Determinants of US Household Debt: New Evidence from the SCF

Rafael Wildauer

May 2016

PKSG

Post Keynesian Economics Study Group

Working Paper 1608

Determinants of US Household Debt: New Evidence from the SCF

Abstract: This paper investigates the factors driving US household borrowing up to 2007. Two popular explanations are tested: First, the expenditure cascades hypothesis based on the assumption of debt-financed expenditures driven by an increasingly polarised distribution of income ('keeping up with the Joneses') and second, the hypothesis of Minskyian households which identifies climbing real estate prices as the decisive factor in household debt accumulation (re-mortgaging in order to cash in on capital gains and rising loan-to-value ratios in property purchases). This paper develops a method for obtaining individual household borrowing figures despite the lack of a panel structure from the Survey of Consumer Finances (SCF); it is the first to use the high quality information the SCF provides to investigate the impact of shifts in the income distribution and asset prices on household borrowing. The findings indicate that it is the interaction between the concentration of income at the top of the distribution and rising real estate prices which explains a large fraction of the increase in household borrowing prior to 2008. Therefore, neither the expenditure cascades hypothesis nor the hypothesis of Minskyian households are, in isolation, sufficient in order to understand household debt accumulation and thus the paper calls for a synthesis: future research should analyse the role of the distribution of income and asset prices together rather than separately.

Keywords: household debt, expenditure cascades, wealth effects, Minsky

JEL classifications: D12, D31, E03, E12

This paper benefited from funding through the INET grant *Rising Inequality as a Structural Cause of the Financial and Economic Crisis.* I thank the following persons for helpful comments and discussions: Engelbert Stockhammer, Javier López Bernardo, Chris Stewart, Michael Carr, Vince Daly and participants of the 18th FMM annual conference. I also thank the SCF staff for providing detailed documentation of the survey and being incredibly fast in answering any questions I had. All remaining errors are mine.

Rafael Wildauer

Kingston University, Penrshyn Road, KT1 2EE Surrey, UK

r.wildauer@kingston.ac.uk

1 Introduction

The financial crisis demonstrated the important role of debt and especially private sector debt for macroeconomic dynamics. Up to 2007 it was widely believed in neoclassical economics that money and credit are mere nominal phenomena and that fluctuations of real quantities such as GDP can be studied by abstracting from the nominal sphere of the economy. Since then the 'credit cycle' has attracted increasing attention in academic and policy debates. Empirical research carried out since the crisis showed that there is an important link between debt and macroeconomic fluctuations. Credit booms for example are a valuable predictor of financial crises (Schularick & Taylor 2012; Borio 2014; Eichengreen & Mitchener 2003) and expanding credit to the household sector and especially to low quality borrowers will eventually trigger mass defaults and adverse growth effects (Bezemer et al. 2014; Mian & Sufi 2009) when asset prices start to decline[1].

While there is a consensus that private debt in general and household debt in particular has to be taken seriously in macroeconomic analysis, there is no consensus about the key drivers of credit booms. Most specifically, why do households decide to take on debt levels which they eventually fail to handle? The textbook version of the life-cycle consumer would use debt only to smooth consumption over her lifespan and never default on it. Thus explanations going beyond the standard consumption model are needed. This paper will investigate two popular explanations for why US households became heavily indebted prior to the financial crisis.

The first focuses on increasing income inequality as a cause for household borrowing: in an environment of well above average income growth for the top tail of the distribution (see Figure A1 in the Appendix), those individuals who fall behind will take on debt in order to maintain spending levels. The maintenance of spending levels relative to higher income classes is important for lower class households because social status is to a great extent perceived through one's expenditures ('keeping up with the Joneses'). Beginning with relatively rich households which fall behind even richer peers, the process of status-driven debt accumulation cascades down the income distribution. The result is a heavily indebted and heavily spending household sector. Thus this so called *expenditure cascades hypothesis* provides a potential explanation of US household behaviour prior to the 2007 crisis. Frank (1985) and Frank et al. (2014) who trace this idea back to Veblen (1899) and Duesenberry (1949) were

---

[1] The less sharp decline in property prices and less aggressive lending to low income households compared to the US are the two main reasons why the UK so far managed to cope with household debt levels well above 100% of disposable income.

among the first to make that argument. Belabed et al. (2013), Kapeller and Schütz (2014), Kim (2013) and Ryoo and Kim (2014) are recent theoretical models of status driven debt accumulation.

The second explanation of household indebtedness emphasizes the role of rising property prices and homeownership rates. Homeownership is the socially desirable norm in the US which is actively supported by policies such as the government-sponsored enterprises. Thus the dream of owning a house coupled with eased access to credit for low income borrowers due to financial deregulation boosted house prices and ownership rates in the 1990s and 2000s. While in 1989 63.9% of US households owned their primary residence, in 2004 already 69.1%[2] did and house prices rose 115% during the same period[3]. With many US households already being owners, rising property prices resulted in substantial capital gains for a large part of the population. Increasingly optimistic households assumed the high and increasing property price level to be a permanent phenomenon and used (second) mortgages to 'realize' these gains. Furthermore, taking on larger mortgages relative to one's income to buy a property became reasonable under the perception of increasing or at least permanently high property prices. The literature on wealth effects (Paiella 2009; Cooper & Dynan 2014) in general and Mian & Sufi (2011) and Greenspan and Kennedy (2008) in particular provide empirical backing for this line of reasoning which will be labelled the *Minskyian households hypothesis* as rising asset prices over a long period of relative stability lead to an over-accumulation of liabilities[4].

With spectacular increases in US property prices and record levels of income inequality prior to the financial crisis, both hypotheses seem to be logically consistent with the rise in US household debt. This paper is the first to assess the explanatory power of both explanations in a unified framework. In order to do so the paper develops a method for deriving annual changes in household liabilities from the Survey of Consumer Finances (SCF), despite that dataset's lack of a panel structure. Having such a measure at hand makes the high quality information on households' income and households' balance sheets in the SCF useable for an econometric study of the determinants of household borrowing.

The paper contributes to the literature in three ways: first, deriving individual annual household borrowing from the repeated cross section data of the SCF allows the author to investigate the

[2] Calculations based on the Survey of Consumer Finances.
[3] The S&P / Case-Shiller U.S. National Home Price Index stood at 73.62 in January 1989 and at 158.67 in December 2004 (https://research.stlouisfed.org/fred2/series/CSUSHPINSA).
[4] Minsky's *Financial Instability Hypothesis* (Minsky 1978) explains financial crises as the result of unsustainable debt accumulation in the corporate sector due to increasing asset prices and a period of relative tranquillity which breeds instability by encouraging overly optimistic and risky behaviour. The term 'Minskyian households' reflects this idea that rising asset prices provide the necessary collateral and confidence for economic actors to accumulate debt up to unsustainable levels but applies it to the household sector instead of the corporate sector as in Minsky's original work.

accumulation of household debt based on the best available survey data on household finances. The unmatched advantage of the SCF compared to other US household surveys is that it relies on tax information in order to adequately capture the top tail of the income and wealth distribution and thus is the only survey which deals in a convincing way with non-observation and differential non-response problems (a more detailed discussion is provided in section 3.1). In addition the SCF provides the most detailed breakdown of the household balance sheet among available surveys and since existing liabilities and assets are crucial in determining current borrowing, not having this kind of information most likely yields biased results. Second, the paper brings together two powerful explanations of US household borrowing by assessing the impact of income inequality and rising property prices in a unified model. Doing so not only avoids potentially serious omitted variable problems but also allows a direct comparison of the effect size of both hypotheses. Thus the paper goes beyond simply reporting the statistical significance of the estimated effects. Third, the paper is the first to separately analyse the groups of borrowing and non-borrowing households. Fundamentally different results for the two groups demonstrate that borrowing and debt repayment decisions are not symmetric. Thus separating them in the analysis yields a better description of household behaviour which is reflected in an improved model fit.

Based on the newly developed measure of household borrowing the paper finds evidence supporting the expenditure cascades hypothesis conditional on homeownership. This means that households are only able to engage in status driven debt-financed expenditures if they have access to collateral. Households borrowed $963 billion[5] more in 2004 compared to 1995 due to the increase in residential real estate prices and the rise of top incomes. In addition the paper also finds that rising real estate prices strongly contributed to household indebtedness via home purchases: households borrowed $584 billion more in 2004 compared to 1995 due to rising costs of purchases. These findings demonstrate that, individually, neither the expenditure cascades hypothesis nor the hypothesis of Minskyian households are sufficient to explain household debt accumulation prior to 2007. Instead a unified approach is needed which allows for both mechanisms.

The rest of the paper is organized as follows: section 2 discusses the related literature, section 3 introduces the data and develops the method used to compute the change in household liabilities from the cross sectional dataset. Section 4 presents the econometric model as well as the definitions of

[5] Expressed in 2013 prices.

reference groups. Section 5 presents the estimation results. Section 6 reports aggregate level effect size computations and the final section concludes.

2 Determinants of Household Borrowing

This section reviews the literature on the expenditure cascades hypothesis and the hypothesis of Minskyian households as explanations for rising household indebtedness. The starting point and point of reference is the life-cycle model of consumption which suggests that households borrow in order to smooth out life-time consumption as a reaction to transitory income and wealth shocks (Romer, 2012; Attanasio and Weber, 2010). As the goal of this paper is to explain the sustained increase in household debt-to-income ratios since the early 1990s, transitory shocks cannot be a valid explanation because borrowing as a reaction would only be transitory as well. In contrast permanent income shocks in a life-cycle model of consumption imply permanent changes of consumption but do not affect borrowing. A permanent reduction in income will result in lower consumption expenditures as rational households would not take on debt as a substitute for income because higher debt levels would reduce future consumption or violate the budget constraint and thus the assumption of no default. Permanent shocks to wealth and most importantly housing wealth in contrast, can have a lasting effect on borrowing behaviour. If households were to consume some part of their real estate capital gains they would have to either sell the entire property or, more conveniently, take on an additional mortgage. Two important implications arise: First, the life-cycle model relies on rational consumers which are able to distinguish between transitory and permanent changes to income and wealth. Second, even permanent changes to income or wealth should only lead to small changes in consumption or borrowing because these gains are consumed over the entire life-time of the household[6]. If the simple model outlined so far is extended by introducing credit constraints, constrained households would react to permanent increases of their housing assets by additional borrowing. These effects would be large even with long optimization horizons. Thus the life-cycle model predicts 'pure' wealth effects to be small due to long optimization horizons whereas wealth effects due to credit constraints might result in large changes of household consumption and borrowing. In contrast to permanent income changes, a change in the distribution of income has no impact on aggregate spending or borrowing because agents are assumed to be homogeneous in their preferences.

With the collapse of the US housing market beginning in 2006 and the following financial crisis the Post-Keynesian economist Hyman P. Minsky and his Financial Instability Hypothesis (Minsky, 1978)

[6] However as Carroll (2001) points out assuming (realistic) planning or optimizing horizons of two to three years instead of two to three decades, results in much larger current effects of permanent changes.

have attracted renewed interest. The powerful notion that periods of stability breed instability by encouraging riskier financial practices was briefly recognized as a theory able to explain major crises in contrast to standard New-Keynesian or Real Business Cycle models. Minsky's work is relevant in the context of household debt accumulation because it provides an explanation of why agents take on too much debt and eventually default. However Minsky's focus in his own work was on the behaviour of corporations and most of the literature trying to formalize the Financial Instability Hypothesis only includes corporate debt in their models[7]. Ryoo (2015) is currently the only paper which applies Minsky's ideas to the household sector and develops a model of endogenous instability[8]: Households borrow in order to finance additional consumption and housing expenditures. Increasing housing demand leads to capital gains which feeds back into additional borrowing due to the improved balance sheet position of households. The result is a prolonged boom followed by a severe recession. The key point is that even if house price appreciations are driven by a growing debt-load and thus cannot be permanent, they encourage household borrowing. Thus a Minskyian model of the household sector does not depend on the questionable distinction between transitory and permanent shocks but predicts positive wealth effects on consumption and offers a powerful tool for understanding households' debt accumulation prior to 2007.

Beyond the Financial Instability Hypothesis, Post-Keynesian economic theory provides another rationale for the role of asset prices in household borrowing. The tradition of stock-flow-consistent (SFC) modelling, going back to Tobin (1969, 1982) and pioneered by Godley and Lavoie (2007) emphasizes the importance of connecting all flows in an economic model to corresponding stocks at the end of the period. An important feature of these models is that economic actors are anchored by so-called stock-flow norms (Godley and Lavoie, 2007, p.75). In the case of households this means they form consumption decisions in line with achieving a desired net-wealth-to-income ratio. These desired wealth-to-income ratios reflect propensities to save and consume out of disposable income and net-wealth respectively and thus are influenced by precautionary saving motives as well as social norms. An increase in (real estate) wealth will lead to a reduction in household savings if the target wealth-to-income ratio is not yet reached and to dissaving in the form of taking on additional debt if the target

[7] For recent examples of corporate Minsky models see Bhattacharya et al. (2015), Keen (2013), Charles (2008), Dos Santos (2005) and the references therein.

[8] Kapeller and Schütz (2014) and Ryoo and Kim (2014) also adopt Minskyian ideas to the household sector but combined that with the expenditure cascades hypothesis and thus these authors are discussed in the next subsection.

is surpassed. On the individual level this prediction translates into positive wealth effects for asset holders and negative wealth effects for non-holders.

With respect to the distribution of income Post-Keynesian economic theory assumes different propensities to consume across the income distribution due to empirically higher saving rates towards the top of the distribution (Dynan et al., 2004). In models without a personal distribution of income a higher propensity to consume out of wages than out of profits is assumed. These assumptions imply that a relative increase of top or profit incomes will lead to a decrease in aggregate consumption spending (Hein and Vogel, 2008; Onaran et al., 2011). A priori it is not clear whether such a redistribution would also increase household borrowing. It would do so if households want to sustain past consumption levels and are not liquidity constrained.

A recent paper emphasizes the importance of the relative position within the income distribution for household spending and borrowing decisions (Frank et al., 2014). Tracing that idea back to Veblen (1899) and Duesenberry (1949), Frank (1985) and Frank et al. (2014) argue that consumption is not only about satisfying one's needs but also about signalling and demonstrating social status ('keeping up with the Joneses'). They use that argument to explain household indebtedness: If status comparison manifests in an upward looking manner, rising income inequality will trigger a debt-financed consumption spree which cascades down the income distribution. The intuition is that households which experience slower income growth relative to their peers at the top of the distribution will start to borrow in order to keep up with the lifestyle of their peers. Since also debt-financed expenditures are relevant for status signalling, the process of borrowing in order to keep up with a life style of richer peers cascades down the income distribution. The pioneering data work of Atkinson, Piketty and co-authors (Atkinson et al., 2011; Piketty, 2014), revealing the sharp rise in income inequality over the last decades, contributed to the popularity of this idea and resulted in a number of papers which develop it in formal theoretical models (Belabed et al., 2013; Kapeller and Schütz, 2014; Ryoo and Kim, 2014; Cardaci, 2014; Nikiforos, 2015).

The empirical literature which seeks to quantify consumption wealth effects largely confirms the existence of such wealth effects (Lehnert, 2004; Bostic et al., 2009; Cooper, 2012; Salotti, 2012; Juster et al., 2005; Dynan, 2010). However since consumption wealth effects only provide indirect evidence for the hypothesis that house price appreciations drive household debt trends, this literature is not discussed here and interested readers are referred to the recent surveys by Cooper and Dynan (2014) and Paiella (2009). All studies which directly assess the impact of asset prices on household borrowing report a positive direct relationship. Hurst and Stafford (2004) for example use PSID data and find that liquidity constrained households are more likely to refinance their mortgage and extract equity compared to unconstrained households. Even though Hurst and Stafford (2004) focus on equity extraction as a reaction to transitory shocks they provide some evidence that house price appreciations are a potential factor which drives up household borrowing due to equity extraction. Haurin and Rosenthal (2006) are interested in the extent to which home equity extraction can explain consumer spending. Their regression of housing capital gains and other controls on household debt yields a positive and statistically positive effect based on data from the SCF and the National Longitudinal Survey of Youth. Dynan and Kohn (2007) have access to the unpublished version of the SCF and argue that house price appreciation was a major reason for rising household indebtedness. Finally Mian and Sufi (2011) analyse the variation across metropolitan statistical areas in household debt growth for homeowners between 2002 and 2006. They demonstrate that house prices matter for household borrowing independent of any potentially omitted third factors such as productivity.

