Keeping Up with the Joneses: Evidence from Micro Data

Enrichetta Ravina*

Northwestern University

January 2005

Abstract

This paper provides evidence that habit persistence is an important determinant of household consumption choices, in a setting that allows for heterogeneity and household-specific interest rates. By estimating Euler equations for a representative sample of U.S. credit card account holders, I find that the strength of the external habit, captured by the fraction of the consumption of the reference group that enters the utility function, is 0.290; while the strength of internal habit, represented by household past consumption, is 0.503. These findings provide empirical support to the theories that explain macroeconomic and asset pricing phenomena by introducing habit persistence in the utility function. The results are robust to the inclusion of the income growth rate and other measures of economic activity in the regression, changes in the specification and the instrument set, and tests of liquidity constraints and precautionary saving motives. I also show that this result carries over in the aggregate, once heterogeneity and market incompleteness are taken into account by aggregating the Euler equations as a weighted average of individual marginal rates of substitution. On the contrary, I find that an econometrician that used per capita consumption, constructed from the same data, and a representative agent framework, would find no evidence of habit persistence.

*I would like to thank Debbie Lucas, Kent Daniel, Bill Rogerson and Paola Sapienza for long discussions, support and guidance. Special thanks go to Donald P. Jacobs and the U.S. credit card company that provided the data. I am also grateful to George Akerlof, Roc Armenter, Gadi Barlevy, Paco Buera, Francesco Caselli, Colin Camerer, John Campbell, Janice Eberly, Martin Eichenbaum, Andrea Eisfeldt, Mike Fishman, Bob Gordon, Johannes Horner, Ming Hsu, Arvind Krishnamurthy, David Laibson, Ron Laschever, Annamaria Lusardi, Sendhil Mullainathan, John Panzar, Marcin Peski, Karl Schmedders, Christopher Polk, Matthew Rabin, Marciano Siniscalchi, Mark Seasholes, Costis Skiadas, Annette Vissing-Jorgensen and seminar participants at Kellogg School of Management (Finance) and Northwestern University for valuable discussions and suggestions. Financial support from the Northwestern University Dissertation Year Fellowship, Banca d'Italia and Fondazione Einaudi is gratefully acknowledged. All remaining errors are mine. Comments are welcome: e-ravina@northwestern.edu.

1 Introduction and Motivation

This paper provides evidence that habit persistence is an important determinant of actual household consumption choices, in a setting that allows for heterogeneity and individual-specific borrowing rates.

It also shows that this result carries over in the aggregate, once the aggregation of individual consumption choices is properly performed, and heterogeneity and nonlinearity of marginal utility are taken into account. On the contrary, I show that an econometrician that used per capita consumption, constructed from the same data, and a representative agent framework, would find no evidence of habit persistence.

Habit formation models have proven very successful in theoretically explaining a variety of dynamic asset pricing phenomena and macroeconomic facts. In the asset pricing literature, they have been employed to explain the equity premium puzzle (Constantinides (1990), Abel (1990, 1999), Campbell and Cochrane (1999)), the procyclical variation of stock prices (Campbell and Shiller (1988)), and the countercyclical variation of stock market volatility (Harvey (1989)). In the macroeconomics literature, habit persistence frameworks explain savings and growth (Carroll, Overland and Weil (2000)), business cycle facts (Boldrin, Christiano and Fisher (2001)), the equity home bias (Shore and White (2002)), and consumption's response to monetary and other shocks (Fuhrer (2000)). However, despite their impressive track record in simulations with aggregate data, the evidence on whether these models reflect actual preferences is mixed. The empirical studies that have addressed this question so far have mostly followed the macroeconomists' approach to aggregate consumption, leaving the micro foundations of the phenomenon largely unexplored.[1]

In this paper, I take a different approach and look into actual household consumption decisions. I estimate a log-linearized Euler equation that incorporates time nonseparabilities and externalities in consumption choices, in a setting characterized by uninsurable income shocks and household-specific borrowing rates.

To measure household consumption, I use a novel panel data set consisting of 2,674 U.S. credit card accounts located in California, over the period between the third quarter of 1999 and 2002. The data provide information on spending and borrowing patterns, the evolution of interest rates, and credit availability, as well as a snapshot of the main economic and demographic characteristics

[1] An exception is Dynan (2000), who investigates habit formation in annual food consumption data from the PSID and reaches negative conclusions. I illustrate in detail the differences between my approach and hers, and the possible reasons for the different findings, later in this section.

of the account holder and the zip code of the area in which she lives. I construct the consumption measure as the sum of all the credit card purchases over the quarter. The main advantage of the data consists in providing a more comprehensive measure of consumption than food, and detailed information on the evolution over time of household-specific financial information. While far from perfect, it allows me to overcome some of the drawbacks of publicly available data sets, as the Panel Study of Income Dynamics (PSID), that contains only information on food consumed at home and out, and the Consumer Expenditure Survey (CEX), that provides a very detailed measure of consumption, but follows every household only up to five quarters, and doesn't contain any household-specific geographic or detailed financial information. The comparison with the U.S. Census and the Survey of Consumer Finances indicates that the sample is representative of the U.S. population for both demographic characteristics and borrowing behavior. Moreover, this measure of consumption exhibits the characteristics we expect to see in household consumption: a hump shaped path over the life cycle and an increasing relationship to income. Another important feature of the data set is the fact that I link the credit card data to the city-level quarterly retail sales data. This allows me to define a more intuitively appealing measure of the external reference point for each household: the consumption of the city in which it lives, rather than the consumption of the entire nation.

I find that the strength of external habit, captured by the consumption of the reference group, is 0.290 (significant at the 5% level), while the strength of internal habit, represented by household past consumption, is 0.503 (significant at the 1% level). These coefficients represent the fraction of city-level aggregate consumption (or own past consumption) that enters the utility function as the reference level to which the household compares itself to. A coefficient of zero would imply that the household is not influenced by the consumption of its neighbors (its own past consumption), and the model collapses to the standard one used in the literature. On the contrary, a coefficient of one would mean that the household only cares about the way its consumption compares to the neighbors' (its past own), and not about the absolute level.

The results are robust to the inclusion in the regression of current and future income growth rates, the change in city-level unemployment rate and measures of the housing market conditions. The inclusion of these variables has the purpose of controlling for the possibility that city or lagged household-level consumption growth capture the effect of some omitted variable and this fact drives the results. I also perform further robustness checks by varying the specification and changing the instrument set. All the tests confirm the economic and statistical significance of the habit persistence coefficients.

Alternative explanations of the findings include the presence of liquidity constraints and precautionary saving motives. These phenomena, like internal habit formation, cause consumption to adjust slowly to changes in income, and therefore induce a positive correlation between current and lagged consumption growth rates. I test the habit formation hypothesis against a liquidity constraints model by including the lagged growth rate of income in the estimation equation, re-estimating the Euler equation on two sub-samples of unconstrained and credit constrained HHs, and by adding a credit constrained indicator directly in the regression. The tests show some evidence of liquidity constraints, in addition to those accounted for by the household-specific borrowing rate, but indicate that they are not the cause of the results. The precautionary motive story has similar implications and is further tested by adding a measure of consumption uncertainty to the regression. Again, the validity of the habit persistence interpretation of the evidence is confirmed.

The findings in this paper contradict those in Dynan (2000), who investigates habit formation in annual food consumption using the PSID and finds no evidence of habit persistence. The differences between this study and hers are many: among them there are the measure of consumption used, the annual versus quarterly frequency of observation, the different estimation equation, which in her case doesn't take the variation of the interest rates into account. However, a closer examination of the two methodologies indicates that the main reason of the different findings is the better quality of the instrument set used in this study, especially due to the availability of household-specific financial information. My instrument sets generates partial R^2s that are more than double and F statistics for the lagged household consumption growth rate above 240 and way out of the dangerous range identified by Stock and Yogo (2002), while this is not the case for Dynan's study. Once I drop the financial variables from the instrument set the endogenous variables are not as well captured as before and, most important, the coefficient of the internal habit drops from 0.60 to 0.13.

Another interesting result and contribution of the paper is the finding that households respond to the price of consumption. The availability of household-specific borrowing rates and financial information allows to study the sensitivity of consumption to individual-specific interest rates: the short run elasticity to the borrowing rate is estimated to be -1.876, and it is statistically significant at the 1% level. Previous studies had difficulties in getting precise estimates of this parameter, as they were forced to use the after tax risk-free rate, which by nature displays very limited cross sectional variability. This magnitude is consistent with the results of Gross and Souleles (2002), who estimate the elasticity of debt, and therefore consumption, to the borrowing rate to be -1.3.

I also examine the aggregate implications of the micro findings and provide a comparison with a representative agent framework. Following Attanasio and Weber (1993), I aggregate household consumption choices and investigate the role of nonlinearity of the marginal rate of substitution (MRS), and the inability of the aggregate studies to account for demographic characteristics, and heterogeneity of preferences, opportunity sets and shocks across households. In particular, the correct way of aggregating across individuals would be to sum their MRS and obtain an average growth rate of aggregate consumption equal to $\Delta\left(\frac{1}{N}\sum_{i=1}^{N} lnc_{i,t}\right)$. On the contrary, the measure available from aggregate data only allows to calculate the MRS of a fictitious individual who consumes per capita consumption: $\Delta ln\left(\frac{1}{N}\sum_{i=1}^{N} c_{i,t}\right)$. By re-estimating an Euler equation similar to that used in the micro data on the correct measure of aggregate consumption, I obtain a habit persistence coefficient of 0.515 (significant at the 5% level). On the contrary, the effect disappears once I estimate the equation using per capita consumption, constructed from the same data. Time-varying skewness of the cross sectional distribution of household consumption plays a role in the differences between the two measures, although the results suggest that omitted demographic variables and other measures of heterogeneity also play a very important role. These findings show the inadequacy of the representative agent as a description of real households consumption choices and of their aggregate implications.

Beside the literature on household-level consumption, asset pricing and macroeconomics, this paper is related to a vast literature that studies the implications of interpersonal effects and time nonseparabilities in settings ranging from status and conformity (Akerlof (1997)) to criminal behavior, welfare choices and labor choice (Case and Katz (1991), Bertrand, Luttmer and Mullainathan (2000), Lalive (2004)), auto purchases (Grinblatt, Kelohariu and Ikaheimo (2004)), addiction (Becker and Murphy (1988)), and stock market participation (Hong, Kubik and Stein (2004)).

The remainder of the paper is organized as follows. Section 2 contains a brief overview of the related literature, while Section 3 describes the data and compares them to those traditionally used in the literature. Section 4 presents the Euler equation that will guide the estimation, describes the empirical strategy, and illustrates the results. Section 5 contains various robustness checks, while Section 6 examines alternative explanations for the results. Section 7 investigates the aggregate implications of the micro findings described in the paper and provides a comparison with a representative agent framework. Section 8, contains a discussion of the relationship between these findings and the CRRA functional form commonly used in the macroeconomic and asset pricing literature and a comparison of the coefficients to the parameters in Abel (1990 and 1999), Constantinides (1990) and Campbell and Cochrane(1999). Section 9 concludes.

2 Related Literature

Habit persistence plays a central role in the asset pricing and macroeconomic literatures that investigates the equity premium puzzle and output persistence. Introducing time non-separabilities and consumption externalities in the utility function reconciles a volatile discount factor with smooth consumption growth by letting marginal utility depend on how far consumption is from the habit level, rather than on its absolute value. Constantinides (1990), Sundaresan(1989), Abel (1990 and 1999), Campbell and Cochrane (1999) and a vast literature thereafter use this specification in a representative agent framework. Using simulations they show that it generates the equity premium and some of the empirical properties of aggregate consumption and asset prices with reasonable values of the parameters. However, despite their impressive track record in simulations with aggregate data, the evidence on whether these models reflect actual preferences is not conclusive and comes mainly fom aggregate data. Heaton (1995) finds evidence of durability at very short horizons and of habit persistence at quarterly frequencies; Constantinides and Ferson (1991) find support for habit formation in quarterly non-seasonally adjusted data; while Eichenbaum, Hansen and Singleton (1988) find evidence of habit persistence in leisure choices, but not in consumption. This paper contributes to this branch of the asset pricing and macroeconomic literature by providing evidence of habit persistence in actual household consumption choices, in a setting that allows for heterogeneity and household-specific borrowing rates.

The papers more closely related to mine are Dynan (2000), Lupton (2003) and Chen and Ludvigson (2003). The first two studies investigate internal habit in the PSID micro data. They both find no evidence of habit persistence in food consumption; while Lupton further analyzes financial decisions and finds support for internal habit formation in that context. The section on the empirical results examines the possible reasons for these findings and concludes that the availability of more detailed household-specific financial information allows me to construct an instrument set that better captures lagged household consumption growth. Conversely, the paper by Chen and Ludvigson analyzes both internal and external habit formation using U.S. aggregate consumption data and finds support for a non-linear form of internal habit. Contrary to their findings, in my data set per capita consumption doesn't display habit formation, even though once the proper aggregation procedure is used aggregate consumption behavior does conserve this feature. Possible reasons for the different findings with regard to per capita aggregate consumption are the estimation technique and the different length of the time period analyzed.

Another paper that investigates the asset pricing implications of household behavior is Chetty and Szeidl (2004), who provide evidence of the importance of consumption commitments and show that their aggregate implications are observationally equivalent to a representative agent that displays habit formation. On the contrary, this paper shows that household-level consumption of nondurables and services itself is characterized by habit formation.

The paper is also related to the microconsumption literature from which it borrows the estimation techniques and the focus on the micro data. This paper builds on Attanasio and Weber (1993), who examine the effects of aggregation on the estimation of the elasticity of intertemporal substitution, and contributes to this literature by investigating the aggregate implications of household behavior in a model with habit persistence and differential borrowing and lending rates. Another contribution related to this area, is the analysis of the sensitivity of consumption choices to household-specific interest rates and the finding that, contrary to the conclusions reached previously from the analysis of the risk-free rate, households do respond to prices. Work in this area includes Gross and Souleles (2002), Ausubel (1999), Attanasio and Weber (1993 and 1995), and Vissing-Jorgensen (2002). The paper also also provides new evidence on the importance of accounting for liquidity constraints and their relevance in household consumption decisions and relates to the papers of Hayashi (1985) and Zeldes (1989), Carroll (2001), and the vast literature on liquidity constraints and precautionary saving motives.

Finally, this paper contributes to the economic literature that incorporates sociological and psychological factors into economic models in order to obtain a more realistic description of human behavior and more accurate predictions and policy implications. Both the economic and the psychological literature stress the importance of interpersonal effects and time nonseparabilities in consumption choices. In the economics realm, Duesenberry (1949) and Veblen (1899) postulate that consumers imitate each others' purchases due to a "status anxiety" and a desire to conform to the expectations of the people in their reference group.[2] At the same time, Pollak (1970) and Ryder and Heal (1973) stress that the utility derived from a certain level of consumption depends not only

[2]This phenomenon is pervasive in everyday life. It is featured in comics ("Keeping Up with the Joneses" by Pop Momand in The New York World), TV shows ("Keeping Up Appearances", BBC) and sociology books (de Botton (2004)). A recent study by Michael Marmot provides evidence that our rank in society can have a significant effect on our health and lifespan (*The Status Syndrome: How Social Standing Affects Our Health and Longevity* (2004)).

on its absolute amount, but also on how it compares to past consumption. A vast literature thereafter has investigated the implications of these features of human behavior in settings ranging from labor choice and criminal behavior, to addiction and stock market participation (Case and Katz (1991), Bertrand, Luttmer, and Mullainathan (2000), Lalive (2004), Grinblatt et al. (2004), Hong, Kubik and Stein (2004). Future research in this direction involves the analysis of the variation in the strength of external habit persistence across cities of different size, once heterogeneity is kept constant. The rationale behind this empirical strategy is that in smaller cities, people have better chances of seeing each others and interacting. However, smaller cities are also usually populated by people more similar to each other. Hence, the necessity of an analysis of the behavior across the size dimension, by keeping the heterogeneity dimension fixed. Another estimation strategy I am exploring is the use of lottery winnings at the zip code level, as an exogenous shock to the external reference point of the households. Section 6.1 contains preliminary results in this area.

3 Data Description

To measure household consumption choices, I use a panel data set consisting of 2,674 U.S. credit card accounts located in California, over the period between the third quarter of 1999 and the third quarter of 2002. The data provide information on spending and borrowing patterns, the evolution of interest rates, and credit availability, as well as a snapshot of the main economic and demographic characteristics of the account holder and the zip code of the area in which she lives.[3] These data are linked to the city-level quarterly retail sales data. The purpose is to obtain a more intuitively appealing measure of the external reference point for each household: the consumption of the city in which it lives, rather than the consumption of the entire nation. The data are also augmented with the Census, BLS and ACCRA information for the area in which the HH lives. This allows me to perform many robustness checks, as it provides information on median house values and rent, median income, unemployment, price level data, and city-level mortgage rates.

The definition and sources of the variables are described in Table I, while summary statistics are presented in Table II. In constructing the sample, I exclude people whose accounts are inactive and those that don't use the card very often, in order to obtain a more meaningful measure of consumption. Following the literature, I also exclude retired account holders, as modeling their consumption and borrowing decisions is very complex and requires the consideration of issues such

[3] The data set has been kindly provided by one of the major U.S. credit card issuers, which would like to remain anonymous.

as bequests, failing health and other specific factors that are different from the scope of this study.[4] Section I of Appendix I provides a detailed description of the sample selection procedure; while the following subsections provide a closer look to the different variables, show that the data set is representative of the U.S. population for both demographic characteristics and borrowing behavior, and illustrate the main features of the consumption measure proposed in the paper.

3.1 Demographic and Financial Characteristics

My sample is representative of the U.S. population in terms of both demographic characteristics and borrowing behavior.

Section B of Appendix I compares the demographics of the account holders (from now on referred to as households, HHs) to those of the U.S. population, reported by the 2000 U.S. Census. Both the distributions of income and age are very alike. The main discrepancy is due to the fact that HHs in the lower range of income and individuals in very young or old age are under-represented in the credit card data set. The data set also contains information on the occupation of the account holders, although it is harder to make a comparison with the Census, given the different classification criteria. An interesting feature related to this variable is that 2.51% of the HHs in the sample are headed by self-employed individuals. This category is somewhat problematic, since they could be using the credit card for their business rather than personal expenditures. In the estimation, I control for this fact by including a dummy variable equal to one if the person is self-employed and zero otherwise. Also, when I exclude entrepreneurs from the analysis the results don't change.

The sample compares well to the U.S. data with regard to household borrowing behavior as well. Section B of Appendix I contains a comparison between my data set and a large multi-issuer credit card data set covering the period between 1995 and 1998 and used by Gross and Souleles (2002). Indebtedness is highly skew in both data sets and all the variables displays similar medians and averages across the two data set.

The other main data set containing information on assets and liabilities of U.S. HHs is the Survey of Consumer Finances. The samples are similar with regard to the percentage of people not paying the balance in full at the end of the month, which is estimated to be 44.4% in the SCF and 45.75% in my sample. However, this survey has been proven to suffer from under-reporting of

[4]Similar selection procedures are followed by Zeldes (1989) and Attanasio and Weber (1993 and 1995).

debt.[5] An advantage of my data set is precisely that it is not a survey and therefore under-reporting and measurement error in the financial variables are not an issue.

3.2 Credit Card Expenditures as a Measure of Consumption

The measure of household consumption that I use is obtained by summing the purchases and cash advances charged on the credit card each quarter.[6] Since the previous section shows that this sample is representative of the U.S. credit card accounts, an important question is whether credit card expenditures are a good measure of household consumption.

Statistics from the Survey of Consumer Finances (SCF) and other sources indicate that in year 2000, 76% of Americans had at least one credit card and charged on it over $1 trillion in purchases, more than they spent in cash.[7] Credit card purchases represent an increasing fraction of U.S. consumer spending, having recently overtaken cash and filled part of the gap with respect to checks. In particular, bank credit cards, retail cards, and debit cards account for roughly 24% of personal expenditures in the United States.[8] Section C of Appendix I provides evidence that on average the HHs in my data set use this credit card conspicuously and are therefore offering a good measure of their expenditures.

Most important, Panels A and B of Figure I show that the measure of consumption constructed in my data set exhibits the characteristics we expect to see in household consumption: a hump shaped path over the life cycle and an increasing relationship to income.

The main advantage of the data consists in providing a more comprehensive measure of consumption than food, and detailed information on the evolution over time of household-specific financial information. While far from perfect, it allows me to overcome some of the drawbacks of publicly available data sets, as the Panel Study of Income Dynamics (PSID), and the Consumer Expenditure Survey (CEX). In particular, the PSID contains only information on food consumed at home and out, which is inadequate for many reasons. First, food is a necessity, its share of expenditure falls with wealth, and it might not represent well the overall consumption basket. Moreover, using food as a proxy for total consumption implicitly assumes separability between food and other commodities, and this has been rejected by numerous studies. In particular, Attanasio and We-

[5]Both Gross and Souleles (2002) and Laibson et al. (2000) point out this problem.

[6]Notice that the data set contains information on the size and timing of balance transfers and that these quantities are excluded from the consumption measure.

[7]Lim, Paul J., and Matthew Benjamin. "Digging Your Way Out of Debt", U.S. News and World Report, (3/19/01).

[8]Gerdes and Walton (2002) provide a description of noncash payments in the U.S, while Zinman (2004) provides an analysis of the credit and debit card markets in the U.S.

ber (1995) analyze the bias induced by using PSID food consumption rather than a more general measure and find it sizeable. Another problem with the PSID measure of consumption is that the way the question about food expenditure is posed leaves a lot of space to the interpretation of the timing of the variable making the choice of the correct timing for the instrumental variables very hard. Finally, Runkle (1991) estimates that up to 76% of the variability of food expenditure is noise. The Consumer Expenditure Survey (CEX) obviates to some of the drawbacks of the PSID by providing a very detailed measure of overall consumption. However, the families are only interviewed up to five quarters and it doesn't contain any household-specific geographic or detailed financial information.

Despite having some important advantages, the data set used in this study has also some shortcomings. The main one is the fact that I observe the income and the demographic characteristics of the HHs only at a certain point in time and not their evolution. In the estimation and robustness sections I address this point by adding aggregate income and future income to the regression to control for the possibility that some relevant information about the time variation of income is missing and my measure of aggregate consumption captures this aspect. Another feature of these data is that we can observe expenditures fluctuate on this card even though they stay relatively stable overall, if the appeal of using this card versus another varies over time. This is addressed by controlling for interest rates and credit line variations, unused portion of the credit line and balance transfers in and out of the card. Also, low activity accounts are excluded and time dummies and account characteristics will capture any aggregate phenomenon and any individual time invariant one respectively. Finally, there is no particular pattern in the way the appeal of this card should vary over time and economic conditions or across HHs. This will add noise to the data and make the estimation harder, but, after controlling for the variables listed above, will not bias the results in any particular direction. Table III compares the mean and standard deviation of my consumption measure to those of both micro-level and aggregate data and shows that they are comparable. The statistics confirm that this measure of consumption is very noisy, with a standard deviation of 2.73 if all the observations are used, down to 0.37 if only those cases in which the growth rate of consumption is between -1.1 and 1.1 are considered. This last figure compares well to the statistics in Zeldes (1989) who uses annual data on food from the PSID and the same selection criterion. A comparison with the CEX data is provided as well: since these are cross sectional averages, I build the same quantity in my data set. The volatility of my individual consumption measure is 0.33, way above the value of 0.06 obtained by Brav et al. (2002) over the period between 1982 and 1996.

