ESSAYS ON STRUCTURAL MODELING OF LIFE CYCLE BEHAVIOR

by

Natalia Khorunzhina

Submitted to the Graduate Faculty of

the Department of Economics in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF ECONOMICS

This dissertation was presented

by

Natalia Khorunzhina

It was defended on

July 12th, 2011

and approved by

Dr. David DeJong, University of Pittsburgh, Economics

Dr. James Feigenbaum, Utah State University, Economics and Finance

Dr. George-Levi Gayle , Carnegie Mellon University, Tepper School of Business

Dr. Robert Miller, Carnegie Mellon University, Tepper School of Business

Dr. Jean-François Richard, University of Pittsburgh, Economics

Dr. Randall Walsh, University of Pittsburgh, Economics

Dissertation Advisors: Dr. David DeJong, University of Pittsburgh, Economics,

Dr. Jean-François Richard, University of Pittsburgh, Economics

Copyright © by Natalia Khorunzhina
2011

ESSAYS ON STRUCTURAL MODELING OF LIFE CYCLE BEHAVIOR

Natalia Khorunzhina, PhD

University of Pittsburgh, 2011

In the economic literature there are divergences on a number of issues between the results obtained with macro- and micro-based models. Habit formation in consumption is one example of such disagreement. Another example is the discrepancy between the theoretical prediction that all investors should participate in stock markets if the equity premium is positive, and empirical evidence that a substantial fraction of individual consumers do not invest in stock markets. A primary goal of my thesis is to try narrowing the gap between the results in these two literatures, looking at the problem from the prospective of households. Chapter 1 develops a nonlinear GMM estimator to investigate the presence of habit formation in household consumption. The estimation results support the existence of habit formation in food consumption. Next, I exploit the property of habit formation preferences to generates the time- and individual-varying Relative Risk Aversion and Intertemporal Elasticity of Substitution to analyze the degree of heterogeneity in these coefficients. I find that these parameters display significant variation across individuals and over time. In Chapter 2 I develop a dynamic structural model of stock market participation and portfolio choice to investigate whether financial education programs can affect consumers' choices and increase participation in financial markets. I estimate the model, where the consumers' decisions regarding stock market participation are influenced by participation costs. The results provide evidence that the participation cost is substantial. The model estimates are then used to conduct simulation exercises to evaluate the effect of financial education programs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 ESTIMATION OF OPTIMAL CONSUMPTION CHOICE WITH HABIT FORMATION AND MEASUREMENT ERRORS

1.1 INTRODUCTION

In this paper we empirically investigate the presence of internal habit formation in household food consumption using data from the Panel Study on Income Dynamics (PSID). We assume that habit formation takes a multiplicative (or ratio) form and that observed consumption is measured with error (i.e., assumed to be classical). The consumption Euler equation belongs to a class of moment conditions for which we develop a method to account for measurement errors without parametric assumptions on the distribution of measurement errors. We then exploit the structure of the Euler equation to develop a nonlinear generalized method of moments (GMM) estimator. We prove identification of all the structural parameters of interest: the time-discount factor, the utility curvature parameter, and the habit formation parameter. A small simulation exercise shows that the proposed estimator performs well in recovering the parameters of interest. The simulation exercise also shows that not accounting for measurement errors leads to underestimation of the utility curvature parameter and the strength of habits. We find that habit formation is an important determinant of food consumption patterns.

In models of habit formation with measurement errors, the Intertemporal Elasticity of Substitution (IES) and Relative Risk Aversion (RRA) are generally not point identified without further assumptions. We develop bounds on the IES and RRA that account for measurement errors and compute asymptotically valid confidence intervals on these bounds

using the parameter estimates from the model. Post estimation analysis also shows that the IES is decreasing and convex in age and displays a U shape in income. Also, the RRA is increasing and concave in age, dome shaped in income, and increasing in the level of education.

Allowing for habit formation in consumption offers a tractable approach to modeling time nonseparabilities in preferences. Time nonseparabilities in preferences have been used to explain a wide variety of macroeconomic and financial phenomena. In the representative agent framework, the empirical evidence largely supports the existence of habit formation in consumption.[1] However, in micro data the evidence of habit formation is inconclusive. The studies of Carrasco, Labeaga, and Lopez-Salido (2005) and Browning and Collado (2007) support habit formation in food consumption, while those of Meghir and Weber (1996) and Dynan (2000) do not. Comparisons between these results are confounded by differences in the specification of preferences, the approximations employed to obtain estimating equations, and differences in the data employed in estimation. Carrasco, Labeaga, and Lopez-Salido apply the test proposed by Meghir and Weber on a different data set of the same frequency. However, the data set is longer, and Carrasco, Labeaga, and Lopez-Salido argue that this is the main reason for their contrasting conclusions. In this paper, we employ the same data set as in Dynan, that is, household food consumption data from the Panel Study on Income Dynamics (PSID). However, we find evidence of significant habit formation in food consumption. We argue that the differences in results stem mainly from the differences in the estimation approach as well as the treatment of measurement errors in consumption and other unobserved heterogeneity. Our results are directly comparable not only to micro studies such as Dynan (2000), but also to the large body of macroeconomic literature.

In our analysis, we assume that individual preferences over food consumption is characterized by the multiplicative habit formation specification introduced in Abel (1990), where consumption services is given by $\tilde{c}_t = c_t/c_t^{\alpha}$.[2] The main alternative to this specification is the difference model of habit that is widely assumed in macroeconomic studies and some

[1] See Fuhrer (2000), Chen and Ludvigson (2008), and Smith and Zhang (2007) for later examples.

[2] In general, consumption services is given by $\tilde{c}_t = c_t/h_t^{\alpha}$ with $h_t = f(c_{t-1}, h_{t-1})$.

important microeconomic studies. In this case, consumption services is given by, for example $\tilde{c}_t = c_t - \alpha c_{t-1}$. The choice of the multiplicative specification is motivated by two related points. First, individual consumption data is more volatile than aggregate consumption data. As a result, while the restriction that consumption services is positive is relatively easy to satisfy in the difference specification using aggregated data, it is likely to be violated in micro data when the difference specification is assumed. The fact that the multiplicative specification of consumption services satisfies this positivity constraint for any pair (c_t, c_{t-1}) makes it more appropriate when employing micro data.

Second, under the multiplicative specification of habits, the model belongs to a class of moment conditions for which we develop a method for controlling for classical measurement errors without imposing parametric assumptions on the distribution of measurement errors. This method also avoids imposing log-linearizations and other assumptions on the Euler equation to obtain an estimable equation.

We show that all parameters of interest are identified from the moment condition, and develop a nonlinear GMM estimator for these parameters. To the best of our knowledge, this paper presents the first exact Euler equation nonlinear GMM method developed to investigate the existence of habit formation without imposing parametric assumptions on the distribution of measurement errors.

In order to investigate the small sample properties of the estimator, we perform a simulation exercise where we solve a life-cycle model with multiplicative habit consumption services. Two sets of estimation procedures are examined. The first is replication of the procedure employed by Dynan (2000), and the second is the procedure developed in this paper. The results from implementing the estimator developed in Dynan (2000) using the simulated data reject the existence of habit formation. On the other hand, the simulation exercise shows that the estimator developed in this paper performs well in recovering the parameters of interest. Also, the results indicate that ignoring measurement errors in food consumption results in a downward bias in the estimate of the habit formation parameter.

The estimator developed in this paper is implemented using household food consump-

tion data from the PSID. The results from the baseline model support the existence of a strong habit in food consumption. We also find that ignoring measurement errors weakens significantly the evidence of habit formation.

A variety of modifications to the baseline model are considered. Three key modifications are: (i) allowing of aggregate effects in the expectations error; (ii) restricting the sample to households that are not liquidity constrained; and (iii) extending the model to allow for external habit formation. We find that estimates of the structural parameters are robust across these specifications. The estimate of the external habit formation parameter is positive and statistically significant, but small relative the the internal habit formation parameter.

Habit formation also generates time- and individual-varying coefficients of relative risk aversion (RRA) and intertemporal elasticity of substitution (IES). We exploit this property and derive inferences about these economic quantities, which constitutes another novelty of the paper. Until recently, little was known about changes in risk aversion and intertemporal substitution across groups of individuals and over time.[3] Because observed consumption is measured with error, the implied household- and time-specific IES and RRA are not observed. Furthermore, the expectation of the IES and the RRA is not point identified. We define bounds for the conditional expectations of the IES and the RRA. We estimate these bounds and report asymptotically valid confidence intervals. For the preferred specification, the 95% confidence interval for the IES is [0.086, 0.189] and the corresponding 95% confidence interval for the RRA is [5.258, 13.093]. We also investigate the patterns of the IES and the RRA via regression analysis of the implied IES and RRA on household demographic characteristics. We find that the IES is decreasing and convex in age, and exhibits a U shape with respect to income. We also find that the RRA is increasing in concave in age, exhibits a dome shape with respect to income, and is increasing in education.

[3]See Allan and Browning (2009) and the references therein for a discussion on the recent developments in investigating individual heterogeneity in the RRA.

1.2 THEORETICAL FRAMEWORK

Household i chooses a sequence of consumption $\{c_{is}, s = t, \cdots, T\}$ to maximize its expected lifetime utility function, given by

$$E_{it} \sum_{s=t}^{T} \beta^{s-t} \phi_{is} \frac{\tilde{c}_{is}^{1-\gamma} - 1}{1-\gamma},$$

where the expectation is taken conditioning on all relevant information for household i at time t; $\beta \in (0, 1)$ is the time-discount factor; γ the utility curvature parameter; and $\tilde{c}_{it}$ denotes consumption services in period t. Consumption services is defined as the ratio between current consumption expenditures and past consumption expenditures geometrically weighted:

$$\tilde{c}_{it} = \frac{c_{it}}{c_{it-1}^{\alpha}},$$

where $0 \leq \alpha \leq 1$ measures the strength of habits; $\alpha = 1$ denotes the highest, while $\alpha = 0$ indicates no habit in consumption.

