our empirical analysis relies on three key features of the IDS data: First, the survey component of the IDS data allows us to measure expectations. Second, the short panel aspect enables an empirical analysis of the subsequent realizations and, thereby, calculation of forecast errors. Third, the structure of the IDS data allows us to link each household's forecast error to the same household's borrowing behaviour.

The sample available to us covers years from 1994 to 2013. The initial annual IDS sample consists of about ten thousand households per year, but due to the timing of measurement of the expectations variables (see below), there are on

[4]Until 2009, each household was interviewed twice. From that year onwards, the same respondents are interviewed in four successive years.

average nearly five thousand households per year in our data. After allowing for some missing data and non-response to certain key questions and after dropping those households where the respondent was younger than sixteen, our baseline sample consists of 91 827 household-year observations. The exact size of the estimating samples varies a bit, depending on which variables are included. The sample we use in the empirical analysis is also trimmed for outliers and cleaned for certain data errors.[5]

We have matched to the IDS data a number of macroeconomic variables, such as the growth rate of gross domestic product (GDP), inflation, unemployment, stock and regional housing prices, and interest rates. The descriptive statistics of the IDS variables and macrovariables used in our empirical analysis are presented in the Appendix (Table A1).

We use the sampling weights provided by Statistics Finland when calculating the descriptive statistics and when implementing our estimations.[6] The results are therefore representative of the population.

3 Measuring expectations and forecast errors

3.1 Expectations and realizations

To be able to quantify the nature and size of households' forecast errors, we need a measure for household i's financial expectation for year t and the associated realization. Our measures of these key quantities are based on the following two

[5] We dropped from the sample all those person-year observations that were - for any of the debt measures we use in our empirical analysis - at the top 99% percentile of the distribution of the measure. We also dropped those person-year observations for which the change in total debt was at the top 99 percentile or at the lowest 1 percentile, when the debt was measured in euros. There also were some errors in the data, as for some person-years the debt or income measures were negative. They were dropped from the estimation sample(s) as well.

[6] The sampling scheme overweights entrepreneurs and high-income households.

questions, asked in the survey component of the IDS:

Expectation, $E_{it|t-1}$, is derived from the first of the two surveys in which household i participates. It is based on question "*How do you think that the financial situation of your household develops during the next 12 months (or during year t)?*". The following response categories are allowed: "*1 = is clearly better*", "*2 = is somewhat better*", "*3 = stays about the same*", "*4 = is somewhat worse*", and "*5 = is clearly worse*". The question refers to year t, but was asked in the survey that primarily concerns year $t-1$.

Actual outcome, A_{it}, is derived from the second of the two surveys (re-interview) in which i participates. It is based on question "*How do you think that the financial situation of your household developed in year t?*". It allows the same response categories as the expectation question. The question refers to year t and was asked in the survey that primarily deals with year t.[7]

We stress three aspects of these questions: First, the question on which $E_{it|t-1}$ is based asks about the future development of household i's economic and financial position. Its wording and allowed categories match exactly with those of A_{it}, which measures the corresponding realization one year later. These questions are very similar to those used by Souleles (2004) to examine the development of the financial condition of US households and their expectations about it (i.e., his variables QFP^r and QFP^e). Like Souleles, we use the match between the two questions to analyse in-sample forecast errors at the level of households. Second, the expectation and realization questions line up timewise quite nicely. The survey is typically conducted in the early part of each year. The forvard-looking expectation question therefore refers to year t and is asked at the beginning of t (in Spring), and the realization question, asked as a part of

[7] Both questions also included categories "*don't want to say*" and "*don't know*". We drop from the main analysis the households who did not answer to both questions on scale from 1 to 5.

the re-interview at the beginning of $t+1$, looks back at t. Third, it is important to note that while households show up in the data for two consecutive periods (until 2009), we can match $E_{it|t-1}$ with A_{it} only for the latter period.[8]

Table 1 displays a cross-tabulation of $E_{it|t-1}$ and A_{it}. The entries are the number of observations falling in each cell and the associated cell probabilities. The table shows that a bit more than half of the observations in the data (57%) can be found from the diagonal cells. Moreover, about 34% of the observations are in the cells that are adjacent to the diagonal. These patterns indicate that expectations are strongly, but not perfectly, correlated with the subsequent outcomes.

[Insert Table 1 about here]

We can also infer from Table 1 that A_{it} vary around $E_{it|t-1}$ symmetrically. This indicates that on average, households have neither made optimistic nor pessimistic forecast errors over the sample period.

The cell probabilities of the table also indicate that the fraction of households who appear to make larger forecast errors in either direction is moderate but not negligible. For example, out of those who expect that their financial situation does not worsen (i.e., those for whom $E_{it|t-1} = 1, 2$ or 3), around 14% find that it actually gets worse (i.e., $A_{it} = 4$ or 5).

3.2 Measuring forecast errors

The matched pair of questions, $E_{it|t-1}$ and A_{it}, allow us to calculate a forecast error for each household. We consider two complementary measures, which combine insights both from Souleles (2004) and Puri and Robinson (2007): Like

[8] This timing means, in particular, that we can compute only one forecast error per household. We can do so for a large number of households over a period of nineteen years.

Souleles, we use the difference between $E_{it|t-1}$ and A_{it} as the basis of our forecast errors (see, e.g., his εFP measure). However, instead of just focusing on the difference, we follow Puri and Robinson and explicitly consider the *nature* of the forecast errors, such as whether the expectations and realizations imply mild or extreme optimism or non-prudent expectation formation.

Our first measure, denoted $FE1_{it}$, allows for 5 categories. They are clearly pessimistic forecast error ($FE1_{it} = 1$ if $A_{it} - E_{it|t-1} \leq -2$), moderately pessimistic forecast error ($FE1_{it} = 2$ if $A_{it} - E_{it|t-1} = -1$), no forecast error ($FE1_{it} = 3$ if $A_{it} - E_{it|t-1} = 0$), moderately optimistic forecast error ($FE1_{it} = 4$ if $A_{it} - E_{it|t-1} = 1$), and clearly optimistic forecast error ($FE1_{it} = 5$ if $A_{it} - E_{it|t-1} \geq 2$).

The logic of this measure is that it characterizes whether the expectation of household i about the development of its financial situation matches with its ex post view of the eventual realization. The optimistic (pessimistic) errors refer to the cases in which a household experiences a negative (positive) surprise relative to what it expected. For example, a household is said to make a clearly optimistic forecast error ($FE1_{it} = 5$) if it initially thought that its financial situation would clearly improve, but if it in the end stayed about the same or worsened.

It is also worth pointing out two features about $FE1_{it}$. First, the middle category ($FE1_{it} = 3$) correponds to the case of the expectation matching with the subsequent perceived realization. Second, the forecast error is allowed to be two-sided, i.e. it can obtain a different or the same sign on the two sides of the middle category.[9] We make use of this property when we study whether households' borrowing behaviour reacts symmetrically or asymmetrically to the optimistic and pessimistic forecast errors.

[9] For example, in Souleles (2004), the effect of forecast errors is monotonic, due to the linear specification used.

Our second measure, denoted $FE2_{it}$, is asymmetric and has four categories. Besides allowing for separate categories for those making pessimistic or no forecast errors, this measure distinguishes between *prudent* and *non-prudent* optimistic errors. The categories of $FE2_{it}$ are pessimistic forecast error ($FE2_{it} = 1$ if $A_{it} - E_{it|t-1} < 0$), no forecast error ($FE2_{it} = 2$ if $A_{it} - E_{it|t-1} = 0$), prudentially optimistic forecast error ($FE2_{it} = 3$ if $A_{it} - E_{it|t-1} > 0$ and $A_{it} \leq 3$), and non-prudentially optimistic forecast error ($FE2_{it} = 4$ if $A_{it} - E_{it|t-1} > 0$ and $A_{it} > 3$).

It is important to be clear what prudent and non-prudent optimistic forecast errors mean. The prudent optimistic forecast error refers to those cases in which a household experiences a negative surprise relative to what it expected but, despite of that happening, its financial situation does not actually get worse. A household is said to experience a non-prudent optimistic forecast error if it both experiences a negative surprise relative to what it expected and if its financial situation actually gets worse. The difference between the prudent and non-prudent errors is that the latter is more important for, and likely to call for a greater response by, financially fragile households (Lusardi, Scheiner and Tufano 2011). In particular, if the financial situation of a household does not actually get worse despite it having made a forecast error, there is less need for adjustment.