One of the possible explanations for these results is that since credit card expenditures are more volatile than total non-durable and services consumption because they are only a fraction of it.

Finally, notice that the assumption underlying the estimation will be that of separability between credit card expenditures and the rest of the consumption basket. Unfortunately, detailed statistics on the type of goods people buy on credit cards are not available. Macroeconomic model of cash and credit good can shed light on the issue of separability.

3.3 City-level Consumption

I choose the reference group of the household to be the city in which it lives. The measure of local aggregate consumption that I employ in the analysis is city-level quarterly per capita taxable sales, from the California Board of Equalization (BOE). The main reason I use this variable is that it is a very comprehensive and natural measure of aggregate consumption and, despite it doesn't capture all consumption categories, it is a good aggregate counterpart for credit card expenditures[9]. Also, retail sales constitute an important component of personal consumption expenditures at the national level.[10] Finally, the measure used here is not a sample, but the total of all reported taxable sales and therefore it doesn't suffer from sampling error.

Table IV contains a comparison of the summary statistics of my measure of local aggregate consumption and personal consumption expenditures on nondurables and services from NIPA. The two measures compare very well in their variability and are highly correlated.

Finally, from the summary statistics displayed in Table II we can see that the standard deviations of income, house values and unemployment rate are very big, reflecting the high cross sectional variations in economic conditions faced by the HHs.

4 Estimation and Empirical Evidence on the Micro Foundations of Habit Persistence

Households face uninsurable income shocks and borrowing rates that depend on their asset position and credit history. The theoretical and calibrated models of intertemporal choice show that dis-

[9] The main categories excluded are necessities (food consumed at home, prescription medicines) and sales for resale. More details on this variable are provided in the Appendix.

[10] This quantity is constructed in the National Income and Product Accounts (NIPA) from a variety of sources among which a monthly sample of national retail sales plays a central role See Wilcox (1992) for a thorough description of the way NIPA personal consumption expenditures are constructed and the implications of these imperfections for empirical work.

regarding borrowing frictions, and especially the wedge between borrowing and lending rates, can lead to very unrealistic predictions in terms of the amounts borrowed and the portfolio allocations of the households.[11] Despite this evidence, empirical analyses of consumption decisions usually assume that HHs can borrow and save at the same rate, due to lack of data. In this paper I depart from these assumptions. In Appendix II, I present a simple yet realistic model of intertemporal choice that incorporates uninsurable income risk and household-specific borrowing rates.[12] The Euler equation derived from this model provides guidance about the variables to consider in the empirical analysis and offers a framework to aid the interpretation of the results:

$$u^c_{i,t} = \beta E_t \left[[u^c_{i,t+1} + \beta \zeta E_Y u^h_{i,t+2}](1 + (R^f_{i,t+1} - 1) 1[Y^H_{i,t+1}])(1 + (R^C_{i,t} - 1) 1[B]) - \zeta u^h_{i,t+1} \right] \quad (4.1)$$

where $1[Y^H_{t+1}]$ is an indicator function that equals one for high realizations of income that will place the HH in the non-borrowing region next period;[13] while $1[B]$ is an indicator function that equals one if in the current period the HH is borrowing.

These results are very intuitive. The household decides how much to consume today versus tomorrow by weighting future utility and different interest rates by the probability that it will actually face them. If for a moment we disregard the effect of the habit stock, we can see that if the HH consumes \$1 less today it looses $u^c(c_t)$ and gains the following: next period the credit card balance will be \$1 lower and so one more dollar will be available for consumption, yielding a utility of $u^c(c_{t+1})$; if, in addition to this gain, the income realization is high enough that the HH is able to repay the balance in full, it will earn the gross risk-free rate on the dollar moved through time and the utility will be $u^c(c_{t+1})R^f_{t+1}$. Analogously, if the HH carries a balance, consuming one dollar less today means that the credit card balance next period will be R^C_t dollars less. The utility deriving from this intertemporal transfer will be $u^c(c_{t+1})R^C_t$ if the HH doesn't have enough resources to pay the balance in full in period *t+1*, and $u^c(c_{t+1})R^C_t R^f_{t+1}$ in case it does and can invest the dollar charged on the credit card at the risk-free rate. The presence of the habit stock

[11]Davis, Kubler and Willen (2004) show in a theoretical model that allowing the HHs to borrow and save at the same low interest rates generates unrealistic predictions in which many are up to they credit limit and borrow money to invest in the stock market. These results are robust to the inclusion of quantity limits on the amount that each HH can borrow, but disappear once a wedge between borrowing and lending rate is introduced.

[12]Borrowing limits and default motives are not modelled directly, but will be accounted for in the estimation. Heuristically, the availability of unsecured debt and bankruptcy protection allows better consumption smoothing and decreases precautionary motives for saving, leading households to consume more and have a smoother consumption path. There is however, also a supply effect: banks and credit providers face a higher probability of default and larger losses and therefore decrease credit availability and charge higher interest rates. Which of these effects prevails is an empirical issue that hasn't been settled yet (Gropp, Scholtz and White (1997), Berkowitz and Hynes (1999)).

[13]See Appendix II for a more detailed explanation.

generates an additional effect due to the fact that when the HH consumes one dollar less today it increases tomorrow's utility not only directly, but also by decreasing the habit level.

4.1 Specification and Estimation Issues

In this section, I estimate the effect of household past consumption and the consumption level of the reference group on individual choices. By using a non-structural specification, I obtain an estimate that is not influenced by any specific restrictions on the form of the utility function. The only principle imposed is that each household chooses consumption with the objective of maximizing its intertemporal utility function.

Following Deaton (1992), the above Euler Equation (4.1) can then be expressed as a second order difference equation in $u^c_{i,t}$, whose solution is given by:

$$u^c_{i,t} = \beta E_t \left[u^c_{i,t+1}(1 + (R^f_{i,t+1} - 1) \Pr[Y^H_{i,t+1}])(1 + (R^C_{i,t} - 1)1[B]) \right] \tag{4.2}$$

This equation holds approximately if the number of lags of consumption entering the habit stock is small relative to the HH lifetime horizon and the HH has static expectations about future interest rates.[14]

Following the consumption literature, I consider a log-linear version of (4.2):

$$\ln u^c_{i,t} = \ln \beta + k + \ln u^c_{i,t+1} + \ln(1 + (R^f_{i,t+1} - 1) \Pr[Y^H_{i,t+1}]) + \ln(1 + (R^C_{i,t} - 1)1[B]) + \varepsilon_{i,t+1} \tag{4.3}$$

where $\varepsilon_{i,t+1}$ contains an expectation error, a multiplicative measurement error in consumption and preference shocks; while k contains second and possibly higher moments of the variables, which, as it is traditional in the literature, are assumed to be constant or uncorrelated with the instruments used in the estimation.[15]

The utility function depends not only on the level of current consumption, but also on own past consumption, the consumption of the reference group and demographic characteristics:

$$u(c_{i,t}, H_{i,t}, \Theta_i, h_{i,t}) = u(c_{i,t} - h_{i,t} - H_{i,t}) \exp(\theta' \Theta_{i,t}) \tag{4.4}$$

[14] Hayashi (1985) obtains a similar result and provides a proof of this statement. The equation holds exactly if the interest rates are constant.

[15] In the robustness section I provide evidence supporting this assumption by showing that including a measure of the variance of consumption among the regressors doesn't change the results.

The variable $h_{i,t}$ represents the internal habit stock and depends on the household's past consumption:

$$h_{i,t} = \zeta c_{i,t-1}$$

The effect of internal habit formation is to make the utility derived from consumption depend on the amount of previous consumption the HH has enjoyed. An implication of this is that in order to mantain the marginal utility of constant, an increase in past consumption needs to be followed by an increase in present consumption, as the stock oh habit to which consumption is compared is higher. Therefore the HH will try to smooth not only consumption levels, but also changes. Another implication of internal habit is that consumption will react slowly to changes in permanent income to avoid the risk of building a habit too quickly.

The variable $H_{i,t}$ represents the external habit level and it captures the complementarity between the consumption of each HH and its reference group. It is modelled as a function of various lags of the aggregate consumption of the area in which the HH lives:

$$H_{i,t} = \alpha_0 C_{i,t} + \alpha_{-1} C_{i,t-1} \tag{4.5}$$

Finally, $\Theta_{i,t}$ represents household demographic characteristics. Micro level consumption studies provide evidence that age, family characteristics and labor supply choices are very important explanatory factors for individual consumption.[16] Following this literature, I condition on the optimal value of these variables by incorporating them in the utility function in the multiplicative way shown in (4.4), if they are time-varying, or by adding them directly to the estimation equation (4.3), if they are constant over time. In particular, $\Theta_{i,t}$ contains age and age squared, as measures of the evolution of family size over the life-cycle, an unobservable HH-specific effect, a time-varying effect that is constant across HHs and an idiosyncratic component orthogonal to the previous two:

$$\Theta_{i,t} = \theta_1 age_{i,t} + \theta_2 age_{i,t}^2 + a_i + t_t + e_{i,t} \tag{4.6}$$

I also include in the estimation equation individual characteristics such as marital status of the HH's head, homeownership, income bracket, occupation and, in some specifications, the median income, house value and unemployment rate in the zip code area of the HH at the end of 1999. Finally, to control for cyclical fluctuations in consumption, I include in the estimation seasonal

[16] See the work of Attanasio and Weber (1993 and 1995), Attanasio and Browning (1993) and Zeldes (1989).

dummies. This specification is equivalent to modelling the discount factor as depending on HH socioeconomic characteristics and the time and seasonal dummies.

The log-linear Euler equation (4.3) can be re-written more extensively as:

$$\begin{aligned}\Delta \ln c_{i,t} = k_1 + \alpha_0 \Delta \ln C_{i,t} + \alpha_{-1} \Delta \ln C_{i,t-1} + \zeta \Delta \ln c_{i,t-1} + \gamma \ln(1 + (R^f_{i,t} - 1) \Pr[Y^H_{i,t}]) \\ + \eta \ln(1 + (R^C_{i,t-1} - 1) 1[B]) + \theta_1 \Delta age_{i,t} + \theta_2 \Delta age^2_{i,t} + \theta_3 maritstatus_i \\ + \theta_4' (socioec.char)_i + \theta_5' (local.char)_i + Seas.Dummies + \varepsilon_{i,t}\end{aligned} \tag{4.7}$$

or, without accounting for the socioeconomic and local area characteristics, as:

$$\begin{aligned}\Delta \ln c_{i,t} = k_1 + \alpha_0 \Delta \ln C_{i,t} + \alpha_{-1} \Delta \ln C_{i,t-1} + \zeta \Delta \ln c_{i,t-1} + \gamma \ln(1 + (R^f_{i,t} - 1) \Pr[Y^H_{i,t}]) \\ + \eta \ln(1 + (R^C_{i,t-1} - 1) 1[B]) + \theta_1 \Delta age_{i,t} + \theta_2 \Delta age^2_{i,t} + Seas.Dummies + \varepsilon_{i,t}\end{aligned} \tag{4.8}$$

where, following Muellbauer (1988) and Dynan (2000), I approximate the expression $\ln u(c_{i,t} - H_{i,t} - h_{i,t})$ with $\ln u(c_{i,t}) - \ln u(H_{i,t}) - \ln u(h_{i,t})$.[17] The utility function is specified to depend linearly on the growth rate of HH and aggregate consumption for purposes of simplicity and flexibility. Section 8 shows that this specification subsumes the widely used CRRA functional form. Notice also that unobserved household heterogeneity in consumption levels is taken into account, since the equation is in first differences and the HH fixed effect a_i contained in (4.3) drops out.[18]

The identification of the parameters is achievied in the cross sectional dimension. The equation is estimated using a GMM procedure with robust standard errors. Household and aggregate consumption, as well as the interest rates are treated as endogenous, either because they are uncertain at the time the HH makes the consumption decision or because they are affected by measurement error.

In particular, Table V shows that the autocorrelation in HH consumption growth rates is negative and consistent with a MA(1) structure induced by the presence of measurement error and taste shocks when true consumption changes are not serially correlated.[19] In my specification, I follow Vissing-Jorgensen (2002) that shows that consistent estimates of the parameters can be obtained if the error is multiplicative in levels and independent of true consumption, returns and

[17] Muellbauer (1988) and Dynan (2000) show that the correlation between this approximation and the exact expression is very high.

[18] Due to the presence of lags of the dependent variable in the equation, the within estimator would lead to inconsistency and first differencing is the best way to account for the household specific component. Chamberlain (1984), Arellano and Bond (1989) and Runkle (1991) discuss this issue in detail.

[19] Similar values of the autocorrelation coefficients are obtained by Hayashi (1985), who analyzes a panel of Japanese household expenditures.

instruments.[20]

Local aggregate consumption is considered endogenous since it is simultaneously determined with household consumption and is not observed perfectly in the current period. Similarly, the risk free interest rate is considered endogenous because it is uncertain at the moment in which the HH makes the consumption decision. Finally, the borrowing rate is endogenous, even though it is known at the beginning of the period, because the household determines when to pay back the balance, and therefore the effective interest rate, contemporaneously with the consumption choice.

The instrumental variables used are the exogenous variables and second and previous lags of state-level disposable income, city-level consumption, city-level unemployment, mortgage and inflation rates, the marginal tax rate for the HH income bracket and some variables aimed at capturing the household asset position at the beginning of the period, such as changes in revolving debt, credit line, amount charged off and a measure of the tightness of credit constraints. Given that household consumption is affected by measurement error I choose not use past lags of the consumption growth rate as an instrument. The basic assumption required for identification is that the instruments are uncorrelated with the preference shocks, measurement error and expectation errors contained in $\varepsilon_{i,t}$. I test this assumption by performing various test of overidentifying restrictions, such as the Hansen J test for the entire instrument set and the difference-in-Sargan statistic on the subset of household-specific instruments. Both tests strongly support the validity of the instruments set by failing to reject the null of no correlation between the instruments and the error term.

Finally, the standard errors are robust to arbitrary correlation and heteroskedasticity within households. The errors are however assumed to be uncorrelated across households, once a common aggregate component is accounted for by the time dummies.

4.2 Empirical Evidence

The main results of the estimation of the loglinearized Euler equation (4.7) are reported in Table VI. I regress household consumption growth on past HH consumption growth, as a measure of internal habit stock, city-level per capita consumption and its first lag, as a measure of the external habit level, the household specific interest rate, HH demographic characteristics and seasonal dummies.

Household and aggregate consumption, as well as the interest rates are treated as endogenous, either because they are uncertain at the time the HH makes the consumption decision or because they are affected by measurement error. The instruments used in the estimation are the second

[20] In the next Section, I provide tests for the validity of these assumptions.

lag of the marginal tax rate, the local unemployment rate, the inflation rate, aggregate disposable income growth rate, mortgage rate, and some individual variables such as lags of the growth rate of debt, amount charged off, automatic credit line changes, and a credit constrained indicator. Aggregate and city-level indicators of economic activity are used to predict the risk-free rate and city-level consumption growth, as they have been shown to be good predictors of these variables in the macroeconomic literature. The individual variables are aimed at capturing the household lagged asset position and resources on hand and are used as predictors of the household-specific lagged consumption and borrowing rate, as they are correlated with the resources and constraints faced by the household. The standard errors are corrected for the non-indipendence of the observations within the same household. In addition, controls for the evolution of city-level prices and seasonal dummies are included in all the regressions.

Column I of Table VI estimates the basic model, while columns II to IV progressively add to the specification family composition, home ownership and occupation. The coefficients capturing the strength of the habit persistence and the sensitivity of household consumption to the interest rates prove extremely stable and maintain their statistical significance across specifications.

All versions of the estimates are consistent with the presence of habit formation in household consumption decisions, both as an external habit, captured by the consumption of the reference group, and as an internal one, represented by household past consumption.

The specification reported in column IV of Table VI shows that after controlling for own consumption and demographic and socioeconomic characteristics, the **effect of the external habit** is significant at the 5% level. This coefficient captures complementarities in consumption and represents the fraction of city-level consumption that enters the utility function as the reference level to which the household compares itself to. To meaningfully interpret the quantitative effect of this variable, some scaling procedure was necessary, as individual consumption is sizably more volatile than aggregate one. Among the available alternatives, I've chosen to scale city-level aggregate consumption by the ratio of the standard deviations of individual and city-level consumption.[21] The strength of the external habit is 0.290 in the specification reported in Column IV and ranges between 0.258 and 0.295 in the other columns of the table. A coefficient of zero would imply that the HH is not influenced by the consumption of its neighbors and the model collapses to the standard

[21] A variation of this procedure scales $\Delta C_{i,t}$ by the ratio of the standard deviations of each HH consumption growth rate and its reference group's consumption growth rate. The external habit coefficient is slightly lower, but continues to be statistically significant at the 5% level. Alternatively, I standardize the city-level consumption growth rate and obtain a habit coefficient with the same statistical significance and an even higher economic significance than those obtained with the two procedures outlined above.

one used in the literature; while, a coefficient of one would mean that the HH only cares about the way its consumption compares to the neighbors' and not about the absolute level.

The **strength of internal habit** is also very high: the coefficient on past household consumption growth is 0.503 and it is significant at the 1% level. These findings provide empirical support to the theories that explain macroeconomic facts and the equity premium puzzle by introducing habit persistence in the utility function. The latter coefficient is of the order of magnitude required in Constantinides (1990) to explain the equity premium puzzle.[22] In section 8, I discuss the relationship between these findings and the CRRA functional form commonly used in the macroeconomic and asset pricing literature and I compare the coefficients to the parameters in Abel (1990 and 1999), Constantinides (1990) and Campbell and Cochrane (1999).

This finding contradicts those in **Dynan (2000)**, who investigates habit formation in annual food consumption using the PSID and finds no evidence of habit persistence. The **differences between this study and hers** are many. The first is the measure of consumption used: she analyzes food, I analyze credit card expenditures. As Dynan admits, food expenditures are an inadequate measure of consumption for many reasons. Section 3.2 in this paper provides an illustration of them. Unfortunately, it is not possible to test directly whether this is the reason of the differences, as a measure of credit card expenditures on food is not available. An indirect indication that this could be part of the story comes from Lupton (2003) who investigates habit formation in the PSID and finds negative evidence when looking at food consumption, but positive one when looking at portfolio decisions. Another difference is related to the fact that Dynan doesn't include an interest rate in all except ine regression and when she does she is forced to use the risk free rate, as it is the only measure available. An extra difference is given the PSID data have an annual frequancy, while mine are quarterly. The studies also differ in the choice of the instrument set: Dynan uses the second lag of income, hours worked and job loss as instruments; I use the second lag of local unemployment rate, the inflation rate, aggregate income growth rate, mortgage rate, and some individual financial variables. The financial variables prove especially valuable in explaining the endogenous variables in the first stage regressions. In order to better identify the reasons of the different findings I aggregate my data at the annual level and examine the sensitivity to the different instrument set and time interval. Table VII contains the results of the analysis: it shows that the

[22] If we sum the coefficients of the external and internal habit we find that the overall importance of habit is 0.793. Despite being somewhat arbitrary, summing the coeffients on the external and internal habit variables is a simple and direct way to obtain a general measure of the importance of habit persistence. The measure of partial correlation between the external and internal habit variables yeilds -0.0078 (p-value=0.330), refuting the possibility of double counting and inflation in the measure of the strenght of habit.

annual frequency doesn't seem to be the cause of the different results, as in columns I and II I find evidence of habit persistence as well, although the model is rejected in Column II. The instruments are the annual counterparts of those used in the baseline analysis in Table VI. As the attached panels show, both the partial R^2s and the F tests indicate that the instument set is better able to describe the variables than Dynan's: the R^2s are way higher and the F statistic is very high, and safe from the Stock and Yogo (2002)'s critique. In Column III I exclude from the instrument set the household-specific financial variables, the lags of the debt growth and the credit constrained indicator. The result is a huge drop in the internal habit coefficient, from 0.60 to 0.10, although still statistically significant. On the contrary, the exclusion of the household-specific interest rate doesn't seem to be the cause of the discrepancies in the results, as the internal habit coefficient remains approximately constant. The difference in findings from the better quality of the instrument set, thanks especially to the household-specific financial information, and only marginally from the consumption measure used.

Another interesting result and contribution of the paper is the finding that households respond to the price of consumption. The availability of household-specific borrowing rates and financial information allows to study the **sensitivity of consumption to individual-specific interest rates**: the short run elasticity of consumption to the borrowing rate, R^C, is estimated to be -1.876 and it is statistically significant at the 1% level. Previous studies had difficulties in getting precise estimates of this parameter, as they were forced to use the after tax risk-free rate, which by nature displays very limited cross sectional variability. This magnitude is consistent with the results of Gross and Souleles (2002), who estimate the elasticity of debt, and therefore consumption, to the borrowing rate to be -1.3.

The other variables in the regression are consistent with the previous findings. The coefficient on the risk-free rate represents the Elasticity of Intertemporal Substitution (EIS), or the willingness of the household to move consumption across time periods in response to changes in the investment opportunity set. The estimates indicate that an increase of 1 percentage point in the risk free rate leads on average to an increase of 0.765% in the consumption growth rate. As in most of the literature, the effect is not precisely measured and cannot be statistically distinguished from zero, due to the small amount of cross sectional variation exhibited by this variable. Nevertheless, the value of the coefficient is very similar to those obtained in previous studies of individual consumption choices.[23] For some very common specifications of the utility function the inverse of the EIS

[23] Vissing-Jorgensen (2002) using the CEX obtains a value of 0.8-1 for bondholders and lower values of 0.3-0.4 for

constitutes the coefficient of relative risk aversion, measuring the curvature of the value function, or the willingness to substitute consumption across different states of nature. When this is the case the results imply a coefficient of risk aversion of 1.31, judged reasonable based on the range of values found by other microdata studies and surveys.

Consistent with the predictions of economic theory, the coefficients on age and age squared indicate that consumption exhibits an hump-shaped path over the life cycle. This result is similar to the findings of Carroll and Summers (1991) and Attanasio and Weber (1995).

Occupation dummies are included in the regression as well, although they are imprecisely measured. Among them, particularly interesting is the positive, although not significant, effect on consumption growth of being self-employed. Since this category of people could be using the credit card for business related expenses, as a robustness check I re-estimate all the regression discarding households headed by a self-employed individual and find that the results don't change.[24] Table VI also shows that the effect of home ownership and income bracket on consumption growth is negative, although small and indistinguishable from zero. These findings could be interpreted as mild evidence of precautionary savings or liquidity constraints: families that own a house or are in a higher income bracket have less need to save for a rainy day and face less limitations to the amount of funds they can borrow. Consequently, they are better able to smooth consumption and, on average, exhibit a lower consumption growth rate.