There are several authors relying on aggregate data to test the expenditure cascades hypothesis (Cynamon and Fazzari, 2015; Stockhammer and Wildauer, 2015; Behringer and Treeck, 2013) yielding mixed results. Since this paper relies on survey data this review focusses on the much richer literature which uses micro data to investigate the role of the relative position of an individual with respect to its peers. A closer look at this literature reveals a striking degree of heterogeneity in the way reference groups are defined. In contrast to upward looking status comparison as required by the expenditure cascades hypothesis, many authors define peer groups based on household characteristics such as age and education (Maurer and Meier, 2008), the region the household lives in (Ravina, 2007), or just compare different categories of consumption goods (Heffetz, 2011). All of these studies do find evidence of (at least) modest positive effects of peer group consumption on household spending. Two recent studies explicitly adopt upward looking preferences and thus are important reference points for this paper. Bertrand and Morse (2013) model consumption for the bottom 80% of households depending not only on individual characteristics such as income or age but also on the consumption of the average household in the 80th percentile of the income distribution. Income percentiles are computed for each US State. Using data form the Consumer Expenditure Survey (CEX) from 1980 to 2008 they find a positive effect of consumption expenditures of rich households on lower income households and interpret these findings as evidence in favour of the expenditure cascades hypothesis. Carr and Jayadev (2014) model the relative position of the individual household in more detail. They use a measure of relative income defined as the proportion of households which are richer than household *i*. Using data from the Panel Study of Income Dynamics (PSID) from 1999 to 2009 Carr and Jayadev (2014) find positive and statistically significant effects of their relative income measure on household borrowing growth and interpreted it as evidence in line with the expenditure cascades hypothesis.

Overall the empirical evidence seems to strongly support the hypothesis of Minskyian households and the importance of asset price dynamics for household borrowing while evidence for the expenditure cascades hypothesis is more mixed. With respect to the latter studies relying on aggregate data are not conclusive about the existence (Behringer & Treeck 2013; Perugini et al. 2015) or non-existence (Bordo & Meissner 2012; Stockhammer & Wildauer 2015) of a positive link between income inequality and (household) borrowing. Furthermore as pointed out by Bertrand and Morse (2013) the existence of a positive link between relative income and household debt is not sufficient for demonstrating that the expenditure cascades hypothesis is able to explain a major proportion of the increase in household debt prior to 2007. Bertrand and Morse (2013) report that based on their results the increase in income inequality from the 1980s onward only explains 1.3 percentage points of the fall in the NIPA savings rate and thus by far is not enough to be regarded as the main cause for increased household indebtedness. In addition despite strong evidence for the important role housing wealth plays for household spending and borrowing decisions, the studies seeking to test the expenditure cascades hypothesis as an explanation for rising household debt, generally ignore the role of assets. Another issue with survey based papers is that both Bertrand and Morse (2013) as well as Carr and Jayadev (2014) use data which only provides limited detail about the top of the income distribution. In particular there is no study on the expenditure cascades hypothesis using the SCF. This is an important drawback of existing studies since income inequality in the US rose due to strongly growing top incomes and thus not including those top households in one's sample will yield misleading results. The next section elaborates on that latter point by discussing the specific advantages of the Survey of Consumer Finances provides for studying the links between household borrowing, asset prices and the distribution of income.

3 Data: The Survey of Consumer Finances

3.1 The SCF and Non-Observation and Non-Response Bias in Surveys

The paper relies on data from the Survey of Consumer Finances between 1995 and 2007. The SCF is a triannual survey undertaken by the Federal Reserve Board (FRB) in cooperation with the Statistics of Income Division (SOI) of the Internal Revenue Service. In each wave between 4,299 (1995) and 4,519 (2004) observations are included. The SCF focuses on household income, assets and liabilities and represents the most detailed source of information about household balance sheets and especially high income household balance sheets. This latter benefit of the SCF stems from the dual-frame sample design consisting of an area-probability sample and a list sample.

About two thirds of the completed cases stem from the area probability sample which is built in three stages. In the first stage metropolitan areas and rural counties are stratified by a variety of characteristics and primary sampling units (PSUs) are selected proportional to their population. At the second stage subareas are selected within PSUs again based on stratification. At the third stage random samples are drawn within these subgroups. This design ensures that each household in the sample has the same probability of being selected. Thus the area-probability sample covers broadly distributed household characteristics, while at the same time limiting the cost of data collection due to the stratified design.

Nevertheless there are two important shortcomings. First, due to the highly skewed distribution of household characteristics like income and wealth an enormous sample size would be needed to gain sufficient observations of wealthy households to obtain an adequate picture of the distribution of these characteristics. The cost of obtaining such a sample would be substantial. For example the relatively large Consumer Expenditure Survey (CEX) is based on less than 14,000 observations representing 124 million households in 2012. This corresponds to a sample of only 0.11‰ of the target population. Even a large survey like the CEX is not sufficient to adequately represent the highly skewed income distribution because there are not enough observations from the top end of the distribution part of the sample. This fact becomes obvious if one compares the average pre-tax income in the 10th decile calculated from CEX data with data published by the IRS. According to the IRS, average income in the top decile in 2011 was $2.1 million[9] compared to $229,771[10] in 2014 according to CEX data. So despite the timing gap and different income concepts the difference is striking. This is the problem of non-observation also demonstrated by simulation exercises in Kennickell (2005) and Eckerstorfer et al. (2015).

Second, there is evidence that the likelihood of participation in (wealth) surveys is negatively correlated with household wealth itself (Kennickell & McManus 1993; Singer 2006; Kennickell 2008) known as systematic non-response. One can think of several reasons why rich households are less willing to participate, ranging from greater concerns about data protection to higher valuation of the time needed to complete the interview. However if non-response is systematically related to

[9] IRS data based on "1979 Income Concept" from SOI Bulletin article - Individual Income Tax Rates and Tax Shares, Table 7.

[10] Summary Table 1110 (http://www.bls.gov/cex/2014/combined/decile.pdf).

household characteristics like wealth or income, any estimates based on samples which do not address this problem will be biased.

In order to deal with these two problems, the SCF relies on the second component of the dual-frame sample: the list sample. The purpose of the list component is to over-sample wealthy households. In order to be able to identify such households prior to data collection an external data source is needed. Due to a cooperation under extremely strict privacy conditions, the list sample is built by using a sample of income tax records. Based on that information the assets of tax units are estimated by capitalizing asset related income streams[11]. Then observations are selected in a two stage process. First, only observations in PSUs which were already selected for the area-probability sample are considered in order to keep the costs of the survey in check. Second, households are stratified based on percentiles of the estimated asset holdings (SCF documents refer to it as a *wealth index*). Then samples are drawn from each strata and strata corresponding to higher estimated net wealth are sampled at higher rates. The details about the sample design of the list sample are not publicly available in order to ensure anonymity of all participants.

Due to non-observation and non-response problems, surveys which do not pay as much attention to their sample design as the SCF does and in particular surveys which do not apply oversampling techniques suffer from serious shortcomings and are in general not able to provide an adequate picture of the income or wealth distribution. Kennickell (2008) demonstrates the contribution of the list sample for the SCF: While net worth at the 90th percentile only increases by 5.5% due to the additional information from the list sample, at the 99th percentile the increase is 74%. Vermeulen (2014) and Eckerstorfer et al. (2015) demonstrate the impact of such a shortcoming for other countries and surveys. The latter paper estimates that aggregate net wealth is underestimated by about one quarter due to non-observation and non-response problems in the case of Austria in the Household Finance and Consumption Survey. Since the current paper investigates the expenditure cascades hypothesis which crucially relies on strong top income growth, taking non-observation and non-response problems seriously is important. By using income tax information, the SCF deals with both issues in a convincing way and thus should be the first choice when investigating phenomena related to the distribution of income.

Like any household survey the SCF faces the problem that participants are unable or unwilling to answer some of the questions. Leaving these missing values as they are would only allow complete

[11] Kennickell (2000) provides a detailed discussion of the details and the two different models used.

case analysis at the cost of losing a considerable amount of information. Instead the SCF imputes missing values based on statistical modelling, referred to as multiple imputation (Rubin 1987; Kennickell 1998). In order to reflect the uncertainty associated with statistical modelling of missing information the process is repeated 5 times, yielding 5 separate datasets, so called *implicates*. Working with SCF data requires to analyse each dataset and then combine the results based on *Rubin's rules* which in the simplest form state that point estimates should be averaged and in order to obtain proper standard errors for these estimates one has to take into account the variation within implicates as well as between implicates.

3.2 The Problem of the Missing Time Dimension in the SCF

A key limitation when using the SCF in econometric analysis is the missing time dimension of the dataset due to its design as a repeated cross section. Unlike with a panel one cannot follow the same household through time. However the detailed information the survey collects about the credit history of each participant allows to infer by how much the outstanding liabilities changed within the year of the survey. Thus one can infer whether an individual household took on debt in the year of the survey, paid back already existing liabilities or pursued neither of these activities. Creating such a measure is key for answering the question whether households which are exposed to rich peers take on debt in order to keep up in spending with these peers. Without knowing whether and by how much households take on debt in a given year one would have to rely on the total amount of outstanding liabilities. The stock of debt however is strongly influenced by past decisions, most importantly real estate purchases, and thus is not a good measure of whether households which are currently exposed to richer peers take on debt. This section presents the steps involved in creating a measure of the within year change in household liabilities.

In order to understand how the change of an individual household's debt level is constructed one has to keep in mind that the SCF covers 10 different debt categories. Participating households are asked about their outstanding liabilities with respect to mortgages (primary residence as well as other properties), lines of credit, credit on land contracts, consumer loans, credit cards, car and vehicle loans, education loans, loans against pension plans and other loans. Based on the specific information the survey collects about all these categories, the paper is able to construct a measure of how much that liability changed within the last year. For four categories there is not enough information to make such an inference: credit card debt, credit lines, loans against land contracts and loans against pension plans. For each household the changes in each category are summed up to obtain the total change in household i's level of debt.

An example will be the best way to demonstrate how it was done. Let's consider the outstanding amount on the first mortgage on the primary residence for household i, which will be denoted $D_{i,t}^{M1}$. In order to understand by how much the outstanding amount changed, households are characterised in two steps. The first step distinguishes whether the mortgage was taken out in the current year ($tB = year$) or prior to the year of the interview ($tB < year$) and how the money was used (use). The use of the money is summarized in four categories: refinancing an earlier mortgage ($use = ref$), extract equity from the property ($use = extr$), extract equity and refinance ($use = extr + ref$) and mortgage taken out with no prior loan ($use = 0$).

In case the mortgage was taken out in the year of the interview ($tB = year$) the information on the use of the loan becomes crucial. If the mortgage was used to refinance an earlier credit ($use = ref$, <u>case 1</u>), the change of that mortgage is defined as the difference between the amount currently outstanding ($D_{i,t}^{M1}$) and the amount initially borrowed (B_i^{M1}): $\Delta D_{i,t}^{M1} = D_{i,t}^{M1} - B_i^{M1}$. The rationale for this definition is that the amount initially borrowed does not constitute a new liability but replaced an already existing one. Thus within this year only the difference between the refinanced amount and the currently outstanding amount represents an actual change in outstanding liabilities. It is important to note that depending on whether the amount initially borrowed is smaller, bigger or equal compared to the amount currently outstanding, the resulting change will be positive, negative or zero. The case of $B_i^{M1} < D_{i,t}^{M1}$ is interpreted as household i being behind on payments and accumulating overdue interest which is added to the total amount due and thus the currently outstanding amount exceeds the amount initially borrowed.

In contrast if the mortgage was used to extract equity from the residence ($use = extr$, <u>case 2</u>), the change in the amount outstanding is defined as the amount extracted (ex_{it}): $\Delta D_{i,t}^{M1} = ex_{it}$ which represents newly accumulated debt. If the mortgage was taken out to extract equity and to refinance an earlier loan ($use = extr + ref$, <u>case 3</u>) the change is defined as the amount extracted plus the difference between the amount currently outstanding and the amount initially borrowed: $\Delta D_{i,t}^{M1} = ex_{it} + D_{i,t}^{M1} - B_i^{M1}$. The difference between the current amount and the initial amount is added because this difference represents the extent to which (new as well as already existing) debt was paid down. If the household had no prior loan or mortgage ($use = 0$, <u>case 4</u>) the change in debt is simply defined as the amount currently outstanding because the amount currently outstanding represents debt accumulated in the current period: $\Delta D_{i,t}^{M1} = D_{i,t}^{M1}$. The reasoning is that in this case $D_{i,t-1}^{M1} = 0$ and thus the change in debt equals the amount currently outstanding.

In most cases however the households did not take out their mortgage in the year of the interview ($tB < year$). Thus for these households the task is to infer the amounts of principal repayment. In a second step households are distinguished based on whether the initial amount borrowed (B_i^{M1}) is bigger, equal or smaller than the amount currently outstanding ($D_{i,t}^{M1}$). In case the amount initially borrowed is bigger than the amount currently outstanding ($B_i^{M1} > D_{i,t}^{M1}$, case 5.1) it is assumed that these households are paying back their mortgage on schedule. Thus the change in debt equals the amount of principal repaid $P_{i,t}^{M1}$. Principal repayment is computed as the difference between the total regular or typical annual payment the household makes and the implicit interest payments based on the reported interest rate for the loan and the currently outstanding amount. Thus in some cases $P_{i,t}^{M1}$ might be negative which corresponds to a situation in which actual payments are less than the interest due. This means that unpaid interest accumulates and increases the outstanding amount and thus the principal 'repayment' is negative. In both cases the change in the outstanding liability is defined as the negative principal repayment $\Delta D_{i,t}^{M1} = -P_{i,t}^{M1}$.

If the amount initially borrowed and currently outstanding are equal ($B_i^{M1} = D_{i,t}^{M1}$, case 5.2) it is interpreted as the household being in a period of no principal repayment as part of an interest-only agreement. The household only pays interest but no principal for a certain period at the beginning of the contract. Thus the outstanding amount on that mortgage did not change in the year of the interview: $\Delta D_{i,t}^{M1} = 0$. Households which report an amount currently outstanding larger than the amount initially borrowed ($B_i^{M1} < D_{i,t}^{M1}$, case 5.3) are assumed to have fallen behind in payments. Due to the accumulated not paid interest the outstanding balance increased and now exceeds the amount initially borrowed. Only a few households report such a constellation, as expected. It is important to note that also households falling into case 5.1 can be behind on payments. In that case however they underwent an earlier period of repayment and struggled with payments only for a relatively short period such that accumulated interest does not exceed earlier repayments. The change in the mortgage balance for households falling into case 5.3 is also defined as the negative principal repayment: $\Delta D_{i,t}^{M1} = -P_{i,t}^{M1}$. In almost all cases households in this category are paying back principal indicating that they are back on some payment plan. This seems convincing because it is unlikely that a lender will tolerate accumulated unpaid interest and the household still being behind in payments. One expects a lender in such a situation to seek control of the house and recover the lent money. Case 5.3 is also compatible however with reverse mortgage where, especially elderly borrowers, take out a mortgage to finance retirement expenses and pay back the entire mortgage and all accrued interest after the property is sold after their death. Also negative amortization (negAM) contracts under which neither principal nor interest payments are made for a short period, normally not more than 5 years, at the beginning of the mortgage contract are consistent with a payment pattern of case 5.3. All cases are summarized in Table 1 which also indicates the number of observations falling into each case in the 2004 wave in the third implicate. 1,984 out of 4,519 households reported an outstanding first mortgage on the primary residence in 2004 (implicate 3). 1,466 of these fell into case 5.1 and thus form the biggest group. The overwhelming majority of households in this group (1,431) is paying down debt. Thus even if at the aggregate level household debt increases, it is only a small portion of households which actually take on debt in a given year. While this is not a surprising result one has to keep it in mind for econometric modelling.

The categorization of households along the two steps is visualized in Figure A2 in the Appendix. By applying a similar logic as in this example to six more debt categories (second and third mortgages on the primary residence, mortgages on other residential property, consumer loans, car and vehicle loans, education loans, other loans for property purchase and home improvements), the paper constructs a measure of the total change in household i's liabilities (ΔD_{it}) by summing up the changes of the individual categories:

$$\Delta D_{it} = \Delta D_{i,t}^{M} + \Delta D_{i,t}^{OM} + \Delta D_{i,t}^{CL} + \Delta D_{i,t}^{VL} + \Delta D_{i,t}^{EL} + \Delta D_{i,t}^{OL} \quad (1)$$

where $\Delta D_{i,t}^{M}$ is the change in all outstanding mortgages on the primary residence (first, second and third: $\Delta D_{i,t}^{M} = \Delta D_{i,t}^{M1} + \Delta D_{i,t}^{M2} + \Delta D_{i,t}^{M3}$), $\Delta D_{i,t}^{OM}$ is the change in outstanding mortgages on other residential properties, $\Delta D_{i,t}^{CL}$ is the change in outstanding unsecured consumer loans, $\Delta D_{i,t}^{VL}$ is the change in outstanding car and vehicle loans, $\Delta D_{i,t}^{EL}$ is the change in outstanding education loans and $\Delta D_{i,t}^{OL}$ is the change in other outstanding liabilities. The definitions and the steps undertaken to define the change in outstanding liabilities for these other categories can be found in Appendix B.