Table 2 consists of three panels. Panel A displays the distribution of the data by $FE1_{it}$ and $FE2_{it}$ and the average change in disposable income, measured in two alternative ways, for each forecast error category of the two measures. Panel B displays the cross-tabulation of $FE1_{it}$ and $FE2_{it}$. Finally, Panel C reports the average change in disposable income by outcome, A_{it}, and expectation, $E_{it|t-1}$, categories.

Reading Panel A, we can see that $FE1_{it}$ classifies 4.9% of the households as having made clearly optimistic forecast errors and 17.3% as having made moderately optimistic forecast errors. The corresponding numbers for the pessimistic errors are very close to these, 3.7% and 17.1%, respectively. This confirms that on average, households' forecast errors have been quite symmetric over the sample period. The numbers for $FE2_{it}$, in turn, show that 13.0% of the households appear to have made non-prudentially optimistic errors. These households both experienced a negative surprise relative to what they expected and their financial situation actually got worse. The fraction of households that made prudentially optimistic forecast errors is 9.2%.

A particular worry about $FE1_{it}$ and $FE2_{it}$ is that they are based on the subjective views of the households and thus that they do not measure anything real. Starting with $FE1_{it}$, the second last column in Panel A shows the average change in disposable annual income (from $t-1$ to t) declines monotonically as we go from the category of clearly pessimistic forecast errors ($FE1_{it} = 1$) to the category of clearly optimistic forecast errors ($FE1_{it} = 5$). This pattern is what we would expect if households make errors in their expectations and particularly if they have been unable to predict the direction of changes in their disposable income.

The table also shows that the mean drop in disposable income among those households that make the largest, clearly optimistic errors is about around 330 euros. This is lower than the threshold of 1500 euros used by Lusardi, Schneider and Tufano (2011) to measure which European households are likely to be financially fragile.

Panel A also shows how shocks in annual disposable income, as measured by a regression residual from an estimated AR(1) process (with year dummies included, using pooled data), vary across the forecast error categories in an expected fashion: the average shock in income becomes monotonically more negative as we go from the category of clearly pessimistic forecast errors to the category of clearly optimistic forecast errors. The last column shows, in particular, that those making largest optimistic errors experience, on average, greatest negative income shocks. For them, the average size of shocks is as high as $\sim$4200 euros.

Finally, Panel A shows that also $FE2_{it}$ works as expected: households making non-prudentially optimistic forecast errors appear to be different from the rest: Their disposable income develops, on average, clearly less favorably. Their income declines on average about 740 euros, whereas in the other categories, disposable income increases (as it should). Those making non-prudentially optimistic forecast errors experience, on average, largest negative income shocks, as among them, the average size of shocks is about 4500 euros.

[Insert Table 2 about here]

The expected (monotonic) patterns documented in Panel A make it clear that $FE1_{it}$ and $FE2_{it}$ coincide with real changes in households' financial condition. It is of course difficult to take a conclusive stance on whether the households are making errors because of biased ex ante expectations or because they have just experienced truly unexpected financial shocks. It is nevertheless interesting to observe that unlike the changes in disposable income, the shocks to annual disposable income, as captured by the regression residual, are on average negative in *all* forecast error categories. This means that the household-level shocks to disposable income that cannot be predicted using households' lagged income (controlling for year dummies) have on average been negative also for those who made pessimistic forecast errors over the nineteen years that our data cover.

Turning then to Panel B, we find that out of those who make clearly optimistic forecast errors, most (87%) make non-prudentially optimistic forecast errors. However, around a third (33%) of those who make non-prudentially optimistic forecast errors make clearly optimistic forecast errors. The two measures are thus not measuring the same thing and are hence complementary.

Finally, Panel C shows that average changes in disposable income are greater in the extreme categories, if we look at the actual outcomes (A_{it}), as compared to the corresponding categories in the expectations ($E_{it|t-1}$). For example, the disposable income decreases, on average, among those households who expect clearly worse financial situation by about 580 euros. On the other hand, the decrease is on average about two thousand euros among the households who report that their financial situation actually became clearly worse.

4 Forecast errors and household borrowing behaviour

4.1 Household borrowing and indebtedness

The debt measure available from the IDS is comprehensive. It includes many kinds of household borrowing, including mortgages and various forms of non-mortgage debt, such as consumer credit, student loans and loans taken up for the purpose of acquiring income. In our baseline estimations, our measure of debt, D_{it}, refers to the total amount of loans that household i has accumulated by period t. The debt measure is available for two consecutive periods for each household.

To obtain our main measure for the indebtedness of households we divide D_{it} by the annual disposable income of household i at t. This gives us a debt-to-income ratio, D_I_{it}. This measure captures how much in debt the household is relative to its current disposable income at the end of year t, i.e., after the realization of the forecast error. An important benefit of this baseline measure is that it allows us to control for the once lagged debt-to-income ratio in the regressions, because each household shows up in our data twice.

The mean of D_I_{it} is 0.55 and its standard deviation is 0.92 (see Table A1 in the Appendix). The median is much lower, because many households have no debt; as we explain below, our estimation results are robust to using a Tobit model, which explicitly allows for clustering at zero.

In our baseline estimations, we regress D_I_{it} on $FE1_{it}$ and, alternatively, on $FE2_{it}$. In all estimations, the omitted caregory refers to those who do not make a forecast error. We consider four different model specifications. First, in Model 1, we regress D_I_{it} on the forecast error categories without any controls. We report the results of this simple model, because it provides a direct way of showing descriptively how much more (or less) debt those households that make forecast errors have relative to those who do not make such errors.

In Model 2, we control flexibly for cohort, age and time effects. To this end, we include long vectors of dummies for the cohort and age effects, which impose only very mild restrictions on the possible lifecycle and cohort patterns in the data.[10] As suggested e.g. by Gourinchas and Parker (2002), we model the time effects as a function of underlying macroeconomic and market condition variables.[11] The set of variables that we use for this purpose is unusually rich and include, for example, measures for real GDP growth, unemployment, inflation, stock returns and their variability, short interest rates and their variability (see

[10] In our baseline specifications, we have 20 age (running from the age of 19 to 79) and 20 cohort (running from year 1921 to 1981) dummies. A single dummy covers hence a period of three years.

[11] Controlling fully non-parametrically for all three effects simultaneously is not possible (see, e.g., Attanasio 1998 and Hall, Mairesse and Turner 2007): There is no general solution to the fundamental identification problem that is due to the fact that calendar year (time) is equal to the year of birth (cohort) plus age.

Table A1 in the Appendix).

Model 2 also controls for regional housing prices, which may boost home equity-based borrowing and consumption (Mian and Sufi 2011). In addition, it includes dummies for education (5 categories), dummies for the region of residence (20 categories), socioeconomic status of the household (9 categories), type of the household (as determined by the Statistics Finland; 5 categories), size of the household (3 categories), marital status (of the head of the household; 4 categories), and dummies for the gender of household head and whether he/she has recently retired. In addition, Model 2 includes a dummy for the spoken language, which is one for swedish-speaking minority, and a dummy for homeowners (owner-occupiers). Because the dummy for homeowners may not be predetermined, we have checked that none of our estimation results depends on us including it in the model.

Model 3 augments Model 2 by additionally including the lagged disposable income of household i. Given that the dependent variable is the debt-to-income ratio, the lagged disposable income controls for the possibility that making forecast errors is releted to the level of income.

Finally, Model 4 includes the lagged dependent variable as a control variable in addition to all the controls that are already included in Model 3. In these estimations, the lagged dependent variable controls for the stock of accumulated debt.

Table 3 reports the OLS results for Models 1-4: Panel A gives the results when $FE1_{it}$ is used and Panel B for $FE2_{it}$. The standard errors are clustered at region-year level. The lower part of the panels reports F-tests for the differences in the estimated coefficients of $FE1_{it}$ and $FE2_{it}$ across the various forecast error categories.

Four main findings stand out from Panel A: First, Model 1 shows that when nothing is controlled for, households that make clearly and moderately pessimistic forecast errors have respectively 0.136 and 0.117 euros more debt for each earned euro of disposable income than those who do not make forecast errors. Households that make moderately and clearly optimistic forecast errors also have greater debt levels than those who do not make forecast errors: They have respectively 0.172 and 0.311 euros more debt for each earned euro of disposable income.

Second, as Models 2-4 show, there is a clear asymmetry: The finding of pessimistic forecast errors being associated with greater indebtedness no longer holds when controls are introduced. However, the result that the households who make optimistic forecast errors are more indebted than households who make no errors is robust across the four models. The F-tests reported in the bottom of the panel show, moreover, that we reject the null hypothesis that the coefficients of clearly and moderately optimistic forecast errors are equal to those of pessimistic forecast errors. This suggests that the households that make clearly and moderately optimistic forecast errors have higher debt-to-income ratios than those who make similar pessimistic forecast errors.

