The Consumption Function Re-Estimated*

Jonathan A. Parker
Princeton University
August 1999

Incomplete and Preliminary

Abstract

This paper uses a structural model of household optimization to organize a non-parametric investigation of household consumption behavior. Using detailed household-level data from the Consumer Expenditure Survey, the level of consumption is estimated as a function of the household's state variables. Nonparametric estimation that corrects for mismeasurement of current liquid assets yields estimated consumption functions that are concave in current liquid assets, as implied by economic theory when consumption is not fully insured.

* Comments welcome. For helpful comments and discussions, I thank Pierre Olivier Gourinchas, Yuichi Kitamura, Grigori Kosenok, and seminar participants at the NBER summer institute EFACR, Princeton, and Wharton. Grigori Kosenok provided excellent research assistance.

Contents

1. Introduction

Since Hall (1978), consumer behavior and asset pricing have been evaluated mainly within the context of testing and estimation of the household inter-temporal first-order condition, or consumption Euler equation:[1]

$$u'(C_t) = \beta E_t \left[\widetilde{R}_t u'(C_{t+1}) \right] \tag{1.1}$$

This equation represents a necessary condition for the canonical structural model of household optimization to be true. As Hall (1978) pointed out the Euler equation of this model with rational expectations implies that (appropriately discounted) the marginal utility of consumption is a Martingale. Thus most studies of consumer behavior over the past 20 years have evaluated the theory by choosing a utility function (perhaps from a parametric class) so that the expectations errors— constructed by using realizations instead of expectations in equation (1.1)— are orthogonal to anything known to the consumer at time t. To the extent that this is possible, we can learn about consumer behavior from the estimated marginal utility function. To the extent that this is not possible, the theory, or its approximation, can be rejected.

Broadly speaking, several stylized facts have emerged from the estimation of consumption Euler equations. First, the simple, canonical model of a representative household is rejected by aggregate data.[2] While this could be due to the failure of this class of optimization problems to capture household behavior, there are other explanations. Preferences of the representative agent may be non-separable among a potentially large number of other activities such as labor supply, work at home, and labor market search to name a few. Alternatively, or additionally, households may in fact be more heterogeneous than the canonical model allows

[1] The notation is standard: $u(.)$ is an inter-temporally separable utility function; C is expenditures on nondurable goods; β is the discount factor; E_t is the expectations operator conditional on all information available in time period t and $\tilde{R}_t$ represents the gross return on investment in any asset between t and $t+1$.

[2] See for example Hansen and Singleton (1983) and Campbell and Mankiw (1989).

and thus aggregation across household is the cause of this rejection.[3] Finally, the assumptions on market completeness may be incorrect, so that aggregation consumption does not satisfy the simple aggregate Euler equation.[4] In order to distinguish among these alternative explanations, one must turn to the study of household behavior instead of aggregate data.

The second stylized fact is that household behavior is not consistent with insurance of risk that affects consumption across households or groups of households (Nelson (1994), Cochrane (1991), and Attanasio and Davis (1996)). This constitutes some evidence in favor of the proposition that the rejection of the aggregate Euler equation is due to the lack of complete markets.

Third, at the household level, where, due to measurement error in consumption, we are forced to test a linearized version of the consumption Euler equation, a simple model is too simple. By testing the Euler equation on household level data, aggregation across heterogeneous preference groups can be ruled out as an explanation for the rejection of the simple model. However, there is still significant debate as to the source of this rejection. Plausible candidate include heterogeneity of preferences (Attanasio and Weber (1995)), omitted constraints (Zeldes (1989a)), or failure of the linearization of the Euler equation to approximate actual behavior (Carroll (1997)). Given this state, current models of aggregate consumption typically proceed as if the simplest canonical model of a representative household were true.

Instead of estimating an Euler equation, this paper (re)turns to studying consumer behavior by estimating household consumption functions.[5] As in the old-style consumption function estimation that was built around Modigliani and Brumberg (1956) and Friedman (1957), and unlike modern Euler-equation testing, this requires a complete and, in practice, sparse specification of the agent's environment. That is, whereas the current literature must make assumptions on structural stabil-

[3] This line is argued by Attanasio and Weber (1993) and Vissing-Jorgensen (1998) for example.

[4] This line is argued by Caballero (1990) and Carroll (1992) for example.

[5] Gross (1994) takes a similar tack to examine firm's investment policy functions.

ity and observability, I must again adopt a large number of zero restrictions. There are however several advantages that make re-estimating consumption functions of interest.

First, unlike the old-style consumption functions, this paper can make use of the advances made during the past twenty years of studying consumption. The estimation is structured around a model built to be consistent with recent empirical findings and is evaluated in part with reference to recent characterizations of consumer behavior under uncertainty. Second, the consumption function is derived from a well-specified household decision problem, making explicit the exclusions that are made in the structural model. While the complete environment of the household must be specified, once it is specified, the consumption function can be used to completely understand consumer behavior rather than simply intertemporal substitution (the best that one can do with an Euler equation alone). For example, to address issues such as the impact of Social Security reform on household consumption and saving, one takes parameters estimated from Euler equations and places them into a completely specified structural model. If one is going to base one's recommendations on a fully-specified model, why not use this in estimation?

Third, the paper employs a large amount of household-level data. Fourth, estimation takes advantage of this large amount of data and recently developed econometric techniques for characterizing conditional expectations without imposition of functional form restrictions. Estimation of the consumption function is nonparametric and allows for measurement error in the key variable of interest. Conditional on the model, the estimated shape allows inference about important features of behavior that were hidden to earlier attempts by the linear approximations employed. Finally, this technique explores a relatively unexploited dimension of the data. The policy rules are estimated purely from repeated cross-sections on the level of household consumption, not from the growth rates of consumption that are typically studied by Euler equation approaches.

I find that, given a plausible but sparse model of the household environment, household consumption functions are qualitatively similar to those calculated numerically from a model in which precautionary saving is important at low asset levels. Consumption functions are found to be quite concave in liquid assets. One of the key points of debate and interest in the current study of consumption is the extent to which uncertainty invalidates the use of linear approximations to the consumption Euler equation. A key feature of uncertainty is that it creates a consumption function is concave in cash on hand (Carroll and Kimball (1996)). Estimates suggest that the concavity is significant in the range of the function that is important for many households.

Finally, the concavity implies that the marginal propensity to consume varies across the relevant population. Households with low levels of assets are estimated to have marginal propensities to consume out of cash on hand that are many times higher than those households with large reserves of liquid assets. For these households, expected future income is nearly irrelevant, and consumption decisions are made mostly based on current available resources. This result echoes that of Carroll (1994), which demonstrates that on average the level of consumption is closely related to current income and liquid assets rather than human wealth.[6]

Estimated normalized consumption functions are not very different across different groups of the population, suggesting that most heterogeneity in behavior is due to asset levels rather than behavior at a given level of assets. Consumption functions do not shift much in response to the aggregate state. Additionally, when normalized by a typical age-specific income, the choice of optimal consumption as a function of asset level does not change much over the life cycle. The one exception is that individual-level income does shift consumption functions significantly.

The remainder of the paper is structured as follows. The next section describes the formal model that is used to structure the investigation of consumption functions. The model is then solved at a given set of parameters and the main charac-

[6] The parallel is not direct since the current research normalizes by the unconditional expectation of income. Carroll (1994) can be thought of as estimating a linearized consumption function and thus estimating the average MPC out of various items across the population.

teristics of implied behavior briefly reviewed. The third section describes the data used in the investigation and uses them to move from theory to a stochastic model. The fourth section discusses the main estimation methodology. The functions are estimated by discretizing the well-measured regressors into states and approximating the function of cash on hand by a polynomial. Instruments are used to correct for the presence of measurement error in assets by applying a procedure developed by Hausman, Ichimura, Newey and Powell (1991). The fifth section presents both raw plots of the data and the main results of estimation. The final section will conclude.

2. The Consumption Problem and Optimal Behavior

Specification of an estimable model of consumer behavior involves two trade-offs.. First, there is a trade-off between realism and tractability. To the extent that household behavior is in reality very complex, a more complex model– that is one that includes more likely relevant state variables– will presumably fit the data more closely. However, given a certain amount of data, increased complexity of the model of household behavior comes at the cost of precise estimation of the impact of an given state variable. The model is chosen to allow the household problem to be impacted by family size, age, liquid assets, current income, the aggregate real interest rate and the growth rate of real GDP.

Second, there is a trade-off between restricting and estimating features of the consumption function. A completely specified household problem would deliver a functional form for the consumption function given any value for the set of parameters. The advantage of this approach is that the tight parametric restrictions involved lend power to the procedure if the model is correctly specified. The shortcoming of a complete description of the model is that it involves specification of features of the consumer problem about which very little is known and therefore estimation may be inconsistent. In the problem of optimal inter-temporal allocation of consumption, one would need to restrict the allowable utility functions, distribution of income shocks, and the value of assets at death at least to parametric classes in order to obtain identification or reasonably precise inference. In this paper, much of the household problem is specified, but the model leaves unspecified the distribution of income shocks, portfolio choice, adjustment costs on durable goods or labor, and the estimation requires only weak assumptions on the value of bequests at death.

2.1. An Organizing Model of Consumer Optimization

Each household (h) chooses at each age (t) an amount of nondurable expenditures ($C_{h,t}$) that provide utility through an intertemporally-separable, increasing, and concave utility function ($u(.)$):

$$Max_{\{C_{h,t}\}} \quad E_{\tau}\left[\sum_{t=s}^{T} \beta^{t-s} u(\digamma_{h,t} C_{h,t}) + \beta^{T+1-s} V_{T+1,\tau+T+1-s}(\digamma_{h,t} X_{h,T+1})\right]$$

where τ indexes time periods, E_{τ} is the expectations operator conditional on all information available at time τ; β is the discount factor; $\digamma$ is a family-size adjustment that normalizes consumption to per-capita terms, X_t is household cash-on-hand; and $V(\cdot)$ captures the possible value of cash on hand remaining at death. Utility from nondurable consumption is assumed to be additively separable from utility from durable consumption or leisure.

Households optimize subject to a budget constraint and given their current levels of assets and income:

$$X_{h,t+1} = \tilde{R}_{h,\tau}(X_{h,t} - C_{h,t} - C^{D}_{h,t,\tau}) + Y_{h,t+1,\tau+1} \tag{2.1}$$

$$Y_{h,s,\tau} \text{ and } X_{h,s} \text{ given.} \tag{2.2}$$

where $\tilde{R}_{h,\tau}$ is the stochastic gross real return on the household's optimal portfolio between τ and $\tau + 1.$, $Y_{h,t,\tau}$ is disposable non-asset income; and $C^{D}_{h,t,\tau}$ is expenditures, including adjustment costs, on durable goods. While again, this exercise seeks to impose minimal structure on the household problem, specification of the evolution of the family size adjustment, the income process, and the real interest rate process are required. Finally, the household faces either a no-Ponzi game constraint or is required to die with positive assets.[7]

A complete modelling of the dynamic process of family size changes in unfortunately beyond the scope of this paper. The process is discrete and unlikely to be simply a first-order Markov process. Thus, I assume that all households have common profiles of family size that are merely shifted up or down by the present family size. This allows the estimation of the choice of consumption per effective member rather than the addition of the necessary state variables to try to capture all expectations concerning the process.[8]

The disposable, non-asset income process represents two separate processes. First, during the working life, which is assumed identical and exogenous for everyone, $Y_{h,t,\tau}$ represents all non-asset income less all taxes and contributions to pensions, public and private. After retirement, the process represents all income, primarily of course that income received from the previous contributions to pensions. It is assumed that all retirement saving returns to the household during retirement as income in this form.