Table 1: Changes in the outstanding amount of the first mortgage on the primary residence

case	step 1	step 2	definition	N	interpretation
1	tB=year ∧ use = ref		ΔD^{M1} = D-B	168	Mortgage taken out to refinance.
2	tB=year ∧ use = extr		ΔD^{M1} = ex	19	Mortgage agreement altered to extract equity.
3	tB=year ∧ use = extr+ref		ΔD^{M1} = ex+D-B	39	Equity extracting and refinancing.
4	tB=year ∧ use = 0		ΔD^{M1} = D	109	No prior loan thus new debt.
5.1	tB<year	B > D	ΔD^{M1} = -P	1,466	Household is repaying.
5.2	tB<year	B = D	ΔD^{M1} = 0	164	No repayment yet. Probably interest only period.
5.3	tB<year	B < D	ΔD^{M1} = -P	19	Fallen behind payment schedule and interest accumulated.
				1,984	

tB corresponds to X802, use to X7137, ex to X7138, B to X804, D to X805 and P is defined as P=X808-(X816/10000*X805) where X808 are transformed to annual payments and X813 is used if X808 is not reported. Observations falling into cases 3 and 4 are dropped prior to 2004 because ex is not observed.

3.3 The SCF compared to other US Data Sets

The raw dataset covers the years 1995 to 2007 in three year intervals and consists of 21,982 observations. Since the SCF is a multiple imputed dataset with 5 imputations the total number of observations in the raw dataset is 109,910. Observations with zero income are dropped (810) as well as observations with a negative beginning of period debt level (2,460). The SCF does not provide information on how much additional money was borrowed on the first mortgage on the primary residence in case of equity withdrawals in the surveys prior to 2004. Thus in 1995, 1998 and 2001 a total of 748 households are excluded, falling into that category[12]. In addition 294 observations which exhibit change-in-debt-to-income ratios below -250% and above 500% are excluded from estimation. This leaves a final sample of 102,958 observations corresponding to roughly 20,592 observations per implicate. Sample sizes reported in regression tables refer to the number of observations per implicate.

The fact that the SCF is seldom used for econometric modelling justifies to briefly point out the most important differences compared to other US household surveys notably the Consumer Expenditure Survey (CEX) and the Panel Study of Income Dynamics (PSID). First, the SCF and CEX are repeated cross

[12] Households which extract home equity but do not have a mortgage prior to the equity extraction are still part of the sample.

sections[13]. While the CEX re-interviews households within the year, the SCF in general does not interview the same households twice except in 2009 when the 2007 sample was re-interviewed. In addition the SCF collects data in a three year interval while the CEX does so at quarterly and weekly intervals. In contrast the PSID is a panel which started in 1968 collecting annual information until 1997 and from 1999 on interviewing households every other year. The PSID was initially designed to study the dynamics of income and poverty and thus started out by oversampling low income families. While the SCF and CEX rely on new samples on an annual and triannual basis, the PSID sample only changes due to births, deaths, marriage occurring in the families originally sampled in 1968 and attrition which means households stop to participate. In 1997 a major sample adjustment took place in order to represent migration to the US. Second, the SCF focusses on income, assets and liabilities but contains almost no information on expenditures. In contrast the CEX focusses heavily on expenditures while the PSID focusses on low-income, poverty and health. It becomes clear that all three data sources were designed for very different purposes.

Based on this brief comparison of three major US household surveys four key advantages of the SCF are identified. First, using tax information to oversample wealthy households provides detailed information about the top tail of the income and wealth distribution. Second, by not going into detail about household expenditures the SCF is able to provide detailed information about households' balance sheets. Even though the CEX and PSID also collect information on assets and liabilities the SCF does it in a much more detailed way. Third, using a new sample for each wave minimizes the problem of attrition. Panel studies in general face the problem that if households drop out of the sample in a non-random way, the sample becomes unrepresentative over time. By drawing a new sample for each wave, repeated cross sections avoid this problem. Fourth, closely related to the previous point, by drawing new samples for each wave the SCF remains representative of the US population. While panel data sets can rely on non-sample information and adjust weights in order to correct for changes in the underlying population, repeated cross sections actually collect information which represents these changes.

These highly attractive features of the SCF come at a cost however. At least two disadvantages are obvious. The first and most important drawback of the SCF is the low frequency at which data is collected. Three year intervals are long periods of time and by definition the SCF only provides limited information about what is going on in between. It is unthinkable of only having reliable information about inflation or GDP growth every three years. Second, the design as repeated cross section prohibits

[13] Technically the CEX is a rotating panel but it follows households only over 5 quarters.

researchers from using panel data methods. Thus SCF data does not allow to control for (time-invariant) heterogeneity and also the dynamics of the household balance sheet cannot be studied in full detail at the individual level. With respect to the last issue, this paper however demonstrates that by using the rich information the SCF provides about the timing of borrowing decision that this problem can be solved at least partially. Third, the SCF does not contain information on household expenditures which would be necessary in order to investigate consumption wealth effects. Due to its focus on income and the household balance sheet however this seems to be a fundamental constraint due to the limited amount of time interviewers have when collecting information.

Despite these drawbacks, when it comes to studying household indebtedness and income inequality the SCF's advantages outweigh its limitations. Much better coverage of the upper tail of the income and wealth distribution and much more detailed information on household assets and liabilities compared to other household surveys are strong reasons to rely on the SCF for such a research question.

4 Estimating a Household-Debt-Accumulation-Function

The specification of a reduced form household borrowing equation relies on the predictions from economic theory discussed in section 2. Most importantly the variable of interest is the annual change in total outstanding liabilities (ΔD_{it}) and not the total stock of debt (D_{it}). The simple reason is that only the change is directly related to the flows of the current period whereas the stock depends on past decisions which are not observed. The change in liabilities is scaled by current household income (Y_{it}) in order to obtain a measure of borrowing which is comparable across the population. In contrast using the growth rate of outstanding liabilities looks appealing at first but debt growth rates are not useful indicators of household debt accumulation. Taking out new debt with very little outstanding liabilities results in misleadingly large growth rates even if the actual amount of new debt is quite small in absolute terms or relative to household income. Beyond that growth rates are not defined for households with no prior liabilities and using a log approximation is highly problematic because it does not work well for large rates[14]. In addition the amount of already existing liabilities is not a particularly relevant criterion in order to assess the extent of a household's indebtedness. Instead income represents the flow out of which the household has to repay principal and interest and thus yields a

[14] These are specific problems of household data. Aggregate data at the state or county level are less prone to these issues because the outstanding debt stock within the county or State most likely will be well above 0 and thus growth rates will yield an indicator of household borrowing which is comparable across counties or States. However in case of a high degree of heterogeneity in income dynamics, also growth rates at the aggregate level might be a misleading indicator.

more informative benchmark. This paper also refuses to use assets to scale household liabilities because asset prices tend to be volatile and in boom phases increase in line with liabilities exposing high leverage only after asset prices collapse. For all these reasons income is used to scale household borrowing. The exact definitions of all variables used, in line with the notation of the official documentation of the SCF, can be found in Table A1 in the Appendix.

When thinking about determinants of household debt dynamics the first category to consider is household income for obvious reasons: income is the main source of funding for most expenditures especially for consumption expenditures and thus influences to what extent borrowing is needed to achieve desired spending levels. Economic theory however differs about the impact of income on household borrowing. While the life-cycle model predicts that households borrow in order to smooth consumption, the implication of cross-section income variation is not immediately clear. If high-income households are more exposed to more regular but smaller shocks relative to their income like temporary declines in stock markets whereas low-income households face larger shocks relative to income like unemployment, one would expect a negative cross section effect of income on the household borrowing to income ratio. Put differently low income households need to smooth larger shocks (relative to their current income levels) and thus need to borrow more relative to current income. Post-Keynesian theory in general and Minsky in particular do not focus on the role of current income for borrowing purposes but emphasize the role of assets. The stock-flow-consistent tradition within Post-Keynesian economics argues that borrowing households are likely to have reached their wealth-to-income targets and thus borrow in order to consume part of that 'excess' wealth. Due to declining marginal propensities to consume towards the top of the income distribution this effect is smaller for high-income households and thus one expects a negative cross section effect. If one thinks about the expenditure cascades hypothesis a negative cross section effect is expected as well because, holding relative income constant, richer households need to borrow less for their desired status-spending.

If one thinks about the role of income on household borrowing in a cross section context it becomes clear that there is an important difference between borrowing households and non-borrowing households. Implicitly the previous paragraph focused on borrowing households. However since only a small proportion of households actually borrow in any current year but borrowing decisions are likely to result in a lengthy period of repayments, an asymmetry arises between the two groups for two reasons. First, if higher income households borrow less relative to their income, this implies that also their repayments are smaller relative to their income given that borrowing and repayment decisions are not asymmetrically influenced by income. However because borrowing is defined as an increase in outstanding liabilities ($\Delta D_{it} > 0$) and repayment as a decrease ($\Delta D_{it} < 0$), income will have a negative impact on borrowing relative to income but a positive impact on repayments. The reason is that the terminology 'smaller repayments' actually refers to a less negative change and thus to a bigger number, resulting in a positive cross section effect. When discussing the effects on the non-borrowing sample positive effects are equivalent to smaller repayments and negative effects to larger repayments. However due to these different signs, analysing borrowing and non-borrowing households in the same sample will yield meaningless averages. Second, there is no reason to believe that household characteristics such as income will influence borrowing and repayment decisions in a symmetric way. For example if high income households choose to have shorter repayment periods one can easily think of examples where the proportion of borrowing relative to income between high and low income households and the proportion of repayments (relative to income) between high and low income households is different. Thus separating borrowing and non-borrowing households will yield differently signed effects of income on household borrowing due to the first argument and even in absolute values these effects are likely to differ due to the second argument. Therefore this paper will investigate the two groups separately.

Beyond income the model incorporates measures of financial wealth (FW_{it}) and most importantly residential real estate assets[15]. The actual purchase of real estate is one of the most important reasons why households take on debt. For that reason this paper distinguishes between housing assets which existed before the beginning of the period and the value of residential real estate purchased in the year of the interview. Thus the measure of housing wealth used in the paper (HW_{it}) consists of the current value of the primary residence and any other residential real estate minus the value and any potential capital gains on residential real estate purchased in the current period. The value at the time of purchase (thus excluding potential capital gains within the year) of residential real estate obtained in the current period (REP_{it}) enters the empirical model separately. For exact definitions readers are again referred to Table A1 in the Appendix. Being able to include detailed information on asset purchases in addition to pre-period stocks is a major advantage of SCF data when it comes to assessing the explanatory power of the Minskyian households hypothesis. Since asset purchases are the most important reason to take on debt for the majority of households not being able to include this kind of information in one's analysis will result in severely biased estimates. Economic theory predicts financial as well as real estate assets to influence household borrowing due to the presence of credit constraints

[15] This means real estate purchased for investment purposes (buy-to-let) is not part of this housing wealth measure because the SCF does not report debt related to these properties separately but reports net values.

and 'pure' wealth effects. Thus for both measures one expects a positive cross section effect on borrowing relative to household income. Due to the negative sign of liability changes in the non-borrowing sample, the effect of housing wealth is expected to be negative in the non-borrowing sample while the effect of current real estate purchases is expected to be close to zero since there are simply not many households being able to purchase without taking on debt

Since the aim of the paper is to assess the role of relative income on household borrowing, the regression model also includes various measures of peer group income ($\tilde{Y}_{it}$). Several definitions are used and the details are discussed in the subsection below. For now it is sufficient to state that under the expenditure cascades hypothesis one expects those households being exposed to higher levels of peer group income, while holding their own income constant, to borrow more in order to keep up with the expenditures of that peer group and thus a positive effect. For the group of non-borrowing households one expects their repayments to be higher if they borrowed more in the past and thus a negative effect. Thus for similar reasons as in the case of income and housing wealth, most likely there is an asymmetry between borrowers and non-borrowers with respect to the expected signs.

Another explanation for household borrowing which is closely related to the expenditure cascades hypothesis is that households borrow in order to sustain consumption levels in a situation of declining income. The argument is that people are unwilling to cut once-achieved standards of living. It is much harder to adjust downwards than upwards. Thus the regression also controls for those households whose income is lower than in a normal year by means of an indicator variable ($dlinc_{it}$). If past expenditure levels are important for current spending decisions those households with abnormally low incomes should borrow more or repay slower. Another interpretation is that households with abnormally low incomes will borrow less (repay quicker) out of a precautionary motive. Beside the relative position within the income distribution Duesenberry stressed past consumption as an important reference point in his *Relative Income Hypothesis* (Duesenberry 1949). The (microeconometric) literature discusses such phenomena under the label of *habit formation* (Dynan 2000; Fuhrer 2000; Ravina 2007). Since the SCF does not have a panel structure the low-income dummy ($dlinc_{it}$) is the only way to control for such effects.

Finally outstanding liabilities at the end of the previous period (D_{it-1}) are part of the empirical model. Since for most households borrowing results in a lengthy period of repayments, the amount still outstanding is an important predictor of these payments. On the other hand, for borrowing households one expects that already highly indebted households are less likely to be able to borrow even more. One potential exception arises if one thinks about households as 'Minskian agents' whose finance structure becomes more and more risky over the boom period, eventually ending up in what Minsky labelled *Ponzi finance* which describes agents not able to repay principal nor interest out of current cash flows and who need to borrow for these payments. This approach predicts higher indebted households to borrow even more. Most likely however such behaviour is observed only over short periods of time and not over 13 years as in the current sample. Based on these considerations the regression is specified in the following way:

$$\frac{\Delta D_{it}}{Y_{it}} = \alpha^{S} + \beta_1^S \operatorname{ihs}(Y_{it}) + \beta_2^S \operatorname{ihs}\left(\tilde{Y}_{it}\right) + \beta_3^S \operatorname{ihs}(HW_{it}) + \beta_4^S \operatorname{ihs}(REP_{it}) + \\ \beta_5^S \operatorname{ihs}(D_{it-1}) + \beta_6^S \operatorname{ihs}(FW_{it}) + \beta_7^S dlinc_{it} + \beta_8^S dyear_t + \boldsymbol{\beta}^S \boldsymbol{X}_t + \varepsilon_{it} \tag{2}$$

where $S = \{B, NB\}$ indicates the subsamples of borrowing and non-borrowing households, $dyear_t$ is a set of year dummies and $\boldsymbol{X}_t$ is a matrix containing household characteristics such as age and age-squared of the household head as well as dummies for the presence of children, for being married or living with a partner, for being part of the labour-force, for having a college degree, for having been denied when applying for credit and a set of race dummies. The motivation of including such a rich set of household characteristics is the inability to control for unobserved heterogeneity as it would be possible with a panel data set.

The set of time dummies as well as the indicator variable for households having been turned down when applying for credit over the last four years are a way to control for shifts in credit supply conditions. If lenders are more willing to hand out larger loans relative to an households' income level or stock of assets then this should affect households homogeneously across time and thus allowing for time varying intercepts should capture that. In contrast if lenders are become increasingly willing to accept customers which were not able to obtain credit in the past, then the effect of such a shift in credit conditions should affect those households which were denied access to credit in the past. A decline in the proportion of households reporting not able to obtain credit in the past then indicates a shift in the supply of credit towards low quality borrowers.

Equation (2) is estimated by pooled OLS using probability weights. Weighted estimation is important because due to oversampling households from the upper tail of the income and wealth distribution would be overrepresented if all observations were implicitly assigned equal weights as in case of unweighted estimation. Standard errors are based on a bootstrap procedure which re-estimates each regression 999 times using a set of replicate weights instead of the initial weights. Replicate weights are part of the SCF dataset and are designed to replicate the sampling process. Since the Fed does not publish the details of their sampling procedure in order to protect the privacy of the survey participants the replicate weights are the only way to obtain standard errors which take stratification and oversampling properly into account. Instead of taking logarithms of variables in Dollar terms, they are transformed using the (unscaled) inverse hyperbolic sine transformation (ihs) which is defined for zero and negative valued observations[16] (MacKinnon & Magee 1990; Friedline et al. 2015).

4.1 Defining Reference Groups

Even though the idea of the expenditure cascades hypothesis that an increasingly polarized income distribution triggers debt financed spending as left-behind households attempt to maintain their perceived social status is quite simple, the task of testing it empirically is not. The main difficulty lies in defining adequate reference groups.

Before going into details it is important to emphasize again that the expenditure cascades hypothesis strictly relies on upward looking status comparison. Only if households compare themselves with others higher up in the income distribution will an increase in income inequality lead to a cascade of debt-financed expenditures. In contrast if households rely on average income as their point of reference, a polarization in the distribution of income would not lead to an increase in spending and borrowing for households above the average. Thus even though the mean as a point of reference is common in the literature (Alpizar et al. 2005; Maurer & Meier 2008; Alvarez-Cuadrado & Vilalta 2012; Alvarez-Cuadrado et al. 2012), it is not consistent with the expenditure cascades hypothesis. For this reason the paper will not use average income as a point of reference.