Third, the households that make clearly optimistic forecast errors have higher debt-to-income ratios than those who make moderately optimistic errors. This finding shows that the magnitude of optimistic errors matters. It also squares with the results of Puri and Robinson (2007), who find that extreme, as opposed to moderate, optimism covaries positively with non-prudent economic behaviour.

Fourth, the economic magnitudes are not negligible. For example, as the estimates for Model 4 show, those making clearly optimistic forecast errors accumulate 0.06 euros more debt per each earned euro than those who do not make forecast errors. This effect is not small, when compared to the mean of

0.55 of the outcome variable.

Panel B shows that those making non-prudentially optimistic forecast errors accumulate more debt per each earned euro. Model 4 shows, in particular, that once the past level of indebtedness is held constant, the households making non-prudentially optimistic errors are more indebted than others and also relative to those who make prudentially optimistic errors.

[Insert Table 3 about here]

The results reported in Table 3 are robust to a number of alternative specifications (see Online Appendix for details). For example, the results hold when we use the Tobit model to account for the cluster of zeros in the dependent variable; if we use the debt-payment-to-income ratio (i.e., mortgage-service to income -ratio) as an alternative outcome measure; if we decompose D_I_{it} into its mortgage debt and non-mortgage debt components; if we replace the lagged disposable income by a measure of shocks to disposable income; and if we add a proxy for households' financial sophistication (Calvet, Campbell, and Sodini 2009) to the estimated models.

4.2 Overindebtedness

We now take a look at how forecast errors covary with the overindebtedness of households. In particular, if the level of indebtedness of the households that make the largest optimistic forecast errors is unsustainable, we expect a positive relation between making such errors and perceived overindebtedness.

Unfortunately, there is no agreement in the literature on what the overindebtedness of households means or how it should be measured. The standard model of consumption and intertemporal allocation (see, e.g., Attanasio and Weber 2010) predicts that households attempt to smooth their consumption intertemporally and that borrowing (or deleveraging) is an important part of such smoothing (Sullivan 2008, Hurst and Stafford 2004). The possibility that households could carry suboptimal levels of debt is rarely explicitly considered (see however Lusardi and Tufano 2015 and Gathergood 2012).

Due to lack of an established practise, we use a number of indicators of overindebtedness (and problems in making payments). Our primary indicator is a dummy that is one if a household reports in the survey component of the IDS the perception that it has too much debt, and is zero otherwise. The measure is similar in spirit to that used by Lusardi and Tufano (2015). In our data, its mean is 0.05 (see Table A1 in the Appendix). The measure is not available over all the sample years. The estimating sample is therefore smaller in what follows.

Table 4 reports the estimates from linear probability models in which the dummy for the perceived overindebtedness is the dependent variable. There are four panels in the table, each of which reports the results from three model specifications: Panel A reports the results for $FE1_{it}$ and Panel B for $FE2_{it}$. The R.H.S. specifications of the three columns in these panels match with Models 1-3 that we used for Table 3. That is, in Model 1, we have no controls. Model 2 controls flexibly for the cohort, age and time effects and also includes the long vector of demographic controls. Model 3 additionally includes the lagged disposable income. Panel C and D repeat the analyses of Panel A and B, but with the twist that now the three models also include the lagged dependent variable as a control.

[Insert Table 4 about here]

The results reported in Panels A and C show that the households that make clearly optimistic forecast errors are more likely than others to report that they are overindebted. There is also evidence that households that make moderately optimistic forecast errors are more likely to report that they are overindebted than those who do not make forecast errors. Panels B and D show that the same holds for the households that make non-prudentially optimistic forecast errors, but not for those who make prudentially optimistic forecast errors.

The economic magnitudes are significant. As Panel C shows, the probability of a household reporting that it has too much debt is about 10% higher for those that make clearly optimistic forecast errors than for those who do not make forecast errors. This is a large difference, considering that the mean of the dependent variable is 5%. From Panel D we see that the corresponding difference is about 8% for those making non-prudentially optimistic forecast errors.

In sum, the results support the view that making large and/or non-prudent optimistic errors covaries positively with unsustainable levels of debt. We do not find similar evidence for pessimistic forecast errors.

To explore the robustness of the findings reported in Table 4, we have re-estimated the models in three ways. First, we replaced the lagged disposable income in Model 3 of each panel by a measure of shocks to disposable income (obtained as the regression residual from a pooled estimation of AR(1) income process, with year dummies included). The results do not change. Second, we included our proxy for the financial sophistication index (Calvet, Campbell, and Sodini 2009) to the most complete model (Model 3) of the panels. To our surprise, it obtained a positive and significant coefficient. Taken literally, this suggests that higher financial sophistication is associated with greater likelihood of households reporting that they are overindebted. This sounds somewhat peculiar and casts doubt on us being able to replicate the index properly. Be that as it may, the inclusion of this index did not change our key finding of optimistic forecast errors being linked to reported overindebtedness. Third, Logit estimations produce very similar results as the linear probability models reported in Table 4.

Table 5 reports results for four alternative indicators of overindebtedness and households experiencing problems with payments: First, we use as the dependent variable an indicator that is one if a household, which has debt, reports in the survey component of the IDS that it has had to agree with its bank to reschedule its debt payments, and that is zero otherwise ("Reschedule"; mean = 0.09). Second, we use an indicator that is one if a household reports that it has had problems in paying its regular (utility etc.) bills ("Bills"; mean = 0.04). Third, we define an indicator that is one if a household, which has debt, reports that it has had problems in meeting its regular debt payments ("Debt payment problems"; mean = 0.07). Finally, the fourth alternative dependent variable is an indicator that is equal to one if a household reports that, given its income from all sources, it has had problems in covering its basic expenditures ("Problems in making ends meet"; mean = 0.07).

In Panels A and B of Table 5, we report the results for $FE1_{it}$ and $FE2_{it}$, respectively, using Model 3 of Table 4. In Panels C and D, we repeat the estimations but so that the lagged dependent variable is included in the models. The size of the estimating sample varies across the columns, because the alternative measures are not available for the same years and/or households.

As can be seen from the table, the results are robust across the panels and columns. Taken together, they support the view that making clearly optimistic forecast errors covaries positively with problems in making payments and ends meet. As Panels C and D of the table show, controlling for the past problems does not change this conclusion. These results also support the view that there is a clear asymmetry: Making pessimistic forecast errors does not covary positively with payment problems.

[Insert Table 5 about here]

To explore the robustness of the findings reported in Table 5, we implemented the same robustness tests that we did for Table 4 (i.e., replaced the lagged disposable income with the measure of shocks to disposable income; included the proxy of the financial sophistication index; and used Logit estimation). The results of Table 5 did not change.

5 Conclusions

This paper reports three key findings: First, households that make the most optimistic forecast errors accumulate greater levels of debt relative to their disposable income. Second, such households are more likely to report that they are overindebted and that they find it difficult to cope with their bills and other payments. Third, there is a notable asymmetry, as similar systematic patterns cannot be found for the households that make pessimistic forecast errors. We obtained these findings by direct measurement of forecast errors at the level of households and by linking each household's forecast error to its borrowing behaviour and indebtedness, using the same data source.

These findings are novel, but in line with Puri and Robinson (2007), who show that extreme optimism covaries positively with non-prudent economic behaviour. Our findings are also consistent with Souleles' (2004) results on how optimistic forecast errors are associated with reduced consumption in the US.

We cannot take a conclusive stance on why optimistic forecast errors are positively associated with greater levels of debt and overindebtedness. It is possible that the households making such errors have made poor borrowing choices based on biased ex ante expectations. It is also possible that they have experienced unexpected financial shocks which they attempt to smooth by borrowing more.

Be that as it may, it seems that the households that make the most optimistic forecast errors are at risk of accumulating greater and potentially unsustainable levels of debt. Such households may already be or may easily become financially fragile. The reason for this is the following: The mean drop in income among those who make the most optimistic forecast errors is lower than the threshold of 1500 euros used by Lusardi, Schneider and Tufano (2011) to measure which European households are financially fragile. And yet we find that those who make clearly or non-prudent optimistic forecast errors are more likely to report being overindebted and having problems in making ends meet. This suggests that making optimistic forecast errors and being financially fragile are connected.