The importance of augmenting the individual utility function with individual-specific taste shifters has been widely accepted in the estimation of optimal consumption choices with micro data. Household-specific "taste shifters" ϕ_{it} are given by

$$\phi_{it} = \exp(\delta' w_{it} + \omega_i),$$

where w_{it} is a vector of exogenous time-varying observed household characteristics and ω_i is a household fixed effect.

We assume that household i is not subject to liquidity constraints and has rational expectations. The first-order necessary condition for the household's optimization problem is

$$E\left[\beta(1 + r_{i,t+1}) MU_{i,t+1} - MU_{i,t} | z_{it}\right] = 0, \tag{1.2.1}$$

where $r_{i,t+1}$ is the rate of return available to household i between periods t and $t+1$, z_{it} denotes the set of all information that is available to household i at time t, and $MU_{i,t}$ represents household i's marginal utility of consumption in period t:

$$MU_{i,t} = \frac{\phi_{it}}{c_{it}} \left(\frac{c_{it}}{c_{it-1}^{\alpha}} \right)^{1-\gamma} - \alpha\beta \frac{\phi_{it+1}}{c_{it}} \left(\frac{c_{it+1}}{c_{it}^{\alpha}} \right)^{1-\gamma}. \tag{1.2.2}$$

Notice that if $\alpha = 0$, MU_{it} in equation (1.2.2) reduces to the marginal utility of time separable models. For $\alpha > 0$, consumption services are negatively related to past consumption levels. This property is shared with difference models of habit formation with positive α (i.e., $\tilde{c}_{it} = c_{it} - \alpha c_{it-1}$). However, $\alpha > 0$ is not sufficient for the multiplicative model to characterize habit formation. The multiplicative model also requires $\gamma > 1$ to exhibit habit formation. Indeed, as long as both $\alpha > 0$ and $\gamma > 1$, the household's marginal utility of consumption in period t is an increasing function of period $t-1$ consumption, yielding a complementarity effect of consumption over time.

Substituting Equation (1.2.2) into Equation (1.2.1) obtains the following moment condition:

$$E\left[\beta(1+r_{it+1})\left(\frac{\phi_{it+1}}{c_{it+1}}\left(\frac{c_{it+1}}{c_{it}^{\alpha}}\right)^{1-\gamma} - \alpha\beta\frac{\phi_{it+2}}{c_{it+1}}\left(\frac{c_{it+2}}{c_{it+1}^{\alpha}}\right)^{1-\gamma}\right) - \frac{\phi_{it}}{c_{it}}\left(\frac{c_{it}}{c_{it-1}^{\alpha}}\right)^{1-\gamma} + \alpha\beta\frac{\phi_{it+1}}{c_{it}}\left(\frac{c_{it+1}}{c_{it}^{\alpha}}\right)^{1-\gamma}\middle| z_{it}\right] = 0 \tag{1.2.3}$$

1.3 THE ESTIMATOR

In order to derive an exact nonlinear GMM estimator of the parameter vector of interest from Equation (1.2.3), we address two key issues: nonstationarity of consumption and measurement errors in consumption. We discuss these issues in turn.

1.3.1 Consumption growth

The estimator derived in this paper is designed for panel data where the number of households N is large relative to the number of time periods T. The asymptotic properties of the estimator are derived assuming that N goes to infinity and T is fixed. We assume that the (time) vector of household observations is drawn independently from a common distribution. This is in contrast to the implementation using time series data, where the typical sufficient condition is stationarity of the variables. However, the resulting estimator may be subject to small sample instability problems in the panel data context if the data are not stationary. Because consumption trends over time, we transform the moment equation (Equation 1.2.3) into one that is expressed in terms of the growth rates of consumption. Let $g_{it} = c_{it}/c_{it-1}$, and $\varphi_{it} = \phi_{it}/\phi_{it-1}$. Then Equation (1.2.3) can be written as

$$E\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g_{it+1}}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\left(1-\alpha\beta\varphi_{it+2}\left(\frac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}\right)-\left(1-\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)\middle| z_{it}\right] = 0. \tag{1.3.1}$$

Notice that by expressing Equation (1.2.3) in terms of growth in consumption, is that the unobserved household fixed effects ω_i are eliminated, so that $\varphi_{it} = \exp(\delta'\Delta w_{it})$.

1.3.2 Measurement error

Given a set of appropriate instruments and the absence of measurement errors, consistent estimators of the parameters α, β, γ, and δ' can be obtained based on the moment condition in Equation (1.3.1). However, the estimation of nonlinear rational expectation models using micro data is complicated by the problem of measurement errors in consumption, which, if ignored, will likely result in inconsistent estimation of the key parameters of interest.

Log- linearization of the Euler equation has the advantage of remaining tractable when accounting for measurement errors in consumption in both time-separable and nonseparable models.[4] However, as discussed in Carroll (2001), a log-linear approximation can result

[4]Log-linearization of the Euler equation allows Dynan (2000) to account for measurement errors in consumption expenditures without additional parameterization while testing for nonseparabilities in individual current and past consumption.

in severe bias of the parameter estimates. Nonlinear GMM estimators based on the Euler equation provide an alternative to log- linearization, but without additional distributional assumptions, the problem of measurement errors remains difficult. Significant progress has been made in accounting for classical measurement error in time-separable models. Ventura (1994) assumes that measurement errors are serially independent and lognormally distributed, while Hong and Tamer (2003) assume that the measurement errors are independent and that their marginal distributions are Laplace with zero mean and unknown variance. After re-parametrization, the approaches in Ventura (1994) and Hong and Tamer (2003) (applied to the time-separable Euler equation) yield similar moment conditions for the estimation of the utility curvature parameter subject to a proper set of instruments. However, the time-discount factor remains unidentified.

Nonseparabilities in preferences add another layer of difficulty into nonlinear estimation with classical measurement errors. Due to the increasing complexity of the moment conditions with habit formation, measurement errors cannot be easily separated from observed consumption. Thus, to our knowledge, there are no studies that apply nonlinear estimators to test for time- nonseparabilities in individual preferences over consumption accounting for measurement errors.[5]

The method developed in this paper for accounting for measurement errors can be applied to a larger class of moment conditions than equation (1.3.1). We therefore first present the general framework and then discuss its application to equation (1.3.1).

1.3.2.1 A method for controlling for classical measurement errors in nonseparable models Consider the following moment condition:

$$E\left[\kappa + \sum_{k=1}^{K} \varphi_k(w_k;\delta) \cdot \tilde{f}_k(x_k;\theta)|z\right] = 0, \tag{1.3.2}$$

[5]Chen and Ludvigson (2008), and Smith and Zhang (2007) undertake the empirical estimation of the consumption model with habit formation using nonlinear exact estimation. But the authors deal with aggregate consumption data where measurement errors issue are not a significant concern.

where: κ is a known constant; δ and θ are the parameters of potential interest; $w = (w_1, \ldots, w_K)$, and $x = (x_1, \ldots, x_K)$ are sets of regressors; and z is a set of instruments. However, the x_k's are not observed by the econometrician. Instead, what is observed is the corresponding set x_k^* of noisy measures of x_k. Assume x_k and x_k^* are related by the equation $x_k^* = x_k \eta_k$, where η_k denotes unobserved measurement errors.

Suppose the data generating process and the functional forms satisfy the following assumptions.

Assumption 1.3.1. *For $k = 1, \cdots, K$:*

1. *the random vector (w_k, x_k, z) is independent from η_k;*
2. *the function $\tilde{f}_k(a \cdot b; \theta)$ satisfies the following separability condition, $\tilde{f}_k(a \cdot b; \theta) = f_k(a; \theta) \cdot h_k(b; \theta)$; and*
3. $\mathcal{A}_k \equiv E[h_k(\eta_k; \theta)] \neq 0$.

Assumption 1.3.1.1 is a typical classical measurement error independence assumption. For example, this independence assumption is also made in Hong and Tamer (2003). However, unlike Hong and Tamer (2003), the method developed here does not impose parametric restrictions on the distribution of measurement errors. Furthermore, we do not require the measurement errors η_k and η_l to be uncorrelated for $k \neq l$. A weaker mean independence restriction is assumed by Hausman et al (1991) and Schennach (2004). However, to identify the parameters of interest, these authors also assumed the existence of an additional noisy measure of the true unobserved regressor. The method developed here does not rely on the existence of auxiliary data sets. The separability restriction of Assumption 1.3.1.2 holds for the Euler equation (1.3.1) and also holds for a variety of other applications. A sufficient condition for Assumption 1.3.1.2 to hold is that $\tilde{f}_k(x_k; \theta)$ is a power function of x_k. Assumption 1.3.1.3 is a regularity assumption which holds in general if measurement error is specified in the multiplicative form. Strictly speaking, $\mathcal{A}_k$ is a function of the parameter vector θ and the distribution of the measurement errors. The actual functional form may be known if the distribution of the measurement errors is assumed to be known up to a set of finite dimensional parameters. The following proposition shows that under Assumption 1.3.1 the

moment condition in equation (1.3.2), which is a function of the unobserved x_ks, can be transformed into an expression which is a function of the observed x_k^*s and the additional parameters $\mathcal{A}_k$.