Instead the following three definitions of reference income, motivated by upward looking status comparison, are used. The first definition assumes that households compare themselves to richer peers at the top of the income distribution. In particular it is assumed that the pth income percentile is the point of reference for all households below that percentile. This approach is similar to Bertrand and Morse (2013) who use the 80th and 90th percentile of the income distribution within the state household i lives in as the cut-off. While the assumption that status comparison takes place within regionally defined communities is plausible, US States are too large to serve as realistic proxies for such communities. This paper follows an alternative approach and defines reference groups based on educational achievement (college or less-than-college) and the racial background (white or black) of

[16] The inverse hyperbolic sine transformation applied is defined as: $ihs(x) = \ln(x + \sqrt{x^2 + 1})$. The attractiveness of this transformation stems from the fact that it can be applied to zero and negative values while the interpretation of ihs transformed data in a regression context is equivalent to log-transformed data since $\frac{\partial ihs(x)}{\partial x} = \frac{1}{\sqrt{x^2+1}}$ which asymptotically approaches $\frac{1}{x}$ as x increases.

households, resulting in 4 groups[17]. The rationale for grouping along educational achievement and race is that both variables are important factors in defining the social sphere in which individuals engage with others through work, residency and leisure activities. The SCF provides four racial categories: white, black, hispanic and other. The reason for excluding the 'other' category in the definition of reference groups is that it is a residual category which identifies not a homogeneous group. However reference groups are motivated by the idea that households engage with and compare themselves to similar peers. The reason for dropping the hispanic category is of technical nature: The number of observations within the college-hispanic cells are too small to reliably calculate within group income distribution measures[18].

In contrast to Bertrand and Morse (2013) not only the 80th and 90th percentiles are used as cut-off points but the percentiles $q = \{99, 95, 90, 80\}$. Since mainly the top 5% of households enjoyed above average income growth in the period 1995 to 2007 including the 99th and 95th percentiles is crucial in analysing the impact of income inequality on household spending and borrowing[19]. This first definition of reference income based on the percentile cut-off is denoted $\tilde{Y}_{it}^{pq}$ where p indicates the use of the percentile as the point of reference and q indicates the cut-off. With reference income defined in that way one expects that households which are exposed to higher levels of top-incomes in their education-race group borrow more in order to finance a similar expenditure level as these richer peers. In order to distinguish peer effects from the effects of education and race a full set of race-education dummies along with all interactions is included in the regression. Thus the effect of $\tilde{Y}_{it}^{pq}$ is identified by the variation over time only. In a pure cross section $\tilde{Y}_{it}^{pq}$ would not be distinguishable from the effects of education and race captured by the set of dummies and their interactions.

The second definition of reference income is very similar but instead of using income percentiles as reference points, the average household income above a chosen percentile is used as the point of reference. Percentiles are again computed for groups based on educational achievement (college and

[17] The SCF includes identifiers for 9 regions, based on groups of US States, in the years 1995 and 1998 but not from 1998 onwards. As a robustness check income percentiles were computed within these regions and compared to the results based on education and race groupings, see Table A4 in the Appendix. Results do not differ systematically between specifications relying on geographically identified groups and race-education identified groups.

[18] For example in 1995 there are only 31 observations in the college-hispanic cell compared to 71 in the college-black cell, 1,602 in college-white, 141 in no-college-hispanic, 285 in no-college-black and 1,839 in no-college-white.

[19] This becomes even more important when grouping households along education and race as income gains are even more concentrated towards the top within these finer groups compared to the nationwide income distribution.

less-than-college) and race (white, black). Compared to the first definition, the average income above a certain percentile also obtains information about the households above that percentile. In particular the evolution of the average income of e.g. the top 5% of all households (i.e. households above the 95th percentile) is a better indicator about the income share of that group and thus the evolution of the income distribution than just the 95th percentile. Percentiles $q = \{99, 95, 90, 80\}$ are used as cut-off points. This second definition of reference income based on averages above a certain percentile is labelled $\tilde{Y}_{it}^{avq}$ where av indicates the use of income averages and q indicates the cut-off. Since for those households with income levels in excess of the cut-off there is no reference point defined, these are excluded from the estimation. This is the case for $\tilde{Y}_{it}^{avq}$ as well as $\tilde{Y}_{it}^{pq}$. Under the expenditure cascades hypothesis the effect of $\tilde{Y}_{it}^{avq}$ on household borrowing is expected to be positive.

The third definition is closely related to the measure of relative income used by Carr and Jayadev (2014). They cluster households based on which US State they live in and then the proportion of households within that group which are richer than household *i* is used as a proxy for household *i*'s relative income. Since US States are too large to serve as meaningful proxies for communities within which status comparison takes place, this paper again relies on education and race based groups g:

$$\tilde{Y}_{it}^{head} = \frac{\sum I[I = 1|Y_{-igt} > Y_{igt}]}{N_{gt}} \tag{3}$$

where $I \in \{0,1\}$ is an indicator variable equal to 1 when a given observation's income is greater than observation i's income. $g \in \{1, \dots, 4\}$ represents the four groups of households clustered by education (college or less-than-college) and race (white or black). Thus $\tilde{Y}_{it}^{head}$ corresponds to the complementary cumulative distribution function of income within group g. As such it is very different from the previous two definitions because it does not provide an absolute measure of within-group income concentration which is comparable across groups. Instead it provides information of the relative position for individual households within the group. $\tilde{Y}_{it}^{head}$ is expected to yield a positive effect on household debt accumulation in the borrowing and in the non-borrowing sample. Higher values of $\tilde{Y}_{it}^{head}$ are associated with households which are relatively poor in their race-education group, thus are exposed to a large number of richer peers and thus are expected to borrow more or repay slower. There are three important drawbacks with this definition of relative income. First, it implies that expenditure cascades become stronger towards the bottom of the within group income distribution. A priori it is not clear whether this should be the case. Instead if households are only partially able to keep up with their richer peers it might even be the case that expenditure cascades decline towards the bottom of the within group distribution. Second, defining relative income in that way does not provide any information about the degree of within group income concentration towards the top, which is central for the expenditure cascades argument. Thus using $\tilde{Y}_{it}^{head}$ is equivalent of focussing on the within group position in the income distribution at the expense of ignoring the absolute degree of inequality within that group compared across groups. Put differently, one assumes that households do not care by how much their peers are richer than themselves but they care only about how many other households are richer (a headcount of richer households, thus the labelling). The expenditure cascades hypothesis however rests on a situation of growing within group income inequality. Third, due to the focus of within group variation this measure does not allow one to assess to what extent rising income inequality contributed to household borrowing across time. It can only be used to explain differences between households.

Summing up: the paper uses three different version of relative income: The first two measure the absolute degree of inequality within education-race groups and are defined as the qth percentile of the education-race group income distribution ($\tilde{Y}_{it}^{pq}$) and the average income above the qth percentile ($\tilde{Y}_{it}^{avq}$). The third measure is defined as the complementary cumulative distribution function of income within education-race groups ($\tilde{Y}_{it}^{head}$) which corresponds to a head count of households richer than household i.

5 Heterogeneity, Reference Income and Housing Wealth Effects in Household Debt Accumulation

Table 2 presents the first set of results. Equation () is estimated for three different samples: the borrowing and non-borrowing subsamples as well as for the full sample (containing borrowing and non-borrowing households). Since all independent variables denominated in monetary terms are subject to the inverse-hyperbolic-sine transformation the estimated coefficients in Table 2 can be interpreted as changes in household borrowing in % of income due to a 1% increase of the corresponding explanatory variable. In order to interpret the results correctly it is important to keep in mind that repayment is ‘negative borrowing’ and thus a positive coefficient in the non-borrowing sample indicates less ‘negative borrowing’ and thus less repayments. As expected, there is a pronounced asymmetry in household behaviour between borrowing and non-borrowing groups as is indicated by the vastly different coefficients reported in columns (1) and (2). Combining borrowing and non-borrowing households in one group as in column (3) yields misleading averages which fail to adequately describe the behaviour of either group. For this reason the borrowing and non-borrowing sample are analysed separately from here on.

Beyond these differences in behaviour, results in Table 2 indicate that with rising income levels, households borrow less relative to their income (columns 1 and 4) and also use a smaller proportion of their income for repayments (columns 2 and 5). Thus lower income households leverage up more, relative to income. These findings are compatible with the life-cycle model if one assumes that higher income households face smaller shocks relative to their income. It is also compatible with the Post-Keynesian stock-flow-consistent modelling tradition which assumes borrowing for consumption purposes will be smaller relative to income for high income household due to lower marginal propensities to consume. Also the expenditure cascades hypothesis predicts a negative impact of income on household borrowing because richer households are less dependent on borrowing in order to finance a given level of status consumption.

With respect to housing, Table 2 reveals that there are highly statistically significant positive effects of housing wealth (HW) and real estate purchases (REP) on household borrowing (columns 1 and 4). It means that holding everything else constant higher levels of housing wealth lead to more borrowing. In addition, purchasing residential real estate in the year of the interview has a positive and statistically significant impact on borrowing. Since purchasing real estate without taking out a mortgage is highly unusual the insignificant effect in the non-borrowing sample is not surprising. Overall these results are very well in line with a life-cycle model incorporating credit constraints, with a Minskian interpretation of the relationship between assets and leverage and also with the notion of households being anchored by assets-to-income ratios as argued by the stock-flow-consistent modelling tradition. Thus the importance of assets and especially housing assets predicted by these theories in line with the hypothesis of Minskyian households is strongly supported by the data.

Financial wealth (FW) has neither a statistically significant effect on borrowing nor non-borrowing households. This might be due to the fact that even though financial wealth should ease credit constraints in a similar way as housing wealth does, banks are much more reluctant to accept financial assets with potentially highly volatile prices as collateral. Total liabilities at the beginning of the period (D_{it-1}) only have a statistically significant effect on the non-borrowing group indicating that higher indebted households use a larger proportion of their income for repayments. The lack of a significant effect of outstanding liabilities on borrowing is unexpected since the life-cycle model which emphasizes net-wealth as well as the stock-flow-consistent modelling tradition predict a negative effect. On the other hand a Minsky inspired interpretation suggests that heavily indebted households might have to rely on new borrowing in order to keep up with payments, predicting a positive effect. If both arguments are valid potentially at different stages of the credit cycle (i.e. the Minsky argument only holding shortly before the peak) they might cancel each other out over longer sample periods leaving a statistically insignificant result.

While it became clear that assets and outstanding liabilities are crucial in determining household borrowing decisions, Table 2 also reveals the importance of the income distribution. Reference income, defined as the average income of the top 1% of households within the education and race groups ($\tilde{Y}_{it}^{av99}$) has a highly significant positive effect in the group of borrowing households (column 1). Thus a 1% increase in average incomes at the top leads to an increase in household borrowing by 0.2% of household income. This result strongly supports the expenditure cascades hypothesis: Households being exposed to a more unequal distribution of income borrow more relative to their household income. For the non-borrowing sample the effect is not statistically different from zero and thus being exposed to richer peers does not slow down repayment efforts. The fact that repayment conditions are agreed on in advance provides an explanation. Put differently households have a hard time to explaining to their bank a delay in payments due to the neighbour buying a bigger car.

In specifications (1) to (3) in Table 2, the effect of average top incomes ($\tilde{Y}_{it}^{av99}$) does not depend on any measure of credit constraints. However it is not clear how households without any relevant assets should be able to secure additional borrowing in order to finance status driven expenditures. Thus it might be the case that only unconstrained households are able to engage in status seeking borrowing cascades. In order to control for that possibility, columns (4) and (5) of Table 2 report the results of estimating equation (2) with the measure of average top income ($\tilde{Y}_{it}^{av99}$) interacted with a dummy for residential real estate ownership ($HW1$ indicating ownership and $HW0$ no ownership). The results do indeed support the hypothesis that reference incomes only affect non-credit constrained households, in this case owners, since the effect for non-owners is statistically not different from 0. At the same time the effects of the other variables remain unchanged compared to the specification without the interaction terms. This is an important finding because it indicates that while the distribution of income does matter via status comparison across households and so called expenditure cascades it also indicates that only those households which possess some form of collateral are actually able to react to this kind of peer pressure by increasing their own expenditures using borrowed money. So household borrowing is determined in a complex interaction of different channels and it seems that in isolation neither the expenditure cascades hypothesis nor the hypothesis of Minskyian households are able to fully explain household borrowing[20]. Only an interaction of inequality and asset prices and thus a conditional version of the expenditure cascades hypothesis can provide such an explanation.

Table 2: Baseline Specification

	(1)	(2)	(3)	(4)	(5)
sample	ΔD/Y>0	ΔD/Y≤0	full	ΔD/Y>0	ΔD/Y≤0
$\tilde{Y}$ cut-off	99th perc	99th perc	99th perc	99th perc	99th perc
$\tilde{Y}$ definition	av. inc.	av. inc.	av. inc.	av. inc.	av. inc.
Y	-0.314	0.042	0.001	-0.314	0.042
	(0.024)***	(0.003)***	(0.005)	(0.023)***	(0.003)***
$\tilde{Y}^{av99}$	0.216	0.009	0.056		
	(0.080)***	(0.006)	(0.019)***		
$HW0\#\tilde{Y}^{av99}$				0.089	0.004
				(0.115)	(0.006)
$HW1\#\tilde{Y}^{av99}$				0.289	0.010
				(0.085)***	(0.007)
HW	0.086	-0.002	0.006	0.090	-0.003
	(0.017)***	(0.001)	(0.003)*	(0.017)***	(0.001)**
REP	0.659	0.011	0.539	0.658	0.012
	(0.047)***	(0.011)	(0.038)***	(0.048)***	(0.011)
FW	0.008	-0.001	0.000	0.008	-0.001
	(0.006)	(0.001)	(0.001)	(0.006)	(0.001)
D_{t-1}	-0.005	-0.024	-0.024	-0.007	-0.024
	(0.009)	(0.001)***	(0.002)***	(0.009)	(0.001)***
$constant$	0.758	-0.555	-0.669	2.607	-0.480
	(1.176)	(0.086)***	(0.269)**	(1.710)	(0.093)***
N	2,229	14,270	16,510	2,229	14,270
av. R^2	0.65	0.27	0.51	0.65	0.27
time effects	yes	yes	yes	yes	yes
household characteristics	yes	yes	yes	yes	yes

Dependent variable: ΔD/Y. All independent variables are subject to the inverse hyperbolic sine transformation. Y~ is defined as the average income above the 99th percentile within education-race groups (college/less-than-college and black/white). Coefficients are estimated by OLS using probability weights. Bootstrapped standard errors are obtained by re-estimating the regression 999 times using a set of 999 replicate weights. The R^2 is the average across all implicates. Stars indicate 1% (***), 5% (**) and 10% (*) significance levels. A full table including the missing household characteristics and time effects is provided in the Appendix (Table A2).

[20] Table A3 in the Appendix presents fully separated specifications for owners and non-owners.

The results so far are based on a specification using $\tilde{Y}_{it}^{av99}$ as the reference income definition. In order to check the robustness of these results Table 3 reports additional specifications incorporating alternative definitions of reference income. Using other definitions than the average income of the top 1% of households has two important implications. First, using cut-off points other than the top 1% serves as a robustness check to the question whether expenditure cascades are triggered by concentration of income at the very top of the distribution. Second, using relative income measures based on percentiles ($\tilde{Y}_{it}^{pq}$) as well as the relative position within the group ($\tilde{Y}_{it}^{head}$) provides a test of other mechanisms through which status comparison might takes place and expenditure cascades occur.

Table 3 demonstrates that using different percentiles as cut-off points for computing the average income of the top groups still yields statistically significant and positive effects on household borrowing (column 1: 95th percentile and column 2: 80th percentile) while the effects of the other variables remain qualitatively unchanged. However only the 99th percentile cut-off used in Table 2, yields statistically significant effects at the 1 % level, indicating that expenditure cascades are triggered from the very top of the income distribution. When instead of the average income above the cut-off only the income at the cut-off is used in the estimations (columns 3 to 5 and 9 to 11), there is no statistically significant effect of reference income on household borrowing. Defining relative income as the proportion of richer households ($\tilde{Y}_{it}^{head}$) yields a negative and statistically highly significant effect on household borrowing (column 6) both for owners and non-owners. Thus households towards the bottom of the within group income distribution borrow less. The negative effect of $\tilde{Y}_{it}^{head}$ in the non-borrowing sample (column 12) indicates that households towards the bottom of the within group income distribution use a larger proportion of their income for repayments. Thus the results for $\tilde{Y}_{it}^{head}$ do not support the expenditure cascades hypothesis. Due to the focus on within group variation of $\tilde{Y}_{it}^{head}$ instead of across group variation, this paper prefers the other two definitions of relative income based on top group averages ($\tilde{Y}_{it}^{avq}$) and percentile cut-off values ($\tilde{Y}_{it}^{pq}$).