Moreover, in an earlier version of our paper, we explored how the forecast errors correlate with the observed characteristics of the households. It seemed that some of the variables that explain variation in the likelihood of making forecast errors, such as low education, are those that according to the prior literature are related to poor financial literacy (Calvet, Campbell and Sodini 2009; Bruine de Bruin et al. 2010). This suggests an interesting avenue for further research, as poor financial and economic literacy may explain why those who make forecast errors accumulate more and perhaps unsustainable levels of debt. There already is some cross-sectional evidence pointing to this direction, as it seems that overindebted households in the UK are financially less literate and more likely to have experienced various economic and financial shocks (Gathergood 2012).

References

[1] Alessie, R. and Lusardi, A. (1997), Saving and income smoothing: Evidence from panel data, *European Economic Review*, Vol. 41, No. 7, pp. 1251-1279.

[2] Attanasio, O. P. (1998), Cohort analysis of saving behaviour by U.S. households, *Journal of Human Resources*, Vol. 33, No. 3, pp. 575-609.

[3] Attanasio, O. P. and Weber, G. (2010), Consumption and saving: Models of intertemporal allocation and their implications for public policy, *Journal of Economic Literature*, Vol. 48, pp. 693-751.

[4] Bruine de Bruin, W., van der Klauw, W., Downs, J., Fischhoff, B., Topa, G. and Armantier, O. (2010), Expectations of inflation: The role of financial literacy and demographic variables, *Journal of Consumer Affairs*, Vol. 44, No. 1, pp. 381-402.

[5] Calvet, L., Campbell, J., and Sodini, P. (2009), Measuring the financial sophistication of households, *American Economic Review*, Vol. 99, pp. 393-398.

[6] Campbell, J. Y., Jackson, H. E., Madrian, B. C. and Tufano, P. (2011), Consumer financial protection, *Journal of Economic Perspectives*, Vol. 25, No. 1, pp. 91-114.

[7] Disney, R. and Gathergood, J. (2013), Financial literacy and consumer credit portfolios, *Journal of Banking and Finance*, Vol. 37, pp. 2246-2254,

[8] Dominitz, J. (1998), Earnings, expectations, revisions, and realizations, *Review of Economics and Statistics*, Vol. 82, pp. 374-388.

[9] Gathergood, J. (2012), Self-control, financial literacy and consumer over-indebtedness, *Journal of Economic Psychology,* Vol. 33, pp. 590-602.

[10] Gourinchas, P-O. and Parker, J. A. (2002), Consumption over the life cycle, *Econometrica*, Vol. 70, No. 1, pp. 47-89.

[11] Hall, B. H., Mairesse, J. and Turner, L. (2007), Identifying age, cohort, and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on French physicists, *Economics of Innovation and New Technology*, Vol. 16, No. 2, pp. 159-177.

[12] Hayashi F. (1985), The permanent income hypothesis and consumption durability; estimates from panel data on households, *Quarterly Journal of Economics*, Vol. 100, pp. 1083-1113.

[13] Hurst, E. and Stafford, F. (2004), Home is where equity is: Mortgage refinancing and household consumption, *Journal of Money, Credit and Banking*, Vol. 36, No. 6, pp. 985-1014.

[14] Jappelli, T. and Pistaferri, L. (2010), The consumption response to income changes, *Annual Review of Economics*, Vol. 2, pp. 479-506.

[15] Johnson, K. W. and Li, G. (2010), The debt-payment-to-income ratio as an indicator of borrowing constraints: Evidence from two household surveys, *Journal of Money, Credit and Banking*, Vol. 42, No. 7, pp. 1373-1390.

[16] Lusardi, A. and Tufano, P. (2015), Debt literacy, financial experiences, and overindebtedness, Journal of Pension Economics and Finance, Vol. 14, Special Issue 4, pp. 332-368.

[17] Lusardi, A., Schneider, D. and Tufano, P. (2011), Financial fragile households: Evidence and implications, *Brookings Paper on Economic Activity*, Spring 2011, pp. 83-134,

[18] Manski, C. (2004), Measuring expectations, *Econometrica*, Vol. 72, No. 5, pp. 1329-1376.

[19] Mian, A. and Sufi, A. (2011), House prices, home equity-based borrowing, and the U.S. household leverage crisis, *American Economic Review*, Vol. 101, pp. 2132-2156.

[20] Pistaferri, L. (2001), Superior information, income shocks, and the permanent income hypothesis, *Review of Economics and Statistics*, Vol. 83, pp. 465-476.

[21] Puri, M. and Robinson, D. (2007), Optimism and economic choice, *Journal of Financial Economics*, Vol. 86, pp. 71-99.

[22] Souleles, N. S. (2004), Expectations, heterogenous forecast errors and consumption: Micro evidence from the Michigan Consumer Sentiment Surveys, *Journal of Money, Credit and Banking*, Vol. 36, No. 1, pp. 39-72.

[23] Stango, V. and Zinman, J. (2009), Exponential growth bias and household finance, *Journal of Finance*, Vol. 64, pp. 2807-2849.

[24] Sullivan, J. X. (2008), Borrowing during unemployment: Unsecured debt as a safety net, *Journal of Human Resources*, Vol. 43, pp. 383-412.

[25] Zeldes, P. (1989), Consumption and liquidity constraints: An empirical investigation, *Journal of Political Economy*, Vol. 97, pp. 305-346.

Tables

Table 1: Cross-tabulation of expectations and actual outcomes

This table reports the cross-tabulation of households' expectations ($E_{it|t-1}$) on the development of their financial situation over the next 12 months (year t) and the actual outcomes (A_{it}), as assessed a year later. The categories take on values "1 = clearly towards better", "2 = somewhat towards better", "3 = stays/stayed about the same", "4 = somewhat towards worse", and "5 = clearly towards worse". The cell entries are the number of households and the corresponding cell probabilities. The data are from Statistics Finland's Income Distribution Statistics (1994–2013).

	Actual outcome (A_{it})					
Expectation ($E_{it\|t-1}$)	1	2	3	4	5	Total
1	889	1 070	570	176	101	2 806
	1.0 %	1.2 %	0.6 %	0.2 %	0.1 %	3.1 %
2	1 270	6 778	6 778	1 600	543	16 969
	1.4 %	7.4 %	7.4 %	1.7 %	0.6 %	18.5 %
3	928	8 394	40 159	6 900	1 474	57 855
	1.0 %	9.1 %	43.7 %	7.5 %	1.6 %	63.0 %
4	151	1 122	5 249	3 816	1 109	11 447
	0.2 %	1.2 %	5.7 %	4.2 %	1.2 %	12.5 %
5	37	282	859	809	764	2 751
	0.0 %	0.3 %	0.9 %	0.9 %	0.8 %	3.0 %
Total	3 274	17 646	53 615	13 302	3 991	91 827
	3.6 %	19.2 %	58.4 %	14.5 %	4.4 %	100.0 %

Table 2: Distribution and cross-tabulation of forecast errors

Panel A reports the distribution of the data by the forecast error categories and average changes in disposable income for each category. The average change in disposable income is measured in two alternative ways: by the annual change in disposable income and by the shocks to income, as measured by the residual of the regression of income on lagged income and year dummies. Panel B reports the cross-tabulation of the two forecast error measures ($FE1_{it}$ and $FE2_{it}$). Panel C reports the average change in income and the innovation in income by actual outcome (A_{it}) and expectation ($E_{it|t-1}$) categories. The categories take on values "1 = clearly towards better", "2 = somewhat towards better", "3 = stays/stayed about the same", "4 = somewhat towards worse", and "5 = clearly towards worse". The data are from Statistics Finland's Income Distribution Statistics (1994–2013).