The **Hansen J statistics** of overidentifying restrictions confirms the validity of the instruments in all the specifications. Further, to make sure that the household specific instruments are orthogonal to the error term, an additional difference-in-Sargan statistic is computed for this group of instruments and the hypothesis of orthogonality is again not rejected. Nonetheless, the concern is that a too big instrument set could decrease the power of the overidentifying restrictions tests and also bias the coefficients toward the inconsistent ordinary least square estimates. To further investigate whether the findings are influenced in an undue way by the choice of the instrument set, I re-estimate the equation by progressively reducing the number of variables used as instruments, till the point in which the system is just identified. The results of this test are reported in Table VIII. Column I reproduces the first column of Table VI for comparison. In Column II, I eliminate the marginal tax rate from the instrument set illustrated above. In Column III I further eliminate the unemployment rate; in Column IV the inflation rate; in Column V all but one lag of aggregate

individuals that invest in the stock market; Attanasio and Weber (1993 and 1995), using the same data set, obtain values between 0.33 and 0.77; while Zeldes (1989), using the PSID, obtains coefficients between 0.35 and 1.44.

[24]The results are available upon request.

income; in column VI the amount charged off. Finally, in Column VII, I eliminate the automatic credit line increases. The resulting instrument set exactly identifies the system and is composed by the first available lag of mortgage rate, aggregate income growth, household debt growth rate and credit constrained indicator.[25] The results of this test are very encouraging, as the coefficients on ΔC_t, Δc_{t-1} and R^C remain approximately the same or increase slightly, and continue to be statistically significant. The coefficient most sensitive to the reduction of the instrument set is that on the risk free interest rate, which happens sometimes to be negative, even though very close to zero. This coefficient is however very imprecisely measured. Also, the $R^{2'}$**s from the first stage regressions** for Table VI are reported at the bottom. The value of the adjusted R^2 are very good and the F statistic shows that the coefficients on the instruments are statistically different from zero and way outside the critical values indicated by Stock and Yogo (2002) in relation to concerns of weak instruments. An exception is constituted by the external habit variable in some of the regressions: the value of the F tests are way outside of the critical interval required by traditional theory, but they are sometimes too low to dismiss the possibility that the variable is weakly instrumented.

5 Robustness Checks

In this section, I investigate the possibility that the findings are affected by an omitted variable problem which causes city-level and past household consumption growth rates to enter the regression with an economically and statistically significant coefficient even if per se they are irrelavant for the optimization problem of the household. To address this issue, I include in the regression the contemporaneous and future income growth rate, to control for the possibility of excess sensitivity of consumption to income and mis-specifications, city-level unemployment rate and house values, to control for city-level economic conditions, and some individual specific financial variables. In addition, I examine the effect of changes in the specification and the instrument set.

An alternative explanation of the results illustrated in the previous section could be that city-level aggregate consumption influences household behavior not because the enjoyment people receive from consumption depends on what their neighbors consume, but because aggregate sales capture some dimension of economic activity that helps households predict their income. In other words, there is a concern that the results constitute another facet of the "excess sensitivity" of consump-

[25] Among these variables the instrument that is most relevant for the prediction of city-level aggregate consumption is the mortgage rate, followed by local unemployment and inflation rate. The coefficient on income growth rate is statistically significant, but extremely small.

tion to income documented by Flavin (1981), Campbell and Mankiw (1989) and a vast literature thereafter. To exclude this possibility, I add the **growth rate of income** to the Euler equation and check whether it is economically and statistically significant and whether it drives out aggregate consumption. This test constitutes a horse race between city-level consumption and income, since the same instrument set is used for both variables. The ideal way to perform the test would be to use household specific income evolution, but unfortunately my data doesn't contain this information. Alternatively, I use California quarterly aggregate income and later in this section I provide evidence that this choice doesn't influence the conclusions. The results of the test are reported in column I of Table IX and show that the coefficient on the external habit variable is unchanged and significant at the 5% level, while income is not statistically significant (p-value=0.319). The coefficients on the other explanatory variables maintain their economic and statistical significance as well.[26] In column II I include a **lead of growth rate of aggregate income** to control for the possibility that households look at city-level consumption to infer information on future levels of their income. Again, the coefficient on ΔC_t is stable and statistically highly significant, while that on the income lead is not (p-value=0.434).

As I have pointed out above, ideally one would like to have household-specific income information, rather than aggregate one. Despite it seems reasonable that aggregate income is a better proxy of individual income than aggregate sales, it could be the case that although sales contain less information on future individual prospects, this information is orthogonal to that contained in aggregate income and therefore relevant in forming expectations. An encouraging piece of evidence related to this issue are the findings of the literature that analyzes microdata to test the Permanent Income Hypothesis: once demographic variables and labor choice are accounted for, individual income doesn't enter the Euler equation significatively.[27]

While indicative, these arguments alone are unconvincing. Other papers have found evidence of excess sensitivity using microdata and the debate is still open.[28] For this reason, I directly investigate whether city-level aggregate consumption provides any information about household income once aggregate income is available. To perform this test I use household-level data from

[26]Examination of the first stage regressions excludes the possibility that this result is due to income being poorly instrumented, as the adjusted R^2 for the income regression is 0.698, higher than that for city-level aggregate consumption, past household consumption and borrowing rate.

[27]Hayashi (1985) finds that income explains only a very small fraction of consumption changes. These findings are confirmed by Attanasio and Weber (1995) who show that once demographic and labor supply variables are controlled for there is no excess sensitivity of consumption to income. See Browning and Lusardi (1996) for a survey of this literature.

[28]For example, Lusardi (1996) investigates excess sensitivity using PSID data and finds a coefficient of 0.4 on expected income growth.

the PSID, which contains information on the evolution of individual income, and match them to city-level aggregate sales data. The regression I estimate is the following:

$$\begin{aligned} \Delta lnincome_{individual,it} &= \alpha + \beta_1 \Delta lnINCOME_{state-level,it} \\ &+ \beta_2 \Delta lnC_{city-level,it} + \beta_3 age_i + \beta_4 age_i^2 + \varepsilon_{it} \end{aligned} \tag{5.1}$$

Dummy variables capturing marital status, occupational choice, seasonal fluctuations and whether the household owns the house in which it lives are included in the regressions as well. The null hypothesis is that once we control for aggregate income, the coefficient on aggregate consumption is small and statistically insignificant. The results reported in Table X confirm this hypothesis and suggest that aggregate consumption doesn't proxy for individual income. The coefficient on city-level sales is neither economically nor statistical significant. On the contrary, the coefficient on aggregate income is big and statistically highly significant. In the second part of Table X, I repeat the same regression using the growth rate of future HH income as the dependent variable. The results are presented in columns IV to VI and confirm the previous findings that city-level consumption doesn't proxy for individual income.

These findings provides support to the inclusion of aggregate income as a proxy for individual one in the estimations performed in column I of Table IX. It indicates that the reason why income is not significant in the regression cannot be ascribed to the fact that I use aggregate rather than individual income, since Table X shows that aggregate income does predict individual one quite well. Some caution should be exercised in interpreting these results as the time period analyzed is not very long, spanning from 1997 to 2001. Unfortunately, a longer time series of city-level consumption data is not available. Nevertheless, the various pieces of evidence presented above, taken together, seem to indicate that city-level aggregate consumption doesn't proxy for individual income.[29]

In column III of Table IX, I add the **variation in city-level unemployment rate** to the regression to further control for local economic activity: this variable is neither economically nor statistically significant, while the coefficients on the habit persistence variables don't change and continue to be significant. Like in the case of the income growth rate, the result is not driven by the fact that variation in local unemployment rate is not adequately captured by the instruments, as the adjusted R^2 of the first stage regression is 0.84.

[29] Similar regressions using the growth rate of individual wealth as the dependent variable provide evidence that local aggregate consumption doesn't proxy for this variable either.

These findings are also confirmed when adding to the regression variables that capture the **zip code-level housing market conditions**, as the median house value and rent (column IV of Table IX). These variables appear to be economically and statistically insignificant and don't affect the magnitue of the coefficients. A drawback of this test is that the housing market variables are fixed at their 1999 level over the period. A more powerful test would be accomplished by controlling for the evolution over time of these variables. This constitutes the object of future investigations, as I am collecting such information.

Overall, the above evidence suggests that aggregate consumption doesn't proxy for income or local economic activity, as adding these variables to the regressions shows that they are not significant and doesn't change the economic and statistical significance of the habit persistence coefficients.

In column V of Table IX, I investigate the effect of changing the specification and **adding an extra lag of household own consumption** to the regression. Theoretical papers usually model internal habit formation by including a big number of lags of consumption in the utility function. Empirical investigations, however, are usually limited to one lag as adding extra ones would require very long lags of the instruments to insure exogeneity.[30] Nevertheless, it is interesting to see the effect of this simplification on the estimates, since HH consumption exhibits autocorrelation over time. The coefficient on the second lag of HH consumption growth is, as expected, statistically significant, although not very big. Most important, the coefficient on the first lag of consumption growth decreases only slightly (from 0.503 to 0.453), while the one on city-level consumption actually increases, and the statistical significance is the same.

In column VI of Table IX, I test for the presence of **aggregate shocks** that are not captured by the interest rate and the measures of economic activity included in previous robustness checks and that might cause the expectation error in (4.7) to be correlated across HHs. Chamberlain (1984) shows that this kind of shocks lead to inconsistent estimates when the time dimension of the data is not very long. However, the problem doesn't arise if the shock can be decomposed into an economy-wide shock and an idiosyncratic one.[31] Following Runkle (1991), I also directly test for the presence of these aggregate shocks by including year dummies in the regression and checking whether they are valid instruments by looking at the difference in the J statistics from the estimation with and without these dummies. The difference is distributed as a χ^2_3 and it assumes a

[30] See Sundaresan (1989) and Constantinides (1990) as examples of the theoretical literature and Ferson and Constantinides (1991) for an example of the empirical one.

[31] This is the assumption made by Zeldes (1989).

value of 7.817, from which I conclude we cannot reject the null hypothesis of no aggregate shocks.

Results that are available upon request also indicate that the findings are robust to changes in the instrument set beyond those reported in Table VIII, tests of household-specific effect in consumption growth, and to the inclusion in the regression of various measures of household financial conditions, as the growth rate of debt, and balance transfers indicators.[32] Column VII of Table IX contains, as an illustrative example, the estimates obtained by **adding the household-specific debt growth rate** to the regression. The coefficients indicate that the ahbit persistenc variables are not capturing the effect on the consumption growth rate of the evolution of the financial variables.

6 Alternative Explanations of the Results

Alternative explanations of the findings include the presence of liquidity constraints, precautionary saving motives and adjustment costs in consumption. These phenomena, like internal habit formation, cause consumption to adjust slowly to changes in income, and therefore induce positive correlation between current and lagged consumption growth rates.

Liquidity constraints are represented by borrowing interest rates that are higher than the lending ones and quantity constraints on the amount of funds that can be borrowed. Both of these aspects can be controlled for in the current setting. The consequences of binding liquidity constraints are that the HH cannot set the consumption at the optimal level and therefore we will observe consumption to be too low today relative to tomorrow, and the multiplier on the borrowing constraint in the Euler equation to be positive. An implication of this fact is that lagged income growth rate appears to be negatively related to current consumption growth.[33] Moreover, liquidity constraints are more likely to be binding for HHs that experience a low level of cash on hand. Therefore, I test the habit formation Euler equation against a liquidity constraints model by including the lagged growth rate of income in the estimation equation, by re-estimating the Euler equation on two group of unconstrained and credit constrained HHs in the spirit of Zeldes (1989), and by adding a credit constrained indicator directly in the regression.

Column I of Table XI shows that including lagged aggregate income growth rate in the regression

[32] Notice that the measure of consumption is unaffected by the presence of balance transfers on the card, as these sums are separately identified and excluded. However, it could be that HHs that make balance transfers behave differently that others in ways that are not controlled by the variables in the regression. This robustness test controls for this possibility.

[33] See Deaton (1992) and Browning and Lusardi (1996) for a review of liquidity constraints models.

doesn't affect the coefficient of the internal habit parameter. In particular, such coefficient equals 0.502, similar to 0.503 in the baseline regression, and it is highly statistically significant (one percent). On the contrary, the coefficient on the income growth rate doesn't have the expected negative sign and cannot be statistically distinguished from zero.

Another way to test for the presence of liquidity constraints consists in splitting the sample between HHs that have low cash on hand, and are likely to be liquidity constrained, and HHs with high cash on hand. The household-specific financial information available in the data set allows me to build a measure of the tighteness of liquidity constraints based on the ratio of revolving debt to the credit limit. Figure II plots the frequency distribution of this indicator. From the graph in panel A we see that the majority of the observations display quite low credit usage, most of the times due to the very high credit limit they enjoy. Nevertheless, some cases of very high usage and therefore binding credit constraints are present in the data. Column II and III of Table XI contain the results of estimating the baseline regression on subsamples constituted by the lowest and highest quartile of the credit constrained indicator distribution. The lowest quartile subsample contains the "unconstrained" HHs for whom the credit constraint measure is zero; while the highest quartile subsample contains the "liquidity constrained" households, for which the measure is above 0.76.[34] If liquidity constraints were the cause of the correlation between current and lagged household consumption growth rate, we would expect to see that the past consumption growth rate matters only in the highly constrained subsample. The estimates indicate that this is not the case: the magnitude of the internal habit coefficient is similar and highly statistically significant in both subsamples: 0.565 in the "unconstrained" sample, and 0.557 in the "liquidity constrained" one. The inclusion of the lagged growth rate of aggregate income in each of the subsample regressions confirms that income is not economically nor statistically significant (columns IV and V of Table XI, respectively). Nevertheless, it slightly decreases the coefficient on the internal habit parameter in the unconstrained case and increases it in the constrained one, indicating some sign of liquidity constraints, although not strong enough to dismiss the effect of habit persistence.

A further robustness test is presented in Column VI and consists in adding the credit constrained indicator directly in the regression, as a proxy of the Lagrange multiplier on the borrowing constraint. This indicator is increasing in the level of credit line usage and therefore should display

[34] I've decided to split the sample in this way, to make sure that the subsample truly contains HHs that are liquidity constrained. However, the results are robust to the choice of other cutoff points. Also notice that the fact that not all the HHs in the highest quartile subsample are up to their credit limits (indicator bigger or equal than one) doesn't preclude the possibility that they act as liquidity constrained HHs, as it is possible that they foresee being liquidity constrained in the near future.

an increasing relationship with the unobserved shadow price or resources. Since it is not possible to directly test this assertion these results should be considered only indicative. The coefficient on this variable has the expected sign, although it is not significantly different than zero.

All the tests illustrated above exclude that the presence of liquidity constraints is an explanation for the results. The **precautionary motives story** has implications that are very similar to the liquidity constraints model, with regard to the correlation of lagged income growth rate and current consumption growth: HHs that have low cash on hand and uncertain future prospects save more and therefore display a higher growth rate of consumption from one period to the next, much like liquidity constrained HHs.[35] Therefore, the tests of the significance of lagged income growth and splitting of the sample contained in column I to V of Table XI speaks to the precautionary saving alternative as well. As an additional robustness check, I perform a more direct test of theis hypothesis by adding a measure of future household-specific uncertainty to the estimating equation. Following Dynan (1993) and Carroll (2001), I measure uncertainty with the square of household consumption growth rate.[36] To address Carroll's concerns about the validity of group-specific instruments, I include in the instrument set individual variables proxying for household cash on hand resources, such as lags of the growth rate of debt, amount charged off, automatic credit line changes, and a credit constrained indicator among the instruments.[37] Column VII of Table XI shows that the coefficient on the square of consumption innovations is neither economically nor statistically significant and its presence in the regression doesn't affect the estimates of the habit persistence parameters or the interest rates sensitivities.

Another alternative explanation of the internal habit result is that households face adjustment costs in consumption that lead them to react slowly to changes in permanent income. Unfortunately, it is not possible to distinguish this hypothesis from that of habit formation.

6.1 The Reflection Problem: Why Do People within a Group Behave Similarly?

A central issue in the social interaction literature is the possibility of disentangling the reasons why people belonging to the same group behave similarly. Manski (1993) identifies three distinct reasons why a similar behavior occurs: the people in the group face the same shocks; they have similar

[35] See Carroll (1997) and Deaton (1992) for a summary of the implications of precautionary saving models and their relation to liquidity constraints models.

[36] Carroll (2001) illustrates why this is a good measure of uncertainty, and Dynan (1993) provides a test of the theory using group specific instruments.

[37] The same author shows with simulations that, even using individual data, the coefficient on this variable is biased, but, despite this, it expected to be and it is highly significant in all the specifications.

characteristics that leads them to behave similarly, and social interactions. The policy implications of these different phenomena can be very different and it is therefore important to identify the mechanisms at work in the various situations. Unfortunately, this task is very hard, given the data usually available to the econometrician. The analysis in this paper faces similar issues. The robustness section has provided evidence that the household in my reference groups behave in a similar way beyond the effect of common shocks, as income growth doesn't appear to influence the consumption growth rates described in the Euler equation. Nevertheless, it is interesting to distinguish between the case in which people living in the same city behave the same because they have similar characteristics or because of social interactions. I plan to address this issue in a setting in which I analyze the variation in the strength of external habit persistence across cities of different size, once heterogeneity is kept constant. The rationale behind this empirical strategy is that in smaller cities, people have better chances of seeing each others and interacting. However, smaller cities are also usually populated by people more similar to each other. Hence, the necessity of the difference in difference framework and of an analysis of the behavior across the size dimension, by keeping the heterogeneity dimension fixed. The measures of city size I use are area and population density, while the measures of heterogeneity are given by income inequality (measured by the interquartile range and the Gini coefficient) and an index of language diversity and ethnic composition.

Another estimation strategy I am exploring is the use of lottery winnings at the zip code level, as an exogenous shock to the external reference point of the households. Preliminary results indicate that lottery winnings *at the city level* influence household consumption, but the results are not significant. A more close examination at the zip code level will better reveal the relevance of the pehenomenon.

7 Aggregation of Individual Choices and Comparison with Representative Agent Studies

In this section I investigate the aggregate implications of the micro findings described in the paper and provide a comparison with a representative agent framework.

The majority of the empirical literature on habit persistence is based on a representative agent and aggregate consumption data. The findings depend on the interval at which consumption is observed and the instrument set. Using monthly data, Eichenbaum, Hansen and Singleton (1988) find evidence of habit persistence in leisure choices, and time-separability in consumption. Heaton (1993 and 1995) finds strong evidence of durability at monthly horizons, and weak evidence of habit persistence at quarterly frequencies; while, more recently, Chen and Ludvigson (2003) find evidence of habit formation in quarterly data. On the contrary, Constantinides and Ferson (1991) find support for habit formation in monthly, quarterly, and annual data.

Letting aside the statistical problems of the aggregate consumption series (Wilcox (1992), Dynan (2000)), the representative agent framework provides an inadequate description of real households consumption choices, as it doesn't capture heterogenous information sets, finite lives, liquidity constraints, and the fact that not all the households hold the same assets. These factors cause the marginal rate of substitution (MRS) to differ across agents and, unless very restrictive conditions are met, lead to a difference between the weighted average of individuals' MRS and the MRS of a fictitious representative agent that consumes per capita consumption.[38] Estimates of the preference parameters obtained within the representative agent framework can therefore be severely biased.

The finance literature has provided evidence and investigated the implications of heterogeneity and incomplete markets for asset prices and the equity premium. Heaton and Lucas (1996) document the importance of limited participation, borrowing constraints and transactions costs; Mankiw (1986) and Constantinides and Duffie (1996) show that in the presence of idiosyncratic income shocks the Euler equations don't depend only on the growth rate of consumption, but also on its cross sectional variability. On the contrary, Krusell and Smith (1998), show with calibrations that in a setting with transitory idiosyncratic income shocks the mean of the wealth distribution is enough to describe the macroeconomic aggregates. While there are contrasting conclusions on the biases induced by the representative agent framework, the empirical work on aggregation issues indicates the crucial importance of accounting for heterogeneity. Brav et al. (2002) show that in order to obtain plausible asset pricing implications the stochastic discount factor needs to be constructed as a weighted average of individual MRS. Attanasio and Weber (1993) illustrate how correct aggregation matters for estimating the EIS and the excess sensitivity of consumption to income. In a different context, Abel and Eberly (2002) document the importance of nonlinearities and cross sectional variability in firm-level qs for predicting aggregate investment.

[38] Grossman and Shiller (1982) show the conditions under which individual choices can be aggregated to a representative agent with the same type of preferences: quadratic utility, infinitely lived consumers (or dynasties) and homogeneous individual information sets containing all the macro variables. If these conditions are not met, only in the case of complete markets heterogenous consumers are able to pool risks and equate their marginal rates of substitution in every state, despite having different marginal utilities of consumption (Constantinides (1982)). These assumptions are hard to defend.

This section contributes to this literature by investigating the aggregate implications of household behavior in a model with habit persistence and differential borrowing and lending rates. The household-level data allow me to estimate the Euler equations using a weighted average of the individual MRS that takes heterogeneity and non linearities into account. From the same data, I can also construct the per capita MRS used in representative agent studies and investigate any differences and their causes. I find that, once I aggregate individual consumption in the proper way, habit persistence carries over in the aggregate and the magnitude of the phenomenon is the same or higher than that found in the micro data. On the contrary, I find that an econometrician that used per capita consumption, constructed from the same data, would not find any evidence of habit persistence.

In particular, the regression that uses the correct aggregation procedure is given by

$$\Delta\frac{1}{N}\sum_{i=1}^{N}(\ln c_{i,t}) = \alpha + \beta\Delta\frac{1}{N}\sum_{i=1}^{N}(\ln c_{i,t-1}) + R_t^f + \varepsilon_t \tag{7.1}$$

while the regression that an econometrician with only per capita consumption would estimate is given by

$$\Delta\ln\left(\frac{1}{N}\sum_{i=1}^{N}c_{i,t}\right) = \alpha + \beta\Delta\ln\left(\frac{1}{N}\sum_{i=1}^{N}c_{i,t-1}\right) + R_t^f + \varepsilon_t \tag{7.2}$$

where $c_{i,t}$ is household consumption, and R_t^f is the risk-free rate.

The above equations don't contain any demographic characteristic or household-specific interest rate, as they are normally not observable in aggregate data. Moreover, given that by definition the representative agent framework is populated only by one consumer, the difference between internal and external reference point dissipates and only the dependence of current consumption growth on last period growth can be tested.

Equation (7.1) displays the proper aggregation procedure, as it contains the growth rate of the mean of the logarithm of individual consumption; while equation (7.2) contains the logarithm of per capita consumption, as only this information is available in aggregate data. Attanasio and Weber (1993) show that, for any distribution of consumption growth, the difference between the two measures represents the change in the Theil's measure of entropy and can be approximated by the first four central moments of the cross sectional distribution:

$$\Delta Theil = \Delta E_i \ln c_{i,t} - \Delta ln(E_i c_{i,t}) = \Delta\ln\left(1 + \frac{1}{2}\mu_{2,t} + \frac{1}{6}\mu_{3,t} + \frac{1}{24}\mu_{4,t}\right) + \widetilde{R} \tag{7.3}$$

where $\mu_{k,t}$ represents the kth central moment and $\widetilde{R}$ is an approximation residual. The formula shows that aggregate consumption studies miss the information about the cross sectional variability of the distribution of consumption growth rates, itsa symmetry, and the extent to which it is flat or peaked relative to a normal distribution. If the cross sectional moments of consumption growth rates change over time, the two measures will move in an asynchronous way and the per capita regression will be biased. The intuition for this result is that if we keep aggregate consumption constant and we just vary its distribution across individuals, the incorrect measure, $\Delta ln(E_i c_{i,t})$, doesn't detect any difference, as only the average consumption enjoyed by the representative agent matters. On the contrary, the correct measure, $\Delta E_i \ln c_{i,t}$, reflects these variations, as they lead to changes in individual MRS and thus in their weighted average.