The effects of housing wealth (HW) and real estate purchases in the current year (REP) do not change while using different measures of reference income. Thus Table 3 demonstrates the robust support of the data for the hypothesis of Minskyian households. However up to now the discussion of the results solely focussed on the statistical significance and the signs of the estimated effects in order to judge whether they are in line with predictions from economic theory. The next step is to assess the economic significance of the estimated effects and to compute comparable effect size measures. This is done in the next section.

Table 3: Additional Reference Income Definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
sample	ΔD/Y>0	ΔD/Y>0	ΔD/Y>0	ΔD/Y>0	ΔD/Y>0	ΔD/Y>0	ΔD/Y≤0	ΔD/Y≤0	ΔD/Y≤0	ΔD/Y≤0	ΔD/Y≤0	ΔD/Y≤0
$\tilde{Y}$ definition	$\tilde{Y}^{av95}$	$\tilde{Y}^{av80}$	$\tilde{Y}^{p99}$	$\tilde{Y}^{p95}$	$\tilde{Y}^{p80}$	$\tilde{Y}^{head}$	$\tilde{Y}^{av95}$	$\tilde{Y}^{av80}$	$\tilde{Y}^{p99}$	$\tilde{Y}^{p95}$	$\tilde{Y}^{p80}$	$\tilde{Y}^{head}$
$\tilde{Y}$ cut-off	95th perc	80th perc	99th perc	95th perc	80th perc	none	95th perc	80th perc	99th perc	95th perc	80th perc	none
Y	-0.307	-0.327	-0.313	-0.307	-0.327	-0.517	0.043	0.043	0.042	0.043	0.043	0.025
	(0.026)***	(0.032)***	(0.023)***	(0.026)***	(0.032)***	(0.073)***	(0.003)***	(0.004)***	(0.003)***	(0.003)***	(0.004)***	(0.005)***
$HW0\#\tilde{Y}$	0.121	0.191	0.047	0.159	0.294	-0.645	-0.004	-0.008	-0.010	-0.012	0.005	0.002
	(0.171)	(0.235)	(0.141)	(0.245)	(0.380)	(0.215)***	(0.010)	(0.014)	(0.009)	(0.013)	(0.027)	(0.015)
$HW1\#\tilde{Y}$	0.203	0.330	0.014	0.052	0.467	-0.666	0.009	0.017	0.004	0.011	0.029	-0.089
	(0.105)*	(0.142)**	(0.071)	(0.164)	(0.305)	(0.197)***	(0.007)	(0.012)	(0.005)	(0.010)	(0.019)	(0.013)***
HW	0.110	0.115	0.095	0.113	0.115	0.095	-0.004	-0.004	-0.003	-0.004	-0.004	-0.008
	(0.018)***	(0.019)***	(0.017)***	(0.018)***	(0.019)***	(0.016)***	(0.001)***	(0.002)***	(0.001)**	(0.001)***	(0.002)***	(0.001)***
REP	0.752	0.857	0.659	0.751	0.856	0.606	0.014	0.017	0.012	0.014	0.017	0.012
	(0.052)***	(0.052)***	(0.047)***	(0.051)***	(0.052)***	(0.045)***	(0.012)	(0.016)	(0.011)	(0.012)	(0.016)	(0.010)
FW	0.007	0.006	0.008	0.007	0.006	0.005	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.001)	(0.001)*	(0.001)	(0.001)	(0.001)*	(0.001)*
D_{t-1}	-0.005	-0.004	-0.007	-0.005	-0.004	-0.009	-0.025	-0.028	-0.024	-0.025	-0.027	-0.024
	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)	(0.009)	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
$constant$	2.238	1.669	3.258	1.839	0.538	6.511	-0.383	-0.331	-0.297	-0.288	-0.496	-0.251
	(2.345)	(3.025)	(1.921)*	(3.127)	(4.612)	(0.945)***	(0.127)***	(0.184)*	(0.122)**	(0.163)*	(0.331)	(0.065)***
N	2,242	2,024	2,477	2,242	2,024	2,722	14,183	12,971	15,931	14,183	12,971	18,358

Dependent variable: ΔD/Y. All $ valued independent variables are subject to the inverse hyperbolic since transformation. Coefficients are estimated by OLS using probability weights. Bootstrapped standard errors are obtained by re-estimating the regression 999 times using a set of 999 replicate weights. Stars indicate 1% (***), 5% (**) and 10% (*) significance levels. A full table including the missing household characteristics and time effects is available upon request.

6 Aggregate Effect Size of Housing Wealth and Relative Income

In order to compare the explanatory power of the expenditure cascades and the Minskyian households hypotheses, contributions to the predicted value of household borrowing are calculated. For example in the case of housing wealth (HW) the contribution of housing wealth to the accumulation of debt in year t aggregated over all households is calculated as:

$$\Omega_t^{HW,S} = \sum_{i=1}^{N_s} \hat{\beta}_3^S \; ihs(HW_{it}) \; Y_{it} \tag{4}$$

where $S \in \{B, NB\}$ indicates that the computation is done for the borrowing (B) as well as the non-borrowing (NB) subsample. Equivalent computations are carried out for all independent variables in the model. Even though the contributions are not very useful for cross section comparisons they are useful in shading light on the questions which household characteristics are most important in order to explain household borrowing over time[21].

Table 4 presents the contributions of the main variables of interest: household income, relative income and household balance sheet items. Also the overall model prediction is reported in the first column. Based on equation (4) contributions are computed for each subgroup, expressed as the difference relative to 1995 and scaled by aggregate income. For example the housing effect reported in Table 4 is computed as:

$$\Delta\Omega_t^{HW} = (\Omega_t^{HW,B} + \Omega_t^{HW,NB} - \Omega_{1995}^{HW,B} - \Omega_{1995}^{HW,NB}) / \boldsymbol{Y}_t \tag{5}$$

where $\boldsymbol{Y}_t$ is aggregate income of the actual sample, excluding households which identify as hispanic or other and households with borrowing to income ratios smaller than -250% and bigger than 500% but including those households above the cut-off points used for the definition of relative income. Panel A of Table 4 is based on the results from columns 4 and 5 of Table 2 using the average income of households above the 99th percentile of the within group income distribution as reference income and Panel B uses the average income of households above the 95th percentile of the within group income distribution (columns 1 and 7 Table 3). The different measures of reference income are the main reason why contributions of individual variables as well as predicted total borrowing differ across the two Panels: Panel B is based on a smaller sample excluding not only the top 1% of the within group income distributions but the top 5%. However the sample size is not the only reason why results vary. Another important reason is that top 1% income shares and thus average income increased much

[21] One way to compare the contribution of individual variables in the cross section would be to demean the regressors prior to estimation. Since the focus here is on changes over time demeaning was not performed because it complicates comparison across time.

stronger than top 5% income shares over the sample period. Thus while absolute effects are not comparable between Panel A and B the relative sizes of the effects are.

Table 4: contributions to household borrowing relative to 1995

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: top 1%							
	total	$HW0 \# \tilde{Y}^{av99}$	$\tilde{Y}^{av99} + HW$	Y	REP	FW	D_{t-1}
1998	0.8%	-0.5%	5.9%	-3.3%	0.6%	0.7%	-1.3%
2001	1.1%	-0.1%	6.2%	-2.3%	1.1%	0.4%	-3.2%
2004	7.0%	1.2%	11.2%	-8.6%	6.8%	1.7%	-3.6%
2007	2.3%	-0.3%	8.7%	-2.9%	3.5%	0.5%	-4.7%
Panel B: top 5%							
	total	$HW0 \# \tilde{Y}^{av95}$	$\tilde{Y}^{av95} + HW$	Y	REP	FW	D_{t-1}
1998	1.0%	-0.8%	3.7%	-1.3%	0.3%	0.3%	-1.0%
2001	0.8%	-0.6%	4.3%	-0.7%	0.3%	0.3%	-2.3%
2004	5.6%	1.4%	6.7%	-4.1%	4.6%	0.8%	-3.0%
2007	2.0%	-0.4%	4.6%	0.1%	2.4%	0.2%	-3.5%

Contributions to household borrowing are computed based on results from columns 4 and 5 of Table 2 for Panel A and from columns 1 and 7 of Table 3 for Panel B using equation (5). Values correspond to changes in household borrowing with respect to 1995 expressed in % of aggregate sample income (1998: \$6,893 bn; 2001: \$8,611 bn; 2004: \$8,584 bn; 2007: \$9,573 bn in 2013 prices). Total refers to the total change in household borrowing. Total is based on all variables and thus is not equivalent to the sum of the variables displayed in the Table. Full tables containing the contributions of all variables in the model can be found in the Appendix (Tables A5 and A6).

Since the effect of reference income is interacted with a housing ownership dummy in the specification used to calculate the contributions, the contribution of changes in housing wealth and reference income cannot be distinguished for the group of owners[22]. For that reason Table 4 presents the effect of reference income for non-owners (column 2) and a combined effect of reference income and real estate wealth for owners (column 3). First Table 4 shows that household borrowing strongly increased in 2004: the difference in household borrowing between 2004 and 1995 amounts to 7% of household income or 600 billion in 2013 prices[23]. The effect of rising within group top incomes contributed to additional household borrowing of 1.2% (\$104 billion with $\tilde{Y}^{av99}$) and 1.4% (\$120 billion with $\tilde{Y}^{av95}$) of household income among non-owners, depending on the relative income measure used. In contrast the combined effect of increasing housing wealth and top incomes stimulated household borrowing

[22] The reason is that the positive effects of reference income and housing wealth need to be summed up with the negative effect of the ownership dummy and there is no meaningful way of separating the effect of reference income and ownership and changes in housing wealth.

[23] All subsequent dollar values are expressed in 2013 prices.

during the same period by 11.2% (\$963 billion with $\tilde{Y}^{av99}$) and 6.7% (\$575 billion with $\tilde{Y}^{av95}$) of household income. With both specifications this partial effect exceeds the total amount of additional borrowing in all years, emphasizing the importance of the interaction of housing and relative income dynamics. In addition to existing assets also real estate purchases in the current period significantly contributed to the leveraging up process of the household sector: by 6.8% (\$ 584 billion with $\tilde{Y}^{av99}$) and 4.6% (\$395 billion with $\tilde{Y}^{av95}$) respectively. The growth in household income reduces household borrowing over the same period by 8.6% (\$740 billion with $\tilde{Y}^{av99}$) and 4.1% (\$352 billion with $\tilde{Y}^{av95}$). Table 4 also demonstrates the crucial role of past liabilities for household borrowing (-3.6% and -3% in 2004). Increasing debt levels slowed down the accumulation of new liabilities. The negative contribution of past liabilities on borrowing almost exclusively works through the group of non-borrowing households: with increasing debt levels households have to use larger proportions of their disposable income to repay principal and interest.

Based on these numbers, the paper confirms earlier findings in support of the expenditure cascades hypothesis but emphasizes the fundamental role of real estate assets in easing credit constraints: Expenditure cascades only materialise if households can rely on some form of collateral for borrowing in order to keep up in status driven expenditure cascades. Thus simple statements of the expenditure cascades hypothesis which ignore the limiting role of credit constraints do not describe the pre-2007 experience of the US very well. The interaction between homeownership and income inequality calls for an integrated analysis of household borrowing which takes relative income as well as household balance sheet dynamics properly into account.

The results presented also strongly emphasize the importance of real estate transactions and past liabilities which increase the intensity of repayments. Both factors are often overlooked when studying household borrowing behaviour.

7 Conclusions

This paper investigates the rise in US household debt levels prior to the financial crisis. Two potential explanations are of particular interest: First, according to the expenditure cascades hypothesis, upward looking status comparison in an environment of increasing income inequality triggers debt-financed consumption sprees which cascade down the income distribution. Against the background of strongly rising top incomes in the United States several authors argue on theoretical (Cardaci 2014; Frank et al. 2014; Kapeller & Schütz 2014; Ryoo & Kim 2014; Nikiforos 2015; Cynamon & Fazzari 2015) and empirical (Belabed et al. 2013; Bertrand & Morse 2013; Carr & Jayadev 2014; Perugini et al. 2015) grounds that expenditure cascades explain the rise in household debt and therefore identify it as a root cause of the financial crisis. The second explanation of interest is the hypothesis of Minskyian households, which argues that rising real estate prices boost household borrowing because of equity extraction by increasingly optimistic owners and the demand for bigger mortgages by (first time) buyers. This latter line of thinking is compatible with much of the literature on wealth effects in the consumption function if one allows for credit constrained households (Paiella 2009; Cooper & Dynan 2014). However the life-cycle model has difficulties in explaining why rational consumers did not anticipate the transitory nature of the surge in house prices. Thus using a Minskyian framework for explaining pre-2007 events appears to be a more realistic description of reality (Minsky 1978; Borio 2014; Ryoo 2015).

By constructing a measure of annual household borrowing from cross sectional data from the SCF, the paper is able to put the unmatched quality and level of detail the SCF provides to work. It means in particular that the results of this paper are based on the only US household survey which deals with the problem of non-observation and non-response of rich households in a convincing way by means of oversampling based on information from the federal tax authority. Furthermore the SCF offers an extremely detailed breakdown of the household balance sheet in terms of liabilities and assets. Thus since the expenditure cascades hypotheses focuses on the role of income inequality while the hypothesis of Minskyian households emphasizes the positive impact of rising asset values on household borrowing, the SCF provides unmatched high quality data to investigate both hypotheses.

The results presented based on this newly created measure of household borrowing indicates that it is the interaction of rising asset prices and the polarization of the income distribution which explains a large part of the increase in household borrowing before the crisis in 2008. In particular, the results support the expenditure cascades hypothesis conditional on homeownership. Thus higher average incomes for the top 1% of households within race (white, black) and education (less-than-college,

college) groups result in higher borrowing-to-income ratios for owners. Households can only engage in status driven borrowing if they have access to collateral and if the value of that asset increases. The magnitudes of these effects are substantial: household borrowing in 2004 was by $963 billion[24] higher compared to 1995 due to the increase in income inequality and the rise in housing wealth. In addition household borrowing in 2004 was higher compared to 1995 by an additional $584 billion due to the increased cost of purchasing properties. Rising household income had a negative impact on borrowing over the same period of $740 billion. In total households borrowed $600 billion more in 2004 compared to 1995. These results strongly suggest that explaining US household borrowing up to 2007 requires information on the shifts in the distribution of income as well as on the increasing value of real estate assets. Using only one part of the explanation without the other misses a large part of the story. Thus this paper calls for a synthesis of these two strands of the literature. Relative income (income inequality) and (housing) wealth should be treated as equally important and plausible mechanisms when analysing household debt.

The implications of the findings presented above are substantial. First, the distribution of income is a relevant factor in shaping macroeconomic outcomes. Indirectly the results challenge a standard assumption in Keynesian models: due to higher marginal saving rates towards the top of the income distribution, a redistribution of income towards the top reduces aggregate consumption. It seems that as long as the financial sector accommodates households' demand for credit a positive relationship between income inequality and consumption is possible. Second, (housing) asset price dynamics are important for aggregate household liability levels. The degree of importance will vary across countries depending on whether an ownership culture like in the US or a renter culture like in Germany prevails, but the asset side of the household balance sheet cannot be ignored when investigating household borrowing. Third, due to the large effects of relative income and housing wealth on household borrowing, the distribution of income as well as house prices are key indicators for monitoring the financial stability of the household sector. Large scale defaults will translate into less consumption spending and thus lower growth rates and also have the potential for critical knock-on effects in the financial sector. Fourth, limiting the impact of expenditure cascades requires policy measures which result in a more equal primary distribution of income in order to tackle the root cause of the problem. Minimum wage legislation, affordable access to high quality education and a clampdown on tax havens might be noteworthy strategies. In addition, increases in the public provision of housing in regions with

[24] Expressed in 2013 prices.

strong demand and weak private supply will keep price increases in check. This will prevent borrowing sprees backed by unsustainable house price booms.

The paper also provides methodological insights for investigating household borrowing behaviour when using survey data. First, with cross sectional data, borrowing and non-borrowing households need to be analysed separately. Borrowing and repayment decisions are driven by different variables: assets for example are more important for borrowing decisions due to their role in easing credit constraints while liabilities have a direct impact on repayments. Furthermore, variables which influence borrowing and repayment decisions do so in different ways, as the example of household income demonstrates. The benefits of a separate analysis also materialize in an increased model fit. Second, for the majority of households the most important borrowing decision is to take out a mortgage in order to buy a property. This means that the analysis of household borrowing requires either panel data which automatically provides information on purchases (as long as assets are observed) or in a cross section context information about the timing of purchases is needed in order to be able to explain the large spikes in household borrowing related to property acquisitions. Third, growth rates of household liabilities are not an informative indicator because household borrowing (due to property purchases) is not a smooth process like consumption and thus identical growth rates can refer to households engaging in very different behaviour. Furthermore due to the relevance of large one-off borrowing decisions, the use of logarithmic differences in order to proxy growth rates is highly inaccurate. Using absolute changes in household liabilities scaled by household income is a more informative and robust approach.