Panel A: Distribution of the forecast errors and changes in income by the error category

	Obs.	Percent	Cum.	Change in income	Shocks to income
$FE1_{it}$				(In thousands EUR)	
Clearly pessimistic	3 378	3.7 %	3.7 %	2.89	-0.44
Moderately pessimistic	15 722	17.1 %	20.8 %	2.32	-0.56
None	52 405	57.1 %	77.9 %	1.41	-1.75
Moderately optimistic	15 856	17.3 %	95.1 %	1.07	-2.34
Clearly optimistic	4 464	4.9 %	100.0 %	-0.33	-4.18
Total	91 827	100.0 %			
$FE2_{it}$				(In thousands EUR)	
Pessimistic	19 101	20.8 %	20.8 %	2.42	-0.54
None	52 405	57.1 %	77.9 %	1.41	-1.75
Prudentially optimistic	8 417	9.2 %	87.0 %	2.87	-0.27
Non-prudentially optimistic	11 904	13.0 %	100.0 %	-0.74	-4.49
Total	91 827	100.0 %			

[Table 2 continued]

Table 2, continued: Panel B: Cross-tabulation of forecast error measures ($FE1_{it}$ and $FE2_{it}$)

	$FE2_{it}$				
$FE1_{it}$	Pessimistic	None	Prudentially optimistic	Non-prudentially optimistic	Total
Clearly pessimistic	3 378	-	-	-	3 378
	3.7 %	-	-	-	3.7 %
Moderately pessimistic	15 722	-	-	-	15 722
	17.1 %	-	-	-	17.1 %
None	-	52 405	-	-	52 405
	-	57.1 %	-	-	57.1 %
Moderately optimistic	-	-	7 847	8 009	15 856
	-	-	8.6 %	8.7 %	17.3 %
Clearly optimistic	-	-	570	3 895	4 464
	-	-	0.6 %	4.2 %	4.9 %
Total	19 101	52 405	8 417	11 904	91 827
	20.8 %	57.1 %	9.2 %	13.0 %	100.0 %

Panel C: Average changes in income by actual outcome and expectation (in thousands EUR)

	Actual outcome				
	1	2	3	4	5
Change in income	7.98	3.64	1.13	-0.52	-2.07
Shocks to income	4.24	0.84	-2.02	-4.07	-5.96
	Expectation				
	1	2	3	4	5
Change in income	5.70	3.44	1.11	-0.13	-0.58
Shocks to income	1.59	0.29	-2.00	-3.43	-4.25

Table 3: Forecast errors and indebtedness

Panels A and B report OLS estimates of the effect of forecast errors ($FE1_{it}$ and $FE2_{it}$ respectively) on the debt-to-income ratio. Standard errors are clustered at the region-year level and reported in parentheses; *** p<0.001, ** p<0.01, * p<0.05. The data are from the Statistics Finland's Income Distribution Statistics (1994–2013).

Panel A: FE1 and debt-to-income ratio				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Clearly pessimistic	0.136***	0.030	0.029	0.002
	(0.021)	(0.018)	(0.018)	(0.014)
Moderately pessimistic	0.117***	0.018	0.018	0.004
	(0.012)	(0.010)	(0.010)	(0.006)
Moderately optimistic	0.172***	0.086***	0.085***	0.031***
	(0.012)	(0.009)	(0.009)	(0.006)
Clearly optimistic	0.311***	0.166***	0.165***	0.059***
	(0.022)	(0.018)	(0.018)	(0.011)
Demographic control variables	No	Yes	Yes	Yes
Macroeconomic variables	No	Yes	Yes	Yes
Lagged income	No	No	Yes	Yes
Lagged debt-to-income ratio	No	No	No	Yes
N	91 827	91 827	91 827	91 827
R-squared	0.009	0.304	0.305	0.733
p-value for F-test of equality of coefficients				
Clearly optimistic vs. clearly pessimistic	0.000	0.000	0.000	0.002
Clearly optimistic vs. moderately optimistic	0.000	0.000	0.000	0.040
Clearly pessimistic vs. moderately pessimistic	0.377	0.524	0.531	0.921
Moderately optimistic vs. moderarately pessimistic	0.000	0.000	0.000	0.001

Panel B: FE2 and debt-to-income ratio				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Pessimistic	0.120***	0.020*	0.020*	0.003
	(0.012)	(0.009)	(0.009)	(0.006)
Prudentially optimistic	0.266***	0.093***	0.093***	0.011
	(0.018)	(0.013)	(0.013)	(0.009)
Non-prudentially optimistic	0.158***	0.109***	0.108***	0.056***
	(0.012)	(0.010)	(0.010)	(0.007)
Demographic control variables	No	Yes	Yes	Yes
Macroeconomic variables	No	Yes	Yes	Yes
Lagged income	No	No	Yes	Yes
Lagged debt-to-income ratio	No	No	No	Yes
N	91 827	91 827	91 827	91 827
R-squared	0.009	0.304	0.305	0.733
p-value for F-test of equality of coefficients				
Non-prudentially optimistic vs. pessimistic	0.012	0.000	0.000	0.000
Prudentially optimistic vs. pessimistic	0.000	0.000	0.000	0.451
Non-prudentially opt. vs. prudentially optimistic	0.000	0.311	0.325	0.000

Table 4: Forecast errors and overindebtedness

Panels A and B report OLS estimates of the effect of forecast errors ($FE1_{it}$ and $FE2_{it}$ respectively) on perceived overindebtedness. Panels C and D report the same models including a lagged dependent variable as a control. Standard errors are clustered at the region-year level and reported in parentheses; *** p<0.001, ** p<0.01, * p<0.05. The data are from the Statistics Finland's Income Distribution Statistics (1996–2013).

Panel A: FE1 and perceived overindebtedness

OLS estimates	Model 1	Model 2	Model 3
Clearly pessimistic	0.027***	0.014	0.014
	(0.008)	(0.008)	(0.008)
Moderately pessimistic	-0.003	-0.004	-0.004
	(0.003)	(0.003)	(0.003)
Moderately optimistic	0.032***	0.024***	0.024***
	(0.004)	(0.004)	(0.004)
Clearly optimistic	0.156***	0.135***	0.135***
	(0.012)	(0.011)	(0.011)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	46 933	46 933	46 933
R-squared	0.031	0.099	0.099

Panel B: FE2 and perceived overindebtedness

OLS estimates	Model 1	Model 2	Model 3
Pessimistic	0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)
Prudentially optimistic	0.007	0.000	0.000
	(0.004)	(0.004)	(0.004)
Non-prudentially optimistic	0.110***	0.095***	0.095***
	(0.008)	(0.007)	(0.007)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	46 933	46 933	46 933
R-squared	0.029	0.098	0.098

[Table 4 continued]

Panel C: FE1 and perceived overindebtedness with lagged overindebtedness as a control			
OLS estimates	Model 1	Model 2	Model 3
Clearly pessimistic	-0.004	-0.009	-0.009
	(0.008)	(0.008)	(0.008)
Moderately pessimistic	-0.005	-0.006	-0.006
	(0.003)	(0.003)	(0.003)
Moderately optimistic	0.025***	0.021***	0.021***
	(0.004)	(0.004)	(0.004)
Clearly optimistic	0.113***	0.104***	0.104***
	(0.010)	(0.010)	(0.010)
Lagged overindebtedness	0.401***	0.361***	0.361***
	(0.016)	(0.016)	(0.016)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	42 448	42 448	42 448
R-squared	0.196	0.220	0.220

Panel D: FE2 and perceived overindebtedness with lagged overindebtedness as a control			
OLS estimates	Model 1	Model 2	Model 3
Pessimistic	-0.005	-0.007*	-0.007*
	(0.003)	(0.003)	(0.003)
Prudentially optimistic	0.001	-0.003	-0.003
	(0.004)	(0.004)	(0.004)
Non-prudentially optimistic	0.086***	0.079***	0.079***
	(0.006)	(0.006)	(0.006)
Lagged overindebtedness	0.403***	0.362***	0.362***
	(0.016)	(0.016)	(0.016)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	42 448	42 448	42 448
R-squared	0.198	0.222	0.222

Table 5: Forecast errors and alternative indicators of overindebtedness

Panels A and B report OLS estimates of the effect of forecast errors ($FE1_{it}$ and $FE2_{it}$ respectively) on alternative indicators of overindebtedness. Dependent variables in models 1–4, respectively, are dummy variables as follows: having new loan repayment schedule ("Reschedule" = 1), having problems in paying regular bills ("Bills" = 1), having problems in meeting regular debt payments ("Debt payment problems" = 1), and having had problems in making ends meet ("Problems in making ends meet" = 1). Panels C and D report the same models including a lagged dependent variable as a control. Standard errors are clustered at the region-year level and reported in parentheses; *** p<0.001, ** p<0.01, * p<0.05. The data are from the Statistics Finland's Income Distribution Statistics (different data periods).