Proposition 1.3.2. *Consider the moment condition 1.3.2. Suppose Assumption 1.3.1 hold. Then*

$$0 = E\left[\kappa + \sum_{k=1}^{K} \varphi_k(w_k;\delta)\cdot \tilde{f}_k(x_k;\theta)|z\right] = E\left[\kappa + \sum_{k=1}^{K} \mathcal{A}_k^{-1}\varphi_k(w_k;\delta)\cdot f_k(x_k^*;\theta)|z\right].$$

The proof of Proposition 1.3.2 follows from the following series of equalities.

$$\begin{aligned} E[\varphi_k(w_k;\delta)\cdot \tilde{f}_k(x_k^*;\theta)|z] &= E[\varphi_k(w_k;\delta)\cdot \tilde{f}_k(x_k\eta_k;\theta)|z] = E[\varphi_k(w_k;\delta)\cdot f_k(x_k;\theta)\cdot h_k(\eta_k;\theta)|z] \\ &= E[\varphi_k(w_k;\delta)\cdot f_k(x_k;\theta)\cdot E[h_k(\eta_k;\theta)|w_k,x_k,z]|z] \\ &= E[\varphi_k(w_k;\delta)\cdot f_k(x_k;\theta)|z]\cdot \mathcal{A}_k. \end{aligned}$$

The first equality follows by the definition $x_k^* = x_k\eta_k$; the second equality follows from Assumption 1.3.1.2; the third equality follows from the law of iterated expectations; and the fourth equality follows from Assumption 1.3.1.1. By Assumption 1.3.1.3, $\mathcal{A}_k$ is invertible which gives us the required result.

1.3.2.2 Accounting for measurement errors in the Euler equation The results in the previous section apply directly to equation (1.3.1). Let true consumption c_{it} be measured with a multiplicative error $\tilde{\eta}_{it}$, so that observed consumption is given by $c_{it}^* = c_{it}\tilde{\eta}_{it}$, where $\tilde{\eta}_{it} > 0$. It is interesting to note that the transformation of the Euler equation into one that is expressed in terms of growth rates also eliminates individual-specific, time-invariant measurement errors in consumption. The assumptions on the measurement errors are therefore presented conditional on these individual-specific effect. Suppose that the measurement errors can be decomposed as $\tilde{\eta}_{it} = \mu_i\eta_{it}$.

Assumption 1.3.3. *Given μ_i and for each t, η_{it} is independent from time vector of consumption, the taste shifters, the information set, the interest rate, and income.*

Define $g^*_{it} = c^*_{it}/c^*_{it-1}$ and $v_{it} = \tilde{\eta}_{it}/\tilde{\eta}_{it-1} = \eta_{it}/\eta_{it-1}$, so that $g^*_{it} = g_{it}v_{it}$. Assumption 1.3.3 corresponds to Assumption 1.3.1.1. Assumption 1.3.1.2 is satisfied by the structure of the moment condition in equation (1.3.1). Because $\eta_{it} > 0$, $v_{it} > 0$ so that Assumption 1.3.1.3 is satisfied by this observation and the structure of equation (1.3.1). Therefore, under Assumption 1.3.3, the moment condition in equation (1.3.1), which is a function of unobserved true consumption, is transformed into a moment condition consisting of observed consumption:

$$E\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\left(\mathcal{A}_1^{-1}-\alpha\beta\mathcal{A}_2^{-1}\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)-\right.$$
$$\left.\left(1-\alpha\beta\mathcal{A}_3^{-1}\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)\middle| z^*_{it}\right] = 0, \tag{1.3.3}$$

where z^*_{it} is a q-dimensional observable subset of z_{it} that can include current and past interest rates as well as observable consumption growth up to time $t-2$. The details of the derivation are in Appendix A.

It is of interest to evaluate the variation in observed food consumption is due to measurement errors. The variance of measurement errors is not identified without additional assumptions. To this end, we will make the following functional form assumption as an *alternative* specification.

Assumption 1.3.4. *Conditioned on μ_i, measurement errors in consumption are serially independent and log-normally distributed:*

$$\ln \eta_{it}|\mu_i \sim N(0, \sigma^2). \tag{1.3.4}$$

Under this additional assumption, we have that

$$\begin{aligned}
\mathcal{A}_1 &= \exp\{\sigma^2\left(\alpha^2(1-\gamma)^2+\gamma^2-\alpha\gamma(1-\gamma)\right)\} \\
\mathcal{A}_2 &= \exp\{\sigma^2\left(\alpha^2(1-\gamma)^2+\gamma^2+(1-\gamma)(1+\alpha)\right)\}, \quad \text{and} \\
\mathcal{A}_3 &= \exp\{\sigma^2\left((1+\alpha+\alpha^2)(1-\gamma)^2\right)\}.
\end{aligned}$$

The details of the derivation can are found in Appendix A.

1.3.3 Identification

This section investigates identification of the parameters from the moment condition in equation (1.3.3). We first consider identification of the parameters of the model without imposing Assumption 1.3.4. To this end, let $\kappa_1 = \mathcal{A}_2/\mathcal{A}_1$, $\kappa_2 = \mathcal{A}_2$, and $\kappa_3 = \mathcal{A}_2/\mathcal{A}_3$. Then noting that $\mathcal{A}_2 > 0$, equation (1.3.3) can be rewritten as

$$E\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\left(\kappa_1 - \alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right) - \left(\kappa_2 - \alpha\beta\kappa_3\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)\middle| z^*_{it}\right] = 0, \tag{1.3.5}$$

Define $x^*_{it+2} = \left(g^*_{i,t+2},\, g^*_{i,t+1},\, g^*_{i,t},\, r_{i,t+1},\, \Delta w_{i,t+2}, \Delta w_{i,t+1}\right)$, $\theta = (\alpha, \beta, \gamma, \kappa', \delta')'$, and $\Delta_2 w_{i,t} = w_{i,t} - w_{i,t-2}$. The following conditions are sufficient for the identification of θ.

Assumption 1.3.5. *1. For at least one t in $\{4, \cdots, T\}$, the conditional density $f(x^*_{i,t}|z^*_{i,t-2})$ of $x^*_{i,t}$ given $z^*_{i,t-2}$ is complete.*

2. Rank($E[(1,\ \Delta_2 w_{i,t})'(1,\ \Delta_2 w_{i,t})]$) = Dim($(1,\ \Delta_2 w_{i,t})'$).

3. $\delta_1 = 1$.

See Newey and Powell (2003) for discussions of Assumption 1.3.5.1. This completeness assumption is also used in Chen and Ludvigson (2006) to prove identification of the parameters characterizing their asset pricing model. The larger is T, the easier it is to find at least one period for which this condition is satisfied. Assumption 1.3.5.3 is a full rank assumption. One consequence of this restriction is that a constant cannot be included in $w_{i,t}$. Furthermore, a variable that changes by a constant amount, such as age, may not be included in $w_{i,t}$. Assumption 1.3.5.3 helps to pin down κ, and δ. More importantly however, this normalization eliminates the trivial solution. In particular, notice that setting $\delta = 0$, $\gamma = 1$, and $\alpha\beta = \kappa_1 = \kappa_2/\kappa_3$, then the term inside the expectation on the right hand side of equation (1.3.5) becomes identically zero. Setting $\delta_1 = 1$ eliminates this trivial solution. However, this assumption does require a priori knowledge of the sign of δ_1. In the next section we discuss the choice of $w_{i,t,1}$. Let θ_0 denote the true parameter vector. The proof of following theorem is presented in Appendix B

Theorem 1.3.6. *Consider equation (1.3.5) and suppose Assumption 1.3.5 hold. Then θ is uniquely identified.*

The intuition behind the identification of the parameters is as follows. First, the normalization of Assumption 1.3.5.3 eliminates the trivial solution. The completeness assumption, the continuity of consumption growth, and the differentiability of the second term on the right hand side of equation (1.3.5) in $g^*_{i,t+2}$ obtains identification of γ, and α. Given this, the rank assumption 1.3.5.2 obtains identification of δ and β. Finally, the independent time variation in the observed consumption process along with the variation in the interest rate process obtain identification of the κs. This discussion highlights the significance of sufficient time variation in the consumption process in identifying the parameters of the model.

If in addition Assumption 1.3.4 is imposed, the σ^2 is identified from the resulting form of $\kappa_2 = \mathcal{A}_2$. It is important to note that σ^2 should be considered a lower bound on the amount of noise present in observed consumption, since the additional variation contributed by μ_i is not contained in σ^2.

We can therefore estimate the unknown structural parameters of interest using equation (1.3.3) as a conditional moment of the form

$$E\left[\rho(x^*_{it+1}, \theta_o)|z^*_{it}\right] = 0, \tag{1.3.6}$$

Define the q-dimensional vector $m_{it}(\theta) := m(x^*_{it+1}, z^*_{it}, \theta) := z^{*'}_{it}\rho(x^*_{it+1}, \theta)$ and the corresponding $q(T-4)$-dimensional moment vector $m_i(\theta) = m(x^*_i, z^*_i, \theta) := (m'_{i3}(\theta), \cdots, m'_{i,T-2}(\theta))'$ Then equation (1.3.6) implies that

$$m(\theta_0) = E[m_i(\theta_0)] = 0. \tag{1.3.7}$$

Let $\hat{m}(\theta) := \sum_{i=1}^N m_i(\theta)/N$ and $\hat{\Omega}(\theta) := \sum_{i=1}^N m_i(\theta)m'_i(\theta)/N$. Then, our estimator for the parameters of interest is defined by

$$\hat{\theta} = \arg\min_{\theta\in\Theta} \hat{m}'(\theta)\hat{\Omega}(\theta)^+\hat{m}(\theta),$$

where $\hat{\Omega}(\theta)^+$ is the generalized inverse of $\hat{\Omega}(\theta)$. Two remarks are in order.