Finally, there remain problems which need to be addressed in future research. First, so far a positive link between (household) debt and the distribution of income is more consistently supported by microeconomic than by aggregate data. It is not yet clear why this is the case, although most of the evidence based on aggregate data relies on panel rather than time series techniques for individual countries. Second, the lack of data is a fundamental problem for studying the links between the distribution of income, asset prices and household liabilities. One specific way to enhance future research would be to conduct high quality surveys relying on oversampling methods using household level tax data. In the first wave of the ECB's Household Finance and Consumption Survey only nine out of fifteen participating countries made an effort to oversample wealthy households and only Spain and France relied on individual tax information to do so (ECB 2013). More effort and dedication from local central banks and tax authorities is needed. In addition central banks should provide more resources for collecting data in order to enable bigger samples and, most importantly, shorter intervals of data collection. Three year intervals as with the SCF and the HFCS leave substantial gaps which would be unimaginable in the case of GDP or inflation.

References

Alpizar, F., Carlsson, F. & Johansson-Stenman, O., 2005. How much do we care about absolute versus relative income and consumption? *Journal of Economic Behavior & Organization*, 56(3), pp.405–421. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0167268104000885 [Accessed February 16, 2014].

Alvarez-Cuadrado, F. et al., 2012. Envy and habits: Panel data estimates of interdependent preferences. *Banco De Espana Documentos de Trabajo*, 1213. Available at: http://intra.merit.unu.edu/publications/wppdf/2012/wp2012-054.pdf [Accessed February 16, 2014].

Alvarez-Cuadrado, F. & Vilalta, M.E.-A., 2012. *Income inequality and saving*, Available at: http://ssrn.com/abstract=2196748.

Atkinson, A.B., Piketty, T. & Saez, E., 2011. Top Incomes in the Long Run of History. *Journal of Economic Literature*, 49(1), pp.3–71.

Attanasio, O.P. et al., 2009. Booms and Busts: Consumption, House Prices and Expectations. *Economica*, 76(301), pp.20–50. Available at: http://doi.wiley.com/10.1111/j.1468-0335.2008.00708.x [Accessed November 28, 2014].

Attanasio, O.P. & Weber, G., 2010. Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy. *Journal of Economic Literature*, 48(3), pp.693–751. Available at: http://www.nber.org/papers/w15756 [Accessed August 22, 2014].

Behringer, J. & Treeck, T. van, 2013. *Income distribution and current account: A sectoral perspective*, Düsseldorf: Macroeconomic Policy Institute. Available at: http://www.boeckler.de/pdf/p_imk_wp_125_2013.pdf [Accessed July 10, 2014].

Belabed, C.A., Theobald, T. & van Treeck, T., 2013. *Income distribution and current account imbalances*, IMK Working Paper 126.

Bertrand, M. & Morse, A., 2013. *Trickle-Down Consumption*, Available at: http://www.nber.org/papers/w18883 [Accessed December 19, 2014].

Bezemer, D., Grydaki, M. & Zhang, L., 2014. *Is Financial Development Bad for Growth?*,

Bezemer, D. & Zhang, L., 2014. *From Boom to Bust in the Credit Cycle: The Role of Mortgage Credit*,

Bhaduri, A. & Marglin, S., 1990. Unemployment and the real wage: the economic basis for contesting political ideologies. *Cambridge Journal of Economics*, 14(4), pp.375–393.

Bordo, M.D. & Meissner, C.M., 2012. Does inequality lead to a financial crisis? *Journal of International Money and Finance*, 31(8), pp.2147–2161. Available at: http://dx.doi.org/10.1016/j.jimonfin.2012.05.006.

Borio, C., 2014. The financial cycle and macroeconomics: What have we learnt? *Journal of Banking and Finance*, 45(1), pp.182–198. Available at: http://dx.doi.org/10.1016/j.jbankfin.2013.07.031.

Bostic, R., Gabriel, S. & Painter, G., 2009. Housing wealth, financial wealth, and consumption: New evidence from micro data. *Regional Science and Urban Economics*, 39(1), pp.79–89. Available at: http://linkinghub.elsevier.com/retrieve/pii/S016604620800063X.

Campbell, J.Y. & Cocco, J.F., 2007. How do house prices affect consumption? Evidence from micro data. *Journal of Monetary Economics*, 54(3), pp.591–621. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0304393206001279 [Accessed December 1, 2014].

Cardaci, A., 2014. Inequality , Household Debt and Financial Instability : An Agent-Based Perspective. , (October), pp.1–30.

Carr, M.D. & Jayadev, A., 2014. Relative Income and Indebtedness: Evidence from Panel Data. *Review of Income and Wealth*, pp.1–14.

Carroll, C.D., 2001. A Theory of the Consumption Function, With and Without Liquidity Constraints. *Journal of Economic Perspectives*, 15(3), pp.23–45. Available at: http://pubs.aeaweb.org/doi/abs/10.1257/jep.15.3.23.

Cooper, D., 2012. House Price Fluctuations: The Role of Housing Wealth as Borrowing Collateral. *Review of Economics and Statistics*, 95(4), pp.1183–1197. Available at: http://dx.doi.org/10.1162/REST_a_00323.

Cooper, D. & Dynan, K., 2014. Wealth Effects and Macroeconomic Dynamics. *Journal of Economic Surveys*, Advance Ac. Available at: http://doi.wiley.com/10.1111/joes.12090 [Accessed January 3, 2015].

Cynamon, B.Z. & Fazzari, S.M., 2015. Inequality, the Great Recession and slow recovery. *Cambridge Journal of Economics*. Available at: http://cje.oxfordjournals.org/cgi/doi/10.1093/cje/bev016.

Duesenberry, J.S., 1949. *Income, Saving and the Theory of Consumer Behavior*, Harvard University Press.

Dutt, A.K., 2006. Maturity, Stagnation and Consumer Debt: a Steindlian Approach. *Metroeconomica*, 57(3), pp.339–364. Available at: http://doi.wiley.com/10.1111/j.1467-999X.2006.00246.x.

Dynan, K., 2010. *Wealth Effects and the Changing Economy*,

Dynan, K.E., 2000. Habit formation in consumer preferences: Evidence from panel data. *American Economic Review*, 90(3), pp.391–406. Available at: http://www.jstor.org/stable/117335 [Accessed November 23, 2014].

Dynan, K.E. & Kohn, D.L., 2007. The Rise in U.S. Household Indebtedness: Causes and Consequences. *Finance and Economics Discussion Series*, 37.

Dynan, K.E., Skinner, J. & Zeldes, S.P., 2004. Do the Rich Save More? *Journal of Political Economy*, 112(2), pp.397–444.

ECB, 2013. *The Eurosystem Household Finance and Consumption Survey Methodological Report for the First Wave*, Frankfurt.

Eckerstorfer, P. et al., 2015. Correcting for the Missing Rich: An Application to Wealth Survey Data. *Review of Income and Wealth*.

Eichengreen, B. & Mitchener, K., 2003. The Great Depression as a Credit Boom Gone Wrong. *BIS Working Papers*, (137). Available at: http://ideas.repec.org/p/bis/biswps/137.html.

Frank, R., 1985. The demand for unobservable and other nonpositional goods. *The American Economic Review*, 75(1), pp.101–116. Available at: http://www.jstor.org/stable/1812706 [Accessed November 10, 2014].

Frank, R.H., Levine, A.S. & Dijk, O., 2014. Expenditure Cascades. *Review of Behavioral Economics*, 1(1-2), pp.55–73.

Friedline, T., Masa, R.D. & Chowa, G.A.N., 2015. Transforming wealth: Using the inverse hyperbolic sine (IHS) and splines to predict youth's math achievement. *Social Science Research*, 49, pp.264–287. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0049089X14001732.

Fuhrer, J.C., 2000. Habit Formation in Consumption and Its Implications for Monetary-Policy Models. *American Economic Review*, 90(3), pp.367–390.

Godley, W. & Lavoie, M., 2007. *Monetary Economics*, New York: Palgrave Macmillan.

Greenspan, A. & Kennedy, J., 2008. Sources and uses of equity extracted from homes. *Oxford Review of Economic Policy*, 24(1), pp.120–144. Available at: http://oxrep.oxfordjournals.org/cgi/doi/10.1093/oxrep/grn003\nfile:///C:/Users/Lidia/Desktop/Doutorado/Artigos/Artigo consumo/Textos e artigos/Greenspan, Kennedy (2008).pdf.

Haurin, D.R. & Rosenthal, S.S., 2006. *House Price Appreciation, Savings, and Consumer Expenditures*,

Heffetz, O., 2011. A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics*, 93(4), pp.1101–1117. Available at: http://www.mitpressjournals.org/doi/abs/10.1162/REST_a_00116 [Accessed December 19, 2014].

Hein, E. & Vogel, L., 2008. Distribution and growth reconsidered: empirical results for six OECD countries. *Cambridge Journal of Economics*, 32(3), pp.479–511. Available at: http://cje.oxfordjournals.org/cgi/doi/10.1093/cje/bem047 [Accessed February 24, 2014].

Holt, R.P.F. & Greenwood, D.T., 2012. Negative Trickle-Down and the Financial Crisis of 2008. *Journal of Economic Issues*, 46(2), pp.363–370. Available at: http://mesharpe.metapress.com/openurl.asp?genre=article&id=doi:10.2753/JEI0021-3624460211 [Accessed December 21, 2014].

Hurst, E. & Stafford, F., 2004. Home is where the equity is: Liquidity constraints, refinancing and consumption. *Journal of Money, Credit and Banking*, 36(6), pp.985–1014.

Juster, F.T. et al., 2005. The Decline in Household Saving and the Wealth Effect. *The Review of Economics and Statistics*, 87(4), pp.20–27.

Kapeller, J. & Schütz, B., 2014. Debt, boom, bust: a theory of Minsky-Veblen cycles. *Journal of Post Keynesian Economics*, 36(4), pp.781–814.

Kennickell, A., 2005. The Good Shepherd: Sample Design and Control for Wealth Measurement in the Survey of Consumer Finances. *Luxembourg Wealth Study Conference*.

Kennickell, A.B., 1998. Multiple imputation in the Survey of Consumer Finances. *Joint Statistical Meetings*. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.9972{&}rep=rep1{&}type=pdf.

Kennickell, A.B., 2008. The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances. *IFC Bulletin*, (28), pp.403–408.

Kennickell, A.B., 2000. Wealth Measurement in the Survey of Consumer Finances: Methodology and Directions for Future Research. *Finance & Economic Discussion Paper, Federal Reserve Board Washington D.C.*

Kennickell, A.B. & McManus, D. a, 1993. Sampling for Household Financial Characteristics Using

Frame Information on Past Income. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 1, pp.88–97. Available at: http://www.federalreserve.gov/econresdata/scf/files/asa93.pdf.

Kim, Y.K., 2013. Household debt, financialization, and macroeconomic performance in the United States, 1951-2009. *Journal of Post Keynesian Economics*, 35(4), pp.675–694. Available at: http://mesharpe.metapress.com/openurl.asp?genre=article&id=doi:10.2753/PKE0160-3477350408 [Accessed December 21, 2014].

Lehnert, A., 2004. *Housing, Consumption, and Credit Constraints*,

MacKinnon, J.G. & Magee, L., 1990. Transforming the Dependent Variable in Regression Models. *International Economic Review*, 31(2), pp.315–339.

Maurer, J. & Meier, A., 2008. Smooth it Like the "Joneses"? Estimating Peer-Group Effects in Intertemporal Consumption Choice. *The Economic Journal*, 118(2006), pp.454–476. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.2007.02129.x/full [Accessed February 16, 2014].

Mian, A. & Sufi, A., 2011. House prices, home equity-based borrowing, and the US household leverage crisis. *American Economic Review*, 101(5), pp.2132–2156.

Mian, A. & Sufi, A., 2009. The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *Quarterly Journal of Economics*, 124(4), pp.1449–1496.

Minsky, H.P., 1978. *The Financial Instability Hypothesis: A Restatement*,

Nikiforos, M., 2015. *A Nonbehavioral Theory of Saving*,

Onaran, Ö., Stockhammer, E. & Grafl, L., 2011. Financialisation, income distribution and aggregate demand in the USA. *Cambridge Journal of Economics*, 35(4), pp.637–661. Available at: http://cje.oxfordjournals.org/cgi/doi/10.1093/cje/beq045 [Accessed February 24, 2014].

Paiella, M., 2009. The Stock Market, Housing and Consumer Spending: a Survey of the Evidence on Wealth Effects. *Journal of Economic Surveys*, 23(5), pp.947–973. Available at: http://doi.wiley.com/10.1111/j.1467-6419.2009.00595.x [Accessed December 13, 2014].

Palley, T.I., 1994. Debt, aggregate demand, and the business cycle: an analysis in the spirit of Kaldor and Minsky. *Journal of Post Keynesian Economics*, 16(3), pp.371–390. Available at: http://www.jstor.org/stable/4538399\nfile:///C:/Users/Lidia/Desktop/Doutorado/Comportamento econ?mico e institui??es/Palley (1994).pdf.

Perugini, C., Holscher, J. & Collie, S., 2015. Inequality, credit and financial crises. *Cambridge Journal of Economics*, pp.1–31. Available at: http://cje.oxfordjournals.org/cgi/doi/10.1093/cje/beu075.

Piketty, T., 2014. *Capital in the Twenty-First Century*, Cambridge: Harvard University Press.

Ravina, E., 2007. *Habit Persistence and Keeping Up with the Joneses: Evidence from Micro Data*, Available at: http://archive.nyu.edu/handle/2451/26455 [Accessed December 20, 2014].

Romer, D., 2012. *Advanced Macroeconomics* 4th ed., New York: McGraw-Hill.

Rubin, D., 1987. *Multiple Imputations for Nonresponse in Surveys*, New York: Wiley & Sons.

Ryoo, S., 2013. Bank profitability, leverage and financial instability: A minsky-harrod model. *Cambridge Journal of Economics*, 37(5), pp.1127–1160.

Ryoo, S., 2015. *Household debt and housing bubble: A Minskian approach to boom-bust cycles*, Ryoo, S. & Kim, Y.K., 2014. Income Distribution, Consumer Debt and Keeping up with the Joneses. *Metroeconomica*, 65(4), pp.585–618. Available at: http://doi.wiley.com/10.1111/meca.12052 [Accessed December 21, 2014].

Salotti, S., 2012. Wealth effects in the US: Evidence from the combination of two surveys. *The Economic and Social Review*, 43(1), pp.67–98. Available at: http://www.esr.ie/article/view/54 [Accessed February 16, 2014].

Dos Santos, C.H., 2005. A stock-flow consistent general framework for formal Minskyan analyses of closed economies. *Journal of Post Keynesian Economics*, 27(4), pp.711–735.

Schularick, M. & Taylor, A.M., 2012. Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, 102(2), pp.1029–1061.

Singer, E., 2006. Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70(5), pp.637–645. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=23631889&site=ehost-live.

Stockhammer, E. & Wildauer, R., 2015. *Debt-driven Growth? Wealth, Distribution and Demand in OECD Countries*,

Tobin, J., 1969. Equilibrium Approach To Monetary Theory. *Journal of Money, Credit and Banking*, 1(1), pp.15–29.

Tobin, J., 1982. Money and Finance in the Macroeconoic Process. *Journal of Money, Credit, and Banking*, 14(2), pp.171–204.

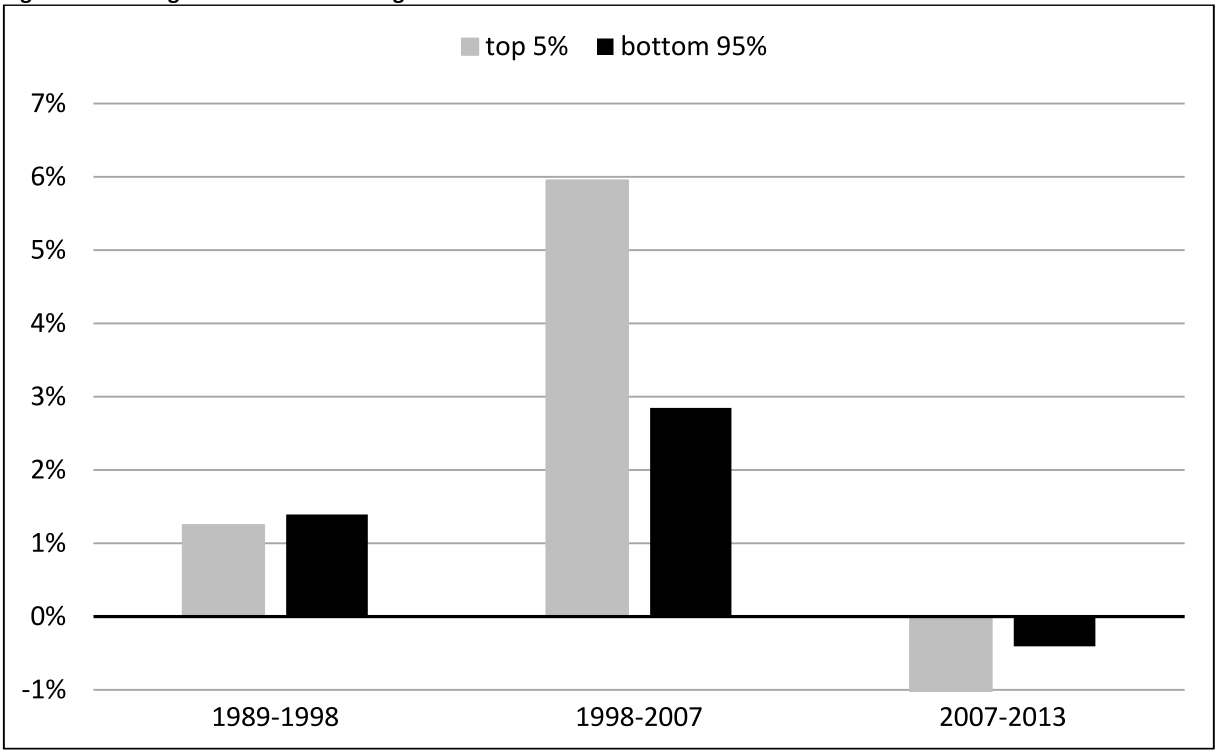
Veblen, T., 1899. *The Theory of the Leisure Class*, Oxford University Press.