Panel A: FE1 and alternative indicators of overindebtedness				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Clearly pessimistic	0.018	0.008	0.024*	0.015
	(0.010)	(0.006)	(0.011)	(0.012)
Moderately pessimistic	-0.008	-0.001	-0.003	0.001
	(0.004)	(0.002)	(0.004)	(0.004)
Moderately optimistic	0.033***	0.020***	0.035***	0.055***
	(0.005)	(0.003)	(0.005)	(0.005)
Clearly optimistic	0.150***	0.103***	0.123***	0.187***
	(0.011)	(0.009)	(0.011)	(0.014)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	46 850	79 438	44 464	43 990
R-squared	0.043	0.087	0.075	0.128

Panel B: FE2 and alternative indicators of overindebtedness				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Pessimistic	-0.004	-0.000	0.001	0.001
	(0.004)	(0.002)	(0.004)	(0.005)
Prudentially optimistic	0.012*	0.001	0.019**	0.007
	(0.006)	(0.004)	(0.006)	(0.006)
Non-prudentially optimistic	0.104***	0.064***	0.088***	0.132***
	(0.007)	(0.005)	(0.008)	(0.007)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	46 850	79 438	44 464	43 990
R-squared	0.041	0.086	0.074	0.130

[Table 5 continued]

Panel C: FE1 and alternative indicators of overindebtedness including a lagged dependent variable as a control				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Clearly pessimistic	0.011	-0.007	0.015	-0.002
	(0.009)	(0.006)	(0.010)	(0.012)
Moderately pessimistic	-0.011*	-0.005*	-0.004	-0.006
	(0.005)	(0.002)	(0.004)	(0.004)
Moderately optimistic	0.026***	0.012***	0.030***	0.043***
	(0.005)	(0.003)	(0.005)	(0.005)
Clearly optimistic	0.131***	0.076***	0.093***	0.143***
	(0.010)	(0.008)	(0.012)	(0.013)
Lagged dependent variable	0.307***	0.388***	0.353***	0.421***
	(0.011)	(0.013)	(0.014)	(0.015)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	42 445	75 373	38 349	39 415
R-squared	0.136	0.245	0.189	0.291

Panel D: FE2 and alternative indicators of overindebtedness including a lagged dependent variable as a control				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Pessimistic	-0.008	-0.006**	-0.001	-0.007
	(0.004)	(0.002)	(0.004)	(0.005)
Prudentially optimistic	0.003	-0.007*	0.015**	0.003
	(0.005)	(0.003)	(0.005)	(0.006)
Non-prudentially optimistic	0.094***	0.048***	0.071***	0.102***
	(0.007)	(0.004)	(0.007)	(0.007)
Lagged dependent variable	0.310***	0.389***	0.354***	0.421***
	(0.011)	(0.013)	(0.014)	(0.015)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	42 445	75 373	38 349	39 415
R-squared	0.136	0.245	0.189	0.293

Appendix

Table A1: Descriptive statistics

Panel A reports descriptive statistics on dependent household-level variables. Panel B reports the same statistics on continuous and indicator household-level control variables. Panel C reports the proportion of households that belongs to each category of control variables. The data of panels A, B and C are from Statistics Finland's Income Distribution Statistics (1994–2013). Panel D reports descriptive statistics on macro-level control variables from Reuters, Statistics Finland, NASDAQ OMX, Federation of Finnish Financial Services and Bank of Finland, including the authors' own calculations.

Panel A: Main dependent variables					
	Mean	Median	St. dev.	Min	Max
Debt-to-income ratio	0.55	0.03	0.92	0.00	6.85
Perceived overindebtedness (dummy)	0.05	0.00	0.22	0.00	1.00
Reschedule (dummy)	0.09	0.00	0.28	0.00	1.00
Bills (dummy)	0.04	0.00	0.20	0.00	1.00
Debt payment problems (dummy)	0.07	0.00	0.26	0.00	1.00
Problems in making ends meet (dummy)	0.07	0.00	0.26	0.00	1.00

Panel B: Continuous and binary household-level control variables					
	Mean	Median	St. dev.	Min	Max
Age	49.29	49.00	17.22	19.00	79.00
Gender (1 = male)	0.59	1.00	0.49	0.00	1.00
Language (1 = swedish-speaking)	0.05	0.00	0.22	0.00	1.00
Owner-occupier (dummy)	0.65	1.00	0.48	0.00	1.00
Retired within a year (dummy)	0.02	0.00	0.14	0.00	1.00
Lagged income (EUR thousand)	32.16	26.27	33.62	0.00	3862.84
Shock to income (EUR thousand)	-1.72	-3.38	18.72	-1914.43	2233.07

[Table A1 continued]

Panel C: Categorical household-level control variables (proportion of households in each category)			
Structure of household		**Level of education**	
Single	0.40	Unknown	0.05
Single parent	0.07	None or comprehensive	0.27
Couple without children	0.29	Secondary-level	0.41
Couple with children	0.24	Lower-degree tertiary	0.19
Other	0.02	Higher-degree tertiary or doctorate	0.09
Married		**Socioeconomic status**	
Single or unknown	0.34	Agricultural employer or entrepreneur	0.02
Married	0.41	Other employer and entrepreneur	0.05
Divorced	0.15	Upper-level white-collar employee	0.16
Widow	0.10	Lower-level white-collar employee	0.17
Size of household (# of children)		Blue-collar employee	0.19
No children	0.75	Student	0.04
One or two children	0.20	Pensioner	0.30
More than two children	0.04	Long-term unemployed	0.06
Age cohort by year of birth		Other	0.01
$<$ 1924	0.05	**Region of residence**	
1924–1926	0.03	Uusimaa	0.27
1927–1929	0.03	Varsinais-Suomi	0.09
1930–1932	0.03	Satakunta	0.05
1933–1935	0.03	Kanta-Häme	0.03
1936–1938	0.04	Pirkanmaa	0.09
1939–1941	0.04	Päijät-Häme	0.04
1942–1944	0.04	Kymenlaakso	0.04
1945–1947	0.06	South Karelia	0.03
1948–1950	0.06	Etelä-Savo	0.03
1951–1953	0.06	Pohjois-Savo	0.05
1954–1956	0.06	North Karelia	0.03
1957–1959	0.06	Central Finland	0.05
1960–1962	0.06	South Ostrobothnia	0.03
1963–1965	0.06	Ostrobothnia	0.03
1966–1968	0.05	Central Ostrobothnia	0.01
1969–1971	0.05	North Ostrobothnia	0.06
1972–1974	0.04	Kainuu	0.02
1975–1977	0.04	Lapland	0.04
1978–1981	0.04	East Uusimaa	0.01
$>$ 1981	0.08	Åland	0.01

Panel D: Macro-level control variables

	Mean	Median	St. dev.	Min	Max
Short-term interest rate	0.03	0.03	0.02	0.01	0.06
Volatility of short-term interest rate	0.14	0.13	0.14	0.00	0.60
Unemployment rate	0.09	0.08	0.02	0.06	0.15
Inflation	0.02	0.02	0.01	0.00	0.04
Real GDP growth	0.02	0.03	0.03	-0.08	0.06
Real house price change	0.03	0.02	0.05	-0.08	0.21
Real share price change	0.09	0.07	0.33	-0.45	0.87
Volatility of share prices	0.40	0.14	0.53	0.02	1.81
Maturity of new housing loans	16.31	16.00	3.91	11.00	21.00

Tables

Table 1: Cross-tabulation of expectations and actual outcomes

This table reports the cross-tabulation of households' expectations ($E_{it|t-1}$) on the development of their financial situation over the next 12 months (year t) and the actual outcomes (A_{it}), as assessed a year later. The categories take on values "1 = clearly towards better", "2 = somewhat towards better", "3 = stays/stayed about the same", "4 = somewhat towards worse", and "5 = clearly towards worse". The cell entries are the number of households and the corresponding cell probabilities. The data are from Statistics Finland's Income Distribution Statistics (1994–2013).

	Actual outcome (A_{it})					
Expectation ($E_{it\|t-1}$)	1	2	3	4	5	Total
1	889	1 070	570	176	101	2 806
	1.0 %	1.2 %	0.6 %	0.2 %	0.1 %	3.1 %
2	1 270	6 778	6 778	1 600	543	16 969
	1.4 %	7.4 %	7.4 %	1.7 %	0.6 %	18.5 %
3	928	8 394	40 159	6 900	1 474	57 855
	1.0 %	9.1 %	43.7 %	7.5 %	1.6 %	63.0 %
4	151	1 122	5 249	3 816	1 109	11 447
	0.2 %	1.2 %	5.7 %	4.2 %	1.2 %	12.5 %
5	37	282	859	809	764	2 751
	0.0 %	0.3 %	0.9 %	0.9 %	0.8 %	3.0 %
Total	3 274	17 646	53 615	13 302	3 991	91 827
	3.6 %	19.2 %	58.4 %	14.5 %	4.4 %	100.0 %

Table 2: Distribution and cross-tabulation of forecast errors

Panel A reports the distribution of the data by the forecast error categories and average changes in disposable income for each category. The average change in disposable income is measured in two alternative ways: by the annual change in disposable income and by the shocks to income, as measured by the residual of the regression of income on lagged income and year dummies. Panel B reports the cross-tabulation of the two forecast error measures ($FE1_{it}$ and $FE2_{it}$). Panel C reports the average change in income and the innovation in income by actual outcome (A_{it}) and expectation ($E_{it|t-1}$) categories. The categories take on values "1 = clearly towards better", "2 = somewhat towards better", "3 = stays/stayed about the same", "4 = somewhat towards worse", and "5 = clearly towards worse". The data are from Statistics Finland's Income Distribution Statistics (1994–2013).