Table XII contains the estimates from (7.1) and (7.2). These regressions try to be as similar as possible to those I've performed on the micro data. As before, household consumption, as well as the risk free rate are treated as endogenous, either because they are uncertain at the time the HH makes the consumption decision or because they are affected by measurement error. The instrument set is similar to the previous one, despite being somewhat reduced, due to the smaller number of observations available. It contains the second lag of aggregated city-level sales, income growth rate, average mortgage rate, and unemployment rate. A final remark relates to the relatively short time dimension of the data set, which should invite some caution in the interpretation of the results.

Column I shows that when estimating the per capita regression (7.2) I find no evidence of habit persistence in consumption choices: the coefficient on past consumption growth rate is actually negative, albeit not statistically different from zero. Similarly, the elasticity of intertemporal substitution, represented by the coefficient on the risk-free rate, is sizeable, but again not statistically significant. Columns II and III illustrate the result of testing for excess sensitivity by adding the growth rate of aggregate income to the regression. In both cases the magnitude and statistical significance of the habit persistence coefficient are unchanged, and confirm the absence of habit persistence. Interestingly, the estimates display signs of excess sensitivity of consumption to income, when lagged income growth rate is added to the regression. This finding is often interpreted as evidence that some households are liquidity constrained or rule-of-thumb consumers, and that per capita regressions are not able to capture the effect of this phenomenon on the cross sectional distribution of consumption growth and therefore the Euler equations.

A different picture arises when the properly aggregated quantities are used. Column IV reports the results of the estimation of (7.1). The habit persistence coefficient is now equal to 0.515 and it is statistically significant at the 5% level. This number is quite sizeable, and provides support to the hypothesis that, once the correct measure of consumption is used, there is evidence of habit formation also at the aggregate level. On the contrary, the estimates of the EIS are still indistinguishable from zero. Tests of excess sensitivity, reported in Columns V and VI, display no sign of it, using either the contemporaneous or lagged income growth rate. The estimates of the habit persistence coefficient mantain statistical significance and actually display an increase in magnitude.

The Hansen J statistic indicates that in all cases considered the overidentifying restrictions don't reject the model. This statistic however is only indicative as it has low power, due to the relatively big number of instruments in comparison to the number of observations available. I have therefore performed some robustness checks by changing the instrument set and found that the habit persistence coefficient is relatively stable when the properly aggregated measure is used, while it varies from -0.8 to 0.15, and it is never significant, when per capita consumption is analyzed.

Subject to the caveat expressed above, the difference between the estimates reported in the first three columns versus those contained in Columns IV to VI indicate the importance of accounting for demographics, liquidity constraints, and heterogeneity of information sets and preference parameters in the analysis of consumption choices and the inadequacy of the representative agent framework in this respect.

Figure III, panel A, shows the evolution of $E_i \ln c_{i,t}$ and $ln(E_i c_{i,t})$ over time. As expected, $E_i \ln c_{i,t}$ is always lower than $ln(E_i c_{i,t})$, the difference being given by the variance and higher moments. Despite the short time period of analysis, we can see that as the economy slows down, both measures of aggregate consumption fall, but the difference between them widens. Fig. III, panel B, confirms that the two series not only diverge in levels, but have also different growth rates, which causes the marginal utilities generated from them to differ. To further investigate the origins of the differences, Figure IV plots each of the higher moments of the cross sectional distribution of household consumption included in (7.3). In particular, panel A shows that the cross sectional standard deviation of consumption growth displays a strong seasonal pattern. From panel B we can see that the skewness is always positive and tends to increase as time passes and the economy gets deeper into recession. Finally, panel C displays the kurtosis and shows that as time passes, the distribution evolves from fairly flat to quite peaked.

To check if there is any way to improve the estimates delivered by per capita consumption studies, I add to equation (7.2) higher moments of the cross sectional distribution. This strategy is employed by Attanasio and Weber (1993) in the estimation of the EIS, and by Abel and Eberly (2002) in a firm level analysis of the q-theory. Columns VII to IX of Table XII report the results of adding to the regression cross sectional standard deviation, the skewness and the kurtosis, respectively. The estimates indicate that adding these moments strongly improves the significance of the EIS, but has no effect on uncovering the habit persistence revealed by the micro data. These results suggest that omitted demographic variables and other measures of heterogeneity also play a very important role. Unfortunately, it is not possible to check directly the effects of omitting these variables in the aggregate regression. A possible solution, proposed by Attanasio and Weber (1993), involves re-estimating the equation on a cohort of individuals, while including averages of the demographic characteristics, as well as letting the preference parameters differ across cohorts. This approach is the subject of future research.

To conclude, despite the short time period analyzed, the results are encouraging and going in the direction expected: if constructed in the proper way, aggregate consumption contains evidence of habit persistence. A big role is also left to demographics and heterogeneity of preference, opportunity sets and shocks across individuals.

8 A Structural Interpretation of the Results

Habit persistence plays a central role in the asset pricing literature that investigates the equity premium puzzle. Historically, the average return on equity has exceeded the risk free rate by more than 600 basis points. Mehra and Prescott (1985) show that a standard representative agent with time-separable isoelastic preferences would need to be extremely risk averse to require such a high premium to hold stocks. Using their framework, Cochrane (2001) illustrates that the postwar U.S. market Sharpe ratio of 0.5 implies a volatility of the stochastic discount factor of 50%. This result can be reconciled with the 1% standard deviation of aggregate consumption growth only by assuming a risk aversion coefficient of at least 50.[39]

[39] If we take into account the low correlation between stock returns and consumption growth, the required risk aversion coefficient is as high as 250. The formula reported in Cochrane (2001) is

$$\frac{\sigma(m)}{E(m)} \geq \frac{1}{\rho_{m,R^e}} \frac{E(R^e)}{\sigma(R^e)} \tag{8.1}$$

where $\sigma(m) = \gamma\sigma(\Delta \ln c)$ is the standard deviation of the stochastic discount factor and γ is the coefficient of relative risk aversion. $E(m)$ is the mean of the stochastic discount factor, ρ_{m,R^e} is the correlation coefficient between the stochastic discount factor and the market return and $\frac{E(R^e)}{\sigma(R^e)}$ is the Sharpe ratio. The risk aversion coefficient is 50 if we assume a perfect correlation between consumption and stock returns and goes up to 250 if we use the historical correlation coefficient of 0.2.

Introducing habit persistence in the utility function reconciles a volatile discount factor with smooth consumption growth by letting marginal utility depend on how far consumption is from the habit level, rather than on its absolute value. Constantinides (1990), Sundaresan (1989), Abel (1990 and 1999), Campbell and Cochrane (1999) and a vast literature thereafter use this specification in a representative agent framework. Using simulations they show that it generates the equity premium and some of the empirical properties of aggregate consumption and asset prices with reasonable values of the parameters.

So far, I have provided evidence that habit persistence is an important determinant of actual household consumption choices in a setting that allows for individual heterogeneity and credit constraints. It is therefore interesting to compare the results from the microdata to the parameters obtained in the representative agent studies.

Equation (4.4), illustrated in Section 4.1, subsumes the isoelastic specification traditionally used in the literature and employed in the studies cited above:

$$u(c_t, h_t, H_t) = \frac{(c_t - h_t - H_t)^{1-\gamma}}{1-\gamma} \tag{8.2}$$

where h_t represents the internal habit level and equals ζc_{t-1}, while H_t represents the external habit level and equals $\alpha_0 \mathrm{C}_t + \alpha_{-1} \mathrm{C}_{t-1}$. The presence of both aggregate consumption and household specific lagged consumption is in the spirit of Abel (1990), while the difference specification of habit persistence follows Constantinides (1990) and Campbell and Cochrane (1999). The main advantage of the difference form is that it generates a coefficient of relative risk aversion that changes over time, consistent with the time-varying countercyclical risk premia observed in the data. One disadvantage is that we need to make sure that consumption is always above habit, otherwise the utility function is undefined.[40]

A difference between my specification and those of the authors cited above is the shorter number of consumption lags entering the habit equations. This feature is common in the empirical analyses of habit models, as adding additional lags would require very long lags of the instruments to insure exogeneity. Campbell and Cochrane (1999) show that the main advantage of a slow moving habit that depends on a long history of past consumption is that it generates persistency in volatility

[40] An alternative is the ratio specification employed by Abel (1990, 1999), Gali (1994) and Carroll (2001), among others. In this case, the argument of the utility function is $\left(\frac{c_t}{h_t H_t}\right)$ and utility is always well defined. However, the coefficient of relative risk aversion is constant and equal to γ and this specification cannot account for time-varying risk aversion. See Campbell, Lo, McKinlay (1997) for a further comparison between difference and ratio models.

and mean reversion in price dividend ratios. Unfortunately, my micro data set doesn't allow a test of this feature of the theoretical models.[41]

Plugging the functional form (8.2) in the Euler equation (4.2), I obtain:

$$\beta\frac{E_t\left[(c_{i,t+1}-h_{i,t+1}-H_{i,t+1})^{-\gamma}\right]}{(c_{i,t}-h_{i,t}-H_{i,t})^{-\gamma}}\widetilde{R}^f_{i,t}\widetilde{R}^C_{i,t}\exp\Delta(\theta'\Theta_{i,t+1})=1 \tag{8.3}$$

where $\widetilde{R}^f_t$ and $\widetilde{R}^C_t$ are abbreviated notations for $(1+(R^f_{i,t+1}-1)\Pr[Y^H_{i,t+1}])$ and $(1+(R^C_{i,t}-1)1[B])$, respectively, and $\exp\Delta(\theta'\Theta_{i,t+1})$ captures the evolution of the demographic characteristics. From this expression, I derive the following estimation equation:

$$\begin{aligned}\Delta\ln c_{i,t}=k_1+\zeta\Delta\ln c_{i,t-1}+\alpha_0\Delta\ln C_{i,t}+\alpha_{-1}\Delta\ln C_{i,t-1}+\alpha_{-2}\Delta\ln C_{i,t-2}+\tfrac{1}{\gamma}\ln(1+(R^f_{i,t}-1)\Pr[Y^H_{i,t}])\\+\tfrac{1}{\eta}\ln(1+(R^C_{i,t-1}-1)1[B])+demogr\&othercontrols+\varepsilon_{i,t}\end{aligned} \tag{8.4}$$

This equation allows a comparison between the parameters described in Section 5.2 and the results in the theoretical papers cited above. The values of the parameters are of the same order of magnitude as those required by these studies to explain the equity premium puzzle. From column IV of Table VI we can see that the strength of internal habit, represented by ζ, is 0.503. This value is within the range of those in Table 1 of Constantinides (1990), which contains various combinations of the parameters that fed into his model generate the equity premium and yield plausible values of the mean and variance of consumption. The strength of external habit is also sizeable: the parameter α_0 ranges between 0.258 and 0.295, depending on the specifications. This specific parameter cannot be compared directly to Campbell and Cochrane's, as they model habit as a non-linear function of past habit and consumption.[42] Nevertheless, some of the key parameters in my model, such as the curvature of the utility function and the coefficient of relative risk aversion, are identical or similar to theirs.

The parameter $\frac{1}{\gamma}$ represents the elasticity of intertemporal substitution and is estimated to be 0.765, well within the range of values deemed plausible by the micro consumption literature. In the standard time-separable isoelastic utility function used by Mehra and Prescott (1985) the EIS is given by $\frac{1}{\gamma}$, while the coefficient of relative risk aversion is given by γ. This relationship generates

[41] In column VI of Table X, I investigate the effect of adding extra lags of consumption to the specification. I find a relatively fast decay for household past consumption and lack of economic and statistical significance for the city-level consumption.

[42] Traditionally, habit models specify the habit level to be a linear function of various lags of consumption. Campbell and Cochrane (1999) introduce a non-linear specification to ensure that consumption never falls below habit and to generate a constant risk-free rate. They pick the parameter that governs the dependence of current habit from past habit to be 0.87, and the parameter that captures the dependence of habit on consumption to be 0.13.

another facet of the equity premium puzzle: if we assume the high values of γ required by the model to account for the equity premium, then the representative consumer should be so averse to substitute consumption intertemporally that we would need very high and volatile risk-free rates to account for the small fluctuations we observe in consumption. Weil (1989) labels this phenomenon *risk-free rate puzzle*. One of the advantages of the habit models is that they decuple the coefficient of relative risk aversion from the elasticity of intertemporal substitution.

By definition, the coefficient of relative risk aversion captures the individual's attitude toward wealth bets. The formula is given by:

$$RRA = \frac{-W_t V^{WW}}{V^W} \tag{8.5}$$

where W is household wealth and V^W and V^{WW} are the first and second derivatives of the value function, respectively.

In order to explain time-varying risk premia we need a risk aversion coefficient that varies over time in a countercyclical fashion. The internal habit model of Constantinides (1990) and the ratio model of Abel (1990, 1999) generate a constant coefficient of relative risk aversion equal to γ. On the contrary, Campbell and Cochrane (1999) display a time-varying risk aversion coefficient that depends on the distance between consumption and the habit level.

My model shares this feature with theirs. The curvature of the utility function is the same and is given by:

$$\xi_t = \frac{\gamma}{\frac{(c_t - h_t - H_t)}{c_t}} \tag{8.6}$$

This quantity represents risk aversion with respect to consumption bets and evolves countercyclically: it is high in recessions, when consumption has fallen toward habit, and low in booms. In order to calculate the RRA coefficient, I replace wealth with resources on hand, as the value function in my model doesn't depend directly on wealth:

$$RRA' = \frac{-Z_t V^{ZZ}}{V^Z} \tag{8.7}$$

An appealing aspect of this measure is that in a framework with credit constrained households, resources on hand might matter more than wealth in determining attitudes toward risk. Moreover, this measure can be interpreted as liquid wealth.

Plugging the functional form (8.2) into (8.7) and rearranging yields:

$$RRA' = \frac{\gamma}{\frac{c_t - h_t - H_t}{c_t}} \left(\frac{\partial c_t}{\partial Z_t} \frac{Z_t}{c_t} \right) \frac{1 + \beta \zeta^2 E_t \left(\frac{c_{t+1} - h_{t+1} - H_{t+1}}{c_t - h_t - H_t} \right)^{-\gamma - 1}}{1 - \beta \zeta E_t \left(\frac{c_{t+1} - h_{t+1} - H_{t+1}}{c_t - h_t - H_t} \right)^{-\gamma}} \tag{8.8}$$

The first term represents the curvature of the utility function and evolves countercyclically over time. The second term represents the elasticity of consumption with respect to resources on hand. Intuitively, risk aversion is higher for households whose consumption fluctuates more in responses to changes in cash on hand, as a decline in their resources has a bigger effect on consumption and therefore utility. This elasticity is expected to be high for HHs that are credit constrained, as a relaxation or tightening of their liquidity constraints generates big responses in consumption.[43] On the contrary, a change in the resources on hand should have a smaller effect on the consumption of HHs that are not constrained. These first two terms are similar to the expression obtained for the RRA coefficient by Campbell and Cochrane (1999), with the difference that they obtain the derivative of consumption with respect to wealth in place of the elasticity to resources on hand. In their specification the latter term is always greater than one; in mine the magnitude of the second term varies across HHs, with wealthy, unconstrained, HHs displaying a lower elasticity and a lower degree of risk aversion.

Finally, the third term reflects the presence of internal habit in the utility function. It shows that HHs care not only about how far consumption is from the habit level, but also about the way this distance evolves over time. Ceteris paribus, if the HH expects the future difference between consumption and habit to be lower than the current one, it becomes more risk averse.

Figure V plots the RRA coefficient for various combinations of the surplus consumption ratio, $S_t = \frac{(c_t - h_t - H_t)}{c_t}$, and the third term in (8.8), shortly referred to as *Ratio*. The rate of time preference β is assumed to be 0.97, while the other parameters are set equal to the values obtained in the estimation: γ equals 1.38, ζ equals 0.503 and the elasticity of consumption to resources on hand is conservatively set to 1.1.[44] The range of values of S_t varies between 0.04 and 0.25, so to include the values proposed by Campbell and Cochrane (1999), 0.057, and Constantinides (1990), who indicates it to be 0.20 at the steady state. The range for the *Ratio* is set between 0.8 and 1.2,

[43] A precautionary motive could attenuate this results by leading credit constrained HHs to save part of an increase in cash on hand to prepare for worse times ahead.

[44] From the formula (8.8) we can see that, ceteris paribus, as the elasticity of consumption to resources on hand decreases so does the RRA coefficient. Values of the elasticity lower than the one chosen will genrate lower values of RRA.

allowing for fluctuations up to 20% in the consumption-habit distance.[45] The RRA coefficient is a decreasing convex function of these two quantities. Keeping the *Ratio* constant, risk aversion decreases as consumption moves away from habit and S_t rises, like in the Campbell and Cochrane's formula. Similarly, keeping S_t constant, RRA decreases as the HH expects a wider distance between consumption and habit in the future in comparison to today. The highest level of risk aversion is attained by a HH that is currently very near to its habit level and expects this gap to reduce even further in the future.

Figure VI displays the relationship between the coefficient of RRA and each of these variables taken separately. It contains for each of them a graph with a low, medium and high value of the other. From these graphs, we can see that the relationship between the RRA and these variables is very stable and displays the same shape across different values. Panels a, b and c show the effect of S_t on the RRA coefficient, for a given value of the *Ratio*. In particular, Panel b shows that if the HH expects the distance between consumption and the habit level to be constant over time (*Ratio*=1), then the RRA coefficient is slightly below 65 for a value of S_t of 0.057 proposed by Campbell and Cochrane (1999) and it is around 18.5 for the value of 0.2 indicated by Constantinides (1990). These values are higher in Panel a, which displays the case of an HH expecting a 20% drop in the consumption-habit distance: the RRA coefficient is 112.87 and 32.17, respectively. It is as low as 49.82 and 14.20 in Panel c, which illustrates the case of a HH with an increasing relative standard of living of 20%. Similarly, Panels d, e and f show the effect of the *Ratio* on the RRA coefficient, for a given value of S_t. From the graphs we can see that the RRA coefficient decreases quite fast as consumption rises far away from habit.

From these graphs we can see that the coefficient of relative risk aversion is quite high for some values of the variables, especially the extreme ones. This feature is similar to Campbell and Cochrane's model. There are however values of the parameters for which the RRA is reasonably low.

It is therefore interesting to examine the sensitivity of the results to changes in the parameters governing the rate of time preference, the fraction of past consumption that enters the habit level and the elasticity of consumption to resources on hand. Results that are not reported indicate that risk aversion decreases if the HH is more impatient, has a lower ζ and its utility is less affected by past consumption, and if the elasticity of consumption to cash on hand is lower. If we choose a *Ratio* of 1.1 and a value of S_t of 0.2, the RRA coefficient can be as low as 8.83, for values of

[45] I have tried bigger ranges for both variables and found that the shape of the function doesn't change.

$\gamma = 1.31$, $\beta = 0.87$, $\zeta = 0.4$ and the elasticity of consumption to resources on hand equal to 0.8.

9 Conclusions

This paper provides evidence that habit persistence is an important determinant of household consumption choices in a setting that allows for heterogeneity and household-specific interest rates. By using actual individual consumption data, I find that the strength of the external habit, captured by the fraction of the consumption of the reference group that enters the utility function, is 0.290; while the strength of internal habit, represented by household past consumption, is 0.503. These findings provide empirical support to the theories that explain macroeconomic and asset pricing phenomena by introducing habit persistence in the utility function. The results are robust to the inclusion of income growth rate and other measures of economic activity in the regression, tests of liquidity constraints and precautionary saving motives. I also show that this result carries over in the aggregate, once the aggregation of individual consumption choices is properly performed, and heterogeneity and nonlinearity of marginal utility are taken into account. On the contrary, I find that an econometrician that used per capita consumption, constructed from the same data, and a representative agent framework, would find no evidence of habit persistence.

Another interesting result and contribution of the paper is the finding that households respond to the price of consumption. The availability of household-specific borrowing rates and financial information allows to study the **sensitivity of consumption to individual-specific interest rates**: the short run elasticity of consumption to the borrowing rate, R^C, is estimated to be -1.876 and it is statistically significant at the 1% level. Previous studies had difficulties in getting precise estimates of this parameter, as they were forced to use the after tax risk-free rate, which by nature displays very limited cross sectional variability. This magnitude is consistent with the results of Gross and Souleles (2002), who estimate the elasticity of debt, and therefore consumption, to the borrowing rate to be -1.3.

References

[1] Abel, Andrew B. (1990), "Asset Prices under Habit Formation and Catching Up with the Joneses." *American Economic Review Papers and Proceedings*, pp. 38-42.

[2] Abel, Andrew B. (1999), "Risk Premia and Term Premia in General Equilibrium." *Journal of Monetary Economics*, Vol. 43, No. 1, p. 3-33.

[3] Abel, Andrew B., and Eberly, Janice C. (2002), "Investment and q with Fixed Costs: An Empirical Analysis", Working Paper.

[4] Akerlof, George A. (1997) "Social Distance and Social Decisions" *Econometrica*, Vol. 65, No. 5, pp. 1005-1027.

[5] Arellano, M., and Bond S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies,* Vol. 58, pp. 277-297.

[6] Attanasio, Orazio, and Browning, Martin. (1993) "Consumption over the Life Cycle and over the Business Cycle." *The American Economic Review*, Vol. 85, No. 5, pp. 1118-1137.

[7] Attanasio, Orazio, and Weber, Guglielmo. (1993), "Consumption Growth, the Interest Rate and Aggregation" *The Review of Economic Studies*, Vol. 60, No. 3, pp. 631-649.

[8] Attanasio, Orazio, and Weber, Guglielmo. (1995), "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *The Journal of Political Economy*, Vol. 103, No. 6, pp. 1121-1157.

[9] Ausubel, Lawrence. (1999) "Adverse Selection in the Credit Card Market," Working Paper, Department of Economics, University of Maryland.

[10] Becker, Gary S., and Murphy, Kevin M. (1988) "A Theory of Rational Addiction" *The Journal of Political Economy*, Vol. 96, No. 4, pp. 675-700.

[11] Berkowitz, Jeremy, and Hynes, Richard. (1999), "Bankruptcy Exemptions and the Market for Mortgage Loans", *Journal of Law and Economics*, Vol. 42, No. 2, pp. 809-830.

[12] Bertrand, Marianne, Luttmer, Erzo, F. P., and Mullainathan, Sendhil, (2000) "Network Effects and Welfare Cultures", *The Quarterly Journal of Economics*, Vol. 115, no. 3, pp. 1019-1055.

[13] Boldrin, Michele, Christiano, Lawrence J., and Fisher, Jonas D.M. (2001) "Habit Persistence, Asset Returns and the Business Cycle." *American Economic Review*, Vol. 91, pp. 149-166.