As to the actual process for household labor income, there is substantial uncertainty among labor economists as to how to model the process for individual labor income. I follow the evidence presented in Baker (1997) for labor income, and posit that the process for the log of household income is a first-order autoregressive process with autoregressive coefficient $0 \leq \rho < 1$ and household-specific

[7] The model could also be expanded to allow for liquidity contraints of the form $X_{h,t} - C_{h,t} \geq \Lambda_{h,t,\tau}$ $\forall t$, where Λ is a lower bound on assets and captures possible liquidity constraints.

[8] The addition of some state variables such as current family size to the problem can be attempted to provide a partial evaluation as to the importance of this assumption.

period-specific mean $\ln P_{h,t,\tau}$.[9]

$$\ln Y_{h,t+1,\tau+1} - \ln P_{h,t+1,\tau+1} = \rho(\ln Y_{h,t,\tau} - \ln P_{h,t,\tau}) + u^{H}_{h,t+1} \tag{2.3}$$

It is assumed that $u_{h,t}$, the household-level shock, is mean zero and uncorrelated across time and across individuals.[10] The mean in turn is assumed to consist of an aggregate component, proportional to the aggregate income process, and a household-specific component, modelled as specific to several subgroups of the population by education and age

$$\ln P_{h,t,\tau} = \ln Y^{A}_{\tau} + \ln \mu^{ed}_{t}$$

where $\ln Y^{A}_{\tau}$ is the aggregate income process, $\ln \mu^{ed}_{t}$ are education-group and age specific household shares of the aggregate income process that are mean zero in all periods.[11] Thus the average of the individual processes delivers the aggregate

[9] Baker (1997) directly compares the fit of the two main models of life-cycle earnings using a panel of 20 years of PSID data. After removing time effects, the paper finds that a stationary model for log income that incorporates individual-level heterogeneity in growth rates and means matches the data better than a model that incorporates a unit root and no heterogeneity in growth rates. Table 4, row 7, page 354 demonstrates that the model (in Baker's notation) $\ln Y_{i,t} = \alpha_t + \beta Exp_{i,t} + \gamma_t Exp^2_{i,t} + \gamma_i + \lambda_i Exp_{i,t} + \rho \ln Y_{i,t-1} + u_{i,t}$ where $u_{i,t}$ is white noise fits the data as well as any considered model (Exp is experience).

[10] u need not be stationary across ages. Although stationarity across time is required in the sense that young households always have the same shock process whether they were young in 1930 or 1950, they can both have different shock distributions from old households.

[11] Baker (1997) shows that individual heterogeneity in levels and growth rates cannot be fully captured by interacting individual education with the quadratic in experience and modelling the constant as a function of a set of characeristics similar to those employed here. This paper does limit the heterogeneity in growth rates and levels to groups, but expands the heterogeneity that is modelled by allowing greater flexibility in growth rates by age (the age-profile) and interacting education level with the age profile. However, the Baker (1997) findings and the fact that the object of interest here is actually household non-asset income, may imply that the degree of individual-level heterogeneity is still insufficient to capture fully that in the data. If this is the case, too much of a household's income will be allocate to transitory income with the result that the marginal propensities to consume out of transitory income will be overestimated.

process. Households then have household-specific, persistent shocks to income, a common or aggregate component to income, age-specific expected growth rates of income, and permanent cross sectional differences in income.[12] The household income process is modelled as a first-order Markov process in $\ln Y_{h,t,\tau} - \ln P_{h,t,\tau}$.

The aggregate state, denoted s, determines both the available investment opportunities and the growth rate of the aggregate income process. The income process is nonstationary and follows

$$\ln Y_{\tau+1}^{A} = \ln Y_{\tau}^{A} + u_{\tau+1}^{A}. \tag{2.4}$$

where the value of the shock, $u_{\tau+1}^{A}$ depends only upon the aggregate state, s. The aggregate state evolves as a first-order Markov process. The state will be taken to be a discrete summary of $\Delta \ln Y_{\tau}^{A}$ and R_{τ}, the risk-free real interest rate. The aggregate and household-level income Markov processes are independent. These assumption imply that, to forecast its income in deviation from $\ln P_{h,t,\tau+1}$, an agent needs only to know the aggregate state, its individual income state, $\ln Y_{h,t,\tau} - \ln P_{h,t,\tau}$, and its deterministic age-profile of income, $\ln \mu_{t}^{ed}$.

Two additional assumptions make the problem homogeneous in the unconditional expectation of income, $P_{h,t,\tau}$. First, the value function for bequests at the time of death is of the following form:

$$V_T(X_{h,t}) = \sum_{j=1}^{\infty} \psi_1^j u(\psi_2^j X_{h,t})$$

for some constants $\{\psi_i^j\}$. This assumption allows for a wider range of functions than (but nests) the strict life-cycle model ($\psi_i^j = 0$, $\forall i, j$) and the usual dynastic altruistic model which generates an infinite horizon problem. Secondly, the utility function is of the constant relative risk aversion form, so that the marginal utility function is homogeneous of degree, $-\frac{1}{\sigma}$ where σ is the intertemporal elasticity of

[12] So a household with a highly educated head has a high expected mean of its lifetime income process and a different typical profile; similarly a household with adults born in 1920 has a low mean lifetime income process.

substitution.[13]

These assumptions allow the entire problem to be reposed in terms of income, consumption, and cash on hand normalized by the group-specific mean in income, $P_{h,t,\tau}$, the average level of income for a given type of household at a given age. Defining $c_{h,t,\tau} = \frac{C_{h,t}}{P_{h,t,\tau}}$, $y_{h,t,\tau} = \frac{Y_{h,t}}{P_{h,t,\tau}}$, and $x_{h,t,\tau} = \frac{X_{h,t}}{P_{h,t,\tau}}$, the Euler equation for the risk-free rate and the constraints on household optimization problem as can be written:

$$\begin{aligned} u'(\digamma_{h,t} c_{h,t}) &= E_s\left[\beta R_\tau u'\left(\digamma_{h,t+1} c_{h,t+1} \frac{Y^A_{\tau+1}}{Y^A_\tau}\frac{\mu^{ed}_{t+1}}{\mu^{ed}_t}\right)\right] \\ x_{h,t+1,\tau+1} &= R_\tau \frac{Y^A_\tau}{Y^A_{\tau+1}}\frac{\mu^{ed}_t}{\mu^{ed}_{t+1}}\left(x_{h,t,\tau} - c_{h,t,\tau}\right) + y_{h,t+1,\tau+1} \\ y_{h,t+1,\tau+1} &= (y_{h,t,\tau})^\rho \exp[u^H_{h,t+1}] \end{aligned} \quad (2.5)$$

From here on, references to variables refer to their normalized counterparts unless otherwise noted, so that, for example, consumption refers to normalized consumption. Note that the "cohort effects" that one must be concerned with in making inference in a less-structural fashion are partly contained in this normalization by $\ln Y^A_{\tau+1}$. The remaining part of any cohort effect is absorbed into the household's characteristics such as assets and share of aggregate income. That is, a big transitory income boom experienced by one cohort and not another will lead to that cohort having a higher level of liquid assets today rather than be captured by the aggregate income term in P.

The solution to this problem is a policy function for consumption as a function of the (normalized) payoff-relevant state variables: cash on hand, income, age and the aggregate states:

$$\digamma_h c_h = F(x_h, t, y_h, ed_h, s). \quad (2.6)$$

[13] If there are liquidity constraints, households with higher means of income, $P_{h,t,\tau}$, can borrow proportionately more:

$$\Lambda_{h,t,\tau} = P_{h,t,\tau}\lambda.$$

For example, if there are liquidity constraints that completely restrict borrowing, then $\Lambda_{h,t,\tau} = \lambda = 0$.

where to simplify, the subscript, h, now simply denotes household-level data and the variables s stands for the aggregate state. The policy function of interest is thus a function of five state variables. The aggregate state, s, gives all the information necessary concerning the variables subscripted only τ; age and education give all the information necessary for all the variables subscripted only t. The only other necessary state variables, after normalization, are the household's current income state and its cash-on-hand.

2.2. Characteristics of Optimal Consumption Behavior

Except for a small number of special cases, such as no uncertainty, characterization of optimal behavior must proceed numerically. This subsection displays the policy rules that arises from the household optimization model just discussed coupled with additional structure and assumptions about parameter values. The simulated rules provide a characterization of the features of the model that are robust across many specifications. Similar models have been solved in the literature and the displayed policy rules are quite similar to those presented in Zeldes (1989b), Deaton (1991), Carroll (1997) and to those estimated in Gourinchas and Parker (1997).

Optimal behavior is characterized under the following assumptions. There is no bequest motive and households are required to die with positive assets. The discount factors is set to 0.95, the intertemporal elasticity of substitution is 0.5, the real interest rate is held constant at 3 percent, and there are no fluctuations in the aggregate state. The initial normalized income is 1 and grows in expectation at 1% per year until retirement when it falls to 60% of its previous level. The serial correlation of the log of household income is set to 0.4 and the standard deviation 0.06. The actual average family-size process is input with consumption normalized to a family size of 2.[14] Finally households live from age 25 to 85 and retire at 65.

Figure 1 displays the consumption function for a household at age 35.[15] On the

[14] The normalization proceeds by counting the first householder with weight 1, the second with weight 0.67 and all further members with weight 0.43.

[15] The value chosen is also that of a low income shock.

horizontal axis is normalized assets, and on the vertical axis normalized consumption. The are several important and robust characteristics of the function. The consumption function in the model is concave in current liquid wealth.[16] The concavity of the function is a feature that is due to uncertainty. Absent uncertainty, the household would be willing to borrow at times during the life, since it could be certain of paying back its debts and thus not risk an infinite marginal utility at or immediately before death. With uncertainty as assumed in the simulation (normally distributed shocks to log income) there is always some probability of an extremely bad sequence of income draws so that any household that borrows may find itself with zero consumption and infinite marginal utility of consumption near death. Thus households do not borrow when liquid assets near zero. Instead they consume nearly all their assets. Since, if another bad income realization occurs in the next period, they will be in the same situation again, consumption in the future is quite volatile. Thus it is technically the precautionary motive that causes the lack of borrowing early in life. Marginal utility is smoothed when assets are low by lowering consumption today and decreasing the uncertainty about future consumption until the standard Euler equation is satisfied.

To the extent that the consumption function is not linear, the marginal propensity to consume changes with the level of assets that the household has on hand. For low levels of assets, the marginal propensity to consume is high, and a transitory shock to income can raise consumption much more than simply by its annuity value. If only the annuity value of the income were consumed, the household would see a large decrease in the risk of low consumption in the next period and a corresponding decrease in its expected marginal utility of consumption that is larger than simply that caused by planning to consume the additional annuity value of the extra wealth.

Figures 2 displays the consumption functions for different ages. When interpreting these policy rules one has to keep in mind that the normalizing factor

[16] This is proved for all but three special cases mentioned later by Carroll and Kimball (1997).

on average grows over working life and declines at retirement. The consumption function is lowest for households at age 60, as they are at their highest earning years and retirement is near. At high asset levels, the oldest households have the highest consumption functions, since they are running down their assets as retirement nears. At low asset levels households of all ages consume roughly their liquid assets.

The household smooths marginal utility in the face of the expected profile of income (the normalization) by accumulating assets so that when the consumption function is at its lowest, assets are on average quite high. By late in life, in the simulations the households accumulate large amounts of liquid assets and the behavior of the typical agent is determined by the consumption function to the right of the graph and beyond. In most data on household assets these levels are not achieved by most households. And indeed this is a characteristic of the liquid wealth holding in the current data. It is important to note however, that at all ages the consumption function remains quite curved for those households unlucky enough not to accumulate many assets.

3. Data

In order to maximize the amount of available data, to have access to measures of the necessary state variables, and to minimize the amount of measurement error in consumption, which will represent the error term in estimation, I employ data from the Family, Member, and Detailed Expenditure files of the Consumer Expenditure Surveys(CEX) for the years 1980 to 1993.