Vermeulen, P., 2014. *How fat is the top tail of the wealth distribution?*,

Zezza, G., 2008. U.S. growth, the housing market, and the distribution of income. *Journal of Post Keynesian Economics*, 30(3), pp.375–401. Available at: http://mesharpe.metapress.com/openurl.asp?genre=article&id=doi:10.2753/PKE0160-3477300304 [Accessed February 24, 2014].

Appendix A: Tables and Figures

Figure A1: Average annual real income growth rates



Source: own computations based on SCF waves 1989 to 2013.

Figure A2: Decision tree, first mortgage on primary residence

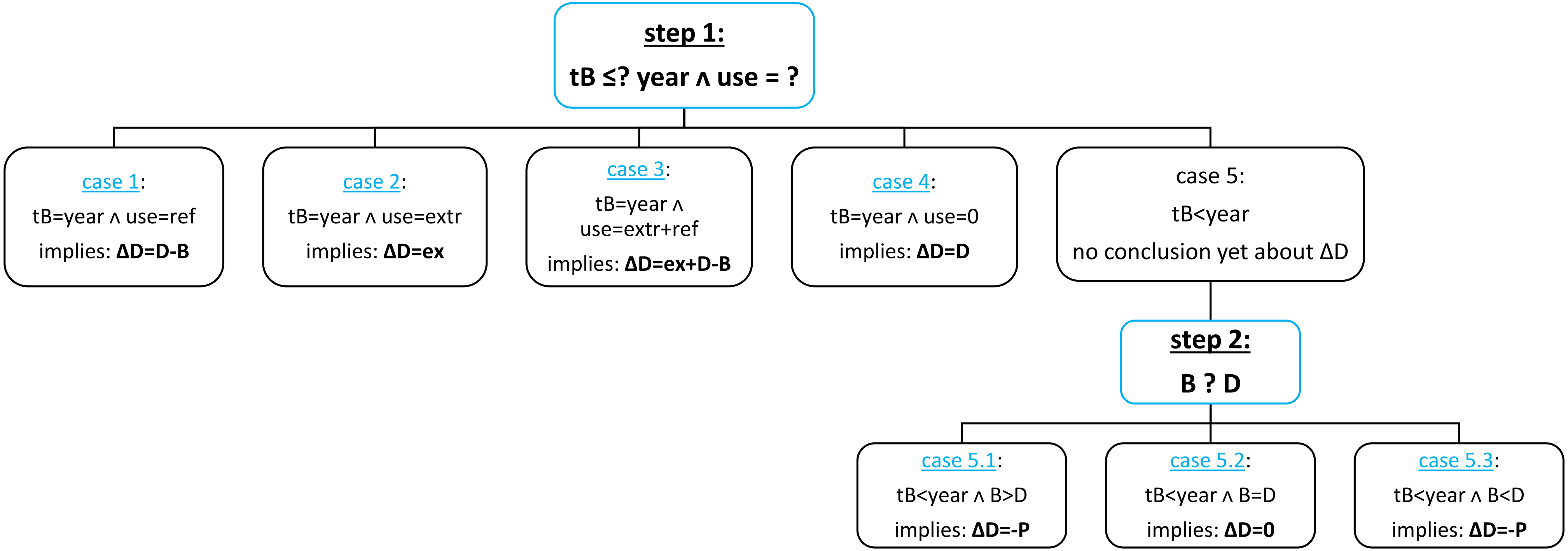


Table A1: Variable description

Variable	Description	SCF code/name
ΔD_{it}	Household borrowing defined as the change in outstanding liabilities within year t. The way how the measure is constructed is discussed in section 3.2 and Appendix B.	
D_{it-1}	Outstanding liabilities at the beginning of the period, derived from the definition: $D_{it-1} \equiv D_{it} - \Delta D_{it}$.	
Y_{it}	Total gross income as reported by the household. Including realized capital gains. Before taxes and other deductions.	Corresponds to the variable 'income' in the summary data set.
HW_{it}	Value of residential real estate, minus the value of residential real estate purchases in the current period and minus any capital gains on these purchases. Thus HW_{it} corresponds to the value of residential real estate the household owned at the beginning of the period.	Residential real estate corresponds to the variables 'houses' and 'oresre' in the summary dataset. For the detailed computations see Stata code.
REP_{it}	Value of real estate purchases in the current period, excluding capital gains on these purchases.	
FW_{it}	Total value of financial assets. Includes: checkings and savings accounts, money market accounts, certificates of deposits, directly held mutual funds, stocks, bonds, quasi liquid pension accounts, savings bonds, cash in life-insurance products, other managed assets and other financial assets.	Coded 'fin' in the summary dataset.
$\tilde{Y}$	Definitions of relative income measures are discussed in section 4.1.	
age_{it}	Age of the household head.	Coded 'age' in the summary dataset.
$kids_{it}$	Dummy variable, 1 indicating the presence of children.	Based on variable 'kids' from summary dataset.
$college_{it}$	Dummy variable, 1 indicating the household head obtaining a college degree.	Based on variable 'edcl' from summary dataset.
$black_{it}$	Dummy variable, 1 indicating household head self-identified as black.	Based on variable 'race' from summary dataset.
$working_{it}$	Dummy variable, 1 indicating household head is part of the labour force.	Coded 'lf' in summary dataset.
$lowinc_{it}$	Dummy variable, 1 indicating household income is lower than in normal year.	Based on X7650 in full data set.
$not\ married_{it}$	Dummy variable, 1 indicating household head is not married or living with a partner.	Coded 'married' in summary dataset.
$turndown_{it}$	Dummy variable, 1 indicating that in the past five years the household had been turned down when applying for credit .	Coded 'turndown' in the summary data set. Based on X407 in full data set.

Table A2, Part1: Complete version of Table 2

	(1)	(2)	(3)	(4)	(5)
sample	ΔD/Y>0	ΔD/Y≤0	full	ΔD/Y>0	ΔD/Y≤0
Y~ cut-off	99th perc	99th perc	99th perc	99th perc	99th perc
Y~ reference definition	av. inc.	av. inc.	av. inc.	av. inc.	av. inc.
Y	-0.314	0.042	0.001	-0.314	0.042
	(0.024)***	(0.003)***	(0.005)	(0.023)***	(0.003)***
Y~	0.216	0.009	0.056		
	(0.080)***	(0.006)	(0.019)***		
HW0 # Y~				0.089	0.004
				(0.115)	(0.006)
HW1 # Y~				0.289	0.010
				(0.085)***	(0.007)
HW	0.086	-0.002	0.006	0.090	-0.003
	(0.017)***	(0.001)	(0.003)*	(0.017)***	(0.001)**
dHW	-1.004	0.006	-0.077	-3.889	-0.074
	(0.200)***	(0.016)	(0.035)**	(1.526)**	(0.108)
REP	0.659	0.011	0.539	0.658	0.012
	(0.047)***	(0.011)	(0.038)***	(0.048)***	(0.011)
dREP	-6.506	-0.150	-5.077	-6.507	-0.153
	(0.563)***	(0.144)	(0.435)***	(0.576)***	(0.144)
FW	0.008	-0.001	0.000	0.008	-0.001
	(0.006)	(0.001)	(0.001)	(0.006)	(0.001)
dFW	0.228	-0.006	0.038	0.188	-0.004
	(0.077)***	(0.007)	(0.011)***	(0.081)**	(0.006)
D_{t-1}	-0.005	-0.024	-0.024	-0.007	-0.024
	(0.009)	(0.001)***	(0.002)***	(0.009)	(0.001)***
dD_{t-1}		0.148	0.215		0.149
		(0.005)***	(0.022)***		(0.005)***
kids	0.035	-0.001	0.011	0.039	-0.001
	(0.023)	(0.002)	(0.004)**	(0.022)*	(0.002)
age	-0.007	-0.001	-0.005	-0.007	-0.001
	(0.005)	(0.000)	(0.001)***	(0.005)	(0.000)
age^2	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)***	(0.000)	(0.000)
N	2,229	14,270	16,510	2,229	14,270
av. R^2	0.65	0.27	0.51	0.65	0.27

Table A2, Part 2: Complete version of Table 2

	(1)	(2)	(3)	(4)	(5)
sample	ΔD/Y>0	ΔD/Y≤0	full	ΔD/Y>0	ΔD/Y≤0
Y~ cut-off	99th perc	99th perc	99th perc	99th perc	99th perc
Y~ reference definition	av. inc.	av. inc.	av. inc.	av. inc.	av. inc.
college	-0.269	-0.012	-0.073	-0.080	-0.016
	(0.111)**	(0.008)	(0.025)***	(0.163)	(0.009)*
HW1#college				-0.297	0.006
				(0.147)**	(0.011)
black	0.218	-0.002	0.044	-0.038	0.000
	(0.100)**	(0.007)	(0.023)*	(0.153)	(0.007)
HW1#black				0.443	-0.012
				(0.149)***	(0.008)
college#black	0.438	0.008	0.094	0.509	-0.006
	(0.098)***	(0.007)	(0.022)***	(0.134)***	(0.010)
HW1#collge#black				-0.145	0.029
				(0.211)	(0.015)*
working	-0.008	-0.001	-0.001	-0.003	-0.001
	(0.048)	(0.003)	(0.008)	(0.047)	(0.003)
lowinc	0.039	-0.017	-0.005	0.039	-0.016
	(0.030)	(0.003)***	(0.007)	(0.029)	(0.003)***
not married	-0.031	0.011	-0.007	-0.033	0.011
	(0.022)	(0.002)***	(0.005)	(0.023)	(0.002)***
turndown	0.035	-0.005	0.023	0.030	-0.004
	(0.023)	(0.003)*	(0.007)***	(0.023)	(0.003)
dum1995	-0.035	0.005	0.011	-0.035	0.005
	(0.029)	(0.003)*	(0.007)	(0.029)	(0.003)*
dum2001	0.013	0.001	0.003	0.013	0.001
	(0.032)	(0.003)	(0.007)	(0.033)	(0.003)
dum2004	0.063	0.004	0.023	0.062	0.003
	(0.032)*	(0.002)	(0.007)***	(0.033)*	(0.002)
dum2007	-0.074	0.003	-0.008	-0.075	0.003
	(0.042)*	(0.003)	(0.009)	(0.044)*	(0.003)
constant	0.758	-0.555	-0.669	2.607	-0.480
	(1.176)	(0.086)***	(0.269)**	(1.710)	(0.093)***
N	2,229	14,270	16,510	2,229	14,270
av. R^2	0.65	0.27	0.51	0.65	0.27

Table A3, Part 1: Owner vs non-owner split

	(1)	(2)	(3)	(4)
sample	ΔD/Y>0	ΔD/Y>0	ΔD/Y≤0	ΔD/Y≤0
Y~ cut-off	99th perc	99th perc	99th perc	99th perc
Y~ definition	av. inc.	av. inc.	av. inc.	av. inc.
restriction	owners	non-owners	owners	non-owners
Y	-0.240	-0.427	0.051	0.019
	(0.025)***	(0.045)***	(0.003)***	(0.003)***
Y~	0.291	-0.008	-0.001	-0.001
	(0.106)***	(0.133)	(0.008)	(0.006)
HW	0.096		-0.006	
	(0.015)***		(0.001)***	
REP	0.475	0.840	0.020	-0.002
	(0.057)***	(0.062)***	(0.017)	(0.003)
dREP	-4.770	-8.405	-0.263	0.014
	(0.656)***	(0.783)***	(0.230)	(0.037)
FW	-0.004	0.010	0.000	-0.001
	(0.006)	(0.011)	(0.001)	(0.001)
dFW	-0.038	0.291	-0.002	0.006
	(0.125)	(0.111)**	(0.012)	(0.005)
D_{t-1}	-0.011	-0.001	-0.024	-0.022
	(0.009)	(0.014)	(0.001)***	(0.001)***
dD_{t-1}			0.158	0.149
			(0.007)***	(0.008)***
kids	0.020	0.009	-0.003	-0.002
	(0.022)	(0.042)	(0.003)	(0.002)
age	-0.006	-0.013	-0.002	0.001
	(0.007)	(0.009)	(0.001)***	(0.000)**
age^2	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)***	(0.000)**
N	1,503	718	10,212	3,667
F-stat	17	98	326	89

Table A3, Part 2: Owner vs non-owner split

	(1)	(2)	(3)	(4)
sample	ΔD/Y>0	ΔD/Y>0	ΔD/Y≤0	ΔD/Y≤0
Y~ cut-off	99th perc	99th perc	99th perc	99th perc
Y~ definition	av. inc.	av. inc.	av. inc.	av. inc.
restriction	owners	non-owners	owners	non-owners
college	-0.356	0.001	-0.002	0.003
	(0.142)**	(0.186)	(0.011)	(0.008)
black	0.452	-0.116	-0.023	-0.008
	(0.125)***	(0.169)	(0.009)**	(0.007)
college#black	0.304	0.412	0.019	0.003
	(0.172)*	(0.129)***	(0.010)*	(0.006)
working	0.019	-0.084	0.003	-0.005
	(0.045)	(0.102)	(0.004)	(0.003)
lowinc	0.101	-0.010	-0.017	-0.005
	(0.038)***	(0.044)	(0.004)***	(0.003)
not married	-0.037	0.005	0.014	0.003
	(0.025)	(0.043)	(0.003)***	(0.002)
turndown	0.002	0.134	-0.007	-0.001
	(0.026)	(0.039)***	(0.003)**	(0.002)
dum1995	-0.028	-0.061	0.004	0.004
	(0.028)	(0.062)	(0.004)	(0.004)
dum2001	-0.018	0.049	0.002	0.004
	(0.035)	(0.068)	(0.003)	(0.004)
dum2004	0.005	0.113	0.005	0.004
	(0.035)	(0.060)*	(0.003)	(0.004)
dum2007	-0.078	-0.036	0.009	0.001
	(0.051)	(0.075)	(0.004)**	(0.004)
constant	-1.887	5.152	-0.456	-0.206
	(1.483)	(2.008)**	(0.115)***	(0.083)**
N	1,503	718	10,212	3,667
F-stat	17	98	326	89

Table A4, Part 1: Robustness check grouping, borrowing sample

	(1)	(2)	(3)	(4)
sample	ΔD/Y>0	ΔD/Y>0	ΔD/Y>0	ΔD/Y>0
Y~ definition	avtop1	avtop1	head	head
grouping	edu-race	region	edu-race	region
Y	-0.291	-0.295	-0.526	-0.454
	(0.041)***	(0.040)***	(0.129)***	(0.105)***
HW0 # Y~	0.092	-0.056	-0.655	-0.407
	(1.150)	(0.145)	(0.369)*	(0.311)
HW1 # Y~	0.109	-0.021	-0.808	-0.657
	(1.188)	(0.113)	(0.345)**	(0.298)**
HW	0.088	0.084	0.076	0.077
	(0.021)***	(0.022)***	(0.021)***	(0.020)***
REP	0.540	0.546	0.496	0.501
	(0.056)***	(0.056)***	(0.054)***	(0.054)***
FW	0.018	0.019	0.015	0.016
	(0.008)**	(0.009)**	(0.008)*	(0.008)*
D_{t-1}	0.007	0.008	0.007	0.008
	(0.010)	(0.010)	(0.009)	(0.009)
region 2	0.117	0.125	0.108	0.088
	(0.058)*	(0.064)*	(0.058)*	(0.061)
region 3	0.131	0.123	0.122	0.090
	(0.059)**	(0.060)*	(0.056)**	(0.059)
region 4	0.055	0.045	0.044	-0.048
	(0.075)	(0.070)	(0.066)	(0.080)
region 5	0.065	0.057	0.053	0.023
	(0.068)	(0.076)	(0.065)	(0.068)
region 6	0.116	0.122	0.106	0.088
	(0.067)	(0.072)	(0.065)	(0.065)
region 7	0.079	0.067	0.059	-0.002
	(0.068)	(0.077)	(0.066)	(0.071)
region 8	0.055	0.022	0.041	0.007
	(0.064)	(0.094)	(0.062)	(0.068)
region 9	0.048	0.060	0.038	0.026
	(0.064)	(0.074)	(0.061)	(0.061)
dum1995	-0.037	-0.057	-0.058	-0.055
	(0.237)	(0.046)	(0.028)**	(0.029)*
constant	2.182	4.358	6.458	5.574
	(16.317)	(2.192)*	(1.659)***	(1.390)***
N	915	914	1,007	1,007

Dependent variable: ΔD/Y. All $ valued independent variables are subject to the inverse hyperbolic since transformation. Coefficients are estimated by OLS using probability weights. Bootstrapped standard errors are obtained by re-estimating the regression 999 times using a set of 999 replicate weights. Stars indicate 1% (***), 5% (**) and 10% (*) significance levels. Full set of results including missing household characteristics can be obtained upon request.