While identification was discussed in terms of conditional moment restrictions, the estimator is defined using unconditional moment restrictions (equation 1.3.7) implied by their conditional counterparts. This is the dominant approach taken in applied work. However, in our case, it is important to know if the identification strategy implemented in this paper using conditional moment restrictions still applies under the unconditional moment restriction. Indeed, if equation (1.3.5) were written in the form of an unconditional moment then identification of the parameters of the model is maintained if the completeness assumption is instead imposed on the joint distribution $f(x^*_{i,t}, z^*_{i,t-2})$.[6]

As suggested by Hansen, Heaton, and Yaron (1996), we apply continuous updating GMM (CUGMM) to obtain estimates of the structural parameters. While CUGMM is known to be somewhat difficult to implement, it has advantages that are pertinent. As stated by Hansen, Heaton, and Yaron, CUGMM alleviates the problem of weak identification of parameters that is common in estimating Euler equations. However this doesn't stop the estimator from approaching the trivial solution, in which case, our experience indicates that the estimator becomes unstable. As discussed, we eliminate this instability by imposing the restriction that the coefficient on lagged income is equal to 1, thus eliminating the choice of $\delta = 0$. This restriction significantly increases the stability of the estimation algorithm up to a point where one might be tempted to replace CUGMM with two-step GMM. However, with approximately 50 moments used in the estimation, CUGMM is also attractive in that it tends to reduce the bias found in efficient two-step GMM with many moments (see Newey and Smith 2000).[7]

[6]Similar to Chen and Ludvigson (2008), an alternative method of estimating the Euler equation is to estimate directly the conditional expectations by employing the sieve minimum distance (SMD) technique (see also Ai and Chen 2003). However, this method is computationally more intensive, and less is known about testing model validity in the SMD framework.

[7]Application of one- and two-step GMM estimators to our model confirmed this result, as the estimated parameters (especially γ) were outside reasonable ranges.

1.4 MONTE CARLO EXPERIMENT

In this section we investigate the finite sample performance of the estimator developed in the previous section, as well as the performance of the estimator developed in Dynan (2000). We do so by conducting a simulation exercise where the life-cycle model presented in Section 1.2 is solved and simulated to obtain a sample of 4000 individuals observed over 24 periods. The details of the solution and simulation method are outlined in Appendix C. The structural parameter values of interest are set as follows: $\gamma = 5$, $\alpha = 0.85$, $\beta = 0.95$. The simulated consumption data is contaminated with measurement error drawn independently over individual and time from a log-normal distribution with variance equal to 20% of the variation in the simulated consumption.

With the simulated data in hand we investigate the performance of the estimator proposed in Dynan (2000). To derive the estimator, Dynan assumes among other things that: (i) interest rates do not vary over individuals and time; (ii) individuals live for infinite period, and (iii) $\Delta \ln(c_t - \alpha c_{t-1}) \approx \Delta \ln(c_t) - \alpha \Delta \ln(c_{t-1})$. We derive a comparable estimator where the first two assumptions are maintained. As shown in Mauelbauer (1988), third assumption requires that consumption does not vary significantly over time. If this and the first two assumptions are valid, then the estimation of the following equation

$$\Delta \ln(c^*_{it}) = \beta_0 + \beta_1 \Delta \ln(c^*_{i,t-1}) + \beta_2 age_t + \beta_3 age^2 + \epsilon_t, \tag{1.4.1}$$

where $\Delta income_{t-1}$ is used as an instrument for $\Delta \ln(c^*_{i,t-1})$ should yield the result the result $\beta_1 = 0.85$. The results from this estimation is presented in Table **??**.

The results from Table **??** indicate that estimation of equation (1.4.1) does a poor job in recovering the habit formation parameter. Indeed the magnitudes of the estimates of the habit formation parameter are similar to those found in Dynan (2000). These results suggest that the assumptions made to obtain equation (1.4.1) are substantial.

Table **??** presents the results from implementing the estimator developed in the previous section using simulated data. Column (1) gives the true values of the "deep" parameters,

Table 1.1: Estimation of equation (1.4.1) using the simulated data

	(1)	(2)
Constant	-0.015	5.650
	(0.042)	(3.562)
$\Delta \ln(c^*_{i,t-1})$	0.077	0.000
	(0.005)	(0.054)
Age	0.004	-0.051
	(0.002)	(0.036)
Age^2	-0.103	0.625
	(0.030)	(0.477)
$\ln(1 + r_t)$	—	137.9
		86.58

Instrument set includes current and past income for model (1), and income and lagged interest rates for model (2). Standard errors in parenthesis.

Table 1.2: Estimation of the Euler equation with habit formation with simulated data[a]

Parameters	Truth	Log-normal ME	Nonparametric ME	Ignoring ME
	(1)	(2)	(3)	(4)
γ	5.00	5.263	5.016	4.802
		(0.324)	(2.600)	(0.021)
β	0.95	0.926	0.967	0.946
		(0.032)	(0.017)	(0.015)
α	0.85	0.887	0.849	0.785
		(0.042)	(0.070)	(0.006)
σ^2	0.06	0.052		
		(0.003)		

[a] The instrument set includes current and past interest rates and current income. Standard errors in parenthesis.

that we try to recover using the proposed estimator. The simulated data set is used to estimate the model: assuming lognormal measurement error (column (2)), assuming non-parametric measurement error (column (3)), and finally, ignoring measurement error while it is present in the data (column (4)).

The results show that both lognormal measurement error model and its non-parametric counterpart works very well when the true measurement error is lognormal. For these two estimators all three "deep" parameter estimates are well inside one standard error range from the truth. Nonparametric measurement error estimator gives larger standard errors for the estimates. The only concern with the lognormal measurement error estimator is the variance of measurement errors σ^2 that is estimated somewhat lower than the corresponding truth. The results in column (4) show that both α and γ are estimated with a significant downward bias. This suggests that while ignoring measurement errors does not affect the estimation of β, it has the effect of biasing downwards the estimates of the habit formation parameter α and the curvature parameter γ.

1.5 DATA

Data on food consumption, as well as income and demographic characteristics of individuals and households are available from the Panel Study of Income Dynamics (PSID). Although it is the longest panel study, and one of the most comprehensive sources of information for studying life-cycle processes and poverty and welfare dynamics, its use for studying consumption involves one drawback: consumption data are available only for food. Fortunately, data on consumption of food as a perishable good are particularly suitable for testing whether this category of consumption can be habit-forming. The annual frequency of observation is also advantageous. As argued in Dynan (2000), if there is any effect of durability in food consumption, it is not likely to last more than a few months.

The main consumption sample that we use consists of data from 1974 through 1987.

Consumption of households consists of expenditures on food consumed at home, away from home, and the value of food stamps. Data on food consumed at home and the value of food stamps are deflated using the consumer price index (CPI) for food at home. Data on food consumed away from home are deflated using the CPI deflator for food away from home. All CPI data are taken from the consumer price index releases of the Bureau of Labor Statistics. Food consumption data are deflated according to the month and year when the interview occurred, while food stamps and data on income are deflated using the CPI for the end of the year before the interview was conducted. In addition, total consumption expenditures are adjusted by the size of household.

We exclude households whose marital status changed or whose head was younger than 22 or older than 65 over the period of estimation. We also exclude observations for which the consumption growth rate was higher than 300% and lower than 33%. It is likely that the extreme outliers in consumption growth rate that we observe in the untrimmed data are due to measurement errors. Thus, the estimated magnitude of the variance of the measurement errors in consumption is to be considered a lower bound after this data trimming. Household characteristics used in estimation as taste shifters include past income, family size and age of the head of household.

As in Shapiro (1984), Runkle (1991), and similar studies, we construct the household-specific real after-tax interest rate as $r_{i,t+1} = R_t(1 - \tau_{i,t+1}) - \pi_{t+1}$, where R_t is the average 12-month Treasury bill for the first half of the preceding year, $\tau_{i,t+1}$ is the household marginal tax rate as reported in the PSID, and π_{t+1} is the CPI deflator for the period of the interview.

The estimation of the moment condition (equation 1.3.5) and of most of its modifications for robustness checks requires data on consumption expenditures for four consecutive years for each orthogonality condition. With the restrictions on data described above, we have an unbalanced panel on 3,402 households covering ten years from 1976 through 1985. This is the baseline sample. We also check the sensitivity of the results to the liquidity constraint assumed in the derivation of the estimator. Therefore, another sample is constructed to exclude households that may be liquidity constrained during the sample period. In this

smaller sample of 1,752 households, we keep only the households that report positive savings over the sample period.

1.6 EMPIRICAL RESULTS

In this section we address several issues while discussing the results obtained from the estimation. The main conclusion from the results are that (i) habit formation plays an important role in explaining household food consumption patterns; and (ii) mis-specifying the distribution of, or not accounting for measurement errors in observed consumption result in a downward bias in the estimates of the habit formation and utility curvature parameters.