3.1. The Consumer Expenditure Survey

The CEX is structured as a rotating panel of households containing information about consumption expenditures, demographics, income and assets, for a large sample of the U.S. population. The Survey is conducted by the Bureau of Labor Statistics in order to construct baskets of goods for use in the bases for the Consumer Price Index. The survey is known to have excellent coverage of consumption expenditures relative to other available surveys, to have noisy data on liquid assets, and to have income information of good quality.[17] The survey interviews about 5000 households each quarter. In a household's first interview, the CEX procedures are explained to them and information is collected so that they can be assigned a population weight. They are then interviewed four more times (once every three months) about detailed consumption expenditures over the previous three months. In interviews two and five, the survey asks the amount of income received during previous 12 months. In the fifth interview, each household is asked about its holdings of four categories of liquid assets and how much these holdings have changed over the past 12 months. Families rotate through the process, so that about 25% of households leave and are replaced in each quarter. About half of all households make it through all the interviews.

The observation period is six months. Each household contributes two data points to the sample based on information collected in its first two and then final two interviews. Any households which are classified as incomplete income reporters or which have any of the crucial variables missing are dropped. The main measure of (un-normalized) consumption employed is nondurable consumption constructed as the sum of expenditures on the following categories of goods across either the first two or last two interviews: food, excluding food as pay and school meals; alcohol; house-furnishings and equipment excluding furniture, major appliances, and floor coverings; apparel and services; transportation excluding new and used vehicle spending and financing; entertainment; personal care; reading; and tobacco and smoking. Due to some extremely unreasonable observations, primarily at the lower tail, the top and bottom one percent of the normalized consumption observations are set to missing and thus the observations dropped from the sample.

Income before retirement is after-tax family income less Social Security taxes,

[17] Income, while reasonably well measured, is not measured with the accuracy of the Surveys that focus explicitly on income, such as the PSID. See Lusardi (1996), Attanasio (1994), and Branch (1994).

(after-tax) asset income and pension contributions. Income measures for the first observation for each household are taken from the first interview and for the second, from the last interview. Income after retirement is total income less only after-tax interest on checking and savings accounts. Liquid assets from the Survey are the sum of the value of all four measured categories of assets: savings accounts; checking, brokerage and similar accounts; U.S. savings bonds; stocks, bonds, mutual funds and other such securities. These are considered missing if one or both of checking and savings account balances are missing. Note that this definition excludes illiquid wealth and that each measure of assets is topcoded at $\$100,000$. Top coded assets are dropped.[18] Cash on hand is constructed as in equation (2.1). Thus to the measure of liquid assets from the Survey is added six months of income for the first household observation and six months of total consumption for the second household observation.

Households are assigned to education cells on the basis of the male reference person, when one exists, otherwise assignment employs the female reference person.[19] The education cells are three: high school degree or less education, more than a high school degree but less than a college degree, and a college or advanced degree. Age is assigned on the basis of the average age of the reference person and spouse, or just that of the reference person if there is no spouse. All households under age 25 are dropped.

Finally, the fact that liquid assets are mismeasured implies that it is necessary to have instruments to perform consistent estimation. As variables which are related to the noisy measure of liquid assets, I employ: a dummy variable for whether the household has positive financial income, the actual financial income with zeros set to an arbitrarily small positive value, a similar pair of variables for interest

[18] As discussed in section (5.2), estimation will focus on a compact range of normalized assets between 0.07 (the first percentile of the distribution) and 3.

[19] The reference person in the CEX is defined as the person who is most responsible for paying the bills. This person is then interviewed, as he or she is probably the most knowledgable about household expenditures.

income. Additionally, unnormalized household income and the GDP growth rate are used as an instruments. Finally, to help predict the normalization factor, I employ a quadratic in age. As will become clear in the methodology section, the instruments are employed separately for each "state" and so are interacted with the well-measured independent variables.

All data is made real by deflating it using the consumer price index (CPI-U) for nondurables. There are $60,459$ observations on $38,532$ households with the required data after trimming the consumption data. The range of normalized liquid assets is from a high of 72 to a low of 0.000177. Normalized consumption runs from a high of 1.58 to a low of 0.107.

3.2. The Data Generating Process

In order to map from the data to the model, the model must be augmented to a data generating process by adding stochastic deviations. In this case the deviations between the model and the data are viewed as coming from mismeasurement in consumption and assets. While the CEX has perhaps the best measures of consumption available in household-level datasets, there is still significant error in the measurement of consumption. Additionally, while the CEX has all the necessary variables for estimation of the household consumption function, there is significant error in the measurement of assets.

Let variables superscripted with * denote true, unobserved values. One does not observe un-normalized consumption and assets, C_h^* and X_h^*, but rather their poorly measured counterparts in the data, C_h and X_h, which are assumed to be related to the truth by a mean-zero error independent of the truth in the following manner:

$$\ln C_h = \ln C_h^* + \omega_h^c; \;\; \ln X_h = \ln X_h^* + \omega_h^x. \tag{3.1}$$

In addition, as discussed subsequently, the process of normalization will introduce additional error so that

$$\ln c_h = \ln c_h^* + \varepsilon_h; \; \ln x_h = \ln x_h^* + \eta_h. \tag{3.2}$$

where $\eta_h \equiv \omega_h^P + \omega_h^x$, and $\varepsilon_h \equiv \omega_h^P + \omega_h^c$ and ω_h^P is the measurement error associated with normalization by the permanent component of income. These measurement errors will not be uncorrelated due both the normalization and due to the fact that, for observations constructed from the last interview information, total consumption is used to construct cash on hand.[20]

The data generating process, or stochastic model, is:

$$\ln c_h = H(\ln x_h^*, t, y_h, ed_h, s) - \ln \digamma_h + \varepsilon_h \tag{3.3}$$

where

$$H(\ln x_h^*, t, y_h, ed_h, s) \equiv \ln\left[F(\exp[\ln x_h^*], t, y_h, ed_h, s)\right]$$

and

$$E[\varepsilon_h | \ln x_h^*, t, y_h, ed_h, s] = 0.$$

As discussed in section 3.1, the dataset also gives additional indicators of the asset level of an individual household. These available instruments, denoted $q_{h,}$ are assumed to be linearly related to the true measure of assets as:

$$\ln x_h^* = q_h' \alpha + v_h = w_h + v_h \tag{3.4}$$

where w_h denotes the optimal linear combination of instruments and v_h is mean zero and independent of the instruments.[21] As is standard, the instruments are assumed to be uncorrelated with the measurement error (including normalization error) in consumption and liquid assets: $E[\varepsilon_h|q_h] = E[\eta_h|q_h] = 0.$[22]

[20] While normalization error is also introduced into the normalized measure of income, since the present analysis discretizes the income state space and since income begins significantly better measured, the present approach assumes that the current state of normalized income is well measured.

[21] Note that the assumption of independence is much stronger than the assumption of zero correlation that is usually invoked for two-stage least squares estimation.

[22] Normalization induces correlation between the measurement error terms so that I allow $E[\eta_h \varepsilon_h|q_h] = \sigma_{\eta\varepsilon}$ and $E[v_h \varepsilon_h|q_h] = \sigma_{v\varepsilon}$.

4. Estimation Technique

In order to estimate consistently the consumption function we must consider three issues. First, the appropriate variables must be normalized. Second, the state-space for the Markov processes must be constructed. Third, because assets are observed only with error, that is one observes $\ln x_h$ instead of $\ln x_h^*$, there is measurement error on the right hand side of equation (3.3) in a variable inside the function of interest. This paper adapts a technique derived in Hausman et al. (1991) that consistently estimates a polynomial function of a single mismeasured regressor. This procedure is extended to estimate to within a small bound of accuracy a non-parametric function.

4.1. Normalization

The income process, equation (2.3) can be rewritten

$$\ln Y_h = \ln P_h + \frac{u_h}{(1-\rho L)} \tag{4.1}$$

$$= \ln Y_\tau^A + \ln \mu_t^{ed} + \frac{u_h}{(1-\rho L)} \tag{4.2}$$

where L is the lag operator. $\ln \mu_t^{ed}$ consists of for each education group, a fifth-order polynomial in age if 65 or younger, a second order polynomial in age if older than 65, a retirement dummy variable, and a retirement dummy variables interacted with a dummy variable for being age $55-65$. Rather than use the actual log of U.S. GDP or disposable income, a complete set of month and year dummies are used to normalize more completely by the aggregate state. These are also interacted with education group.

Equation (4.1) is estimated on all households with the dependent variable being the household disposable income, defined above. The normalization is run separately on each educational attainment group. Given the set of plausible assumptions made about the income process, the fitted value of equation (4.1) upon cross-sectional data gives consistent estimation of the values on of $\ln P_h$.[23] Then

[23] In fact, there are two consecutive observations for many households, so that there will be

$\widehat{P_h} \equiv \exp[\widehat{\ln Y_h}]$ and normalized variables are constructed as

$$c_h \equiv \frac{C_h}{\widehat{P_h}} \text{ and } x_h \equiv \frac{X_h}{\widehat{P_h}}.$$

The measurement error from normalization, denoted $\omega_{h,t}^P$ in section (3.2) is defined as $\ln \widehat{P_h} - \ln P_h$.

4.2. Discretization of the Empirical State Space

The next required step is to discretize the state-space of the well-measured regressors. This is done as follows. The space of normalized income is divided into discrete groupings so that, y_h, the current income state, is constructed by splitting $\frac{Y_h}{\widehat{P_h}}$ into a grid that has both even spacing and some parity in counts in each cell. Normalized income is split into a 7 state Markov process in logs that is symmetric and has roughly equal bins. Age is split into four groups defined as: $25-45$, $45-55$, $55-65$ and not retired, and 65 or retired and older than 55. Education is the three educational grouping discussed above. Finally, the aggregate state is dependent on the growth rate of GDP growth, measured in real 1992 (chain weighted) prices, and the real interest rate, measured as the return on 3-month treasury bills less the inflation rate from the CPI for nondurables. There are four possible aggregate state corresponding to GDP growth of less than or greater than 2.5%, and a real interest rate of less or greater than 3%.

There is no need to discretize the family size adjustment, $\ln \digamma_h$.

Given that the state-space is now discrete, the function of interest can be estimated by estimating $H(\ln x_h^*, t, y_h, ed_h, s) - \ln \digamma_h$ for each set of states separately and transforming it to recover the function of interest.

4.3. Nonparametric Estimation with Measurement Error

This subsection first lays out the problems that measurement error presents for nonlinear regression even with the availability of instruments. Next, the Hausman

correlation between each household's observations. This does not alter the consistency of the procedure as the number of households goes to infinity.

et al. (1991) solution for a problem with a known-order, polynomial approximation is presented. Finally I discuss inference when the function of interest is nonparametric.

4.3.1. The Measurement Error Problem

Consider first a parameterization of the unknown function of interest $H(\ln x_h|S,\Phi)$ where Φ is a finite parameter vector of known order that is to be estimated and S summarizes the states of age, income, education, and aggregate economy. For now the linear term, $\ln \digamma_h$, is ignored. Without measurement error, estimation could proceed by a non-linear estimation technique such as the generalize method of moments. However in the presence of measurement error this option disappears. The situation can be expressed as

$$\ln c_h = H(\ln x_h|S,\Phi) + \varepsilon_h - \Xi_h \tag{4.3}$$

where $\Xi_{h,t}$ is the additional right-hand-side term due to mismeasurement of liquid assets.

$$\begin{aligned}
\Xi_h &\equiv H(\ln x_h|S,\Phi) - H(\ln x_h^*|S) && (4.4)\\
&\cong \frac{\partial H(\ln x_h^*|S,\Phi)}{\partial \ln x_h^*}\ln\left(\frac{x_h}{x_h^*}\right) && (4.5)\\
&= \frac{\partial H(\ln x_h^*|S,\Phi)}{\partial \ln x_h^*}\eta_h && (4.6)
\end{aligned}$$

where the approximation comes from a first-order Taylor approximation of $H(\ln x_h|S,\Phi)$ around $\ln x_h^*$ and the last equality from the definition of the measurement error in assets. Since the measurement error η_h appears in both the error term of equation (4.3), Ξ_h, and in the regressor assets, $\ln x_h$,

$$E[\varepsilon_h - \Xi_h|\ln x_h, S] \neq 0,$$

so that Non-Linear Least Squares (NLLS) or Generalized Method of Moments (GMM) estimation based on $E[H(\ln x_h|S,\Phi)(\varepsilon_h - \Xi_h)] = 0$ is inconsistent.[24]

[24] The statement that the expectation is conditional upon s is a slight abuse of notation.