Panel A: Distribution of the forecast errors and changes in income by the error category

	Obs.	Percent	Cum.	Change in income	Shocks to income
$FE1_{it}$				(In thousands EUR)	
Clearly pessimistic	3 378	3.7 %	3.7 %	2.89	-0.44
Moderately pessimistic	15 722	17.1 %	20.8 %	2.32	-0.56
None	52 405	57.1 %	77.9 %	1.41	-1.75
Moderately optimistic	15 856	17.3 %	95.1 %	1.07	-2.34
Clearly optimistic	4 464	4.9 %	100.0 %	-0.33	-4.18
Total	91 827	100.0 %			
$FE2_{it}$				(In thousands EUR)	
Pessimistic	19 101	20.8 %	20.8 %	2.42	-0.54
None	52 405	57.1 %	77.9 %	1.41	-1.75
Prudentially optimistic	8 417	9.2 %	87.0 %	2.87	-0.27
Non-prudentially optimistic	11 904	13.0 %	100.0 %	-0.74	-4.49
Total	91 827	100.0 %			

[Table 2 continued]

Table 2, continued: Panel B: Cross-tabulation of forecast error measures ($FE1_{it}$ and $FE2_{it}$)

	$FE2_{it}$				
$FE1_{it}$	Pessimistic	None	Prudentially optimistic	Non-prudentially optimistic	Total
Clearly pessimistic	3 378	-	-	-	3 378
	3.7 %	-	-	-	3.7 %
Moderately pessimistic	15 722	-	-	-	15 722
	17.1 %	-	-	-	17.1 %
None	-	52 405	-	-	52 405
	-	57.1 %	-	-	57.1 %
Moderately optimistic	-	-	7 847	8 009	15 856
	-	-	8.6 %	8.7 %	17.3 %
Clearly optimistic	-	-	570	3 895	4 464
	-	-	0.6 %	4.2 %	4.9 %
Total	19 101	52 405	8 417	11 904	91 827
	20.8 %	57.1 %	9.2 %	13.0 %	100.0 %

Panel C: Average changes in income by actual outcome and expectation (in thousands EUR)

	Actual outcome				
	1	2	3	4	5
Change in income	7.98	3.64	1.13	-0.52	-2.07
Shocks to income	4.24	0.84	-2.02	-4.07	-5.96
	Expectation				
	1	2	3	4	5
Change in income	5.70	3.44	1.11	-0.13	-0.58
Shocks to income	1.59	0.29	-2.00	-3.43	-4.25

Table 3: Forecast errors and indebtedness

Panels A and B report OLS estimates of the effect of forecast errors ($FE1_{it}$ and $FE2_{it}$ respectively) on the debt-to-income ratio. Standard errors are clustered at the region-year level and reported in parentheses; *** $p<0.001$, ** $p<0.01$, * $p<0.05$. The data are from the Statistics Finland's Income Distribution Statistics (1994–2013).

Panel A: FE1 and debt-to-income ratio				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Clearly pessimistic	0.136***	0.030	0.029	0.002
	(0.021)	(0.018)	(0.018)	(0.014)
Moderately pessimistic	0.117***	0.018	0.018	0.004
	(0.012)	(0.010)	(0.010)	(0.006)
Moderately optimistic	0.172***	0.086***	0.085***	0.031***
	(0.012)	(0.009)	(0.009)	(0.006)
Clearly optimistic	0.311***	0.166***	0.165***	0.059***
	(0.022)	(0.018)	(0.018)	(0.011)
Demographic control variables	No	Yes	Yes	Yes
Macroeconomic variables	No	Yes	Yes	Yes
Lagged income	No	No	Yes	Yes
Lagged debt-to-income ratio	No	No	No	Yes
N	91 827	91 827	91 827	91 827
R-squared	0.009	0.304	0.305	0.733
p-value for F-test of equality of coefficients				
Clearly optimistic vs. clearly pessimistic	0.000	0.000	0.000	0.002
Clearly optimistic vs. moderately optimistic	0.000	0.000	0.000	0.040
Clearly pessimistic vs. moderately pessimistic	0.377	0.524	0.531	0.921
Moderately optimistic vs. moderarately pessimistic	0.000	0.000	0.000	0.001
Panel B: FE2 and debt-to-income ratio				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Pessimistic	0.120***	0.020*	0.020*	0.003
	(0.012)	(0.009)	(0.009)	(0.006)
Prudentially optimistic	0.266***	0.093***	0.093***	0.011
	(0.018)	(0.013)	(0.013)	(0.009)
Non-prudentially optimistic	0.158***	0.109***	0.108***	0.056***
	(0.012)	(0.010)	(0.010)	(0.007)
Demographic control variables	No	Yes	Yes	Yes
Macroeconomic variables	No	Yes	Yes	Yes
Lagged income	No	No	Yes	Yes
Lagged debt-to-income ratio	No	No	No	Yes
N	91 827	91 827	91 827	91 827
R-squared	0.009	0.304	0.305	0.733
p-value for F-test of equality of coefficients				
Non-prudentially optimistic vs. pessimistic	0.012	0.000	0.000	0.000
Prudentially optimistic vs. pessimistic	0.000	0.000	0.000	0.451
Non-prudentially opt. vs. prudentially optimistic	0.000	0.311	0.325	0.000

Table 4: Forecast errors and overindebtedness

Panels A and B report OLS estimates of the effect of forecast errors ($FE1_{it}$ and $FE2_{it}$ respectively) on perceived overindebtedness. Panels C and D report the same models including a lagged dependent variable as a control. Standard errors are clustered at the region-year level and reported in parentheses; *** p<0.001, ** p<0.01, * p<0.05. The data are from the Statistics Finland's Income Distribution Statistics (1996–2013).

Panel A: FE1 and perceived overindebtedness			
OLS estimates	Model 1	Model 2	Model 3
Clearly pessimistic	0.027***	0.014	0.014
	(0.008)	(0.008)	(0.008)
Moderately pessimistic	-0.003	-0.004	-0.004
	(0.003)	(0.003)	(0.003)
Moderately optimistic	0.032***	0.024***	0.024***
	(0.004)	(0.004)	(0.004)
Clearly optimistic	0.156***	0.135***	0.135***
	(0.012)	(0.011)	(0.011)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	46 933	46 933	46 933
R-squared	0.031	0.099	0.099

Panel B: FE2 and perceived overindebtedness			
OLS estimates	Model 1	Model 2	Model 3
Pessimistic	0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.003)
Prudentially optimistic	0.007	0.000	0.000
	(0.004)	(0.004)	(0.004)
Non-prudentially optimistic	0.110***	0.095***	0.095***
	(0.008)	(0.007)	(0.007)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	46 933	46 933	46 933
R-squared	0.029	0.098	0.098

[Table 4 continued]

Panel C: FE1 and perceived overindebtedness with lagged overindebtedness as a control			
OLS estimates	Model 1	Model 2	Model 3
Clearly pessimistic	-0.004	-0.009	-0.009
	(0.008)	(0.008)	(0.008)
Moderately pessimistic	-0.005	-0.006	-0.006
	(0.003)	(0.003)	(0.003)
Moderately optimistic	0.025***	0.021***	0.021***
	(0.004)	(0.004)	(0.004)
Clearly optimistic	0.113***	0.104***	0.104***
	(0.010)	(0.010)	(0.010)
Lagged overindebtedness	0.401***	0.361***	0.361***
	(0.016)	(0.016)	(0.016)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	42 448	42 448	42 448
R-squared	0.196	0.220	0.220

Panel D: FE2 and perceived overindebtedness with lagged overindebtedness as a control			
OLS estimates	Model 1	Model 2	Model 3
Pessimistic	-0.005	-0.007*	-0.007*
	(0.003)	(0.003)	(0.003)
Prudentially optimistic	0.001	-0.003	-0.003
	(0.004)	(0.004)	(0.004)
Non-prudentially optimistic	0.086***	0.079***	0.079***
	(0.006)	(0.006)	(0.006)
Lagged overindebtedness	0.403***	0.362***	0.362***
	(0.016)	(0.016)	(0.016)
Demographic control variables	No	Yes	Yes
Macroeconomic variables	No	Yes	Yes
Lagged income	No	No	Yes
N	42 448	42 448	42 448
R-squared	0.198	0.222	0.222

Table 5: Forecast errors and alternative indicators of overindebtedness

Panels A and B report OLS estimates of the effect of forecast errors ($FE1_{it}$ and $FE2_{it}$ respectively) on alternative indicators of overindebtedness. Dependent variables in models 1–4, respectively, are dummy variables as follows: having new loan repayment schedule ("Reschedule" = 1), having problems in paying regular bills ("Bills" = 1), having problems in meeting regular debt payments ("Debt payment problems" = 1), and having had problems in making ends meet ("Problems in making ends meet" = 1). Panels C and D report the same models including a lagged dependent variable as a control. Standard errors are clustered at the region-year level and reported in parentheses; *** p<0.001, ** p<0.01, * p<0.05. The data are from the Statistics Finland's Income Distribution Statistics (different data periods).