Table 1.3 presents the parameter estimates. The results from estimating the model without distributional assumptions on measurement errors are presented in columns (1)-(4). Column (1) shows results for the basic model estimated with the full sample, while column (3) reports the results for the same model estimated with the subsample of liquidity-unconstrained households. Column (2) shows the results for the model augmented with the aggregate shocks. Column (4) reports the results for the model extended to allow for external habits. Columns (5)-(7) present the estimation results under the assumption that measurement errors in consumption are log-normally distributed. Again, column (5) shows results for the model estimated with the full sample, column (7) reports the results for the same model estimated with the subsample of liquidity unconstrained households and column (6) shows the results for the model estimated with the full sample and augmented with the aggregate shocks. Column (8) reports the results for the model that ignores the existence of measurement errors. Demographic characteristics as taste shifters are included in all of the specifications that we consider. A more detailed version of Table 1.3 can be found in the appendix of the paper. In all specifications, except where measurement errors are ignored, the J test fails to reject the model at the 5% level of significance, indicating that not accounting for the measurement errors results in a significant model misspecification.

Table 1.3: Estimation of the Euler equation with habit formation[1]

Parameters	Nonparametric ME				Log-normal ME			No ME
	Internal habit			External habit				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
γ	5.279	5.089	4.950	6.024	6.899	3.522	3.029	3.126
	(0.057)	(0.107)	(0.099)	(0.071)	(2.460)	(0.662)	(0.366)	(0.368)
β	0.987	0.973	0.929	0.983	0.955	0.897	0.956	0.801
	(0.068)	(0.013)	(0.043)	(0.025)	(0.023)	(0.049)	(0.030)	(0.170)
α	0.831	0.801	0.807	0.800	0.780	0.573	0.799	0.585
	(0.016)	(0.028)	(0.028)	(0.011)	(0.092)	(0.128)	(0.115)	(0.252)
a				0.473				
				(0.014)				
σ^2					0.042	0.072	0.036	
					(0.012)	(0.008)	(0.002)	
% noise					35	45	32	
Some savings			yes				yes	
Aggregate shocks		yes				yes		
J statistic	29.1	23.0	25.6	31.4	33.5	23.2	30.6	619.0
p value	0.933	0.902	0.978	0.858	0.873	0.936	0.936	0.000
N	3,402	3,402	1,754	3,402	3,402	3,402	1,754	3,402

[1] Number of time periods T = 10. Standard errors in parenthesis.
[2] A blank response in this row indicates that the full sample is used in estimation.

1.6.1 Habit formation

As discussed in Section 2, habit formation exists in the multiplicative model if $\gamma > 1$ and $\alpha > 0$, where α measures the strength of habit formation. In all specifications, γ is estimated to be significantly greater than 1 and α is estimated to be significantly greater than 0. Thus the estimation results support the existence of habit formation in individual food consumption.

In the specifications where measurement errors are nonparametrically accounted for, the estimates of α are precise and found to be between 0.80 and 0.83, indicating significant strength in habits. Restricting the distribution of measurement errors to the log-normal family results in point estimates of α between 0.57 and 0.80. This suggests that the assumption of a lognormal distribution for measurement errors is likely a misspecification that results in a downward bias in the estimate of α. Furthermore, when measurement errors in consumption are not accounted for, the estimate of α fall to 0.59, suggesting that ignoring measurement errors in food consumption results in a significant downward bias in the estimate of the strength of habit. Notice that the same conclusion was drawn from the simulation exercise.

The results indicate that allowing for aggregate expectations errors or restricting the sample to families that are not liquidity constrained does not significantly affect the estimates of the strength of habit.

The basic model is specified to be consistent only with habit being internal to the individual. As an extension of this basic model, we allow for external habits in individual consumption by augmenting the definition of consumption services as follows:

$$\tilde{c}_{it} = \frac{c_{it}}{c_{it-1}^{\alpha} C_{it-1}^{a}},$$

where C_{it} is period t average consumption of households that belong to income group that household i belongs to, and $0 \leq a \leq 1$ measures the strength of external habits. Aggregate consumption is constructed for 4 different income groups with roughly the same number of

households in each of them. We assume that measurement errors in aggregate consumption C are averaged out. Then the external habit counterpart of equation (1.3.5) is:

$$E\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}G^a_{it}}\right)^{1-\gamma}\left(\kappa_1-\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}G^a_{it+1}}\right)^{1-\gamma}\right)-\right.$$
$$\left.\left(\kappa_2-\alpha\beta\kappa_3\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}G^a_{it}}\right)^{1-\gamma}\right)\middle| z^*_{it}\right] = 0,$$

where $G_{it} = C_{it}/C_{it-1}$.

The estimation result is reported in column (4) and indicates that, in addition to internal habits, external habits are significant in explaining household food consumption patterns. However, the strength of external habit is significantly smaller that the strength of internal habit, with a estimated to be 0.47 while α is estimated to be 0.80. Therefore, while internal habit formation has the dominant effect, external habit formation also plays an important role in explaining consumption patterns.

Meghir and Weber (1996) suggests that the finding of habit formation in food consumption may be explained by nonseparabilites in preferences over food and other consumption goods. Carrasco at al.(2005) find that this result of Meghir and Weber is largely due the presence of time invariant unobserved heterogeneity that Meghir and Weber were unable to control for due to the small length of the panel used in their estimation. Our estimation method does control for time invariant heterogeneity. Furthermore, if nonseparabilites in preferences over food and other consumption goods represents a significant misspecification in our model, then it is likely that the J test would reject our preferred specifications, as it rejects the specification that ignores measurement errors.

1.6.2 Utility curvature

The estimates of the utility curvature parameter, γ, after accounting for measurement errors nonparametrically are found to be between 4.95 and 5.28 for the model with only internal habits. In the model where external habit are also accounted for, γ is estimated to be 6.02.

This latter estimate is similar to Fuhrer (2000), who assumed external habits and estimated γ to be 6.11. These parameters are estimated precisely.

When the distribution of measurement errors is restricted to be log-normal, the estimates of γ become less precise, ranging between 3.03 (with standard error of 0.37) to 6.90 (with standard error of 2.46) Not accounting for measurement errors results in a significant downward bias in the estimate of γ, falling to 3.13. This is intuitive, as we would expect a smaller utility curvature parameter with more volatile consumption. By accounting for measurement errors, we control for the excess noise in the observed consumption series.

1.6.3 Discount factor

The discount factor is estimated precisely to be between 0.90 and 0.99 in all specifications except when measurement errors are ignored. Not accounting for measurement errors results in lower and less precise point estimate for the discount factor, falling to 0.8.

1.7 INTERTEMPORAL ELASTICITY OF SUBSTITUTION AND RISK AVERSION

In this section, we analyze the intertemporal elasticity of substitution (IES) and the relative risk aversion (RRA) that is implied by the estimates obtained in the previous section. In the presence of habit formation, the IES and the RRA varies across individual and time. Specifically, Appendixes D and E show that the inverse IES and the RRA take the following form:

$$\frac{1}{IES_{it}} = \gamma - \frac{\alpha\beta(1-\gamma)\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}}{1-\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}} - \frac{\alpha^2\beta(1-\gamma)\varphi_{it+2}\left(\frac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}}{1-\alpha\beta\varphi_{it+2}\left(\frac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}}, \quad (1.7.1)$$

$$RRA_{it} = \frac{\gamma - \alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma} - \alpha^2\beta(1-\gamma)\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}}{1-\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}}. \quad (1.7.2)$$

Three important facts about IES and RRA are made clear by observing equations (1.7.1) and (1.7.2). First, the inverse IES and the RRA are higher for habit forming consumers than for non-habit forming consumers. Second, the habit formation model implies heterogenous IES and RRA. Recent studies that allow for heterogeneity in risk aversion find significant variation in aversion to risk across different groups of individuals. Third, the habit formation model breaks the tight inverse relationship between the IES and the RRA of iso-elastic specification of preferences. These observations imply that the habit formation model is able to explain varying IES and RRA across groups of individuals in ways that the iso-elasticity models cannot.

Because true consumption is not observed, the IES and RRA are generally not observed. Furthermore, the conditional expectation of the inverse IES and the RRA do not conform to the transformation employed to derive the estimator because their functional forms do not satisfy the conditions used to separate true consumption from measurement errors. Therefore, the expectation of the (inverse) IES and the RRA are in general not directly recoverable from equations (1.7.1) and (1.7.2). However, it is possible to construct bounds for the conditional expectation of these quantities given the set of instruments z. The proof of the following proposition is given in Appendix D.

Proposition 1.7.1. *Suppose Assumption 1.3.3 holds. Then*

$$\begin{aligned}\tfrac{1}{\gamma} \geq E\left[IES_{it}|z_{it}\right] \geq \quad & \Bigg(\gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j |z_{it}\right] \\ & -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)^j |z_{it}\right]\Bigg)^{-1},\end{aligned}$$

and

$$\gamma \leq E\left[RRA_{it}|z_{it}\right] \leq \gamma - (1+\alpha)(1-\gamma)\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j |z_{it}\right],$$

with strict inequalities if $\alpha > 0$, $\beta > 0$ and $\gamma > 1$.

Proposition 1.7.1 can be used to construct bounds for the unconditional expectation of the IES and the RRA.

As shown in the next proposition, the assumption that measurement errors are independent and log-normally distributed results in point identification of the expectation of the RRA. However, the expectation of the IES is still only partially identified, but typically with narrowed bounds relative to the more general case above.

Proposition 1.7.2. *Suppose Assumptions 1.3.3 and 1.3.4 hold. Then*

$$\begin{aligned}\tfrac{1}{\gamma} \geq E\left[IES_{it}|z_{it}\right] \geq \quad & \Bigg(\gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j|z_{it}\right] \\ & -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)^j|z_{it}\right]\Bigg)^{-1},\end{aligned}$$

with strict inequalities if $\alpha > 0$, $\beta > 0$ and $\gamma > 1$, and

$$E\left[RRA_{it}|z_{it}\right] = \gamma - (1+\alpha)(1-\gamma)\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j|z_{it}\right],$$

where $A_3 = \exp\{\sigma^2\left((1+\alpha+\alpha^2)(1-\gamma)^2\right)\}$.