Further, however, even with instruments, the problem seems intractable by standard procedures such as Non-Linear Two-Stage Least Squares or GMM based on the orthogonality of the instruments and the residual in equation (4.3), $E[w_h(\varepsilon_h - \Xi_h)] = 0$. Unless $H(\ln x_h|S,\Phi)$ is linear in assets, Ξ_h is also a function of $\ln x_h^*$ which, if the instruments are legitimate ($\alpha \neq 0$), is correlated with the instruments so that

$$E[\varepsilon_h - \Xi_h|q_h, S] \neq 0.$$

Note that the only case in which the function of interest can be estimated consistently is if it is linear in $\ln x_h^*$, so that the first-derivative in equation (4.5) is not a function of $\ln x_h^*$ and the necessary orthogonality condition is met. But this is just (linear) Two-Stage Least Squares (TSLS).

4.3.2. Series Estimation with Measurement Error

Suppose that the function of interest were a power series polynomial of finite and known dimension. Writing the consumption function for a given state as such a polynomial,

$$H(\ln x_h^*|S) = \sum_{j=0}^{K} \beta_j(S)(\ln x_h^*)^j \tag{4.7}$$

consider estimation of $\{\beta_j(S)\}_{j=1}^{K}$ given that the independent variable is mismeasured.

Hausman et al. (1991) (HNIP) propose a consistent estimator for this situation as follows. While $\ln x_{h,t}^*$ is only observed with error, so that the orthogonality condition for equation (3.3) still fail, the orthogonality conditions are met for the following pair of polynomial regressions employing the optimal instrument, w_h:

$$\begin{aligned} \ln c_h &= \sum_{j=0}^{K} \gamma_j(S)(w_h)^j - \ln \digamma_h + e_h^c \qquad (4.8) \\ (\ln x_h \ln c_h) &= \sum_{j=0}^{K+1} \delta_j(S)(w_h)^j - w_h \ln \digamma_h + e_h^{xc} \end{aligned}$$

Thus one can consistently estimate $\gamma_j(S)$ and $\delta_j(S)$ by simple OLS. One can then recover the parameters of interest by deriving the relationship between $\beta_j(S)$ and these estimated parameters. Substituting equation (3.4) and (4.7) into the estimating equation (3.3) yields

$$\begin{aligned}\ln c_h &= \sum_{j=0}^{K} \beta_j(S)(w_h + v_h)^j - \ln \digamma_h + \varepsilon_h \\ &= \sum_{j=0}^{K} (w_h)^j \left[\sum_{l=j}^{K} \binom{l}{j} \beta_l(S)(v_h)^{l-j} \right] - \ln \digamma_h + \varepsilon_h.\end{aligned}$$

This is simply equation (4.8) if one defines

$$\gamma_j(S) = \left[\sum_{l=j}^{K} \binom{l}{j} \beta_l(S)(v_h)^{l-j} \right]. \tag{4.9}$$

Taking expectation of both sides of equation (4.9) gives a set of $K+1$ relationships among the $K+1$ γ's and the K moments of v_h and the $K+1$ structural parameters $\beta_j(S)$. A similar methodology can be used to derive a set of relationships among the $\delta_j(S)$, the $\beta_j(S)$ and again the nuisance parameters, $E[(v_h)^{l-j}]$. In fact, using equation (4.9) and its partner, one can not only recover the parameters of interest, but they are overidentified.

The HNIP estimation procedure proceeds in three steps. First, one estimates α and its covariance matrix by OLS. Second, constructing the optimal instrument, one recovers the reduced form parameters $\gamma_j(S)$ and $\delta_j(S)$ consistently from estimation of equation 4.8. The covariance matrix of the reduced form parameters is recovered by interpreting these first two stages as a two-step GMM estimator. Finally, one recovers the structural parameters by minimum chi-squared procedure since the dimension of the reduced form parameter set is larger than that of the structural parameters to be estimated. The use of minimum chi-squared also provides a mapping from the reduced form covariance matrix to that of the structural parameters.[25][26]

[25] Complete details of estimation, inference and testing are in HNIP, with the following exceptions. [Formulae corrections here.]

[26] Not yet employed: Since the model is overidentified, the procedure provides a test of the model. To the extent that the larger set of parameters estimated in the reduced form do not lie to far outside the closest point in the subspace spanned by the possible strucutral parameters the model is accepted.

4.3.3. Nonparametric Estimation with Measurement Error

While one tack to estimating the function of interest would be simply to choose to consistently estimate an N-th order approximation to the true function of interest and follow HNIP exactly, I instead use a data-dependent order-selection criterion and choose to estimate the function consistently to within a fixed accuracy, ϵ.[27]

First, provided convergence to the true function is uniform, it will be the case that there exists a finite-order approximation that comes within ϵ of the true function at every point over a compact range. Second, this order can be estimated consistently by appropriately reducing the residuals from the estimation of equations (4.8), denoted $\hat{e}_h^i$, as

$$\tilde{e}_h^c \equiv \frac{\hat{e}_h^c}{|\hat{e}_h^c|} Max\{|\hat{e}_h^c| - \epsilon, 0\}, \ \tilde{e}_h^{x,c} \equiv \frac{\hat{e}_h^{x,c}}{|\hat{e}_h^{x,c}|} Max\{|\hat{e}_h^{x,c}| - \epsilon, 0\}. \tag{4.10}$$

The Bayesian Information Criterion provides a consistent estimate of the true model by the usual procedure given the adjusted residuals. The chosen order, K, for polynomial approximation to the function of interest minimizes

$$BIC(K) \equiv \left| \begin{array}{cc} \frac{1}{N}\sum(\tilde{e}_h^c)^2 & \frac{1}{N}\sum \tilde{e}_h^c \tilde{e}_h^{x,c} \\ \frac{1}{N}\sum \tilde{e}_h^c \tilde{e}_h^{x,c} & \frac{1}{N}\sum(\tilde{e}_h^{x,c})^2 \end{array} \right| + 2K\frac{\ln N}{N}$$

where N is the number of observations.[28]

Given a consistent estimate of the model order, inference upon the parameters of interest at the chosen model, is independent of model selection. Thus the HNIP procedure can be applied with the one modification that the estimate is actually

[27] Using different approaches, Gourinchas (1997) and Hausman, Newey and Powell (1995) extend the HNIP procedure for the case where one has a single second measure of the mismeasured regressor and no other covariates to estimation of a nonparametric function and of a finitely parameterized model repectively.

[28] Alternatively, one could employ a data-independent order selection as worked out for the standard series estimation in Andrews (1991) and Newey (1997). Given the correct growth rate of the order of the polynomial in relation to the sample size, the estimated function is consistent for the unknown function when the true polynomial order is infinite. As shown in Newey (1997), the rate at which the order of the polynomials in 4.8 can grow is bounded by the requirement that $\frac{K^3}{N} \to 0$.

a narrow band rather than a line and the standard errors must be increased by ϵ. In practice $\epsilon = 0.001$.

5. The Consumption Function

5.1. Univariate Kernel Regressions of Consumption on Assets

Before turning to estimation of the function of interest, I estimate several Kernel regressions on the normalized data, ignoring a subset of the covariates. These results complement the subsequent findings by showing that similar results are obtained using a local smoother and without the unconventional correction for measurement error that the dual regressions in equation (4.8) represent.

To proceed, the consumption data is normalized by a family size adjustment. Consumption is divided by a weighted family size in which the first householder counts with weight 1, the second with a weight of 0.67, and all further members with weight 0.43.[29] In the subsequent estimation that corrects for measurement error, this adjustment is not used to divide consumption but instead is included in estimation as a linear regressor to capture $\ln \mathcal{F}_h$.

Figure 3 displays a simple centered moving average of the data which includes one percent of the data.[30] The range of assets plotted runs from a high of 3 to a low of the first percentile of the asset distribution. The bandwidth is kept narrow as a simple way to make clear that bias from oversmoothing is not generating the picture and also give some sense as to the sampling uncertainty involved.[31] The top panel employs the log of normalized assets, as will be employed in the regressions that corrects for measurement error, while the lower panel employs normalized assets in levels, as will be shown since it is the function of interest. In the levels

[29] This is nearly the OECD adult equivalent scale, which differs only in requiring that the second member be an adult in order to receive the higher weight.

[30] A tricubic Kernel gives similar results so the most transparent Kernel and weighting is employed.

[31] Sampling uncertainty will be given a more full treatment in the next section.

case there is a clear concave function visible that looks quite like the consumption function from the simulated model. In the logs, there is also a large range of the graph for which the estimated function appears concave in cash on hand.[32]

However, it is quite possible that the lack of covariates is what is generating the results. That is, Figure 3 presents $\widehat{F(x, \bar{t}(x), \bar{y}(x), \overline{ed}(x), \bar{s}(x))}$, the way consumption changes as one varies liquid assets and allows the other state varies to covary with liquid assets. If for example, high income households have high liquid assets today and a high expected income in the future, then the slope of the "consumption function" would be overestimated by the present method. The marginal propensity to consume does seem to be a quite high one-third for low levels of liquid assets. A slightly more complicated story could also generate concavity where household level consumption functions had none. Another problem, of course, is that of mismeasurement of assets, although this would seem to bias away from finding a high marginal propensity to consume and it is less clear how this bias might generate concavity.

To partially address the issue of covariates, and again largely as a way of characterizing the raw data rather than of making structural inference, Figures 4 through 8 present consumption as a function of liquid assets across each state space individually.[33] Controlling for any one of these covariates individually does not change the shape of the consumption rule that we observe. Looking at the figures in turn, we can also make a tentative study of the way behavior changes with the considered covariate. Controlling only for age of the household, consumption functions look quite similar, however there is evidence of an inverse relationship between age and consumption as a function of assets (adjusted for family size and normalized by income). Also, it is clear that higher income households have higher consumption

[32] There is little change in these figures if one drops all observations in which total consumption is used to calculate liquid assets.

[33] Bandwidths are adjusted for each graph. Figure 4 has a bandwidth of 5 percent of the sample, while Figure 5 has a bandwidth of 10 percent, Figure 6 has a bandwidth of 5 percent, and Figures 7 and 8 have bandwidths of 8 percent.

given any level of liquid assets. This is quite consistent with households expecting higher income in the future and so being willing to consume more today. Higher educated households consume more at any given level of liquid assets. Finally, looking separately across four real interest rate levels and four real GDP growth rate, the aggregate current state does not affect consumption functions much at all.

While a useful way to examine the correlations in the data, these figures do not merit a discussion of what one might conclude about household behavior. Instead, I now turn to estimation to which more weight can be given.