[14] Brav, Alon, Constantinides, George M., and Geczy, Christopher C. (2002), "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence." *The Journal of Political Economy*, Vol. 110, No. 4, pp. 793-824.

[15] Browning, Martin, and Lusardi, Annamaria. (1996), "Household Saving: Micro Theories and Micro Facts." *Journal of Economic Literature*, Vol. 34, pp. 1797-1855.

[16] Campbell, John Y., and Cochrane, John H. (1999), "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *The Journal of Political Economy*, Vol. 107, No. 2, pp. 205-251.

[17] Campbell, John Y., Lo, Andrew W., and MacKinlay, A. Craig. (1997), *The Econometrics of Financial Markets*, Princeton University Press.

[18] Campbell, John Y., and Mankiw, Gregory N. (1989), "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence." in O.J. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual 1989*, Cambridge, Mass. MIT Press, pp. 185-216.

[19] Campbell, John Y., and Shiller, Robert J. (1988) "Stock Prices, Earnings, and Expected Dividends." *The Journal of Finance*, Vol. 43, pp. 661-676.

[20] Carroll, Christopher D. (1997), "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *The Quarterly Journal of Economics*, Vol. 112, no. 1, pp. 1-56.

[21] Carroll, Christopher D. (2001), "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second Order Approximation)," *Advances in Macroeconomics*, forthcoming.

[22] Carroll, Christopher D., Overland, Jody R., and Weil, David N. (2000) "Saving and Growth with Habit Formation," *American Economic Review*, 90(3), 341-355.

[23] Carroll, Christopher D., and Summers, Lawrence H. (1991), "Consumption Growth Parallels Income Growth: Some New Evidence." in B. D. Bernheim and J. B. Shoven (eds.) *National Savings and Economic Performance*, Chicago University Press for NBER, pp. 305-343.

[24] Case, Anne, and Katz, Lawrence F. (1991), "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." *NBER Working Paper* No. W3705.

[25] Chamberlain, Gary. (1984), "Panel Data", in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, Vol. 2, North-Holland, Amsterdam, pp.1247-1318.

[26] Chen, Xiaohong, and Ludvigson, Sydney C. (2003), "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Models", Working Paper.

[27] Chetty, Raj, and Szeidl, Adam. (2004) "Consumption Commitments and Asset Prices", Working Paper.

[28] Christiano, Lawrence J., Eichenbaum, Martin, and Marshall, David. (1991), "The Permanent Income Hypothesis Revisited" *Econometrica*, Vol. 59, No. 2, pp. 397-423.

[29] Cochrane, John H. (2001), *Asset Pricing*, Princeton University Press.

[30] Constantinides, George M. (1982), "Intertemporal Asset Pricing with Heterogeneous Consumers and Without Demand Aggregation." *The Journal of Business*, Vol. 55, No. 2, pp. 253-267.

[31] Constantinides, George M. (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle." *The Journal of Political Economy*, Vol. 98, No. 3, pp. 519-543.

[32] Constantinides, George M., and Duffie, Darrell. (1996), "Asset Pricing with Heterogeneous Consumers." *The Journal of Political Economy*, Vol. 104, No. 2, pp. 219-240.

[33] Davis, Steven J., Kubler, Felix, and Willen, Paul. (2004), "Borrowing Costs and the Demand for Equity over the Lifecycle", Working Paper.

[34] Dawsey, Amanda E., and Ausubel, Lawrence M. (2002) "Informal Bankruptcy", Working Paper.

[35] Deaton, Angus. (1992), *Understanding Consumption*. Oxford: Clarendon Press ; New York : Oxford University Press.

[36] de Botton, Alain (2004), *Status Anxiety*, Pantheon Books.

[37] Duesenberry, James S. (1949), *Income, Saving, and the Theory of Consumer Behavior*, Harvard University Press, Cambridge, Massachusetts.

[38] Dynan, Karen E. (1993), "How Prudent are Consumers?" *The Journal of Political Economy*, Vol. 101, No. 6, pp. 1104-1113.

[39] Dynan, Karen E. (2000), "Habit Formation in Consumer Preferences: Evidence from Panel Data." *The American Economic Review*, Vol. 90, No. 3, pp. 391-406.

[40] Eichenbaum, Martin S., Hansen, Lars P., and Singleton, Kenneth J. (1988), "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty." *The Quarterly Journal of Economics*, Vol. 103, pp. 51-78.

[41] Ferson, Wayne E., and Constantinides, George M. (1991), "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests." *Journal of Financial Economics*, Vol. 29, pp.199-240.

[42] Flavin, Marjorie. (1981), "The Adjustment of Consumption to Changing Expectations About Future Income." T*he Journal of Political Economy*, Vol. 89, No. 5, pp. 974-1009.

[43] Fuhrer, Jeffrey C., (2000), "Optimal Monetary Policy in a Model with Habit Formation." FRB Boston Series, paper no. 00-5.

[44] Gali, Jordi. (1994), "Keeping up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices." *Journal of Money, Credit and Banking*, Vol. 26, No. 1, pp. 1-8.

[45] Gerdes, Geoffrey R., and Walton, Jack K. II, (2002), "The Use of Checks and Other Noncash Payment Instrument in the United States." *Federal Reserve Bulletin*, pp. 360-374.

[46] Grinblatt, Mark, Keloharju, Matti, and Ikaheimo, Seppo. (2004), "Interpersonal Effects in Consumption: Evidence from the Automobile Purchases of Neighbors." *NBER Working Paper 10226.*

[47] Gropp R., Scholz, J.K., and White, M. (1997), "Personal Bankruptcy and Credit Supply and Demand." *The Quarterly Journal of Economics*, vol. 112, pp. 217-252.

[48] Gross, David B., and Souleles, Nicholas S. (2002), "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data" *The Quarterly Journal of Economics*, Vol. 117, No. 1, pp. 149-186.

[49] Grossman, Sanford J., and Shiller, Robert J. (1982) "Consumption Correlatedness and Risk Measurement in Economies with Nontraded Assets and Heterogenous Information." *The Journal of Financial Economics*, Vol. 10, No. 3, pp. 195-210.

[50] Hall Robert E., and Mishkin, Frederic S. (1982), "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households." *Econometrica*, Vol. 50, No. 2, pp. 461-482.

[51] Harvey, Campbell R. (1989), "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *The Journal of Financial Economics*, Vol. 24, pp. 289-317.

[52] Hayashi, Fumio. (1985*a*), "The Effect of Liquidity Constraints on Consumption: A Cross-Sectional Analysis." *The Quarterly Journal of Economics*, Vol. 100, No. 1, pp. 183-206.

[53] Hayashi, Fumio (1985*b*). "The Permanent Income Hypothesis and Consumption Durability: Analysis Based on Japanese Panel Data." *The Quarterly Journal of Economics*, Vol. 100, No. 4, pp. 1083-1113.

[54] Heaton, John. (1993), "The Interaction Between Time-Nonseparable Preferences and Time Aggregation." *Econometrica*, Vol. 61, No. 2, pp. 353-385.

[55] Heaton, John. (1995), "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications." *Econometrica*, Vol. 63, No. 3, pp. 681-717.

[56] Heaton, John, and Lucas, Deborah. (1996), "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing." *The Journal of Political Economy*, Vol. 104, No. 3, pp. 443-487.

[57] Hong, Harrison, Kubik, Jeffrey D., and Stein, Jeremy C. (2004), "Social Interactions and Stock-Market Participation." *The Journal of Finance*, Vol. 49, no. 1, pp. 137-163.

[58] Krusell, Per, and Smith, Anthony. (1998) "Income and Wealth Heterogeneity in the Macroeconomy" *The Journal of Political Economy*, Vol. 106, No. 5, pp. 867-896.

[59] Laibson, David, Repetto, Andrea, and Tobacman, Jeremy. (2000), "A Debt Puzzle", *Working Paper, Harvard University.*

[60] Lalive, Rafael. (2004), "Social Interactions in Unemployment", Working Paper, University of Zurich.

[61] Lim, Paul J., and Benjamin, Matthew. (2001) "Digging Your Way Out of Debt", *U.S. News and World Report.*

[62] Lupton, Joseph P. (2003), "Household Portfolio Choice and the Habit Liability: Evidence from Panel Data." *Working Paper, Federal Reserve Board.*

[63] Lusardi, Annamaria. (1996), "Permanent Income, Current Income, and Consumption: Evidence from Two Panel Data Sets." *Journal of Business Economics and Statistics*, Vol. 14, no. 1, pp. 81-90.

[64] Mankiw, Gregory N. (1986), "The Equity Premium Puzzle and the Concentration of Aggregate Shocks." *Journal of Financial Economics*, Vol. 17, pp. 211-219.

[65] Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies*, Vol. 60, No. 3, pp. 531-542.

[66] Marmot, Michael. (2004), *The Status Syndrome: How Social Standing Affects Our Health and Longevity,* Times Books, Henry Holt & Company.

[67] Mehra, Rajnish, and Prescott, Edward C. (1985), "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, Vol. 15, pp. 145-161.

[68] Muellbauer, John N.J. (1988), "Habits, Rationality and Myopia in the Life Cycle Consumption Function" *Annales d'Economie et de Statistique*, Vol. 9, pp. 47-70.

[69] Pollak, Robert A. (1970), "Habit Formation and Dynamic Demand Functions", *The Journal of Political Economy*, Vol. 78, No. 4, pp. 745-763.

[70] Runkle, David E. (1991), "Liquidity Constraints and the Permanent-Income Hypothesis." *Journal of Monetary Economics*, Vol. 27, pp. 73-98.

[71] Ryder, Harl E., and Heal, Geoffrey, M. (1973), "Optimum Growth with Intertemporally Dependent Preferences." *The Review of Economic Studies*, Vol. 40, pp.1-33.

[72] Shore, Stephen H. and White, Joshua S., (2002) "External Habit Formation and the Home Bias Puzzle" Workin Paper.

[73] Sundaresan, Suresh M. (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth." *The Review of Financial Studies*, Vol. 2, no. 1, pp. 73-89.

[74] Shea, J. (1997), "Instrument relevance in multivariate linear models: a simple measure." *Review of Economics and Statistics*, Vol. 79, pp. 348-352.

[75] Stock, James H., and Yogo, Motohiro, (2002), "Testing for Weak Instruments in Linear IV Regression." *NBER Working Paper*, No. T0284.

[76] State Board of Equalization. *Taxable Sales in California,* various issues.

[77] State Board of Equalization. (1999) *Sales and Use Taxes: Exemptions and Exclusions*, No. 61.

[78] Veblen, Thorstein. (1899), *The Theory of the Leisure Class. An Economic Study of Institutions*, Random House.

[79] Vissing-Jorgensen, Annette. (2002), "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *The Journal of Political Economy,* Vol. 110, No. 4, pp. 825-853.

[80] Weil, Philippe. (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle." *Journal of Monetary Economics*, Vol. 24, pp.401-421.

[81] Wilcox, David W. (1992), "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work." *The American Economic Review*, Vol. 82, No. 4, pp. 922-941.

[82] Zeldes, Stephen P. (1989), "Consumption and Liquidity Constraints: An Empirical Investigation" *The Journal of Political Economy*, Vol. 97, No. 2, pp. 305-346.

[83] Zinman, Jonathan. (2004) "Why Use Debit Instead of Credit? Consumer Choice in a Trillion Dollar Market." Working Paper.

Appendix I

IA Sample Selection

The data set used in the analysis is a random sample of credit card accounts, active and not delinquent as of July 1999. Each account is followed for 13 quarters. The unit of observation is the account in a given quarter. The card could be used by more than an individual and therefore I will refer to the decision maker behind it as household.

Although the original sample covers the entire U.S. territory, I restrict the analysis to accounts located in California, the only state for which retail sales, my measure of local aggregate consumption, are available at the city level and quarterly frequency. For the same reason, although the credit card data are available at monthly frequencies, the interval at which the data are analyzed is quarterly since information on retail sales and other local and aggregate variables is available at such intervals. This choice has also the benefit of reducing the noise that plagues individual monthly consumption data.

In constructing the sample, I exclude people whose accounts are inactive and those that don't use the card very often, in order to obtain a more meaningful measure of consumption. In particular, I exclude accounts on which no expense larger than $50 is charged in any quarter. The choice of this cutoff point is somewhat arbitrary and it is meant to compromise between the needs of keeping a good number of accounts in the sample and at the same time excluding those that are not representative of the credit card expenditures of the account holder. The average transaction amount on a credit card is $87. Following the literature, I also exclude retired account holders and people living in military areas, because their expenditures are influenced by special conditions and required specific modelling that is outside the scope of this paper.[1] Overall, I observe 2,674 accounts between the third quarter of 1999 and the third quarter of 2002.

Extreme outliers. Some of the series present a lot of variability. In order to address this issue, part of the previous literature excludes observations in which the growth rate of consumption is too large. Zeldes (1989) for example excludes observations in which the growth rate of expenditures is bigger than 1.1; Brav et al. (2002) exclude observations for which the growth rate goes from less than 1/2 to more than 2 or it is bigger than 5; Vissing-Jorgensen (2002) excludes observations for which the growth rate goes from less than 0.2 or more than 5.

I choose to keep all the observation because the distributions of expenditure growth rates is very symmetric and the cutoff points seem somewhat arbitrary. I have also tried to exclude observations whose growth rate is bigger than 5 and I get very similar results.

The retail sales data exhibit a lot of variability as well. I have tried winsorizing the data at the 5% and 95% cutoff points and obtained similar estimates.

Missing data. I exclude from the sample, for a given quarter, any observation with missing data on any of the variables included in the basic regression (including income). One of the reasons for missing observations is bankruptcy. In particular, 0.92% of the accounts ends up bankrupt at the end of the sample period. This makes the panel unbalanced, but doesn't biased the results, since in the estimation I control the financial health of the account through debt outstanding, amount charged off and credit constrained indicators.

IB Comparison with the U.S. Census 2000: Analyis of Demographic Characteristics and Borrowing Behavior

This section compares the demographics and borrowing behavior of the account holders to those of the U.S. population.

Table AI and Figure AI compare the distributions of income and age in the data set to those in the U.S. Census 2000. Both the distributions are very alike. The main discrepancy is due to the fact that HHs in the lower range of income and individuals in very young or old age are under-represented in the credit card data

[1] Similar selection procedures are followed by Zeldes (1989) and Attanasio and Weber (1993 and 1995).

set. However, this phenomenon is to be expected given the nature of the variables analyzed: as documented in the Survey of Consumer Finances, HHs with lower income and whose head is younger than 35 or older than 65 are less likely to hold bank-type credit cards. For the same reason, the proportion of HHs that own the house in which they live is 74.88% in my data set, compared to the 66.2% in the Census. Finally, the percentage of married people is slightly less than the national average, 44.76% versus 52.4%, even though the sizeable amount of missing data, 35.27%, could be the cause of the difference.

The data set also contains information on the occupation of the account holders. Unfortunately, the quality is not very good, as 64.4% of the observations are missing. This fact and differences in classification criteria make it impossible to compare the statistics with the Census.[2] It is interesting that 2.51% of the HHs in the sample are headed by a self-employed individual. Table AII illustrates the breakdown of the data set by occupation: the "professional/technical" category is the most widely represented, accounting for 13.76% of the observations, followed by the "administrative/managerial" with 7.59%. An interesting feature related to this variable is that 2.51% of the HHs in the sample are headed by self-employed individuals. This category is somewhat problematic, since they could be using the credit card for their business rather than personal expenditures. In the estimation, I control for this fact by including a dummy variable equal to one if the person is self-employed and zero otherwise. Also, when I exclude entrepreneurs from the analysis the results don't change.

The sample compares well to the U.S. data with regard to HHs borrowing behavior as well. Panel A of Table AIII contains a comparison between my data set and a large multi-issuer credit card data set covering the period between 1995 and 1998 and used by Gross and Souleles (2002). Indebtedness is highly skew in both data sets; while the median debt outstanding is $0 and $70 respectively, the average debt outstanding is $1,486 for my data set and similar for Gross and Souleles. The figure more than doubles if we consider only those HHs that have debt outstanding, reaching $3,400. Both the mean and median credit limits are higher in my data set, probably reflecting the increase in credit availability over the period. Similarly, average and median interest rates are lower in the latter period, due to a trend in the reduction of interest rates that is reflected in the statistics on the rate changes as well.

The other main data set containing information on assets and liabilities of U.S. HHs is the Survey of Consumer Finances (Table AIII Panel B). This source provides a partition of average and median debt outstanding categorized according to demographic characteristics such as age, income percentiles and housing status. The samples are similar with regard to the percentage of people not paying the balance in full at the end of the month, which is estimated to be 44.4% in the SCF and 45.75% in my sample. Unfortunately, the different classification criteria cause some difficulties in comparing the statistics, as the SCF reports the total debt outstanding on all the credit cards available for the HH. Moreover, this survey has been proven to suffer from under-reporting of debt.[3]

IC Credit Cards Expenditures as a Measure of Consumption

Credit card purchases represent an increasing fraction of U.S. consumer spending, having recently overtaken cash and filled part of the gap with respect to checks[4]. In particular, bank credit cards, retail cards, and debit cards account for roughly 24% of personal expenditures in the United States. On average, the typical credit card purchase is about $87 in value, 112% higher than those made in cash.

The quantity analyzed in consumption studies is non-durable goods and services. Credit card purchases are likely to capture this type of expenditures well, although statistics on the type of goods bought with this method of payment are not available. Data from the Consumer Expenditure Survey (CEX), show that in 2001 the average annual expenditure was $38,045. Following the selection criteria of Attanasio and Weber (1995), I construct an estimate of consumption of non-durable goods and services by excluding from the aggregate

[2]The categories in which my dataset classifies occupation are very different than the Census and comprise voices, such as "student" and "housewife", that are not part of the labor force and therefore not considered in the Census.

[3]Both Gross and Souleles (2002) and Laibson et al. (2000) point out this problem.

[4]Gerdes and Walton (2002) provide a description of noncash payments in the U.S.

consumer durables, housing, health and education expenditures. This gives a quarterly expenditure between \$4,765 and \$5,053, depending on the classification criteria. According to the statistics presented above, 24% of this amount, between \$1,143 and \$1,212, is paid with credit card.

The average expenditure on the accounts I analyze is \$701.19. This means that on average the expenditures analyzed cover between 13.88 and 14.72 percent of non-durable goods and services consumption. This figure represents good news, because it indicates that on average the HHs in my data set use this credit card conspicuously and are therefore offering a good measure of their expenditures.

Appendix II

In this Section I model the consumption decisions of a household that borrows through a credit card and saves in a savings account. In each period t, household i chooses consumption $c_{i,t}$ and the amount $P_{i,t}$ of the credit card balance to pay back, with the objective of maximizing the expected value of a lifetime utility function:

$$\underset{\{c_{i,t};P_{i,t}\}_0^{T-1}}{Max} E_t \sum_{t=0}^{T-1} \beta^t U(c_{i,t}, h_{i,t}, H_{i,t}, \Theta_{i,t}) \tag{0.1}$$

subject to:

$$A_{i,t+1} = (A_{i,t} + Y_{i,t} - P_{i,t}) R^f_{i,t} \tag{0.2}$$

$$B_{i,t+1} = (B_{i,t} - P_{i,t} + c_{i,t}) \widetilde{R}^C_{i,t} \tag{0.3}$$

$$c_{i,t} \leq A_{i,t} + Y_{i,t} + \overline{B}_i - B_{i,t} \tag{0.4}$$

$$h_{i,t+1} = \zeta c_{i,t} \tag{0.5}$$

$$A_{i,t} \geq 0, \quad B_{it} \leq \overline{B_i}, \quad c_{i,t} > 0 \tag{0.6}$$

where $h_{i,t}$ is the level of the habit stock that HH i derives from its own past consumption, $H_{i,t}$ is the consumption of the reference group and $\Theta_{i,t}$ are demographic and socioeconomic characteristics; $A_{i,t}$ is the amount of money in the savings account at the beginning of period t, which earns the risk-free rate, $R^f_{i,t}$; $Y_{i,t}$ is the income realization; $B_{i,t}$ is the credit card balance at the beginning of period t, $\overline{B}_i$ the credit limit on the card, and $\widetilde{R}^C_{i,t}$ the gross interest rate the HH is charged on the balance outstanding. The value of $\widetilde{R}^C_{i,t}$ depends on whether or not the HH pays the balance in full and is given by the following expression:

$$\widetilde{R}^C_{i,t} = \begin{array}{ll} R^C_{i,t} > R^f_{i,t} > 1 & \text{if } P_{i,t} < B_{i,t} \\ 1 < R^f_t & \text{if } P_{i,t} \geq B_{i,t} \end{array}$$

Equation (0.2) describes the evolution of the savings account balance: at the end of period t, the account contains the initial funds, plus that period income, minus the credit card payment. Analogously, (0.3) shows that the credit card balance at the beginning of period t+1 consists of the unpaid balance from the previous period, $B_{i,t} - P_{i,t}$, and any new expenditure charged on the card. Equation (0.4) states that consumption cannot exceed the sum of the resources on hand and the unused part of the credit line. Equation (0.5) describes the evolution of the habit stock of the HH, which is assumed to depend on last period consumption only in order to make the empirical analysis more tractable. The specification of the evolution of local aggregate consumption and the demographic characteristics is not necessary for the derivation of the optimal consumption rule and is left for later.[5] Finally, equations (0.6), (**??**) and (**??**) represent the no short sales constraint, the borrowing limit and the condition that consumption must be strictly positive, respectively.

I solve the maximization problem delineated above by expressing it in recursive form. To make the model more tractable and help focusing on the key elements of the problem, I make the following simplifying assumptions.

[5] $H_{i,t}$ and $\Theta_{i,t}$ are therefore suppressed in the rest of the section for notational simplicity.

Assumption 1 *Household income lies between* $\underline{Y_i}$ *and* $\overline{Y}_i$ *and evolves according to* $Y_{i,t} = Y_i + \epsilon_{i,t}$, *where* Y_i *is HH's i mean income and* $\epsilon_{i,t}$ *is i.i.d. and has mean 0 and variance* σ^2.

This assumption implies that in any given period the income realization doesn't depend on previous history. Despite being restrictive, it is aimed at capturing the presence of uninsurable income shocks that causing fluctuations around the income level expected by the household or, alternatively, random unexpected expenses. Exploring the effect of more general income processes constitutes material for future research.

Assumption 2 *The credit limit is set to be* $\overline{B_i} = \sum_0^T \frac{\overline{Y_i}}{R_{i,t}^f}$, *the present value of an income stream in which the highest possible income* $\overline{Y_i}$ *is realized in every period.*

Assumption 3 *There is no default.*

As a first step toward the solution, I note that the optimal payment rule for the HH is to employ all the resources on hand to pay back as much as possible of the outstanding credit card balance. The intuition is that since the household will be charged for $c_{i,t}$ only in period t+1, paying the balance doesn't subtract any resource from current consumption. On the contrary, since there is a wedge between borrowing and lending rates, a bigger payment increases the wealth of the HH because it reduces the amount on which it is charged the high interest rate. Therefore, the only choice variable left for the household is $c_{i,t}$.