To see that the bounds on $E\left[IES_{it}|z_{it}\right]$ defined in Proposition 1.7.1 are typically wider than the one defined in Proposition 1.7.2 under the independent log-normal assumption, notice that $\mathcal{A}_3 > 1$ so that $j > 1$, $\mathcal{A}_3^{j^2} > \mathcal{A}_3^j$. Each additional term in the infinite sums in Proposition 1.7.2 is smaller than the corresponding term in Proposition 1.7.1.

Given the parameter estimates of the model, the bounds on the IES and the RRA can be approximated with high precision by replacing the infinite sum with a finite approximation. Consistency of these estimators of the bounds would depend on allowing the order of approximation to increase with the sample size. The results of Ai and Chen (2003) can be used to prove this consistency conjecture. However, this is beyond the scope of the current paper and is left to future work. With this in hand, asymptotically valid confidence sets can be defined for these bounds. Construction of these confidence sets are found in Horowitz and Manski (2000) and Imbens and Manski (2004).

Table 1.4: 95% confidence intervals for the IES, inverse IES and RRA

	Computed using the parameter estimates in Table 1.3, column:			
	(1)	(2)	(3)	(4)
IES	[0.086, 0.189]	[0.086, 0.197]	[0.084, 0.202]	[0.080, 0.166]
RRA	[5.167, 13.093]	[4.879, 13.020]	[4.756, 13.021]	[5.885, 13.074]

Table 1.4 presents 95% confidence intervals for the expectation of the IES and the RRA with estimated parameters taken from selected specifications in Table 1.3. To estimate the bounds involving infinite sums, a fifth-order approximation of the infinite sum is used. 20 bootstrap draws from the estimated asymptotic distribution of the estimated parameters are used to compute the standard error of these bounds. The columns of Table 1.4 are labeled (1), (2), (3) and (4) to correspond to the columns of Table 1.3, that it, the resulting parameter estimates from estimation of the model with non-parametric measurement errors .

The estimated bounds for the IES support typical findings in the literature on estimation of wealth and consumption models. In the habit formation framework Naik and Moore (1996) report the IES close to our estimates. Similar range for the IES is found over households or certain cohorts of individuals (see Barsky, Juster, Kimball and Shapiro (1997)). However larger values of the IES are also reported in the literature.[8]. It is important to note that as we investigate household food consumption, the estimated bounds for the IES seem to be

[8]See, for example, Attanasio and Weber (1993), Atkeson and Ogaki (1996), Vissing-Jorgensen (2002)

reasonable: the consumption of food is likely to be rather inelastic.

The presented bounds on the RRA suggest that the estimates for this parameter are somewhat higher than the prevalent estimates in the range between 1 and 4 in empirical studies of consumption models with no habit formation. Although, recent developments in empirical estimates of risk aversion suggest that this parameter as high as between 4 and 16 can be found if one accounts for individual heterogeneity in risk aversion.[9]

To investigate the existence and significance of heterogeneity in the IES and the RRA, we regress the calculated individual-time specific IES and RRA on a set of regressors. The caveat in this regression analysis is that observed consumption instead of true consumption is used to compute the individual-time specific IES and RRA. Because observed consumption is contaminated with measurement errors it is possible that the estimated coefficients are biased. As is well known, if the dependent variable in a linear regression model is contaminated with an additive measurement error component, the coefficients are still unbiased and consistent under standard assumptions. However, in our cases, the dependent variable is contaminated by measurement errors in a nonlinear way. Therefore, the above argument does not carry over to our framework and we should expect that the coefficients estimates from linear regression will be biased. It is therefore important to understand the direction of potential bias introduced by measurement errors in observed consumption. The additional noise introduced to the dependent variable is independent of the other variables in the regression function so the expected effect of this additional noise is that it biases the estimated coefficients towards zero. Therefore, the regression results should be interpreted as being biased against detecting statistically significant explainable heterogeneity in the IES and the RRA.

Table 1.5 reports the results from the regression of observed IES and RRA on a constant, lagged income, lagged income squared, a dummy for high school graduate (HG), a dummy for college graduate (CG), age, and age squared. Observed IES and RRA are computed using the estimated parameters from columns (1) - (4) of Table 1.3 and the columns labelled

[9] Alan and Browning (2010) provide a discussion on the recent literature and show the estimation results supporting this range of the risk aversion parameter.

Table 1.5: Regression of estimated RRA and IES on income and age

	RRA				IES			
	Parameter estimates in Table 1.3, column:				Parameter estimates in Table 1.3, column:			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Constant	5.840	6.550	8.043	9.331	0.112	0.112	0.107	0.095
	(1.789)	(1.613)	(1.958)	(2.063)	(0.007)	(0.007)	(0.009)	(0.007)
Income_{t-1}	-0.597	-1.062	-2.341	-1.996	-0.001	0.001	0.008	0.005
	(0.765)	(0.696)	(0.840)	(0.881)	(0.003)	(0.003)	(0.004)	(0.003)
Age_t	0.373	0.323	0.248	0.290	-0.001	-0.001	-0.001	-0.001
	(0.089)	(0.080)	(0.097)	(0.102)	(0.0003)	(0.0003)	(0.0004)	(0.0003)
Age_t^2	-0.004	-0.003	-0.002	-0.003	0.00001	0.00001	0.00001	0.00001
	(0.001)	(0.0009)	(0.001)	(0.001)	(0.000004)	(0.000004)	(0.000005)	(0.000004)
R^2	0.0053	0.0044	0.0046	0.0036	0.0046	0.0033	0.0029	0.0033

(1), (2), (3), and (4) of Table 1.5 correspond to the columns of Table 1.3.

The regression results indicate that the IES is increasing and convex in both age and income, while the RRA is decreasing and concave in both age and income. These results are consistent across the different specifications. However, while the effect of age on the IES and the RRA is monotonic throughout the sample, the effect of income on the IES is U shaped with a turning point at approximately \$41,000 for columns (1), (2), and (4), and \$35,000 for column (3). Also, the effect of income on the RRA is dome shaped with turning point at approximately \$36,000 for columns (1), (2), and (4), and \$34,000 for column (3). The results also display modest evidence that the RRA is increasing in the education of the head of the household. However, we do not find that education is a significant determinant of the IES.

In terms of individual actions, the results are interpreted as follows. Given income and education, the intertemporal allocation of food consumption by households with older heads is less responsive to intertemporal changes in food prices. Also, given income and education,

households with older heads are more likely to dedicate a larger percentage of their income to precautionary savings than their younger counterparts. Because the PSID contains the full age profile of interest (ages 22 - 65) in each wave, these results cannot be attributed to cohort effects. Among the other studies that investigate the effect of age on the IES or the RRA, very few recover significant patterns. However, Eisenhauer and Ventura (2003) do find similar increasing and concave patterns of the RRA in age.

Given education and age, households with moderate income (roughly between $30,000 and $40,000) are less responsive to intertemporal changes in food prices, and are more likely to save for precautionary motives. We are not aware of other studies that recover these nonlinear patterns of the IES and RRA with respect to income. In the work of Guiso and Piaella (2001) and Eisenhauer and Ventura (2003) there is some evidence that the RRA is increasing in income. However, these studies do not consider higher order effects of income on the RRA.

The results in Table 1.5 provides modest evidence that the RRA is increasing in the level of education of the head of the household. As expected, these coefficients are larger and more significant when income is omitted from the regression. Allan and Browning (2010) also find that less educated households tend to be less risk averse using a different estimation strategy. Given the iso-elastic preferences assumption of their model, Allan and Browing (2010) also interpreted their result as indicating that more educated households are more prudent. In our framework, the tight link between risk aversion and prudence is broken by habit formation. Thus, this conclusion cannot be extended to our framework.

It is worth reiterating that the results concerning heterogeneity of the IES and the RRA should be interpreted with the qualification that the observed IES and RRA are contaminated with measurement errors in a nonlinear way. Under the assumption that measurement errors are independent of income, education, and age, the result is that the additional noise introduced reduces the explanatory power of these variables. The results that are consistent and precisely estimated across the considered specifications are that: the IES is decreasing and convex in age; the RRA is increasing and concave in age; the IES depicts a U shape in

income; and the RRA depicts a dome shape in income. The other result that is consistent across specifications, but less precisely estimated is that the RRA is higher for households with more educated heads.

1.8 CONCLUSION

Habit formation in preferences have been used to explain a wide variety of macroeconomic phenomena. However, at the level of micro data, the evidence of habit formation in consumption is mixed. Previous micro studies that investigate habit formation using standard preferences impose arguably strong assumptions in order to obtain an estimating equation. The misspecifications that are due to these assumptions are likely to result in significant biases in the estimates. This intuition is confirmed in our simulation exercise. This paper develops a new exact nonlinear GMM estimator that account for measurement errors without the need for these assumptions. As a result of estimation, we find that habit formation is an important determinant of food consumption patterns.

Another advantage of the method developed in this paper is that, to the best of our knowledge, it is the first Euler equation GMM method that allows for the analysis of individual heterogeneity in two key economic parameters, the IES and the RRA. Other methods that analyze heterogeneity in these parameters do so either in experimental frameworks, survey studies, or by methods where the expectations errors are parametrically specified. The method presented here allows for comparable analysis using nationally representative data set such as the PSID, and without the need to specify the distribution of the measurement errors.