Table A4, Part 2: Robustness check grouping, non-borrowing sample

	(5)	(6)	(7)	(8)
sample	ΔD/Y≤0	ΔD/Y≤0	ΔD/Y≤0	ΔD/Y≤0
Y~ definition	avtop1	avtop1	head	head
grouping	edu-race	region	edu-race	region
Y	0.040	0.040	0.015	0.013
	(0.003)***	(0.003)***	(0.005)***	(0.005)***
HW0 # Y~	0.008	0.035	-0.039	-0.057
	(0.016)	(0.052)	(0.017)**	(0.014)***
HW1 # Y~	0.011	-0.002	-0.120	-0.140
	(0.014)	(0.041)	(0.017)***	(0.017)***
HW	-0.006	-0.007	-0.011	-0.010
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
REP	-0.008	-0.003	-0.002	-0.001
	(0.005)	(0.005)	(0.004)	(0.004)
FW	0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
D_{t-1}	-0.025	-0.026	-0.026	-0.026
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
region 2	-0.011	-0.011	-0.011	-0.015
	(0.008)	(0.008)	(0.007)	(0.008)*
region 3	0.005	0.003	0.001	-0.005
	(0.008)	(0.009)	(0.008)	(0.008)
region 4	-0.006	-0.008	-0.009	-0.025
	(0.009)	(0.010)	(0.009)	(0.009)**
region 5	-0.010	-0.016	-0.018	-0.023
	(0.011)	(0.010)	(0.010)*	(0.009)**
region 6	-0.001	-0.002	-0.004	-0.007
	(0.009)	(0.009)	(0.009)	(0.009)
region 7	-0.005	-0.009	-0.010	-0.019
	(0.010)	(0.010)	(0.010)	(0.010)*
region 8	0.009	0.002	0.001	-0.005
	(0.010)	(0.010)	(0.009)	(0.009)
region 9	0.006	0.008	0.007	0.006
	(0.012)	(0.011)	(0.010)	(0.010)
dum1995	0.005	0.003	0.000	-0.001
	(0.005)	(0.007)	(0.003)	(0.003)
constant	-0.525	-0.871	-0.118	-0.078
	(0.241)**	(0.696)	(0.066)*	(0.057)
N	5,460	4,919	6,415	6,415

Dependent variable: ΔD/Y. All $ valued independent variables are subject to the inverse hyperbolic since transformation. Coefficients are estimated by OLS using probability weights. Bootstrapped standard errors are obtained by re-estimating the regression 999 times using a set of 999 replicate weights. Stars indicate 1% (***), 5% (**) and 10% (*) significance levels. Full set of results including missing household characteristics can be obtained upon request.

Table A5: Effect Size based on average income of top 1% of households within edu-race groups

	total	Y	HW0#Y~	Y~ + HW	REP	FW	D	time	kids
1998	0.77%	-3.26%	-0.51%	5.86%	0.62%	0.67%	-1.25%	0.10%	0.05%
2001	1.10%	-2.29%	-0.14%	6.16%	1.12%	0.40%	-3.16%	0.35%	0.05%
2004	6.99%	-8.62%	1.19%	11.21%	6.80%	1.67%	-3.65%	1.39%	0.13%
2007	2.32%	-2.86%	-0.33%	8.66%	3.46%	0.53%	-4.65%	-0.61%	0.05%

	age	college	black	college#black	working	lowinc	not married	turndown
1998	-0.82%	-0.64%	-0.09%	0.05%	-0.01%	0.00%	0.01%	0.02%
2001	-0.96%	-0.58%	0.18%	-0.04%	-0.03%	-0.03%	0.05%	0.01%
2004	-2.00%	-1.25%	0.12%	0.08%	-0.04%	-0.07%	0.02%	0.01%
2007	-1.30%	-0.72%	0.08%	0.02%	-0.03%	-0.05%	0.05%	0.00%

Table A6: Effect Size based on average income of top 5% of households within edu-race groups

	total	Y	HW0#Y~	Y~ + HW	REP	FW	D	time	kids
1998	0.97%	-1.28%	-0.76%	3.72%	0.35%	0.27%	-1.04%	0.36%	0.02%
2001	0.75%	-0.72%	-0.64%	4.26%	0.33%	0.26%	-2.29%	0.48%	0.03%
2004	5.62%	-4.08%	1.39%	6.70%	4.62%	0.83%	-2.99%	0.99%	0.05%
2007	2.02%	0.10%	-0.44%	4.62%	2.43%	0.20%	-3.53%	-0.22%	0.02%

	age	college	black	college#black	working	lowinc	not married	turndown
1998	-0.57%	-0.14%	-0.05%	0.07%	-0.02%	0.00%	0.02%	0.01%
2001	-0.83%	-0.17%	0.12%	-0.07%	-0.03%	-0.03%	0.04%	0.02%
2004	-1.55%	-0.39%	0.05%	0.10%	-0.07%	-0.06%	0.03%	0.00%
2007	-0.99%	-0.20%	0.05%	0.00%	-0.03%	-0.05%	0.06%	0.00%

Appendix B: Changes in outstanding liabilities

This appendix describes how the change in outstanding liabilities was defined for the other debt categories used in this paper beyond the first mortgage on the primary residence which was discussed in section 2. These other debt categories are second and third mortgages on the primary residence, mortgages on other residential properties, consumer loans, car and vehicle loans, education loans, other loans for property purchase and home improvements. There are four categories on which the SCF collects information but timing information is to scarce to make accurate inferences about the change in the year of the interview. These categories are omitted and correspond credit card debt, credit lines (including home equity lines), loans against land contracts and loans against pension plans. This appendix is best understood after reading section 2.1 as many definitions are the same.

Second and third mortgage on primary residence

Information on the second ($D_{i,t}^{M2}$) and third mortgage ($D_{i,t}^{M3}$) is not as comprehensive as for the first mortgage ($D_{i,t}^{M1}$) which is discussed in section 2. In particular whether the mortgage was refinanced or not is not asked for second and third mortgages and thus a different way to defining the change in the outstanding liability is applied. The key difference is the assumption that all second and third mortgages taken out in the year of the interview are new debt and are not refinanced. The definition of cases for second and third mortgages taken out prior to the year of the interview is equivalent to the categorization of the first mortgage. Cases are summarized in Table B1 which also provides the exact codes of the variables used. D and ΔD stand for $D_{i,t}^{M2}$ or $D_{i,t}^{M3}$ and $\Delta D_{i,t}^{M2}$ or $\Delta D_{i,t}^{M3}$ respectively. The reported numbers of observations refer to the second mortgage in implicate 3 of the 2004 wave.

Table B1: Changes in the outstanding amount of the second and third mortgage on the primary residence

case	step 1	step 2	definition	N	interpretation
1	tB = year		ΔD = D	25	Assumption is that second and third mortgages are not refinanced but new debt.
2.1	tB < year	B > D	$\Delta D^M = -P$	64	Household is in the process of repaying.
2.2	tB < year	B = D	$\Delta D^M = 0$	16	No repayment yet. Interest only scheme or behind schedule and only able to pay interest rates.
2.3	tB < year	B < D	$\Delta D^M = -P$	3	Household is behind on payments and interest accumulated.
				108	

tB is defined as X902 (X1002), B corresponds to X904 (X1004), D is X905 (X1005) and P=X908-(X916/10000*X905) and P=X1008-(X1016/10000*X1005). Terms in brackets refer to the third mortgage.

Mortgages on other residential property ($D_{i,t}^{OM}$)

Mortgages on other residential properties refer to mortgages on vacation homes and other property which is used for residential purposes besides the main residency. Property which is not used for residential property is treated like a business asset and any liabilities on such assets are netted out with the current value. The SCF introduces this convention and it is also used for the purpose of this paper. Since information on how the money was used is not available the year of the purchase of the property (tp) is used as a second best option. It is assumed that mortgages taken out in the current year represent new debt if the property was bought in the same year, otherwise they are refinanced. The justification for this assumption is that one the one hand it only affects a small number of observations and on the other hand households with other residential properties are wealthier and less likely to rely on cash out refinancing out of these additional properties. Using the year of the purchase as additional information adds one more step to the analysis. The SCF asks about up to three mortgages on residential property ($D_{i,t}^{OM} = D_{i,t}^{OM1} + D_{i,t}^{OM2} + D_{i,t}^{OM3}$). All cases for all three mortgages of this category are summarized in Table B2. D and ΔD stand for $D_{i,t}^{OM1}$, $D_{i,t}^{OM2}$ or $D_{i,t}^{OM3}$ and $\Delta D_{i,t}^{OM1}$, $\Delta D_{i,t}^{OM2}$ or $\Delta D_{i,t}^{OM3}$ respectively. The number of observations refers to $D_{i,t}^{OM1}$ in implicate 3 in 2004.

Table B2: Changes in the outstanding amount of mortgages on other residential property

line	step 1	step 2	step 3	definition	N	interpretation
1	tB = year	tp=year		ΔD = D	45	Household used mortgage to buy property thus new debt.
2	tB = year	tp<year	B > D	ΔD = D-B	15	Assumption: mortgage was refinanced and already some repayment occurred.
3	tB = year	tp<year	B = D	ΔD = 0	12	Assumption: mortgage was refinanced but no repayment yet.
4	tB = year	tp<year	B < D	ΔD = D-B	0	Irrelevant. No observations fall into this case in any year.
5.1	tB <year		B > D	ΔD = -P	206	Households paying back.
5.2	tB < year		B = D	ΔD = 0	35	Not paying back yet (due to interest only period).
5.3	tB < year		B < D	ΔD = -P	2	Household behind on schedule and interest rate accumulates.
					315	

tp indicates year property was bougth. tB is defined as X1713 (X1813/X1913), B corresponds to X1714 (X1814/X1914), D is X1715 (X1815/X1915) and P=X1718-(X1726/10000*X1715). All stocks and payments are multiplied by the ownership share (X1705/X1805/X1905). Terms in brackets refer to the second and third residential property beyond the main residence.

Cases 1 and 5.1 to 5.3 are identical to the definitions for second and third mortgages on the primary residence. Cases 2 to 3 are different in that the payment is not inferred from the reported payments and interest rates of the household but rather by the difference between the amount initially borrowed and currently outstanding. The reason is that it is easier to remember outstanding amounts and the amount initially borrowed compared to changing interest rates and payments. Accordingly if the amount initially borrowed is different from the amount currently outstanding the difference is defined as principal repayment. In case the current amount exceeds the amount initially borrowed (case 4), this would indicate that the household fell behind in payments within the first year of the mortgage. In 2004 not a single observation fell into this unlikely case.

Consumer loans

The SCF collects information about up to six unsecured consumer loans[25]. The change in the outstanding amount is derived identically to the cases of second and third mortgages. The reason is that for consumer loans are usually not refinanced. Also the way the SCF asks about them in the interview does not allow for the possibility of refinancing (`How much was borrowed or financed, not counting the finance charges?')[26]. So the assumption that consumer loans taken out in the year of the interview represent new debt seems uncontroversial. The definitions are summarized in Table B3, the number of observations corresponds to the third implicate in the 2004 wave.

Table B3: Changes in the outstanding amount of unsecured consumer loans

case	step 1	step 2	definition	N	interpretation
1	tB = year		ΔD = D	110	Taking on new debt.
2.1	tB <year	B > D	ΔD = -P	160	Paying back an existing consumer loan.
2.2	tB < year	B = D	ΔD = 0	32	Interest only period or household struggles with payments.
2.3	tB < year	B < D	ΔD = -P	6	Households behind schedule and interest accumulated.
				308	

tB is defined as X2713, B corresponds to X2714, D is X2723 and P=X2718-(X2724/10000*X2723). If X2724 is not observed X2719 is used, both expressions are transformed to annual frequencies. tB for the next five consumer loans corresponds to X2730/X2813/X2830/X2913/X2930, B for the next five corresponds to X2731/X2814 etc.

[25] Consumer loans are based on item X2730 as long as the loan was not taken out for investment purposes in non-residential real estate (X2710==78).
[26] Refers to item X2714 in the SCF codebook.

Car and vehicle loans

Loans to purchase cars and other vehicles are categorized identical to consumer loans. Debt taken on in the current period is assumed to be new debt. Definitions and number of observations for each case in the 2004 wave (implicate 3) are displayed in Table B4.

Table B4: Changes in the outstanding amount of car and vehicle loans

case	step 1	step 2	definition	N	interpretation
1	tB = year		ΔD = D	284	Taking out a car loan in the current year thus new debt.
2.1	tB <year	B > D	ΔD = -P	932	Paying back an earlier car loan.
2.2	tB < year	B = D	ΔD = 0	3	Not paying on the loan.
2.3	tB < year	B < D	ΔD = -P	0	Accumulating interest failed to pay.
				1219	

tB is defined as X2208, B corresponds to X2209, D is X2218 and P=X2213-(X2219/10000*X2209). If X2213 is not observed X7537 is used, both expressions are transformed to annual frequencies. tB for the next five car loans corresponds to X2308/X2408/X2509/X2609/X7157, B for the next five corresponds to X2309/X2409 etc.

Education loans

The SCF collects information about up to six education or student loans. While the changes in the outstanding amount are defined in the same way as for car and consumer loans, the number of observations in each case in 2004 (implicate 3) demonstrates the different way in which student loans are paid back. In most student loans principal as well as interest payments only start after graduation. This applies to federal as well as private student loans. Federal student loans come in subsidized and unsubsidized form. Federal student loans normally have lower interest rates than private ones and in addition, interest payments are paid for by the government in the case of subsidized loans. Thus except in the case of subsidized student loans, it is the normal situation after graduation that students face an outstanding amount which exceeds the amount initially borrowed due to accumulated interest. This is reflected by the high number of observations falling into case 2.3 compared to consumer or car loans. For student loans it is also much more common to observe that the outstanding amount equals the amount initially borrowed even if borrowing occurred in some year prior to the interview. The reason is that interest on subsidized federal student loans does not accumulate but is paid for by the government.

Table B5: Changes in the outstanding amount of education loans

case	step 1	step 2	definition	N	interpretation
1	tB = year		ΔD = D	53	Students taking out a student loan in the current year thus new debt.
2.1	tB < year	B > D	ΔD = -P	254	Student not enrolled anymore and paying back the loan.
2.2	tB < year	B = D	ΔD = 0	125	Still enrolled, no accumulation of interest rates and not started to pay back.
2.3	tB < year	B < D	ΔD = -P	60	Non subsidized or private loan, interest accumulated, now in process of repaying.
				492	

tB is defined as X7804, B corresponds to X7805, D is X7824 and P=X7815-(X7822/10000*X7824). If X7815 is not observed X7817 is used, both expressions are transformed to annual frequencies. tB for the next five education loans corresponds to X7827/X7850/X7904/X7927/X7950.

Other loans

Other loans represent loans taken out to buy the primary residence or to undertake home improvements. They are categorized as `other´ because they are not owed to a bank or a financial institution but to relatives or to the seller of the property. Thus they refer to loans which are part of rather informal arrangements. In 2004 (implicate 3) a very limited number of households reported such liabilities and their importance for the aggregate picture is marginal.

Table B6: Changes in the outstanding amount of other loans

case	step 1	step 2	definition	N	interpretation
1	tB = year		ΔD = D	5	Taking out new debt.
2.1	tB <year	B > D	ΔD = -P	16	In the process of repaying.
2.2	tB < year	B = D	ΔD = 0	10	No repayments yet.
2.3	tB < year	B < D	ΔD = -P	0	Irrelevant.
				31	

tB is defined as X1034, B corresponds to X1035, D is X1044 and P=X1039-(X1045/10000*X1044). If X1039 is not observed X1040 is used, both expressions are transformed to annual frequencies. tB for home improvement loans is X1205.