The solution of the problem delineated in (0.1)-(**??**) is complicated by the presence of a discontinuity in the value function at the point in which the HH switches from lending at the risk-free rate to borrowing at a higher interest rate. To solve this problem I use as state variable to be the amount of resources on hand, $Z_{i,t} = A_{i,t} + Y_{i,t} - B_{i,t}$. The problem can be then defined over two different regions: the first, region B, in which the HH has negative resources on hand and thus borrows at a high interest rate; and the second, region $\overline{B}$, in which it has enough resources to pay the balance in full, put some money into the savings account and earn the risk-free rate. Since the point of discontinuity coincides with the point in which the HH switches regions, within each region continuity and differentiability are satisfied and standard solution techniques can be applied.

More precisely, the problem of a HH that in period t is not borrowing can be expressed as follows:[6]

$$V_t^{\overline{B}}(Z_t, h_t) = \max_{\{c_t\}} u(c_t, h_t) + \beta \left[\int_{-Z_t R_t^f + c_t}^{\overline{Y}} V_{t+1}^{\overline{B}}(Z_{t+1}, h_{t+1}) f(Y) dY + \int_{\underline{Y}}^{-Z_t R_t^f + c_t} V_{t+1}^{B}(Z_{t+1}, h_{t+1}) f(Y) dY \right] + \lambda_t (Z_t + \overline{B} - c_t) \quad (0.7)$$

$$\text{subject to:} \qquad Z_{t+1} = Z_t R_t^f + Y_{t+1} - c_t \quad (0.8)$$

$$h_{t+1} = \zeta c_t \quad (0.9)$$

where $V_t^{\overline{B}}(Z_t, h_t)$ is the value function; the expression in brackets is the expectation of the future value function, taken with respect to next period income realization; and the last expression is the product of the Lagrange multiplier and the resource constraint. Equations (0.8) and (0.9) describes the evolution of the state variables. The cutoffs of the two integrals show that whether next period the HH falls in one region rather than the other is determined by the resources accumulated from the previous period, $A_{i,t}$, and the realization of the income shock: if $Y_{i,t+1}$ is high enough that the HH is able to repay the credit card balance in full, then it keeps staying in the non-borrowing region; otherwise, it will start borrowing. When the HH makes its consumption decision in period t, it knows that this choice will determine the credit card balance and therefore the likelihood of staying in the non-borrowing region. As time passes and the HH accumulates or decumulates resources, the probability of falling into a certain region gets smaller and smaller, independently from the income realization and the consumption choice.

[6]From now on I suppress the subscript i, for notational simplicity.

The formulation of the problem of a HH that enters period t with an unpaid balance is very similar to (0.7).The only difference is the law of motion of the state variable:

$$V_t^B(Z_t,h_t) = \max_{\{c_t\}} u(c_t,h_t) + \beta\left[\int_{(-Z_t+c_t)R_t^C}^{\overline{Y}} V_{t+1}^{\overline{B}}(Z_{t+1},h_{t+1})f(Y)dY + \int_{\underline{Y}}^{(-Z_t+c_t)R_t^C} V_{t+1}^{B}(Z_{t+1},h_{t+1})f(Y)dY\right] \\ + \lambda_t(Z_t+\overline{B}-c_t) \tag{0.10}$$

$$\text{subject to:} \qquad Z_{t+1} = Y_{t+1} - (-Z_t+c_t)R_t^C \tag{0.11}$$

$$h_{t+1} = \zeta c_t \tag{0.12}$$

The maximization problem delineated above generates the following Euler equations for a HH that at period t is in the non-borrowing region:

$$\begin{aligned} u^c(c_t,h_t) = \beta \int_{-Z_tR_t^f+c_t}^{\overline{Y}} [u_{t+1}^c R_{t+1}^f + \beta R_{t+1}^f \zeta E_Y(u_{t+2}^h) - \zeta u_{t+1}^h] f(Y)dY \\ +\beta\int_{\underline{Y}}^{-Z_tR_t^f+c_t}[u_{t+1}^c + \beta\zeta E_Y(u_{t+2}^h) - \zeta u_{t+1}^h]f(Y)dY \end{aligned} \tag{0.13}$$

and for one that is in the borrowing region:

$$\begin{aligned} u^c(c_t,h_t) = \beta R_t^C \int_{(-Z_t+c_t)R_t^C}^{\overline{Y}} \left[[u_{t+1}^c R_{t+1}^f + \beta R_{t+1}^f \zeta E_Y(u_{t+2}^h)]R_t^C - \zeta u_{t+1}^h\right] f(Y)dY \\ + \int_{\underline{Y}}^{(-Z_t+c_t)R_t^C} \left[[u_{t+1}^c + \beta\zeta E_Y(u_{t+2}^h)]R_t^C - \zeta u_{t+1}^h\right] f(Y)dY \end{aligned} \tag{0.14}$$

These results are very intuitive. The household decides how much to consume today versus tomorrow by weighting future utility and different interest rates by the probability that it will actually face them. If the HH consumes \$1 less today it looses $u^c(c_t)$ and gains the following: next period the credit card balance will be \$1 lower and so one more dollar will be available for consumption, yielding a utility of $u^c(c_{t+1})$; if, in addition to this gain, the income realization is high enough that the HH is able to repay the balance in full, it will earn the gross risk-free rate on the dollar moved through time and the utility will be $u^c(c_{t+1})R_{t+1}^f$. These events realize with probabilities $\int_{\underline{Y}}^{(-Z_t+c_t)R_t^C} f(Y)dY$ and $\int_{(-Z_t+c_t)R_t^C}^{\overline{Y}} f(Y)dY$, respectively.

Analogously, in the case of (**??**), consuming one dollar less today means that the credit card balance next period will be R_t^C dollars less. [7] The utility deriving from this intertemporal transfer will be $u^c(c_{t+1})R_t^C$ if the HH doesn't have enough resources to pay the balance in full in period t+1, and $u^c(c_{t+1})R_t^C R_{t+1}^f$ in case it does and can invest the dollar charged on the credit card at the risk-free rate.

The presence of the habit stock generates an additional effect due to the fact that when the HH consumes one dollar less today it increases tomorrow's utility not only directly, but also by decreasing the habit level. This effect is given by $-\zeta u_{t+1}^h > 0$. [8] The extra dollar consumed in period t+1 will increase the habit stock of period t+2 at a cost in term of utility given by $\beta\zeta E_Y(u_{t+2}^h)$ or $\beta\zeta R_{t+1}^f E_Y(u_{t+2}^h)$, depending on the HH asset position.

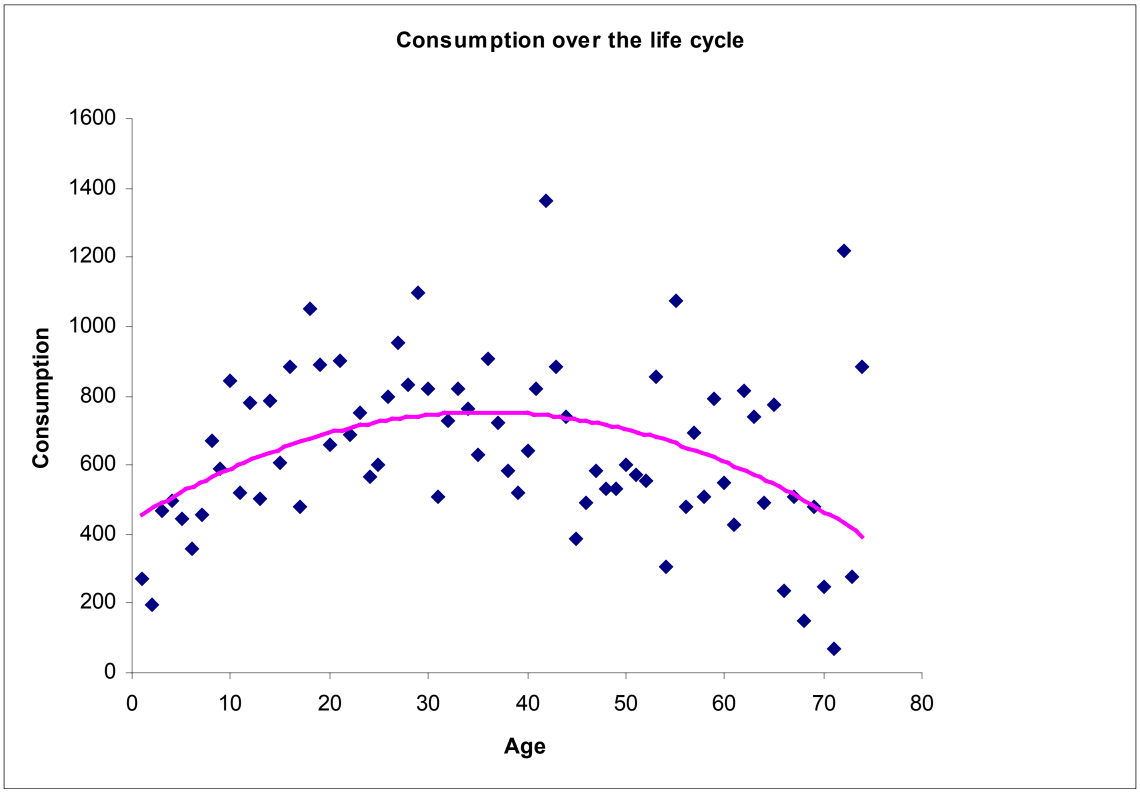
It is also possible to draw a parallel with the Euler equation traditionally obtained in the literature, where the HHs are assumed to be able to borrow and lend at the risk-free rate. In particular, if in case $\overline{B}$ we assume that the resources on hand are high enough that the HH will always pay the balance in full, then (0.13) collapses to the usual Euler equation:

$$u^c(c_t) = \beta\left\{\int_{\underline{Y}}^{\overline{Y}} u^c(c_{t+1})R_{t+1}^f f(Y)dY\right\} \tag{0.15}$$

[7]Since in period t the HH was not able to pay the balance completely it was charged interest on both the unpaid portion of the balance and any new purchases.

[8]Recall that an increase in the habit stock decreases utility and therefore $u^h < 0$.

Fig. I
Panel A



Panel B

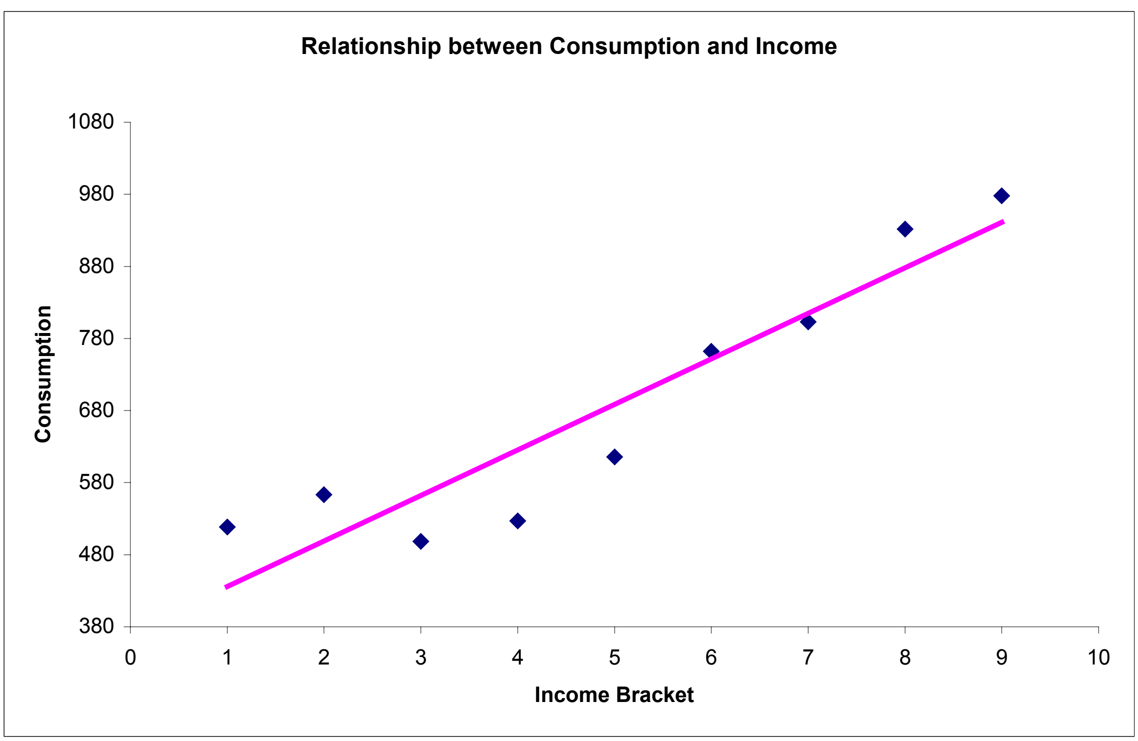
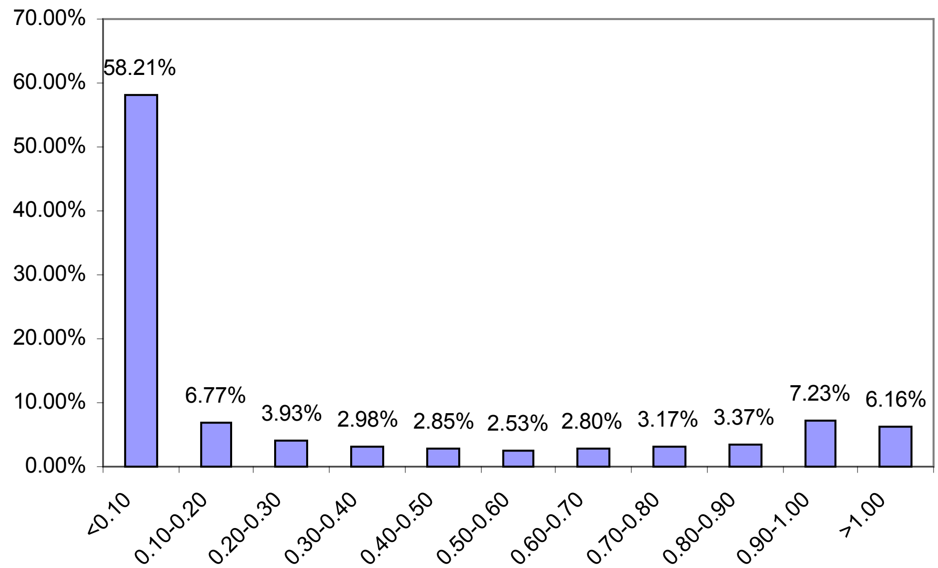


Figure I: plot of average quarterly credit card expenditures against age (panel A) and income bracket (panel B).

Fig. II
Distribution of credit usage



This figure plots the frequency distribution of the ratio between debt and available credit line. From the graph we see that the majority of the observations display quite low credit usage, most of the times due to the very high credit limit they enjoy. Nevertheless, some cases of very high usage and therefore binding credit constraints are present in the data.

Fig. III – Panel A
Evolution of correctly aggregated and per-capita aggregate consumption

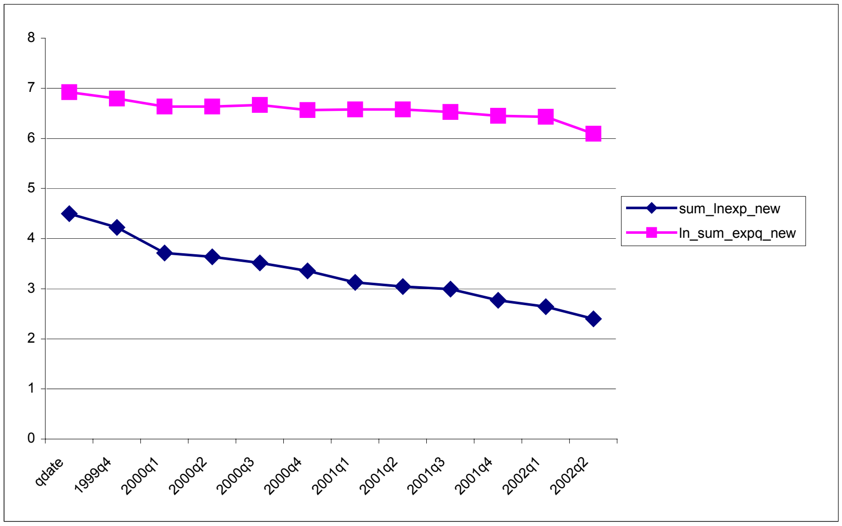
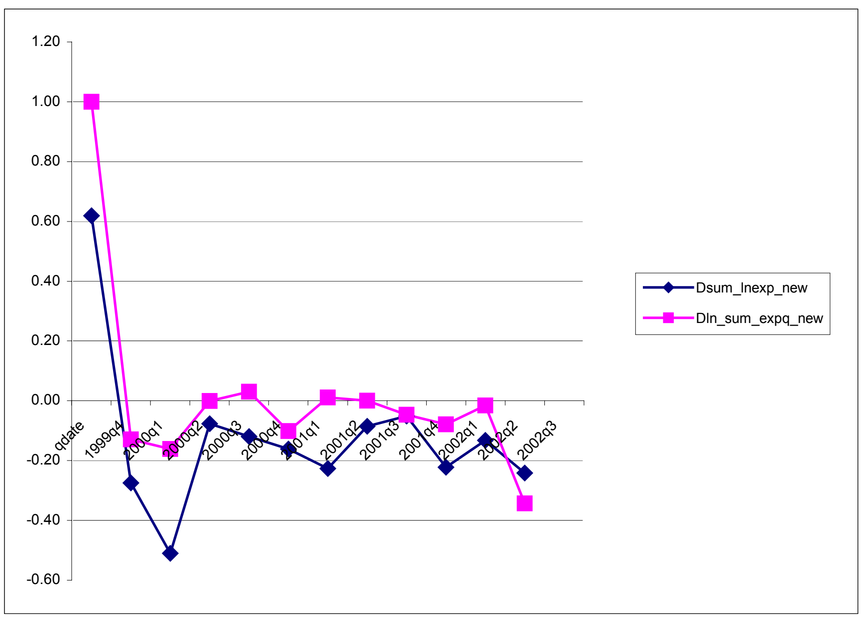
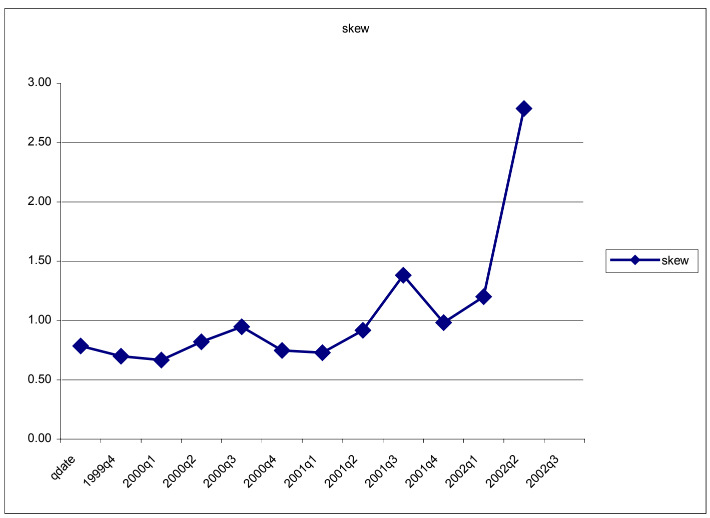


Fig. III – Panel B
Evolution of the growth rates of correctly aggregated and per-capita aggregate consumption

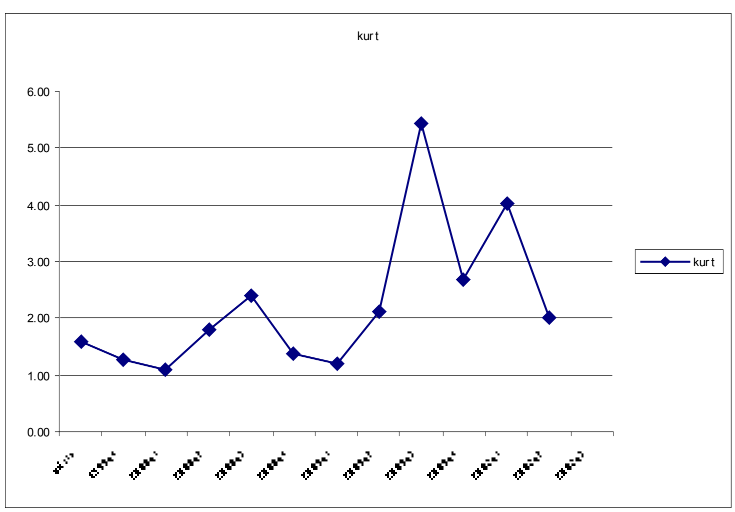


Source: credit card data set

Fig. IV
Evolution of standard deviation, skewness and kurtosis of the cross sectional distribution of household consumption growth rates



Evolution of the cross sectional skewness of household consumption growth



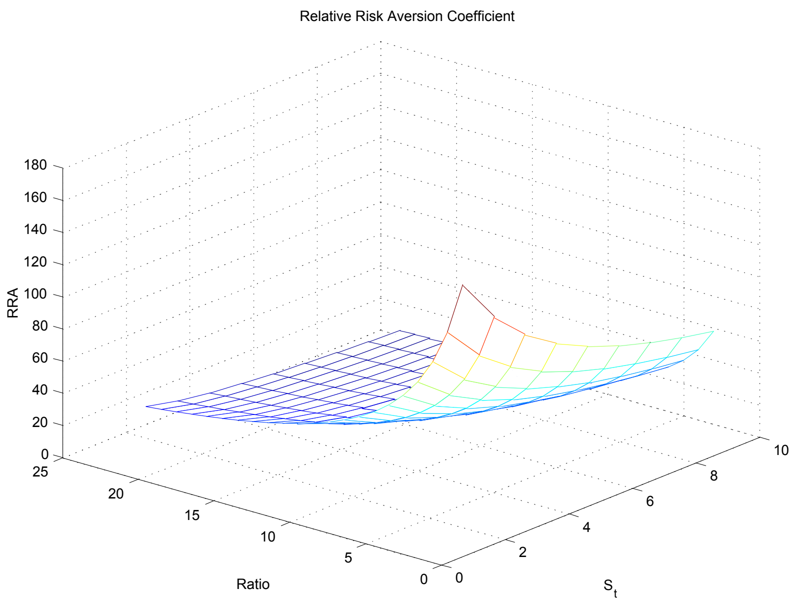


Figure V: Relative risk aversion coefficient as a function of the surplus consumption ratio $S_t = \frac{\gamma}{\left(\frac{c_t - h_t - H_t}{c_t}\right)}$ and $Ratio$=$\frac{E_t(c_{t+1} - h_{t+1} - H_{t+1})}{c_t - h_t - H_t}$.Value of the parameters: $\gamma = 1.38$, $\beta = 0.97$, $\zeta = 0.504$, elasticity=1.1.