There are some extensions to the model presented in this paper that can be pursued in future work. One such consideration is to extend the model to allow for more flexible patterns of habit formation. In this paper, we extended the baseline *internal* habit formation model to allow for *external* habit. Another possibility is to allow for more general model

specifications of internal habits. The current model assumes that internal habit is a function of only the last period consumption. The model and estimation method can be extended to include additional lags, but at the cost of a smaller number of time periods in which to recover parameters that dictate consumption patterns. Another potentially fruitful direction for future work is to estimate the parameters of the model using consumption goods and services other than food consumption. One would then be able to test if the implied IES and RRA vary across these consumption groups.

That being said, this paper proposes a direct GMM estimator of Euler equations with nonseparabilities, which can also be used to investigate individual heterogeneity in the IES and RRA. We find that the IES is decreasing and convex in age, and that the RRA is increasing and concave in age, and increasing in education. These findings are consistent with those of other recent studies. The new findings are that the IES is U shaped in income and the RRA is dome shaped in income. These findings warrant further analysis because of, among other things, their implications for idiosyncratic consumption and savings responses to various economic policy interventions such as interest rate taxation and fiscal policies.

2.0 DYNAMIC STOCK MARKET PARTICIPATION OF HOUSEHOLDS

2.1 INTRODUCTION

Over the past decade the limited stock market participation phenomenon has received growing attention in empirical and theoretical work. Despite the theoretical prediction that all investors will participate in stock markets if the equity premium is positive, empirical evidence shows that a substantial fraction of consumers do not participate in stock markets either directly or indirectly (via mutual funds and similar institutions). One of the leading explanations to the phenomenon is that it is costly to invest in stocks, and these costs arise from investor's inertia, time and effort that consumers have to spend to obtain and process financial knowledge and information, follow the situation on financial markets, pay sign-up fees and file the necessary paperwork.

The profile of an average stock holder differs from the one of a non-participant in that the former is older and more educated. It is also shown that stockholders are more educated financially. Apparently, one of the ways to increase stock market participation is to provide financial education to consumers who are not aware about all financial products available. The importance of promotion of financial education was acknowledged a decade ago with launching financial education curricula via a number of economic policy incentives.[1] Further, the issue of limited stock market participation was reinforced by a growing concern in the literature and in economic policy debates (see Lusardi, van Rooji and Alessie (2007) and Guiso

[1]Examples are The Economic Growth and Tax Relief Reconciliation Act, and Money Smart, a program that is launched by The Federal Deposit Insurance Corporation.

and Japelli (2005)). These debates resulted in a number of financial education programs that are designed to further increase and promote financial literacy among consumers.[2]

The objective of the paper is to investigate whether the financial education programs can influence consumers' financial choices and increase participation in financial markets. I use micro data to estimate a dynamic model of stock market participation, where consumers' decisions regarding stock market participation are influenced by participation costs. Participation costs serve as a channel, through which the financial education programs can affect the consumers' investment decisions. Financial education and counseling alleviates the burden of consumers' time and effort needed to make financial decisions and reduces the participation costs.[3] I show that participation costs are influenced by past stock market participation experience and by consumers' age.

Stock market participation costs are not observed by researchers, which comprises the major difficulty in estimating it. Empirical studies of Paiella (1999), Vissing-Jorgensson (2002) and Andersen and Nielsen (2011) provide evidence that small fixed participation costs have a potential to explain the observed low stock market participation rate. In theoretical simulations Haliassos and Michaelides (2003) and Gomes and Michaelides (2003) show that the non-participation in stock markets can be rationalized by existence of an entry cost or participation cost. However, there is no agreement in the literature on how large these participation costs may be. Also, there is a distinction in the literature between per period participation costs and fixed entry costs. Empirical studies estimate the former, while the latter is calibrated in theoretical simulations of investor's life cycle paths.

The up to date empirical estimates of per period fixed participation costs can only provide the lower bound and suggest that the true (unobserved) participation cost exceeds these levels in reality. Attanasio and Paiella (2011) and Paiella (2007) estimate the lower bound to the forgone gains of holding an incomplete portfolio in units of non-durable consumption as low

[2]The latest incentive is The Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010.

[3]Consumers do not necessarily need to go to financial consultants in order to receive financial advice. One of the ways to increase one's financial awareness is through the interaction with co-workers and neighbors. As an example, Duflo and Saez (2000) and Hong, Kubik and Stein (2004) provide evidence on peer effect as a determinant for participation.

as 0.4% and 0.7-3.3% of consumption per year.[4] On the other hand, theoretical simulations of stock market entry cost calibrate it as a fixed share of expected annual income. Haliassos and Michaelides (2003) obtain a wide range for entry costs from 3% to 34% of mean annual income. Alan (2006) finds stock market entry cost about 2% of annual permanent income. I consider these studies as a guidance on the possible magnitudes of participation costs and take one step further to explore the heterogeneity in it. Further, I show how participation costs change as a consumer (a household in my study) grows older, obtains stock market participation experience and finally, receives educational services that may affect his financial literacy and awareness.

In developing the model I relate to a literature that focuses on the determinants of heterogeneity in stock market participation and portfolio choice. Bertaut and Starr-McCluer (2002), Guiso and Japelli (2002), Banks and Tanner (2002), and Brunnermeier and Nagel (2008) provide evidence that wealth, age and education are important factors that affect risky asset investment decisions. The choice and intensity of stock market participation are also affected by past participation and investment choices as well as expectations of the future returns. Likewise, the portfolio choice is not independent from past investment choices. This dynamic dependence is confirmed empirically by Alessie, Hochguertel, and van Soest, 2002, 2004, and Munoz (2006) among others. Similar to this strand of literature, I allow for controlling for observed heterogeneity represented by investor's wealth, age and education. However, econometric framework distinguishes this paper from above empirical studies: the estimated model is fully structural and allows for endogenous investment choices both in the participation decision and the share of wealth invested in stocks.

The estimation results provide evidence that the participation cost, measured as a share of income can be substantial. The average stock market participation cost is estimated to be about 5% of labor income. However, it is not constant over the life cycle. The results show that the participation cost is decreasing with education and age. When age is allowed

[4]These estimates translate into $72 bound for Attanasio and Paiella (2011) and $130 bound for Paiella (2007). Vissing-Jorgensen (2002) finds that per period fixed cost as low as $260 can explain the behavior of majority of nonparticipants.

to enter participation cost in a quadratic form, participation cost becomes increasing in concave in age, with the turning point around age 35-40. Participation cost is smaller for consumers who invested in stock market in the past. The model estimates are used to conduct simulation exercises to investigate how the participation decision is affected when consumers are provided financial consulting. Their participation rate is greater compared to that of baseline consumers.

2.2 THE MICROECONOMIC PICTURE: DATA AND DISCUSSION

2.2.1 Data

The analysis of direct stock market participation is a part of a bigger problem of investment in risky assets. Diversification of extended financial portfolio, which includes other risky assets (e.g. investment in business or farm), may drive households away from direct stock market participation. In addition, changes in housing can provide another reason to withdraw savings from stock markets. Finally, deep changes in family composition (for example, new head of household) can influence changes in household portfolio. These factors are important and deserve to be explored in detail on their own. However, in the current study I exclude the influence of these factors, and consider households that satisfy the following restrictions: (i) do not invest in business and / or farm; (ii) do not experience changes in housing; (iii) have the same head of household for the whole observation period.

Data on household liquid wealth, income and demographic characteristics are taken from the Panel Study of Income Dynamics (PSID). Wealth supplement of the PSID is available every 5 years between 1984 and 1999. After 1999 the data become available every other year. In my analysis I use the later data starting from 1999 through 2007. Table 2.1 shows summary statistics for the data that I use in my analysis.

Wealth and Income

The value of households' stock holdings is observed in the PSID as "non-IRA stock

holdings".[5] Total liquid wealth is computed as a sum of non-IRA stock holdings, money in checking or savings accounts, money market funds, certificates of deposit, government savings bonds, or Treasury bills, bond holdings and money in other savings or assets, such as bond funds. The value of liquid wealth that in not invested in stocks is considered as a risk-free asset. Total household income consists of labor income, financial income (interest, dividends, income from trust fund) and other money inflow (child support, help from relatives, rent, worker's compensation) of head of household and the spouse. Top coded observations on wealth and income are excluded from the sample. I also remove extreme outliers by excluding observations for wealth and income at 99-percentile from above and 1-percentile from below. Data on income and wealth are deflated using the CPI for the end of the year before the interview was conducted. CPI deflator is taken from the consumer price index releases of the Bureau of Labor Statistics.

Table 2.1 shows that the difference in both income and wealth between participants and non-participants is dramatic. On average non-participants have lower income. The accumulated wealth of nonparticipants is substantially lower as well. However larger average income and wealth of stock market participants are accompanied by larger standard deviations.

Demographics

Demographic characteristics include age, education and occupation of the head of household, as well as marital status and family size. Individual consumption and savings behavior differs depending on in what stage of life a consumer finds himself: prime working age or retirement. This study mostly considers prime age consumers and excludes households whose head is younger than 22 or older than 65 over the period of estimation. I exclude households whose marital status changed over the sample period.

Table 2.1 shows that there are differences in demographic characteristics of participants and non-participants. On average, stock market participants are 2-3 years older and more educated. Their occupation is on average more related to management, business operation

[5]The PSID separately provides information on total amount held in individual retirement accounts and an approximate split of the amount in IRA between interest- and dividend-earning assets. However, the exact value of IRA in stock holdings is not available.