Panel A: FE1 and alternative indicators of overindebtedness				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Clearly pessimistic	0.018	0.008	0.024*	0.015
	(0.010)	(0.006)	(0.011)	(0.012)
Moderately pessimistic	-0.008	-0.001	-0.003	0.001
	(0.004)	(0.002)	(0.004)	(0.004)
Moderately optimistic	0.033***	0.020***	0.035***	0.055***
	(0.005)	(0.003)	(0.005)	(0.005)
Clearly optimistic	0.150***	0.103***	0.123***	0.187***
	(0.011)	(0.009)	(0.011)	(0.014)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	46 850	79 438	44 464	43 990
R-squared	0.043	0.087	0.075	0.128

Panel B: FE2 and alternative indicators of overindebtedness				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Pessimistic	-0.004	-0.000	0.001	0.001
	(0.004)	(0.002)	(0.004)	(0.005)
Prudentially optimistic	0.012*	0.001	0.019**	0.007
	(0.006)	(0.004)	(0.006)	(0.006)
Non-prudentially optimistic	0.104***	0.064***	0.088***	0.132***
	(0.007)	(0.005)	(0.008)	(0.007)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	46 850	79 438	44 464	43 990
R-squared	0.041	0.086	0.074	0.130

[Table 5 continued]

Panel C: FE1 and alternative indicators of overindebtedness including a lagged dependent variable as a control				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Clearly pessimistic	0.011	-0.007	0.015	-0.002
	(0.009)	(0.006)	(0.010)	(0.012)
Moderately pessimistic	-0.011*	-0.005*	-0.004	-0.006
	(0.005)	(0.002)	(0.004)	(0.004)
Moderately optimistic	0.026***	0.012***	0.030***	0.043***
	(0.005)	(0.003)	(0.005)	(0.005)
Clearly optimistic	0.131***	0.076***	0.093***	0.143***
	(0.010)	(0.008)	(0.012)	(0.013)
Lagged dependent variable	0.307***	0.388***	0.353***	0.421***
	(0.011)	(0.013)	(0.014)	(0.015)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	42 445	75 373	38 349	39 415
R-squared	0.136	0.245	0.189	0.291

Panel D: FE2 and alternative indicators of overindebtedness including a lagged dependent variable as a control				
OLS estimates	Model 1	Model 2	Model 3	Model 4
Pessimistic	-0.008	-0.006**	-0.001	-0.007
	(0.004)	(0.002)	(0.004)	(0.005)
Prudentially optimistic	0.003	-0.007*	0.015**	0.003
	(0.005)	(0.003)	(0.005)	(0.006)
Non-prudentially optimistic	0.094***	0.048***	0.071***	0.102***
	(0.007)	(0.004)	(0.007)	(0.007)
Lagged dependent variable	0.310***	0.389***	0.354***	0.421***
	(0.011)	(0.013)	(0.014)	(0.015)
Demographic control variables	Yes	Yes	Yes	Yes
Macroeconomic variables	Yes	Yes	Yes	Yes
Lagged income	Yes	Yes	Yes	Yes
N	42 445	75 373	38 349	39 415
R-squared	0.136	0.245	0.189	0.293

Appendix

Table A1: Descriptive statistics

Panel A reports descriptive statistics on dependent household-level variables. Panel B reports the same statistics on continuous and indicator household-level control variables. Panel C reports the proportion of households that belongs to each category of control variables. The data of panels A, B and C are from Statistics Finland's Income Distribution Statistics (1994–2013). Panel D reports descriptive statistics on macro-level control variables from Reuters, Statistics Finland, NASDAQ OMX, Federation of Finnish Financial Services and Bank of Finland, including the authors' own calculations.

Panel A: Main dependent variables					
	Mean	Median	St. dev.	Min	Max
Debt-to-income ratio	0.55	0.03	0.92	0.00	6.85
Perceived overindebtedness (dummy)	0.05	0.00	0.22	0.00	1.00
Reschedule (dummy)	0.09	0.00	0.28	0.00	1.00
Bills (dummy)	0.04	0.00	0.20	0.00	1.00
Debt payment problems (dummy)	0.07	0.00	0.26	0.00	1.00
Problems in making ends meet (dummy)	0.07	0.00	0.26	0.00	1.00
Panel B: Continuous and binary household-level control variables					
	Mean	Median	St. dev.	Min	Max
Age	49.29	49.00	17.22	19.00	79.00
Gender (1 = male)	0.59	1.00	0.49	0.00	1.00
Language (1 = swedish-speaking)	0.05	0.00	0.22	0.00	1.00
Owner-occupier (dummy)	0.65	1.00	0.48	0.00	1.00
Retired within a year (dummy)	0.02	0.00	0.14	0.00	1.00
Lagged income (EUR thousand)	32.16	26.27	33.62	0.00	3862.84
Shock to income (EUR thousand)	-1.72	-3.38	18.72	-1914.43	2233.07

[Table A1 continued]

Panel C: Categorical household-level control variables (proportion of households in each category)			
Structure of household		**Level of education**	
Single	0.40	Unknown	0.05
Single parent	0.07	None or comprehensive	0.27
Couple without children	0.29	Secondary-level	0.41
Couple with children	0.24	Lower-degree tertiary	0.19
Other	0.02	Higher-degree tertiary or doctorate	0.09
Married		**Socioeconomic status**	
Single or unknown	0.34	Agricultural employer or entrepreneur	0.02
Married	0.41	Other employer and entrepreneur	0.05
Divorced	0.15	Upper-level white-collar employee	0.16
Widow	0.10	Lower-level white-collar employee	0.17
Size of household (# of children)		Blue-collar employee	0.19
No children	0.75	Student	0.04
One or two children	0.20	Pensioner	0.30
More than two children	0.04	Long-term unemployed	0.06
Age cohort by year of birth		Other	0.01
$<$ 1924	0.05	**Region of residence**	
1924–1926	0.03	Uusimaa	0.27
1927–1929	0.03	Varsinais-Suomi	0.09
1930–1932	0.03	Satakunta	0.05
1933–1935	0.03	Kanta-Häme	0.03
1936–1938	0.04	Pirkanmaa	0.09
1939–1941	0.04	Päijät-Häme	0.04
1942–1944	0.04	Kymenlaakso	0.04
1945–1947	0.06	South Karelia	0.03
1948–1950	0.06	Etelä-Savo	0.03
1951–1953	0.06	Pohjois-Savo	0.05
1954–1956	0.06	North Karelia	0.03
1957–1959	0.06	Central Finland	0.05
1960–1962	0.06	South Ostrobothnia	0.03
1963–1965	0.06	Ostrobothnia	0.03
1966–1968	0.05	Central Ostrobothnia	0.01
1969–1971	0.05	North Ostrobothnia	0.06
1972–1974	0.04	Kainuu	0.02
1975–1977	0.04	Lapland	0.04
1978–1981	0.04	East Uusimaa	0.01
$>$ 1981	0.08	Åland	0.01

Panel D: Macro-level control variables	Mean	Median	St. dev.	Min	Max
Short-term interest rate	0.03	0.03	0.02	0.01	0.06
Volatility of short-term interest rate	0.14	0.13	0.14	0.00	0.60
Unemployment rate	0.09	0.08	0.02	0.06	0.15
Inflation	0.02	0.02	0.01	0.00	0.04
Real GDP growth	0.02	0.03	0.03	-0.08	0.06
Real house price change	0.03	0.02	0.05	-0.08	0.21
Real share price change	0.09	0.07	0.33	-0.45	0.87
Volatility of share prices	0.40	0.14	0.53	0.02	1.81
Maturity of new housing loans	16.31	16.00	3.91	11.00	21.00