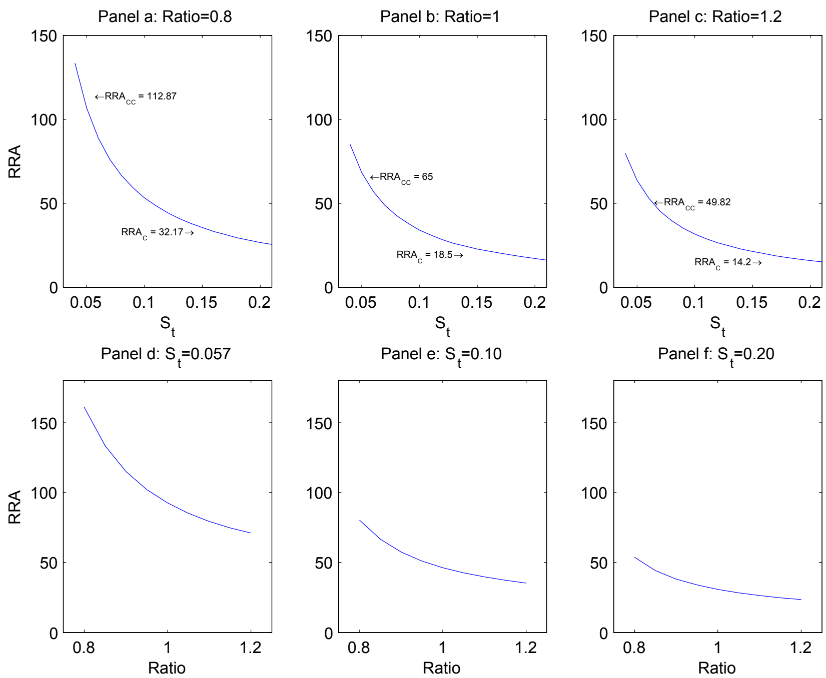


Figure VI: Relative risk aversion coefficient as a function of the surplus consumption ratio $S_t = \frac{\gamma}{\left(\frac{c_t - h_t - H_t}{c_t}\right)}$ and $Ratio = \frac{E_t(c_{t+1} - h_{t+1} - H_{t+1})}{c_t - h_t - H_t}$ taken separately. Panel a shows the relationship between RRA and S_t when $Ratio$=0.8; RRA_{CC} points to the value of the RRA coefficient for the S_t proposed by Campbell and Cochrane (1999), RRA_C points ot the value of the RRA coefficient for the S_t proposed by Constantinides (1990). Panel b shows the relationship between RRA and S_t when $Ratio$=1; Panel c shows the relationship between RRA and S_t when $Ratio$=1.2. Panel d shows the relationship between RRA and $Ratio$ when S_t=0.057; Panel e shows the relationship between RRA and $Ratio$ when S_t=0.10; Panel f shows the relationship between RRA and $Ratio$ when S_t=0.20.

Table I
Data Sources and Variables Description

Variable	Description	Source
Δc_{it}	**Household consumption growth rate** Amount charged on the credit cardduring the quarter for purchases and cash advances. I construct growth rates by taking the difference in the logarithms of this variable between time t and (t-1)	Credit card data set
ΔC_{jt}	**External Reference Point: Citi-level consumption growth rate** city level quarterly taxable sales, defined as "the dollar amount of California retail transactions, excluding those transactions specifically exempt from the California Sales and Use tax". This measure excludes prescription medicines, sales of nontaxable items such as some food products consumed at home and prescription medicines, and taxable transactions disclosed by BOE audits. A detail description is available in "Publication Number 61, Sales and Use Taxes: Exemptions and Exclusions" (March 2003). People not living in a city are associated to the nearest one by geographic matching based on distance between zip code's centroids. The average distance between the centroids of the originating and matching point for HHs not living in a city is 1.1 miles, while the median is 0.4 miles. I construct per capita sales by dividing this quantity by the city population available from Current Population Survey at annual intervals.	State Board of Equalization (BOE)
	Financial Variables	
R^C_{it}	**Household-specific Borrowing Rate** Interest charged on the debt outstanding on the credit card, calculated by taking the ratio of the total charges incurred in a period (finance charges, late charges and over the limit charges) to the balance outstanding.. It is different from the stated APR, because it takes into account compounding and the effective period of time over which the money is borrowed. This results in a more accurate measure of the cost of borrowing	Credit card data set
R^f_t	**Risk Free Rate** 3-month T-bill rate. It represents the risk free at which the HHs are supposed to invest the funds that are left after paying the balance on the credit card.	Federal Reserve Bank of St. Louis (FRED)
Credit constrained indicator$_{it}$	Ratio of debt outstanding to credit limit.It captures credit availability	Credit card data set
Δdebt$_{it}$	**Household-specific Debt Growth Rate** Amount of the credit card balance unpaid and on which the HH is charged interest. Two measures of debt are considered: debt, which represents the overall debt on the card, including balance transfers from other cards. And debt2, which excludes balance transfers and assumes that of any debt outstanding the HH first repays the newly generated debt and only after that the one transferred from other cards.	Credit card data set
Δcredit_line$_{it}$	Household total credit limit on the card	Credit card data set
Δcharged_off$_{it}$	**Amount charged off.** Amount of debt outstanding that the credit card issuer will not be able to recoup and thus writes down as a loss on the account. The reason is bankruptcy, both formally recognized by a court or informal. According to this measure 2.68% of the observations have a positive amount charged off, corresponding to 0.92% of the HHs.	Credit card data set
Balance Transfer$_{it}$	Amount of debt outstanding on another credit card and transferred to this one, or transferred from this card to another. The total number of balance transfers is 562, equal to 1.6% of the observations. The measure of consumption doesn't include balance transfers.	Credit card data set

	Demographic Variables	
Age	Age of the main account holder as of July 1999.	Credit card data set
Marit_status	Dummy variable equal to one if the main account holder is married as of July 1999	Credit card data set
Homeowner	Dummy variable equal to one if the HH owns the house it lives in.	Credit card data set
Income bracket	Income category the HH belongs to	Credit card data set
Occupation dummies	dummy variables indicating the occupation of the primary card holder	Credit card data set
Self_empl	Dummy variable equal to one if the HH head is self-employed.	Credit card data set
Marginal Tax Rate	Marginal tax rate faced by a family or single individual in a given income bracket	Internal Revenue Service (IRS)
	Zip Code-Level Economic Variables	
Median House Value	Median value of the house for a specific sample of owner-occupied houses in the 2000 U.S. Census. The value is the respondent's estimate of how much the property would sell for if it were for sale.	2000 U.S. Census
Median Rent	Median rent asked in the zip code area. No adjustment is made for the inclusion of utilities and fuel.	2000 U.S. Census
	City-Level Economic Variables	
Mortgage Rate_t	Average mortgage rate faced by people living in the city in a given quarter.	American Chamber of Commerce Research Association (ACCRA)
Δln U_rate_t	Quarterly average of the monthly MSA level unemployment rate	Bureau of Labor Statistics (BLS)
Inflation Rate_t	**All Urban Consumers Price Index, base 182:84, not seasonally adjusted.** The quantities are deflated using the BLS Consumer Price Index of the MSA to which they belong (Los Angeles, San Francisco-San Jose or the index for the West Region).	Bureau of Labor Statistics (BLS)
	State-Level Economic Variables	
$\Delta \text{lnincome}_t$	Growth rate of quarterly per-capita disposable income in current dollars	California Department of Finance

Table II
Summary Statistics

		Mean	Median	Std. Dev.
Individual Consumption		701.19	114.92	1596.92
	growth rate	-0.12	0.00	2.73
Aggregate Consumption		2367.22	2239.58	975.53
	growth rate	0.01	0.03	0.19
R^{C}_{it}		16.95%	0.00%	114.65%
R^{f}_{it}		3.88%	4.42%	1.71%
Inflation rate		0.79%	0.87%	0.46%
DEMOGRAPHIC VARIABLES				
Age.		46	45	15
Marital Status		44.76%		
Home Owner/ Renter		74.88%		
Marginal Tax Rate		28.44%	28.00%	7.37%
CHARACTERISTICS of the CONTRACT				
Debt		$1,060.05	$0.00	$2,048.60
Credit Line		$8,823.51	$10,000.00	$3,733.95
Charged Off		$152.33	$0.00	$1,058.07
	dummy	0.027	0.000	0.160
Credit Constrained		0.33	0.07	0.41
	dummy	0.31		
Balance Transfers		$132.64	$0.00	$1,004.97
	dummy	0.016	0	0.126
ZIP CODE-LEVEL DATA				
Median House Value		$258,980.80	$222,700.00	$146,479.10
Rent.		$791.10	$752.00	$245.95
CITY-LEVEL DATA				
Mortgage Rate		7.32%	7.21%	0.59%
Unemployment rate		5.07%	4.43%	3.22%
STATE-LEVEL DATA				
Personal Income		$26,792.17	$26,906.59	$895.74

Table III
Mean and Std. Dev. of Individual Consumption

In this Table I compare the mean and standard deviation of my measure of consumption to data from the publicly available data sets traditionally used in the literature: the Panel Study of Income Dynamics (PSID), the Consumer Expenditure Survey (CEX) and aggregate data.

		My data set		Comparison data set	
		Mean	Std.Dev.	Mean	Std.Dev.
Individual Data					
	All	-0.120	2.730		
	-3.3/3.3	0.005	0.870		
	-1.1/1.1	-0.002	0.368	0.000	0.32*
Cross Sectional Data					
	All	0.014	0.330	-0.01*	0.06*
Aggregate Data					
	All	0.000	0.013	0.003***	0.009***

* Comparison data set: PSID, data from Zeldes (1989)

* Comparison data set: CEX, data from Brav et al. (2002)

* Comparison data set: aggregate data, from Constantinides et al. (1991)

Table IV
Aggregate Local Consumption
Comparison with NIPA Aggregate Consumption

In this table I compare the city-level sales data that I use to construct the measure of the external reference point to the aggregate data traditionally used in the literature: the aggregate consumption from the National Income and Product Accounts (NIPA). The main difference between the two measures is the lack of housing services in the California aggregate sales

	California Aggregate Sales	Non-durables and Services (NIPA)
Mean	3674.394	5659.72
Median	3669.729	5737
Std. Dev.	402.2181	404.1349
Correlation Coefficient	0.8583	

For a definition of the California aggregate sales see Table I

Table V
Autocorrelations of Household Consumption

	Δc_t	Δc_{t-1}	Δc_{t-2}	Δc_{t-3}
Δc_t	1			
Δc_{t-1}	-0.3863	1		
Δc_{t-2}	-0.0281	-0.3866	1	
Δc_{t-3}	-0.0212	-0.0317	-0.3914	1

Table VI
Basic Estimation

In this Table I present the results of the estimation of the Euler equation:

$$\Delta lnc_{i,t}=k1+\alpha_0\Delta lnC_{i,t}+\alpha_{-1}\Delta lnC_{i,t-1}+\zeta\Delta lnc_{i,t-1}+\gamma ln(1+(R_{i,t}^{f}-1)Pr[Y_{i,t}^{H}]) +\eta ln(1+(R_{i,t}^{C} -1)1[B])+\theta_1\Delta age_{i,t}+\theta_2\Delta age_{i,t}+\varepsilon_{i,t}$$

Column I of Table VI estimates the basic model, while columns II to IV progressively add to the specification family composition, home ownership and occupation.

The standard errors are corrected for the non-independence of the observations within the same household. In addition, controls for the evolution of city-level prices and seasonal dummies are included in all the regressions.

	(I)	(II)	(III)	(IV)
R^C_t	-1.727***	-1.845***	-1.842***	-1.876***
	(0.000)	(0.000)	(0.000)	(0.000)
R^f_t	0.448	0.770	0.765	0.824
	(0.537)	(0.381)	(0.384)	(0.349)
ΔC_t	0.258***	0.295**	0.294**	0.290**
	(0.005)	(0.024)	(0.024)	(0.026)
Δc_{t-1}	0.530***	0.501***	0.502***	0.503***
	(0.000)	(0.000)	(0.000)	(0.000)
ΔC_{t-1}	0.007	0.014	0.014	0.013
	(0.195)	(0.176)	(0.176)	(0.189)
age		0.004	0.004	0.004
		(0.334)	(0.363)	(0.305)
age^2		-0.000	-0.000	-0.000
		(0.141)	(0.152)	(0.127)
marital status		0.022	0.023	0.023
		(0.399)	(0.375)	(0.386)
homeowner		-0.018	-0.017	-0.017
		(0.656)	(0.676)	(0.668)
income bracket		-0.003	-0.004	-0.002
		(0.524)	(0.470)	(0.705)
self_empl			0.085	0.071
			(0.158)	(0.252)
Occupation dummies			Yes	Yes
Seasonal dummies	Yes	Yes	Yes	Yes
# Obs	2220	1432	1432	1432
Hansen J statistic	8.725	7.712	7.730	7.853
(pvalue)	0.463	0.441	0.562	0.441
C statistic+	1.021	3.748	3.761	3.747
(pvalue)	0.907	0.563	0.439	0.549
Adj. R-squared	0.214	0.226	0.226	0.225

Robust p values in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

\+ Instruments tested: lags of debt outstanding, amount charged off, change in credit line and credit constraints measure.

A description of the variables is reported in Table I.

Instrument set: marginal tax rate, the local unemployment rate, the inflation rate, aggregate disposable income growth rate, mortgage rate, and some individual variables such as lags of the growth rate of debt, amount charged off, automatic credit line changes, and a credit constrained indicator.

	(I)	(II)	(III)	(IV)
	Adj. R^2	Adj. R^2	Adj. R^2	Adj. R^2
R^C_t	0.0635	0.0602	0.0605	0.0618
R^f_t	0.9543	0.9543	0.9543	0.9543
ΔC_t	0.2351	0.3666	0.3666	0.3661
Δc_{t-1}	0.0281	0.0282	0.0281	0.028
	p-value	p-value	p-value	p-value
R^C_t	0.0000	0.0000	0.0000	0.0000
R^f_t	0.0000	0.0000	0.0000	0.0000
ΔC_t	0.0000	0.0000	0.0000	0.0000
Δc_{t-1}	0.0000	0.0000	0.0000	0.0000

Table VII
Comparison with Dynan (2000)

This Table illustrates the comparison between my results and Dynan (2000). In particular, I re-estimate the same regression as in Col (V) of Table VI using an annualized measure of credit card consumption and of the other variables:

$$\Delta \ln c_{i,t} = k1 + \alpha_0 \Delta \ln C_{i,t} + \alpha_{-1} \Delta \ln C_{i,t-1} + \zeta \Delta \ln c_{i,t-1} + \gamma \ln(1+(R_{i,t}^{f}-1)\Pr[Y_{i,t}^{H}]) + \eta \ln(1+(R_{i,t}^{C}-1)1[B]) + \theta_1 \Delta age_{i,t} + \theta_2 \Delta age_{i,t} + \varepsilon_{i,t}$$

The instrument set includes second lag of local unemployment rate, the inflation rate, aggregate income growth rate, mortgage rate, household-specific debt growth rate and credit constrained indicator.
Col (I) presents the baseline estimation; Col (II) investigates the effect of estimating the above regression without the external habit, as Dynan's estimation doesn't contain this variable. Col (III) drops the HH-specific financial variables from the instrument set; while Col (IV) drops the interest rate from the estimation equation, as Dynan considers an Euler equation with constant interest rates. The standard errors are corrected for the non-independence of the observations within the same household. In addition, controls for the evolution of city-level prices and seasonal dummies are included in all the regressions.

	(I)	(II)	(III)	(IV)
R^C_t	-0.076	-0.116	-2.100	
	(0.487)	(0.246)	(0.352)	
ΔC_t	0.110		1.055***	0.985***
	(0.931)		(0.000)	(0.000)
Δc_{t-1}	0.602	0.662***	0.109	0.139**
	(0.422)	(0.000)	(0.179)	(0.048)
Age	-0.001	-0.001	0.007	0.000
	(0.837)	(0.622)	(0.367)	(0.910)
age^2	0.000	0.000	-0.000	-0.000
	(0.829)	(0.584)	(0.346)	(0.872)
marital status	0.006	-0.000	0.045	0.010
	(0.740)	(0.974)	(0.273)	(0.529)
self_empl	-0.031	-0.031	-0.014	-0.038
	(0.519)	(0.151)	(0.839)	(0.290)
homeowner	-0.001	-0.005	-0.035	0.006
	(0.964)	(0.790)	(0.539)	(0.815)
income bracket	-0.002	-0.001	-0.008	-0.003
	(0.559)	(0.539)	(0.287)	(0.439)
Occupation dummies	Yes	Yes	Yes	Yes
Seasonal dummies	Yes	Yes	Yes	Yes
# Obs	1330	1450	1336	1336
Hansen J statistic	29.145	29.067		1.616
p-value		0.060		0.204
Adj. R-squared	-0.490	0.712	0.386	0.441

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
A description of the variables is reported in Table I.
Instrument Set: second lag of local unemployment rate, the inflation rate, aggregate income growth rate, mortgage rate, household-specific debt growth rate and credit constrained indicator.

First stage Regressions Results

Variable	Column	Partial R^2	F test	P-value
R^C_t	(I)	0.0846	130.9	0.0000
ΔC_t		0.14	230.61	0.0000
Δc_{t-1}		0.1587	267.17	0.0000
R^C_t	(II)	0.0768	126.08	0.0000
Δc_{t-1}		0.162	292.84	0.0000
R^C_t	(III)	0.0014	3.89	0.0086
ΔC_t		0.0744	226.34	0.0000
Δc_{t-1}		0.0814	249.38	0.0000
ΔC_t	(IV)	0.0744	226.34	0.0000
Δc_{t-1}		0.0814	249.38	0.0000

Table VIII
Effect of a smaller IV set

In this table I perform the same regressions as in Column I of Table VIII (Basic Results) with a progressively smaller instrument set. The IVs used in the estimation in the paper are: marginal tax rate, various lags of local unemployment rate, inflation rate, aggregate disposable income growth rate, mortgage rate, and some individual variables such as lags of the growth rate of debt, amount charged off, automatic credit line changes, and credit constrained indicator.
In Column II, I eliminate the marginal tax rate from the IV set illustrated above. In Column III I further eliminate the unemployment rate; in Column IV the inflation rate; in Column V all but one lags of aggregate income; in column VI the amount charged off; finally, in Column VII, I eliminate the credit line increases. The IV set I am left with exactly identifies the system and is composed by the first available lag of mortgage rate, aggregate income growth, household debt growth rate and credit constrained indicator.
The results show that restricting the IV set doesn't change the coefficients on ΔC_t and Δc_{t-1}, or their significance. The coefficient that is most sensitive to the shrinking of the IV set is that on the risk free interest rate, which happens sometimes to be negative, even though very close to zero and very imprecisely measured.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
R^C_t	-1.727***	-1.743***	-1.742***	-1.750***	-1.738***	-1.752***	-1.568***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^f_t	0.448	0.408	-0.002	0.362	-0.009	-0.024	0.322
	(0.537)	(0.576)	(0.998)	(0.670)	(0.992)	(0.981)	(0.762)
ΔC_t	0.258***	0.266***	0.370***	0.283**	0.354**	0.352**	0.365**
	(0.005)	(0.004)	(0.001)	(0.036)	(0.031)	(0.034)	(0.029)
Δc_{t-1}	0.530***	0.527***	0.530***	0.529***	0.526***	0.523***	0.575***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ΔC_{t-1}	0.007	0.007	0.010*	0.007	0.009	0.009	0.010
	(0.195)	(0.202)	(0.083)	(0.247)	(0.190)	(0.215)	(0.188)
Observations	2220	2220	2220	2220	2220	2220	2220
Hansen J statistic	8.725	8.292	3.322	2.286	0.819	0.816	-
p-value	0.463	0.405	0.650	0.683	0.664	0.366	-
Adj. R-squared	0.214	0.211	0.158	0.203	0.168	0.169	0.156

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

A description of the variables is reported in Table I.

Table IX
Robustness Checks

In this Table I analyze the robustness of the results to the inclusion of various measures of economic activity to the regression: the state-level income growth rate (Col (I)), the lead of the state-level income growth rate (Col (II)), change in city-level unemployment rates (Col (III)), housing market conditions (Col. (IV)). I also check the effect of adding an extra lag of the consumption growth rate (Col. (V)), year dummies to control for aggregate shocks (Col. VI)), and the growth of household debt (Col. (VII)).
The standard errors are corrected for the non-independence of the observations within the same household. In addition, controls for the evolution of city-level prices and seasonal dummies are included in all the regressions.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
R^C_t	-1.878***	-1.862***	-1.871***	-1.833***	-2.054***	-1.849***	-1.322***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)
R^f_t	2.116	1.213	0.550	0.711	0.266	3.113	0.189
	(0.174)	(0.230)	(0.527)	(0.432)	(0.777)	(0.548)	(0.835)
ΔC_t	0.271**	0.268**	0.442**	0.290**	0.311**	0.266**	0.295**
	(0.037)	(0.045)	(0.016)	(0.030)	(0.020)	(0.039)	(0.024)
Δc_{t-1}	0.501***	0.507***	0.509***	0.503***	0.453***	0.508***	0.445***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ΔC_{t-1}	0.014	0.012	0.021	0.014	0.016	0.013	0.014
	(0.164)	(0.251)	(0.107)	(0.184)	(0.105)	(0.234)	(0.161)
Age	-0.000	-0.000	-0.000	-0.004	-0.000	-0.000	0.004
	(0.109)	(0.135)	(0.188)	(0.324)	(0.261)	(0.150)	(0.419)
Age^2	0.023	0.025	0.021	0.000	0.027	0.025	-0.000
	(0.375)	(0.360)	(0.466)	(0.138)	(0.348)	(0.366)	(0.199)
Marit_status	-0.014	-0.015	-0.028	0.026	-0.039	-0.014	0.023
	(0.724)	(0.708)	(0.498)	(0.336)	(0.361)	(0.722)	(0.401)
Homeowner	-0.002	-0.002	-0.003	-0.012	-0.000	-0.002	-0.025
	(0.692)	(0.712)	(0.650)	(0.767)	(0.956)	(0.701)	(0.538)
Income bracket	0.068	0.072	0.080	-0.004	0.092	0.071	-0.002
	(0.262)	(0.248)	(0.201)	(0.466)	(0.158)	(0.248)	(0.764)
Self_empl	-1.878***	-1.862***	-1.871***	0.065	-2.054***	-1.849***	0.057
	(0.000)	(0.000)	(0.000)	(0.288)	(0.000)	(0.000)	(0.331)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Δlnincome	0.020						
	(0.319)						
Δlnincome _lead		0.012					
		(0.434)					
Δln U_rate			-0.003				
			(0.465)				
Median House Value				0.000			
				(0.705)			
Rent				0.000			
				(0.544)			
$(\Delta c_t)^2$				-0.013			
				(0.527)			
Δc_{t-2}					0.170***		
					(0.000)		
$\Delta lndebt_t$							-0.157*
							(0.060)
Year dummies						Yes	
Seasonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	1432	1432.	1432	1432	1432	1432	1432
Hansen J statistic	6.986	7.245	4.546	7.890	12.705	5.397	9.695
(pvalue)	0.538	0.524	0.576	0.545	0.176	0.524	0.287
C statistic+	5.224	3.204	2.892	3.933	5.665	3.204	2.487
(pvalue)	0.265	0.510	0.715	0.415	0.226	0.798	0.647
adj. R-squared	0.230	0.231	0.176	0.219	0.249	0.232	0.264

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
+ Instruments tested: lags of debt outstanding, amount charged off, change in credit line and credit constraints measure.
A description of the variables is reported in Table I.
Instrument set: marginal tax rate, the local unemployment rate, the inflation rate, aggregate disposable income growth rate, mortgage rate, and some individual variables such as lags of the growth rate of debt, amount charged off, automatic credit line changes, and a credit constrained indicator.