5.2. The Distribution of Assets

Before estimating the consumption function, one must choose a compact range over which to estimate the nonparametric function of interest. Figure 9 plots a density estimate of the normalized liquid asset data.[34] The data are not weighed, so that this does not represent the population distribution in the U.S., but rather where the data lie that will be used to identify the function of interest. The top panel present the density of the raw data in logs. Clearly there are large areas of the observed data over which the density is extremely sparse. The compact range that is chosen for estimation is shown in the bottom two panels, first in logs and then in levels. The bottom one percent of the data are dropped and then normalized asset levels above three are also dropped. The latter choice is motivated mostly by the observation that the theory suggests that, if there is curvature in the function of interest, it will lie in the low end of the asset distribution. Given the shape of the theoretical consumption functions presented in Figures 1 and 2, a polynomial approximation will have great difficulty fitting a straight line for high asset levels and a highly curved function for low asset levels. Thus I focus on low asset levels at

[34] The Kernel estimate uses an Epanechnikov kernel ($\widehat{f_K} = \frac{1}{Hb}\sum_{h=1}^{H}\frac{3}{4\sqrt{5}}(1-\frac{1}{5}z^2)$) and the bandwidth, b, is chosen as $\frac{0.9}{H^{1/5}}Min\{\sigma_x, \frac{\iota_x}{1.349}\}$ where σ_x is the variance of x and ι_x is the interquartile range of x.

the cost of some data. The bottom two panels of Figure 9, also show what regions of the asset variable are most important for the results.

Figure 9 is interesting in its own right however. While the asset variables are mismeasured, the extent of mismeasurement is not insufficient to allow some inference to be made regarding the true distribution of assets. The mismeasurement however will increase the spread of the distribution of assets, and indeed it is unrealistically high in the top panel. The graphs show that most households hold a level of liquid assets which is less than their "typical" annual income. The distribution is also highly right skewed in levels, but not far from symmetric in logs. Estimates of wealth that include pension wealth, housing wealth, and Social Security wealth tend to generate much larger amounts of wealth, especially at retirement when the typical household has a ratio of wealth to income of roughly 10.[35] In the present data, since most of the wealth items discussed are assumed to be captured by the income profile and thus by age and income, the asset levels are correctly much lower.[36] Also, the present data have higher wealth distributions for older households than for younger, although the difference is smaller in the normalized data since the normalizing factor depends on age and rises over most of the working life.

Although as noted, the distribution is mismeasured and unweighted, the top panel of Figure 9 gives a crude impression as to what range of the consumption functions are most relevant for behavior. Were the data to show that most households held large amounts of liquid wealth, then consumption functions that looked like those from the simulations would demonstrate that curvature in the consumption function is of little practical import for behavior. The fact that many households have normalized liquid assets that are less that their expected annual income (unconditional on their income history) suggest that if consumption functions from the data look like those from the simulation, precautionary saving is most relevant

[35] This particular estimate is based on Venti and Wise (1993).

[36] The Survey of Consumer Finances suggests that the ratio is around 2 (see for example Venti and Wise (1993)), which is still quite a bit larger than in the present data.

for consumption behavior.

5.3. Consumption Functions

Figure 10 displays the average of consumption functions across the cells of the state-space, that is $\overline{F(x)} = \frac{H}{H(y,t,s)} \sum_{\{y,t\}} \widehat{F}(x, y, t, s)$. Any cell with less than 100 observations is not used in estimation and the polynomial approximation that the procedure selects is of order three. The averaging weights the function for each state with a weight proportional to the number of observations in that state. The top panel displays the average consumption function the middle panel its first derivative, and the last panel its second derivative. All graphs display 95 percent confidence intervals which are recovered by bootstrapping from the parameters and asymptotic distribution of the estimated function $\widehat{H}(\ln x_h^*, y_h, t, s)$.

The estimated policy function is quite similar to that recovered from simulation from the theoretical model in section 2.2. The estimated average consumption function is increasing over its entire range, with a slope running from a high of over one half at the low asset levels to about 0.02 at an asset level of 2.5. It is statistically significantly different from zero at asset levels below 2. The marginal propensity to consume at high asset levels is surprisingly consistent, although perhaps a little low, for what one would consider a reasonable marginal propensity to consume in a model without any precautionary saving effect. That is, in an infinite horizon model with log utility and no uncertainty, the marginal propensity to consume out of assets is $\frac{R-1}{R}$. Estimating the marginal propensity to consume could also be done using Euler equation methods for high-asset households under the assumption that the precautionary term for these high-asset households can be ignored. Indeed this finding validates that approach.

The more interesting finding is that the average marginal propensity to consume for $x \in [0.15, 0.5]$ is around 0.4 and quite precisely estimated. At low asset levels, households are estimated to spend significantly more out of cash on hand than out of permanent income or total wealth. A lump-sum tax cut for such households would lead to significant increases in spending even if it were expected and even if it were accompanied by future lump-sum tax increases. In the present data, and indeed other sources of information on the asset distribution of households, most young households[37] and a non-trivial portion of older households have normalized assets below 0.5.

Consumption as a function of liquid assets is concave in theory in the presence of uncertainty except in three special cases (Carroll and Kimball (1996)): if preferences are quadratic; second if the only source of uncertainty is in the return to savings and the household has CRRA preferences; and finally if the only source of uncertainty is in income and the household has CARA preferences. The bottom panel of Figure 10 tests the prediction of most LCH/PIH models under uncertainty by testing whether the consumption function is concave. The second derivative of the consumption function is everywhere negative and a two standard deviation band around the second derivative does not include zero over the range of the estimated function. This provides significant confirmation that the household consumption function is indeed concave.

Figure 11 displays the consumption function for households age $55-65$ (the third state), with income between 25 and 50 percent below the unconditional mean (the second state), with high education (the third state) and when the aggregate state such that GDP growth is high and the real interest rate low (the third state).[38] That is, this is a picture of one consumption function of interest, $\widehat{F}(x,3,2,3,3)$. The same set of conclusions can be drawn with the important exception that the inference is not as precise. Importantly, the concavity is no longer significant, and is estimated to be much more limited. There are also cells for which the concavity

[37] Including mine.

[38] Additionally, neighboring groups are included but downweighted by a factor of four. Cells that are neighboring along more than one dimension are downweighted by a factor of four raised to the dimension that they are neighboring along. Thus retired households with state 4 incomes, and in the employed aggregate state and education group are downweighted by 16 while similar young households with state 3 incomes are only downweighted by a factor of 4. Young households are not included irregardless of their other states.

is more economically significant than that displayed for the average. There are also cells in the data for which the consumption functions are much worse.[39] Typically "much worse" means one of two things. For some cells the chosen polynomial order is quite high and the estimated function extremely volatile and unreasonable. These are almost exclusively cells for which there is not much data and can be eliminated by increasing the number observations required to do estimation on a cell. The second way in which the functions can be "much worse" is that they can decline at higher asset levels. This decline is not statistically different from zero, in all cases examined. Further, this unpleasant feature is a symptom of the curvature at low assets and the fact that the estimation procedure is a global smoother. Re-estimating the consumption function for such a cell using only asset levels greater than unity often does not replicate this unpleasant feature of the estimated function, confirming the suspicion that it was due to the nature of the data at low asset levels.[40]

Figures 12 through 15 display average consumption functions for different state variables set at different levels. At every state, the average consumption function is quite curved. Only across education groups do the consumption functions differ much across states when households have few assets. Indeed, only across education and income groups are there much differences in consumption functions even at higher asset levels. At high asset levels there is in general more heterogeneity across states, however interpretation must be limited by the fact that the polynomial estimator may impose features at high asset levels that are driven in part by the denser data at low asset levels.

Figure 12 shows consumption functions by age. Unlike in the kernel estimates, it is not at all clear that young households consume more of their liquid assets than

[39] In "much worse" cells the chosen order of the polynomial is 7 or higher and the estimated function highly nonlinear to the extent of being nonsense. These cells can largely be eliminated by requiring a greater number of observations in each cell.

[40] Sometimes it is precisely the lack of data at high asset levels that creates a problem and here estimation on the bottom 90 percent of the data generally eliminates the decline.

the old. This is only barely true only at low asset levels. Unlike in the simulations, the age $55 - 65$ households consume a large fraction of their income. At higher asset levels, as in the simulations, the oldest groups are the ones that have the highest consumption functions. Thus, one sees the elderly running down their high asset levels but being somewhat frugal with their liquid assets relative to the young at low asset levels.

Estimated consumption functions depend a lot on the current household income state. Figure 13 shows the consumption functions by income for the first, third, fifth and seventh income states: $\{y < -0.50; -0.25 \leq y < -0.10; 10 \leq y < 0.25, 0.50 \leq y\}$. Consumption is highest for the highest income state and lowest for the lowest income state. Households with higher income expect higher income in the near future with high confidence.

Figure 14 displays how the estimated consumption functions respond to the aggregate state. There is little impact of the aggregate state on consumption at low asset levels, however the consumption functions are lower at any GDP growth rate for high real interest rates. At high asset levels the pattern is reversed: the consumption functions are higher for any GDP growth rate for high real interest rates. This is puzzling but again may in part be due to the dense data at low asset levels driving inference about behavior at high asset levels. In any case this is consistent with low elasticities of consumption to the real interest rate as estimated in Euler equation approaches. With respect to GDP growth, the relative levels also make little sense. There is no tendency to estimate a higher consumption level when GDP growth is higher. This is most likely due to the fact that for the household and in the CEX data, the impact of GDP growth on household income growth is small.

Finally, figure 15 displays the consumption functions for different education groups. The more educated households have lower consumption functions are medium to high levels of assets, consistent with higher education groups saving more and/or with their consuming a lower share of nondurables out of total consumption. At low asset levels however the pattern is reversed, consistent with high education households having better access to credit. Alternatively they may be able to access some of the wealth that has been modelled formally as illiquid, particularly housing wealth.

In sum, the figures that display consumption functions for different individual states, averaged across other states consistently show increasing and concave consumption as a function of liquid assets.

5.4. Impulse Response Functions

To be done.

6. Conclusion and Implications

To be written.

References

Andrews, Donald W. K., "Asmptotic Normality of Series Estimators for Nonparmaetric and Semiparametric Regression Models," *Econometrica*, Mar 1991, *59* (2), 307–345.

Attanasio, Orazio, "Personal Saving in the US," in James M. Poterba, ed., *International Comparisons of Household Saving*, University of Chicago Press, 1994.

__ **and Guglielmo Weber**, "Consumption Growth, the Interest Rate and Aggregation," *Review of Economic Studies*, 1993, *60*, 631–49.

__ **and** __ , "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, Dec 1995, *103* (6), 1121–57.

__ **and Steven J. Davis**, "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy*, Dec 1996, *104* (6), 1227–62.

Baker, Michael, "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 1997, *15* (2), 338–75.

Branch, E. Raphael, "The Consumer Expenditure Survey: a comparative analysis," *Monthly Labor Review*, Dec 1994, pp. 47–55.

Caballero, Ricardo J., "Consumption Puzzles and Precautionary Saving," *Journal of Monetary Economics*, 1990, *25*, 113–36.

Campbell, John Y. and N. Gregory Mankiw, "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," in Oliver Jean Blanchard and Stanley Fischer, eds., *N.B.E.R. Macroeconomics Annual*, Cambridge: MIT Press 1989, pp. 185–215.

Carroll, Christopher D., "The Buffer Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 1992, *2*, 61–135.

__ , "How Does Future Income Affect Current Consumption," *Quarterly Journal of Economics*, 1994, *109* (1), 111–147.

__ , "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, Feb 1997, *107* (1), 1–56.

__ **and Miles S. Kimball**, "On the Concavity of the Consumption Function," *Econometrica*, Jul 1996, *64* (4), 981–92.

Cochrane, John H., "A Simple Test of Consumption Insurance," *Journal of Political Economy*, Oct 1991, *99* (5), 957–76.

Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica*, 1991, *59*, 1121–1142.

Friedman, Milton, *A Theory of the Consumption Function*, Princeton, Princeton University Press, 1957.

Gourinchas, Pierre-Olivier, "Non Linearities in Exchange Rates: A Non-Parametric Instrumental Variables Approach," June 1997. manuscript, Stanford GSB.

__ **and Jonathan A. Parker**, "Consumption Over the Lifecycle," Dec 1997. Social Science Research Insitute Working paper 9722.