Table X
Does Local Aggregate Consumption Proxy for Individual Income?

In order to answer this question I use household-level data from the PSID, in which individual income is provided.[1] I investigate whether local aggregate consumption provides any information about household income once aggregate income is available. The regression is the following:

$$\Delta \mathbf{lnincome}_{\text{individual, it}} = \alpha + \beta_1 \Delta \mathbf{lnINCOME}_{\text{aggregate, it}} + \beta_2 \Delta \mathbf{lnC}_{\text{citylevel, it}} + \beta_3 \mathbf{age}_i + \beta_4 \mathbf{age}^2_i + \varepsilon_{it}$$

Dummy variables capturing marital status, occupational choice, seasonal fluctuations and whether the individual owns the house in which he lives are included in some of the regressions as well.

The null hypothesis is that once we control for aggregate income, the coefficient on aggregate consumption is small and statistically insignificant. The results below confirm this hypothesis and suggest that aggregate consumption doesn't proxy for individual income.

I have also tried the above regression using the growth rate of future HH income as the dependent variable. The results are presented in columns IV to VI and show that neither aggregate income nor consumption are good predictors of future individual income. The regression is the following:

$$\Delta \mathbf{lnincome\_lead}_{\text{individual, it+1}} = \alpha + \beta_1 \Delta \mathbf{lnINCOME}_{\text{aggregate, it}} + \beta_2 \Delta \mathbf{lnC}_{\text{aggregate, it}} + \beta_3 \mathbf{age}_i + \beta_4 \mathbf{age}^2_i + \varepsilon_{it}$$

Notice should be given to the fact that the time period analyzed is not very long, spanning from 1997 to 2001. Unfortunately a longer time series of city-level consumption is not available.

Dependent Variable	**Δlnincome$_{\text{individual, it}}$**			**Δlnincome_lead$_{\text{individual, it+1}}$**		
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
ΔlnC	0.045	0.066	0.051	0.093	0.057	0.053
	(0.550)	(0.377)	(0.492)	(0.290)	(0.474)	(0.508)
ΔlnINCOME	2.697***	2.660***	2.887***	-0.674	-0.334	-0.525
	(0.009)	(0.009)	(0.005)	(0.726)	(0.849)	(0.768)
Age	-0.018**	-0.023***	-0.026***	0.001	-0.000	-0.002
	(0.041)	(0.010)	(0.005)	(0.940)	(0.967)	(0.886)
Age^2	0.000*	0.000**	0.000***	-0.000	0.000	0.000
	(0.075)	(0.022)	(0.006)	(0.989)	(0.915)	(0.818)
Marital status	-0.057	-0.076	-0.072	-0.077	-0.094	-0.112*
	(0.241)	(0.144)	(0.172)	(0.243)	(0.153)	(0.098)
Homeowner dummy		Yes	Yes		Yes	Yes
Occupation dummies			Yes			Yes
Observations	921	891	891	426	413	413
R-squared	0.017	0.022	0.036	0.019	0.017	0.027

p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Sample: households in the PSID living in California. Time span: 1997 to 2001.[2]
Variables description and definitions: *Δlnincome$_{individual, it}$* is the quarterly growth rate of household total income; *ΔlnINC* is the quarterly growth rate of aggregate California per capita income (Source: California DOF); *ΔlnC* is the quarterly growth rate of city-level aggregate taxable sales and constitutes my measure of aggregate consumption (Source: California DOF); *Age* is the age of the head of the household; *Marital status* is a dummy equal to 1 if the head of the HH is married and 0 otherwise; the *Homeowner dummy* equals 1 if the HH owns the house in which it lives and 0 otherwise; the *Occupational dummies* are dummies that categorize HH heads in main occupational areas and are built to be as similar as possible to those used in the rest of the paper.

[1]Unfortunately, this dataset doesn't contain a good measure of consumption; for other drawbacks of the PSID see Section 3.2.
[2] This is the period for which data on taxable sales are available on the California Department of Finance website.

Table XI
Alternative Explanations: Liquidity Constraints and Precautionary Saving Motives

This Table contains tests of the habit persistence hypothesis against the liquidity constraint and precautionary saving motive alternatives. The baseline regression to which the results are compared is:

$$\Delta lnc_{i,t}=k1+\alpha_0\Delta lnC_{i,t}+\alpha_{-1}\Delta lnC_{i,t-1}+\zeta\Delta lnc_{i,t-1}+\gamma ln(1+(R_{i,t}^{f}-1)Pr[Y_{i,t}^{H}])+\eta ln(1+(R_{i,t}^{C}-1)1[B])+\theta_1\Delta age_{i,t}+\theta_2\Delta age_{i,t}+\varepsilon_{i,t}$$

.Col (I) adds the lagged growth rate of income to the regression; Col. (II) and (III) re-estimate the regression on two sub-samples of unconstrained and credit constrained HHs; col (IV) and (V) perform the same regression as (II) and (III) adding the lag of income growth rate. Col (VI) adds a credit constrained indicator. Col (VII) adds the square of consumption growth to the regression.
The standard errors are corrected for the non-independence of the observations within the same household. In addition, controls for the evolution of city-level prices and seasonal dummies are included in all the regressions.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
R^C_t	-1.879***	1.788	-1.109	-0.141	-2.395	-0.603	-1.695***
	(0.000)	(0.390)	(0.829)	(0.939)	(0.638)	(0.423)	(0.000)
R^f_t	0.414	-2.313	3.906	-6.201*	7.733	0.161	1.543
	(0.694)	(0.556)	(0.350)	(0.099)	(0.124)	(0.870)	(0.287)
ΔC_t	0.274**	0.644**	0.186	0.267	0.233	0.240*	0.290**
	(0.035)	(0.035)	(0.674)	(0.392)	(0.598)	(0.063)	(0.024)
Δc_{t-1}	0.502***	0.565***	0.557***	0.416***	0.693***	0.498***	0.504***
	(0.000)	(0.000)	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)
$\Delta lnincome_{t-1}$	0.012			0.081**	-0.076		
	(0.473)			(0.048)	(0.124)		
ΔC_{t-1}	0.014	0.023	-0.362***	0.016	-0.396***	0.011	0.014
	(0.165)	(0.307)	(0.005)	(0.269)	(0.003)	(0.253)	(0.184)
Age	0.005	-0.015	0.001	-0.006	0.006	0.001	0.004
	(0.276)	(0.241)	(0.963)	(0.596)	(0.708)	(0.900)	(0.339)
Age^2	-0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.111)	(0.384)	(0.968)	(0.842)	(0.736)	(0.775)	(0.149)
Marit_status	0.023	0.026	-0.048	0.009	-0.038	0.012	0.026
	(0.372)	(0.759)	(0.490)	(0.902)	(0.590)	(0.623)	(0.372)
Homeowner	-0.015	-0.072	-0.148	-0.021	-0.159	0.007	-0.017
	(0.713)	(0.529)	(0.247)	(0.840)	(0.236)	(0.839)	(0.691)
Income bracket	-0.002	0.015	0.027*	0.008	0.024	0.001	-0.003
	(0.698)	(0.437)	(0.094)	(0.585)	(0.173)	(0.822)	(0.653)
Self_empl	0.068	-0.188	0.243	-0.031	0.287	0.063	0.067
	(0.266)	(0.473)	(0.347)	(0.885)	(0.271)	(0.177)	(0.296)
Credit constr. indic.						0.062	
						(0.327)	
$(\Delta c_t)^2$							-0.013
							(0.527)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Seasonal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	1432	773	347	773	347	1432	1432
Hansen J statistic	7.491	5.400	3.298	3.780	1.028	7.426	7.567
Pvalue	0.485	0.483	0.654	0.779	0.906	0.406	0.472
C statistic	5.174	0.493	0.000	0.079	0.002	2.911	3.540
p-value	0.270	0.249	0.990	0.286	0.966	0.491	0.477
Adj. R-squared	0.230	0.229	0.282	0.165	0.240	0.255	0.233

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

A description of the variables is reported in Table I.
Instrument set: marginal tax rate, the local unemployment rate, the inflation rate, aggregate disposable income growth rate, mortgage rate, and some individual variables such as lags of the growth rate of debt, amount charged off, automatic credit line changes, and a credit constrained indicator.

Table XII
Aggregate Consumption Regressions

This Table investigates the aggregate implications of household level consumption choices. Columns (I), (II), and (III) present the results of estimating an aggregate Euler equation based on per capita consumption: (a) $\mathbf{\Delta ln\Sigma c_{i,t}= \alpha+\beta\Delta ln\Sigma c_{i,t\text{-}}1 + R^f_t + \varepsilon_t}$. Columns (IV), (V), and (VI) presents the results of estimating the same Euler equation using correctly aggregated data: (b) $\mathbf{\Delta\Sigma lnc_{i,t}= \alpha+\beta\Delta\Sigma lnc_{i,t\text{-}}1 + R^f_t + \varepsilon_t}$. Finally, Columns (VII), (VIII), and (IX) illustrate the effect of adding moments of the cross sectional distribution of consumption growth rates to regression.
The standard errors are corrected for the non-independence of the observations within the same household. In addition, controls for the evolution of city-level prices and seasonal dummies are included in all the regressions.

	Per-capita Consumption Aggregation Method			**Correct Aggregation Method**			**Per-capita Consumption Aggregation Method plus Moments**		
Dependent variable	$\Delta ln\Sigma c_t$			$\Delta\Sigma lnc_t$			$\Delta ln\Sigma c_t$		
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
R^f_t	1.361	1.468	0.872	-0.112	-0.018	-0.042	2.401*	1.454**	1.444**
	(0.186)	(0.144)	(0.190)	(0.818)	(0.974)	(0.926)	(0.064)	(0.038)	(0.040)
$\Delta ln\Sigma c_{i,t-1}$	-0.685	-0.739	-0.494				0.139	-0.791**	0.233
	(0.205)	(0.191)	(0.164)				(0.942)	(0.033)	(0.704)
$\Delta\Sigma lnc_{i,t-1}$				0.515**	0.727**	0.538*			
				(0.042)	(0.039)	(0.089)			
$\Delta lnincome_t$		0.002			-0.131				
		(0.996)			(0.404)				
$\Delta lnincome_{t-1}$			0.471*			0.147			
			(0.061)			(0.344)			
Δstddev/2							36.669		
							(0.489)		
Δskewness/6								2.278	
								(0.571)	
Δkurtosis/24									-4.954
									(0.116)
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	10	10	10	10	10	10	10	10	10
Hansen J statistic	2.244	2.061	4.108	2.286	1.213	1.210	0.002	3.532	0.370
p-value	0.326	0.151	0.043	0.319	0.271	0.271	0.965	0.060	0.543
Adj. R-squared	0.380	0.363	0.685	0.285	0.009	0.170	-.241	0.479	0.174

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

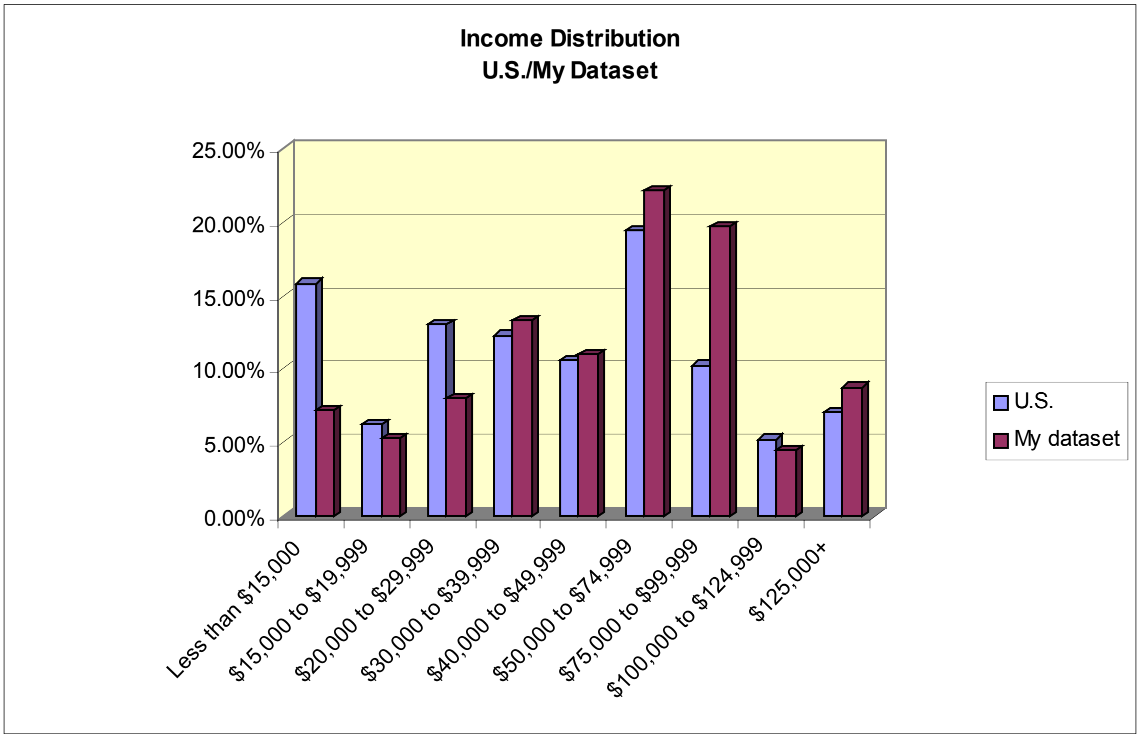
A description of the variables is reported in Table I. Δstddev, Δskewness, and Δkurtosis are the standard deviation, skewness and kurtosis of the cross sectional distribution of consumption.
Instrument set: second lag of aggregated city-level sales, income growth rate, average mortgage rate, and unemployment rate.

Table AI
Demographic Characteristics
Comparison with U.S. Census data

This Table contains a comparison between the breakdown by age and income of my data set versus the U.S. Census.

	Age	
	My dataset CA	*U.S. Census*
20 to 24 years	7.11%	9.44%
25 to 34 years	17.61%	19.85%
35 to 44 years	24.91%	22.47%
45 to 54 years	24.01%	18.75%
55 to 59 years	8.12%	6.70%
60 to 64 years	5.12%	5.38%
65 to 74 years	9.16%	9.15%
75 to 84 years	3.48%	6.15%
85 years and over	0.49%	2.11%
Total	**100%**	**100%**
	Income	
	My dataset CA	*U.S. Census*
Less than $15,000	7.21%	15.85%
$15,000 to $19,999	5.33%	6.25%
$20,000 to $29,999	8.04%	13.02%
$30,000 to $39,999	13.32%	12.27%
$40,000 to $49,999	11.04%	10.62%
$50,000 to $74,999	22.12%	19.46%
$75,000 to $99,999	19.74%	10.23%
$100,000 to $124,999	4.44%	5.20%
$125,000+	8.75%	7.09%
Total	100%	100%
	Other Demographic Characteristics	
	My dataset CA	*U.S. Census*
Home owner	74.68%	66.20%
Renter	9.69%	33.80%
Missing	15.63%	0.00%
Married	44.76%	52.40%
Single	19.97%	47.60%
Missing	35.27%	0.00%

Figure AI



Source: U.S. Census and credit card data set.

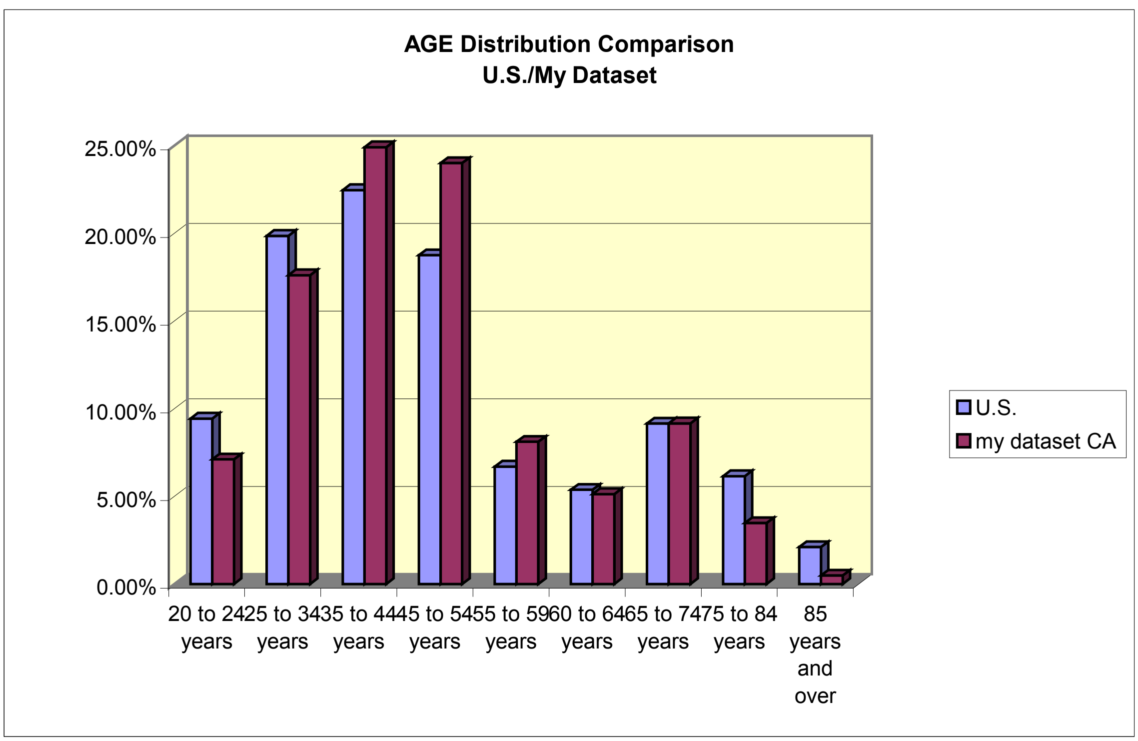


Table AII
Occupation:
Comparison with U.S. Census 2000

My Dataset	
Occupation	*Percent*
Administrative/Managerial	7.59%
Clerical/White_Collar	3.25%
Craftsman/Blue_Collar	3.93%
Farmer	0.11%
Housewife	0.67%
Military	0.19%
Professional/Technical	13.76%
Sales/Service	1.91%
Self_Employed	2.51%
Student	1.65%
Missing	64.44%
Total	100%

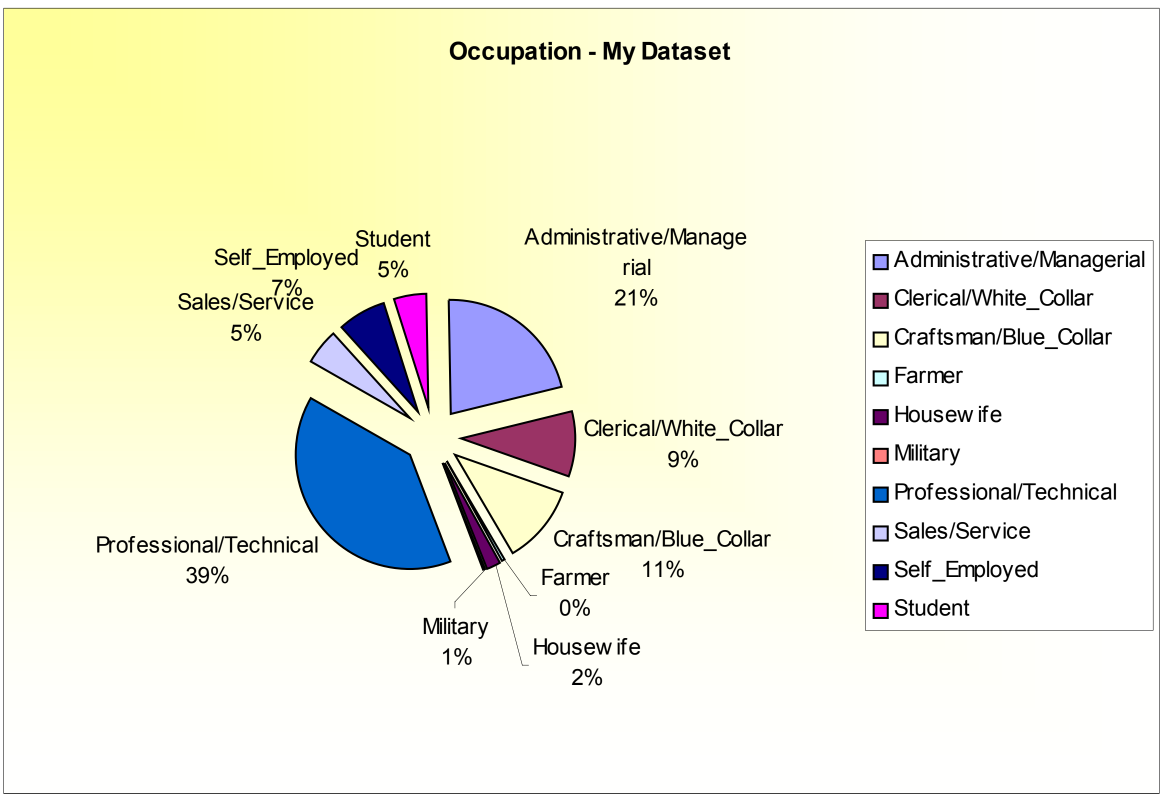


Table A III
Panel A
Financial Characteristics
Comparison with multi-issuer credit card dataset

	My dataset*		Gross and Souleles (2002)**	
	Mean	Median	Mean	Median
debt	$1,486.10	$0	$1,349.00	$70
debt|debt>0	$3,408.36	$2,821	$2,809.00	$2,120
credit limit	$8,823.51	$10,000	$6,207.00	$5,000
D credit limit	$154.15	$0	$76.80	$0
if D credit limit~=0	$1,181.65	$1,000	$1,985.00	$1,000
interest rate	16.13	14.99	16.60	17.20
D interest rate	-0.109	0	0.036	0
if D interest rate~=0	-1.012	-2.25	0.914	0.25

* The period analyzed is Aug.1998-Jul.2002
** Source: Gross and Souleles (2002), Table I. The period analyzed is Jan. 1995-Jan.1998

Panel B
Financial Characteristics
Comparison with the Survey of Consumer Finances

	My dataset		SCF	
	Mean	Median	Mean	Median
All	$3,408.36	$2,821	$4,100	$1,900
by Age				
Less than 35	$2,952	$2,312	$4,000	$2,000
35-44	$3,578	$2,967	$4,300	$2,000
45-54	$3,880	$3,202	$4,200	$2,300
55-64	$3,092	$2,463	$4,100	$1,900
65-74	$3,276	$2,810	$5,200	$1,000
older than 75	$3,720	$3,233	$1,900	$700
by Income Percentiles				
Less than 20	$3,039	$2,516	$2,100	$1,000
20-39.9	$3,285	$2,814	$2,800	$1,200
40-59.9	$3,820	$3,207	$3,700	$2,000
60-79.9	$3,551	$2,850	$4,700	$2,300
80-89.9	$3,215	$2,643	$7,200	$3,800
90-100	$3,836	$3,401	$6,600	$2,800
by Housing Status				
Home Owner	$3,559	$3,033	$4,500	$2,100
Renter	$3,405	$2,826	$3,400	$1,200