Gross, David B., "The Investment and Financing Decisions of Liquidity Constrained Firms," 1994. MIT mimeo.

Hall, Robert E., "The Stochastic Implications of the Life Cycle Permanent Income Hypothesis," *Journal of Political Economy*, 1978, *86*, 971–87.

Hansen, Lars Peter and Kenneth J. Singleton, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 1983, *91*, 249–68.

Hausman, Jerry A., Hidehiko Ichimura, Whitney K. Newey, and James L. Powell, "Identification and Estimation of Polynomial Errors-in-Variables Models," *Journal of Econometrics*, 1991, *50*, 273–95.

__ **, Whitney K. Newey, and James L. Powell**, "Nonlinear Errors in Variables Estimation of some Engel Curves," *Journal of Econometrics*, 1995, *65*, 205–33.

Lusardi, Annamaria, "Permanent Income, Current Income, and Consumption:

Evidence from Two Panel Data Sets," Jan 1996, *14* (1), 81–90.

Modigliani, Franco and Richard Brumberg, "Utility Analysis and the Consumption Function: an Interpretation of Cross-Section Data," in Kenneth K. Kurihara, ed., *Post-Keynesian Economics*, New Brunswick, N.J. Rutgers University Press 1956, pp. 388–436.

Nelson, Julie A., "On Testing for Full Insurance Using Consumer Expenditure Survey Data," *Journal of Political Economy*, Apr 1994, *102* (2), 384–94.

Newey, Whitney K., "Convergence Rates and Asymptotic Normality for Series Estimation," *Journal of Econometrics*, 1997, *79*, 147–168.

Venti, Steven and David Wise, "The Wealth of Cohorts: Retirement Saving and the Changing Assets of Older Americans," 1993. NBER Working Paper 4600.

Vissing-Jorgensen, Annette, "Limited Stock Market Participation," 1998. MIT Mimeo.

Zeldes, Stephen P., "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, Apr 1989, *97*, 305–346.

__ , "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *Quarterly Journal of Economics*, 1989, *104*, 275–98.

Appendices

To be written

A. Data

More details on the CEX and its use to be added..

B. Inference

More details on HNIP and inference, including order selection.

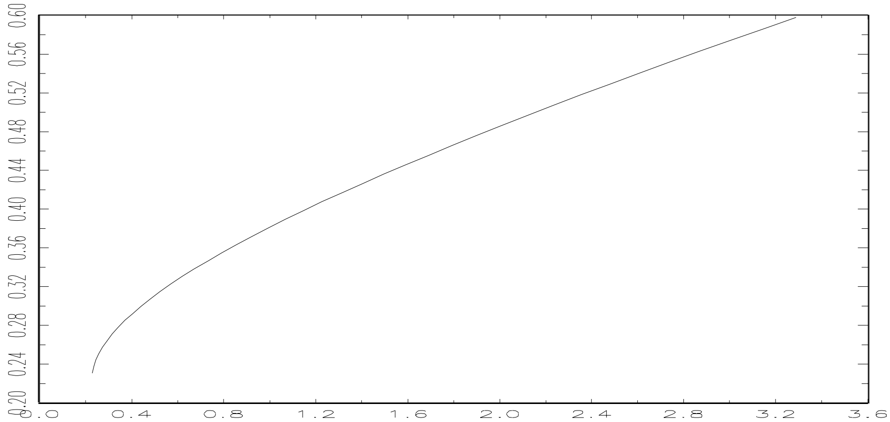

Figure 1: Simulated Consumption Function at Age 35

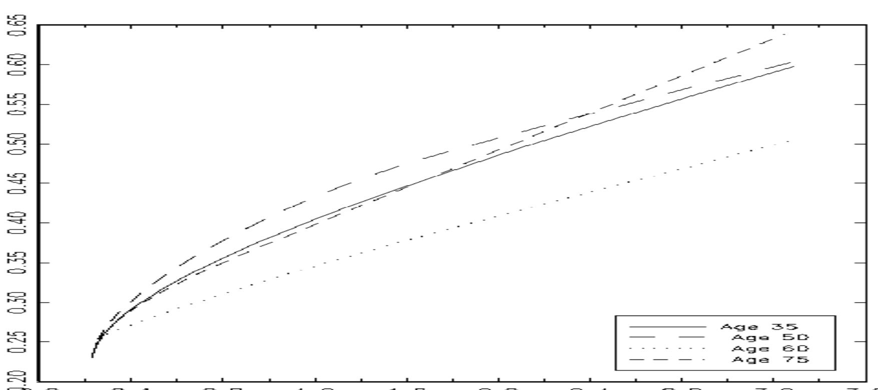

Figure 2: Simulated Consumption Functions by Age

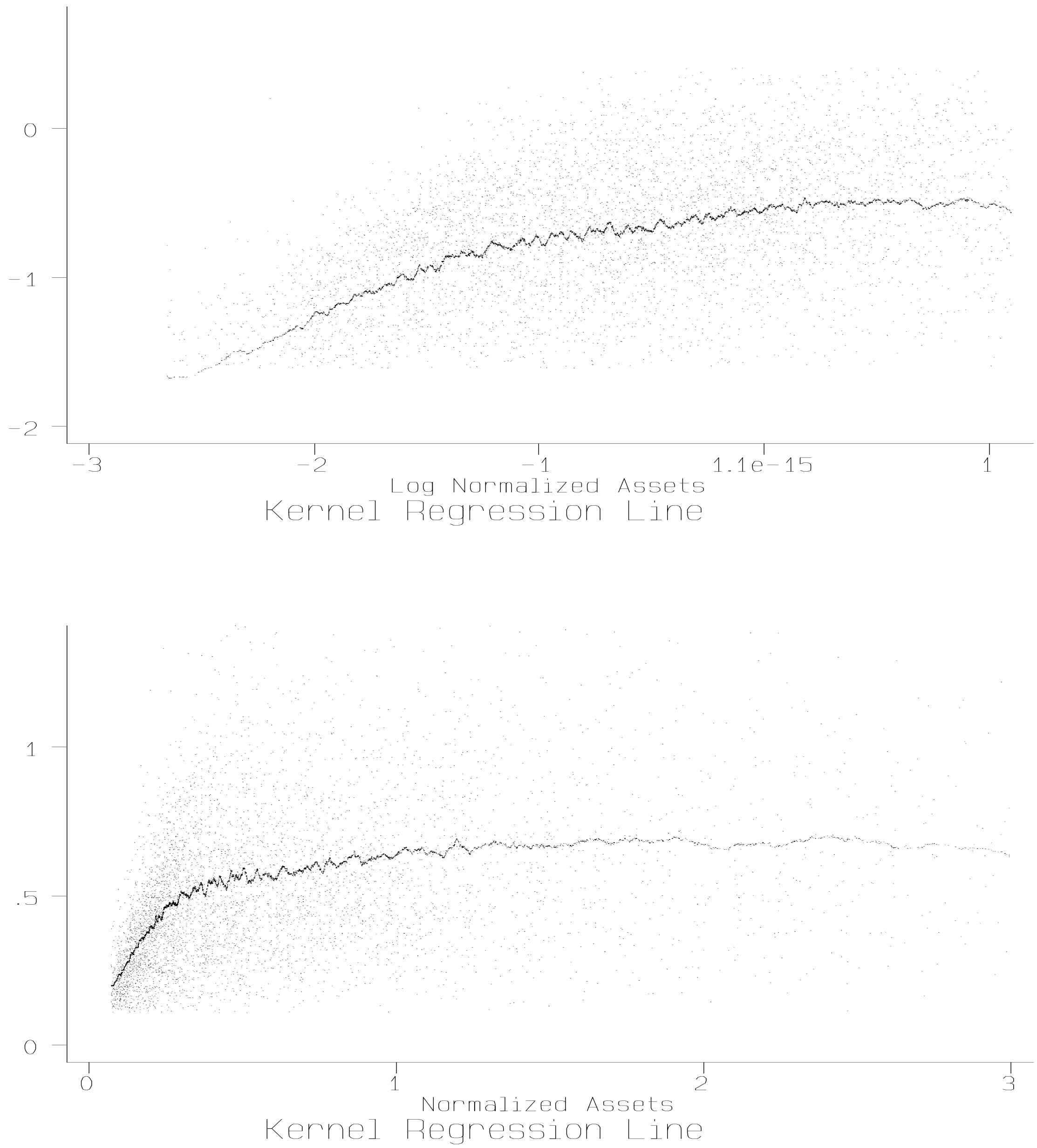

Figure 3: Kernel Estimates of Consumption Function without Covariates

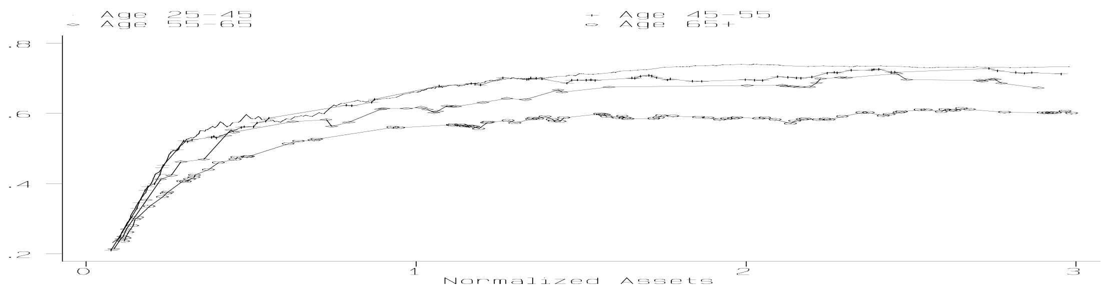

Figure 4: Kernel Estimates of Consumption Functions by Age

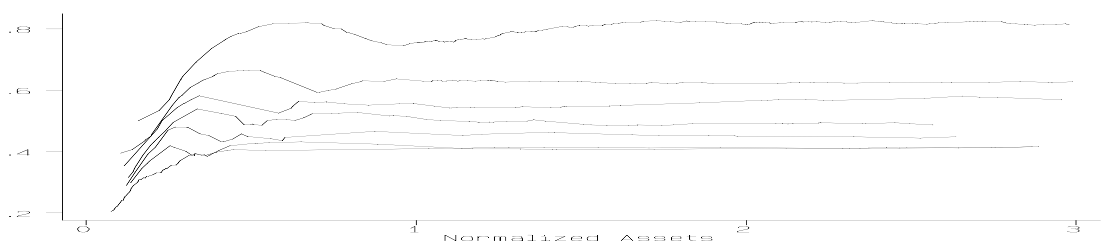

Figure 5: Kernel Estimates of Consumption Functions by Income Level

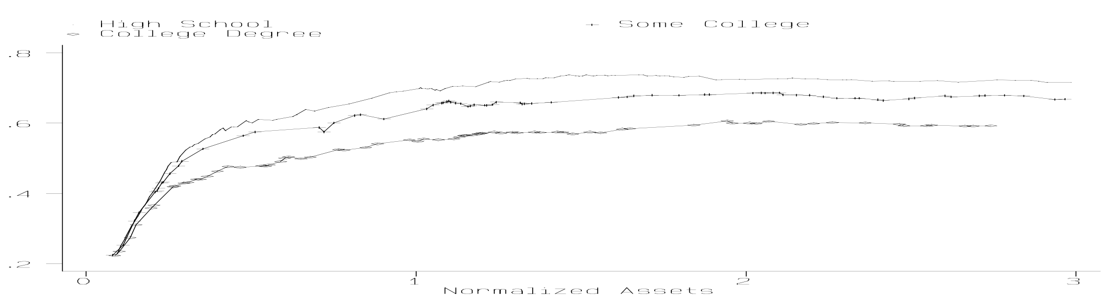

Figure 6: Kernel Estimates of Consumption Functions by Education Level

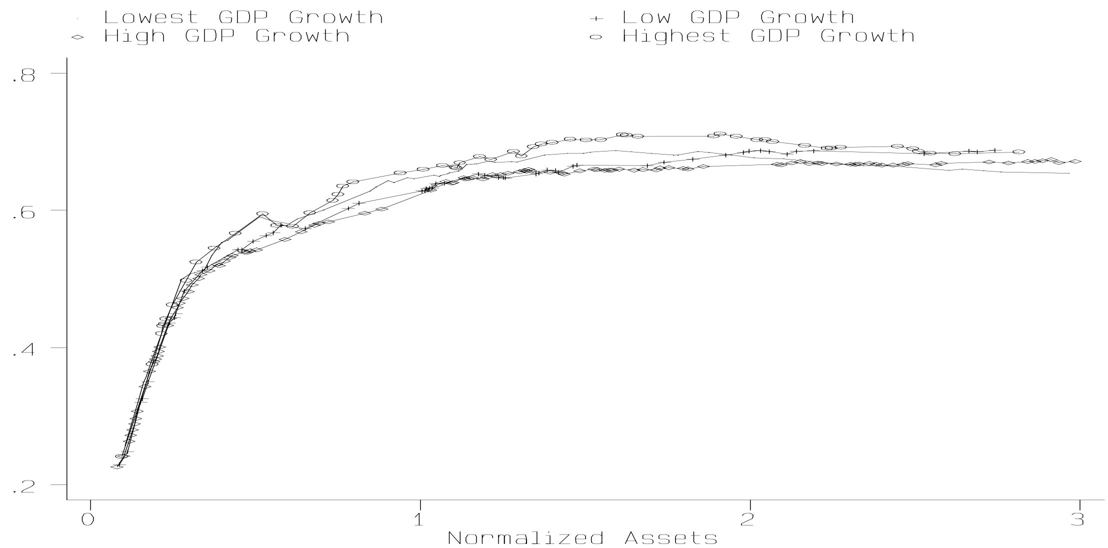

Figure 7: Kernel Estimates of Consumption Functions by GDP Growth

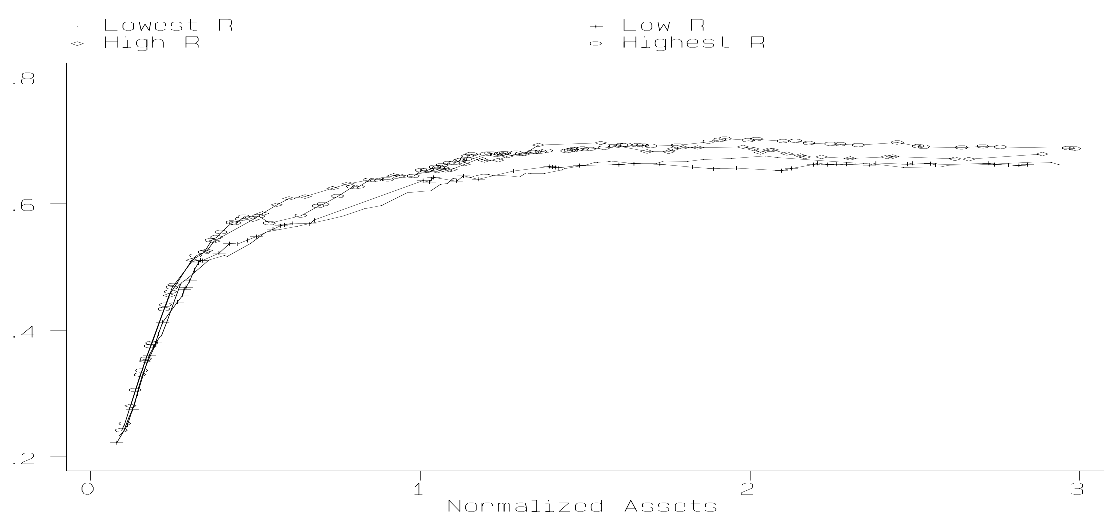

Figure 8: Kernel Estimates of Consumption Functions by Real Interest Rate

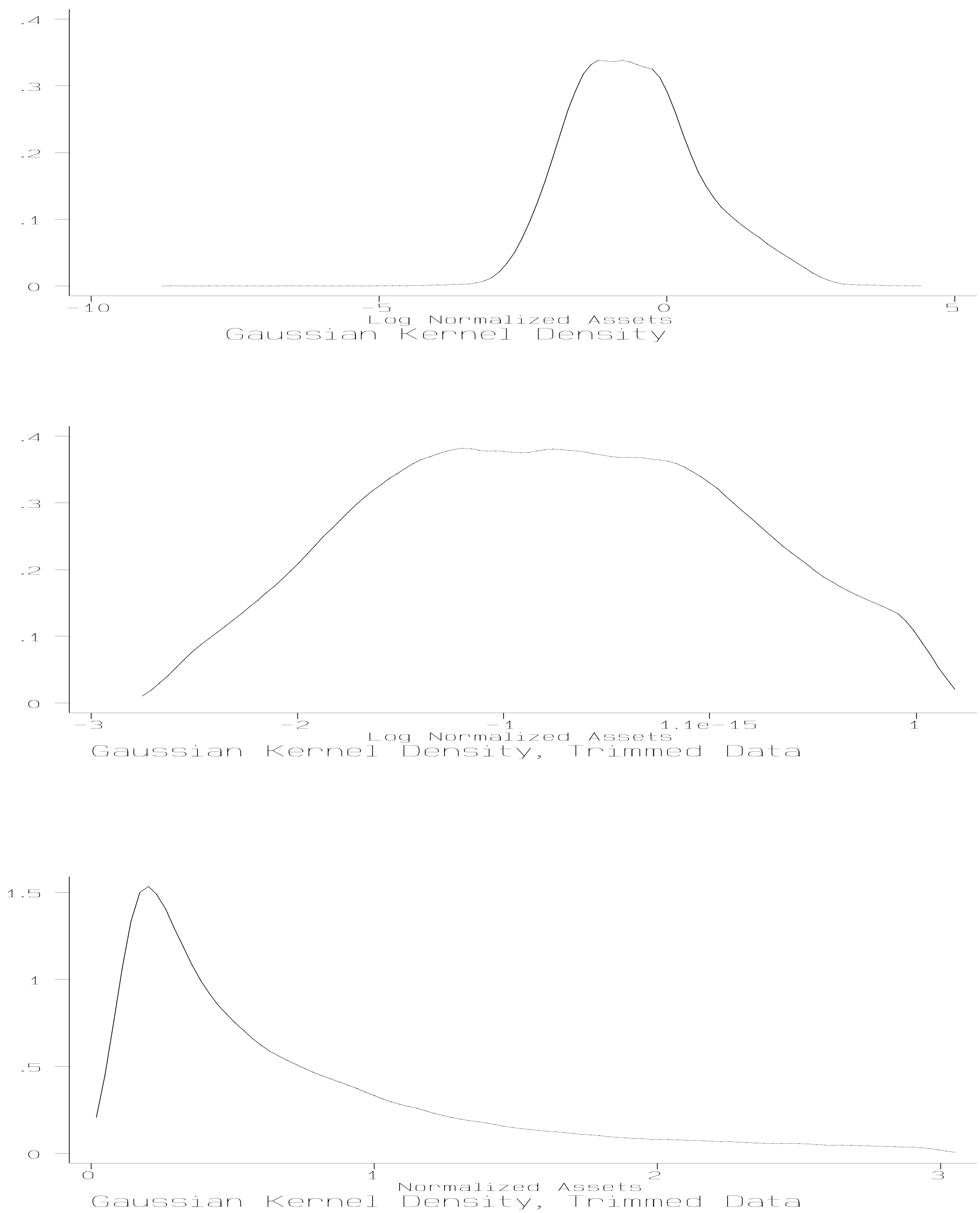

Figure 9: The Distribution of Liquid Assets

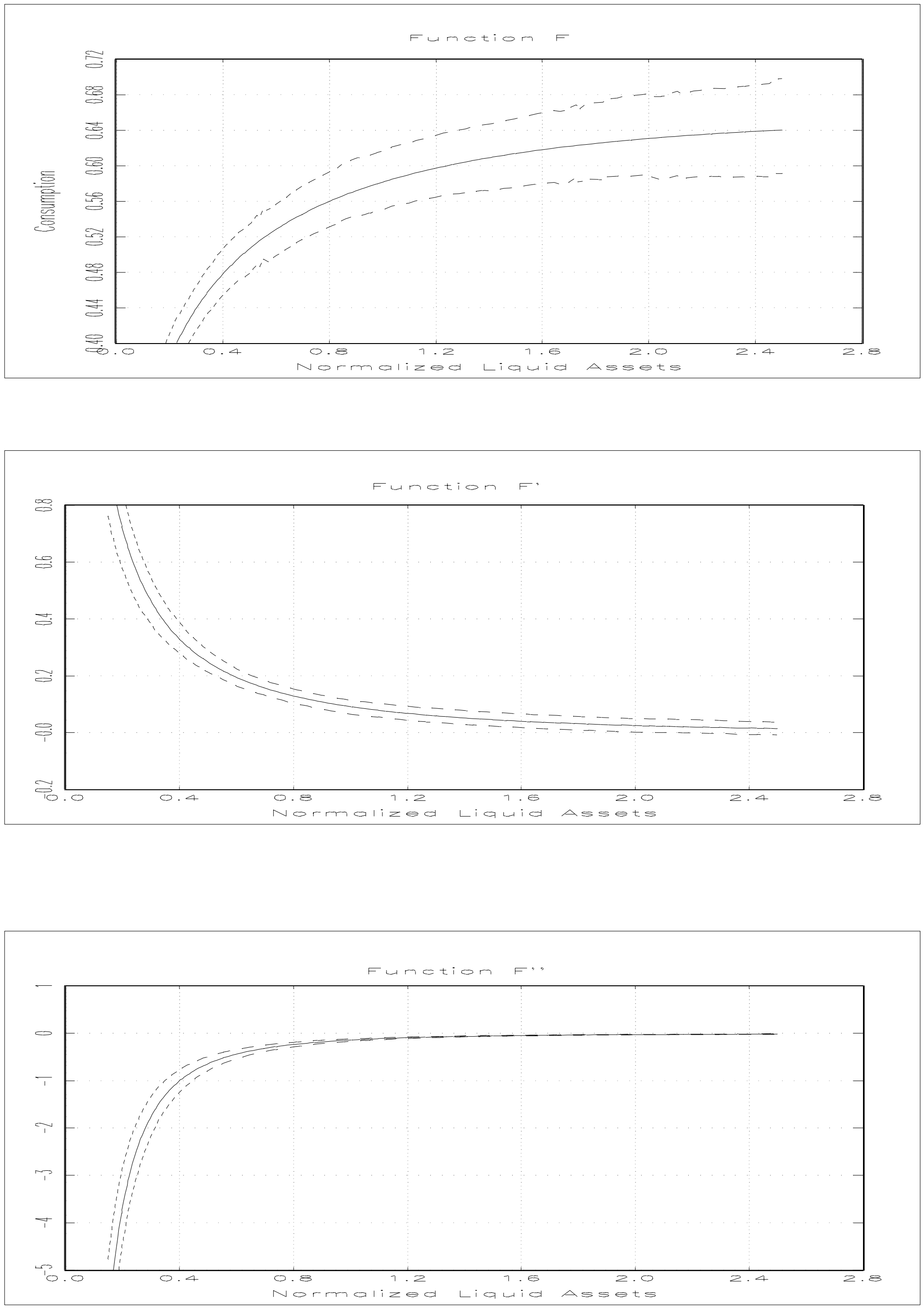

Figure 10: The Average Consumption Function and Its Derivatives

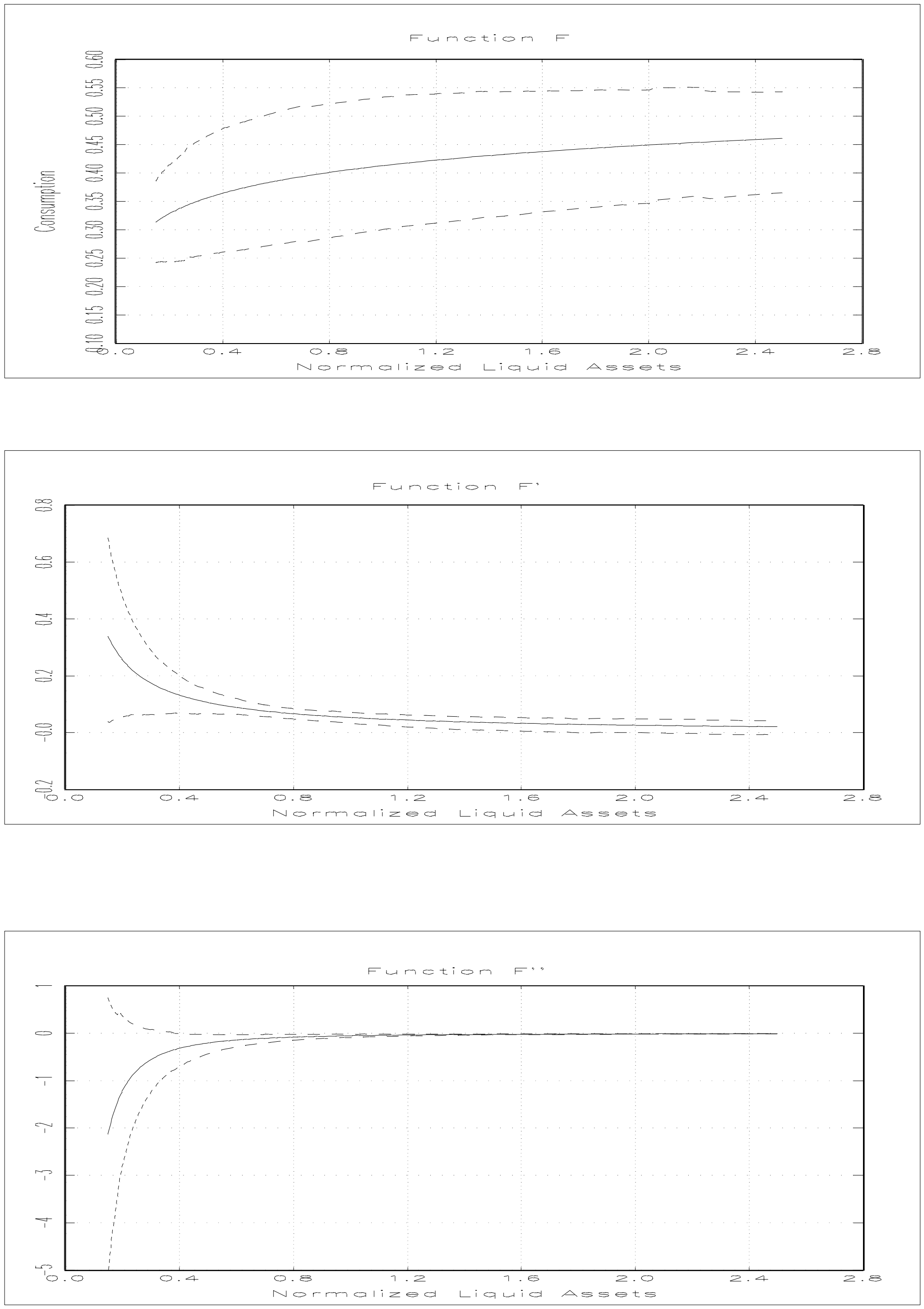

Figure 11: A Consumption Function and Its Derivatives

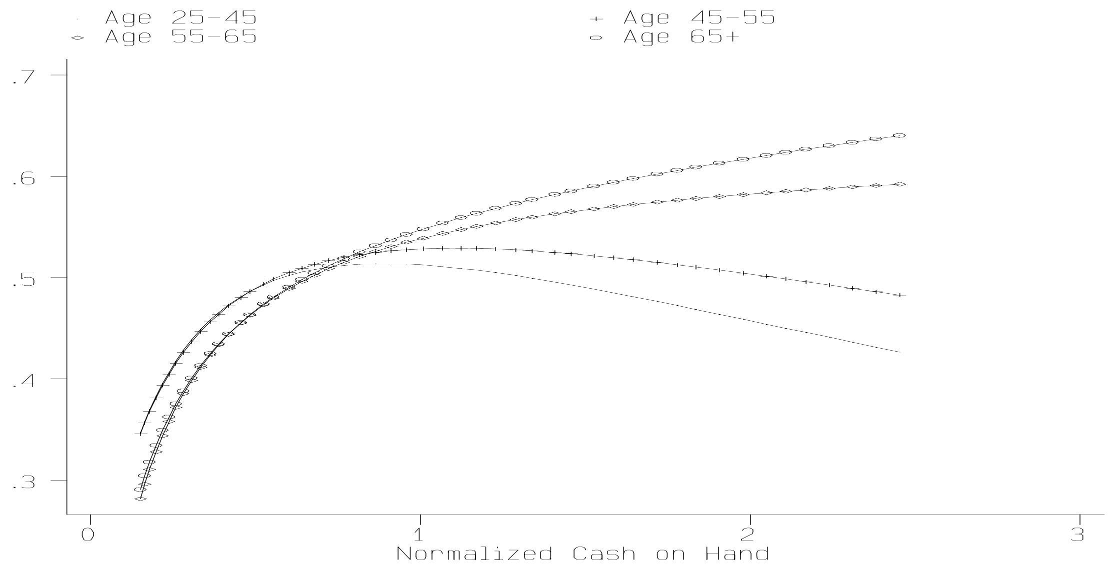

Figure 12: The Consumption Function by Age

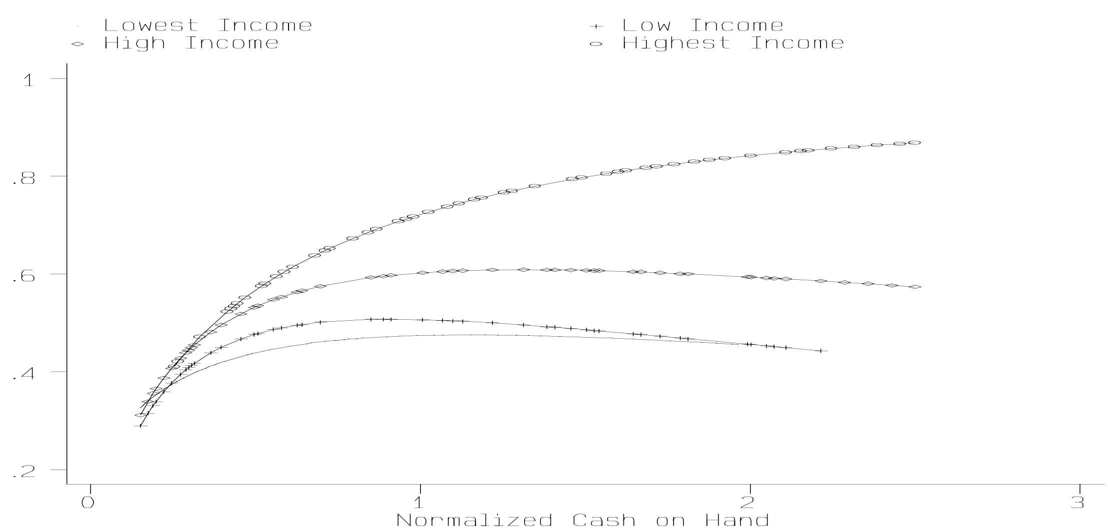

Figure 13: The Consumption Function by Current Household Income Level

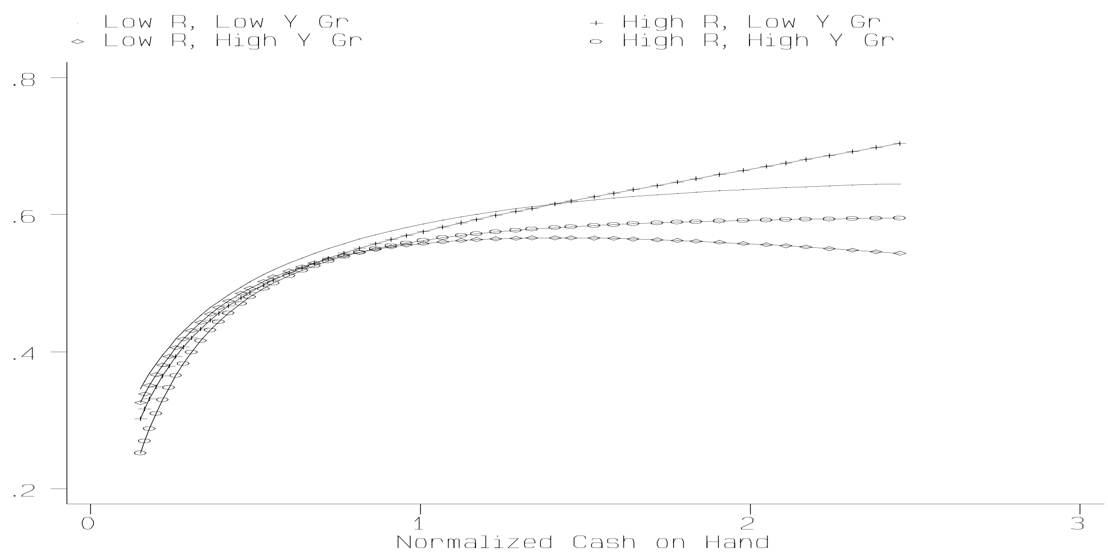

Figure 14: The Consumption Function by the Aggregate State

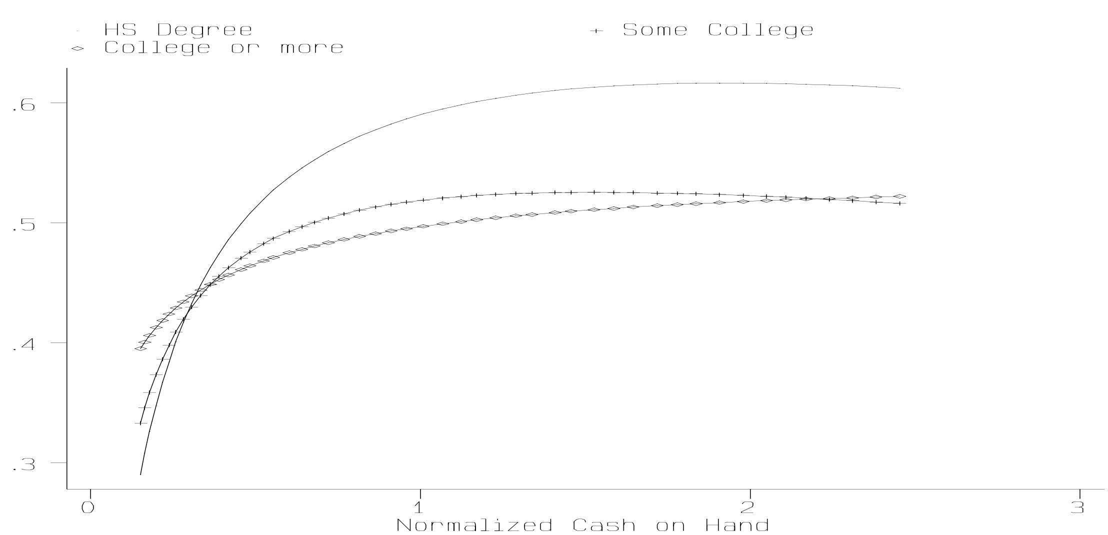

Figure 15: The Consumption Function by Education