Table 2.1: Summary Statistics

Year		1999	2001	2003	2005	2007
Observations		1,332	1,424	1,382	1,481	1,524
Income and Wealth[a]						
Income		39,420.6	39,532.9	38,748.0	40,040.0	40,252.8
	Non-participants	34,716.9	35,363.8	35,698.1	36,913.2	36,965.0
		(19,885.0)	(21,087.4)	(21,198.6)	(22,709.3)	(22,290.0)
	Participants	50,740.6	50,243.1	46,938.1	49,565.4	51,530.6
		(29,623.8)	(27,695.0)	(29,134.3)	(30,019.2)	(29,623.2)
Wealth		21,525.4	21,974.9	23,793.2	22,864.0	21,859.5
	Non-participants	10,407.1	10,725.3	11,440.3	10,338.6	10,532.2
		(23,649.3)	(25,346.4)	(27,277.2)	(22,870.9)	(25,020.8)
	Participants	48,283.5	50,874.2	56,965.1	61,022.1	60,714.7
		(58,981.7)	(64,580.2)	(71,231.2)	(74,778.4)	(71,578.9)
Stock holders		0.29	0.28	0.27	0.25	0.22
Share of wealth in stock		0.61	0.58	0.53	0.55	0.57
Stock holdings of participants		32,883.1	32,979.9	32,451.9	37,838.5	37,270.7
Demographic data						
Age		43.3	44.3	45.1	46.1	46.0
	Non-participants	42.6	43.7	44.6	45.4	45.4
	Participants	44.8	46.0	46.5	48.3	48.0
Family size		3.03	3.00	2.93	2.94	2.89
	Non-participants	3.10	3.02	2.97	2.99	2.92
	Participants	2.87	2.94	2.82	2.80	2.78
Education		13.7	13.6	13.6	13.7	13.7
	Non-participants	13.3	13.3	13.3	13.4	13.4
	Participants	14.6	14.6	14.5	14.6	14.7
Marital status		0.70	0.70	0.70	0.71	0.71
	Non-participants	0.69	0.67	0.67	0.68	0.68
	Participants	0.74	0.78	0.76	0.79	0.79
Occupation[b]		n.a.	n.a.	0.16	0.17	0.16
	Non-participants			0.13	0.15	0.14
	Participants			0.23	0.23	0.26

[a]Standard deviation is reported in parenthesis.
[b]Equals to 1 if related to management, business operation or financial specialist; 0 otherwise.

or financial specialist. There are more married individuals among stock market participants compared to non-participants. The difference in family composition is not substantial.

Rates of Return

The uncertainty about returns on risky assets plays an important role in the dynamic decisions about participation and portfolio composition. Additional information is needed to assess the realizations of the household-specific rates of return on stock holdings and risk-free assets.

The PSID provides data not only on the value of stock holdings, but also the amount that was invested in or taken out of stocks between periods $t-1$ and t. Then the household specific rate of return on stock holdings R^s_{it} can be computed using the household data on total value of stock holdings and sales (purchases) of stocks from the constraint $S_{it} = s_{it} + R^s_{it} S_{it-1}$ where S_{it} is the total value of stock held by household i, s_{it} is the value of purchases or sales of stock between periods $t-1$ and t for household i. Unfortunately, the data in wealth supplement of the PSID are known to suffer from systematic underreporting of trades. The "forgotten" trades affect the computed idiosyncratic rates of return and often make them either unrealistically large or small. This caveat makes using the computed R^s_{it} as a rate of return on household's stock holdings problematic. However, I still use the information on the computed R^s_{it} to construct an indicator on whether a household received a high, moderate or low return on stock holdings in the current period, conditional on the participation in the previous period. I construct the indicator for high, moderate or low rate of return by matching the household specific rate of return on stock holdings R^s_{it} with the distribution of the market index rate for the period. I use SP500 as such index. The real risk-free rate R^f_t is constructed based on seasonally adjusted deflated average 6-month Treasury bill.

The summary statistics for the risk-free rate and rates of return on stock holdings are reported in Table 2.2. The time period that is considered in the paper can be characterized as years of turbulence in financial markets. It is reflected in the rates of return on market indices that vary substantially over the reported period. In spite of the deficiencies in the reported stock holdings, the median for the computed household-specific rate of return R^s

admits reasonable values and follows the trend of market indices.[6] One can see that there are two "regimes" in the empirical distribution of the computed rates of return. Over time period from 1999 to 2002 smaller values of returns were prevalent. These years were most affected by financial markets turbulence over the observed period. In 2003 to 2007 small, medium and large returns showed similar weights. This is consistent with the period of stability in financial markets.

Data Sets Used in Estimation

For the analysis that follows I take into account the data restriction discussed above and construct 3 data samples: data sample used in estimation of the main model; data sample with extended time frame for estimation of the earnings' equation and data sample with extended number of individuals to estimate the conditional choice probabilities.

The first stage of analysis involves the estimation of the individual effects. To reduce the bias in individual effects one needs a data sample with larger time dimension. Individual fixed effects are estimated using the earnings' equation. Unlike the wealth data, labor income is reported in the PSID in all time periods well beyond the period for which wealth supplement is available. To estimate the earnings' equation, I take the extended data set starting from 1979 through 2007. To be consistent with the main model I use data on odd years only as the main model uses data on odd years as well. Using the age and income restrictions, described above, I construct the extended data set of 7,744 households that contains data on household's labor income, age, education and family size.

The second stage of the analysis estimates conditional choice probabilities that also require a large number of observations, especially so on the cross section dimension of the panel. This data set is constructed using wealth variables, among others, therefore has time dimension from 2001 (accounting for one lag) through 2007. To ensure larger number of observations, I drop the age and marital status restrictions discussed above. Namely, I use households of all ages and disregard the changes in marital status while still controlling for marital status among other family characteristics. The conditioning set includes the

[6]The average household specific return is not so meaningful due to the substantial number of computed rates that are either very large or very small, depending on the time period.

Table 2.2: Rates of return

	1999	2000	2001	2002	2003	2004	2005	2006	2007
R^f	0.020	0.026	0.020	-0.008	-0.008	-0.019	0.000	0.024	0.000
Nasdaq	0.581	-0.246	-0.242	-0.307	0.547	-0.021	0.082	0.056	-0.063
SP500	0.070	-0.044	-0.165	-0.249	0.305	0.016	0.043	0.105	-0.079
R^{s*}	-0.029		-0.114		0.067		0.024		
$\mathcal{F}_t^L$	0.428		0.427		0.344		0.318		
$\mathcal{F}_t^M$	0.277		0.283		0.327		0.371		
$\mathcal{F}_t^H$	0.295		0.290		0.329		0.311		

* R^s is median of ex post annual rate between two consecutive time periods, calculated from the wealth supplement of the PSID and reported for stock market participants only. $\mathcal{F}_t^k$ for $k = \{L, M, H\}$ is the probability of the individual return of stock R^s be one standard deviation lower,in between, or greater than the average return on SP500 for the observed time period, calculated from the data. Returns are inflation adjusted.

following covariates: past wealth, past share in wealth invested in stocks, return on household portfolio, income, age, education, occupation, family size, marital status as well as estimated individual effects and time dummies from the earnings' equation. All in total, I have 10,708 observations available over 4 time periods.

Finally, the main sample is used to estimate preference parameters and parameters of the participation cost. The main data sample is constructed as an unbalanced panel of 299 households observed over 3 periods: 2001, 2003 and 2005. Two other time periods (1999 and 2007) we lost while taking one lag and one lead. These households are observed participating in stock markets in the current period. Two comments are due to mention. First, while the main model is estimated with stock market participants, due to the way the estimator is constructed, the information on non-participants is fully used in the estimation of the conditional choice probabilities and is, therefore, incorporated into the main sample through CCP's. Second, the base of the estimator is an identity equation that holds equally for both participants and non-participants. Therefore, self selection issue cannot be a problem.

The set of model covariates includes past and current wealth, past and current share of wealth invested in stocks, current and future income, current demographic characteristics, as well as the estimates of the conditional choice probabilities and transition probabilities. Transition probabilities $\mathcal{F}_t^k$ for $k = \{L, M, H\}$ are computed from the data as reported in Table 2.2. The instrument set to form orthogonality conditions include variables from the state vector: past share of wealth invested in stock, past portfolio allocation, return on household portfolio, income as well as family characteristics.

2.2.2 Regression Analysis

Table 2.1 shows that the observed characteristics of households in the sample differ significantly between stock market investors and non-investors. The difference between participants and non-participants is especially striking along the dimension of liquid wealth, income, age and education. In this section I analyze the probability of investing in stocks and the share of wealth invested in stocks. I estimate Heckman sample selection model to determine the

Table 2.3: Regressions for stock market participation and share of wealth invested in stocks, Heckman two-step estimator.

Variable	Participation	Share
History of participation in the past	1.337^{*} (0.050)	
Income	0.069^{*} (0.026)	0.010 (0.008)
$\text{Income}^2/1000$	-2.841 (1.991)	-0.235 (0.570)
Wealth_{t-1}	0.068^{*} (0.012)	0.004 (0.004)
$\text{Wealth}^2_{t-1}/1000$	-1.244^{**} (0.486)	-0.007 (0.134)
Age	0.032^{***} (0.019)	0.005 (0.008)
$\text{Age}^2/1000$	-0.215 (0.154)	-0.045 (0.061)
Education	0.078^{*} (0.011)	0.001 (0.004)
Male	0.040 (0.084)	-0.008 (0.032)
White	0.444^{*} (0.065)	-0.007 (0.031)
Married	-0.019 (0.039)	-0.042 (0.031)
Family size	-0.009 (0.026)	-0.013 (0.015)
Number of children	-0.006 (0.047)	0.029 (0.018)
Constant	-3.977^{*} (0.622)	0.547^{***} (0.294)
Mills		-0.047^{**} (0.021)

*,** and *** denote 1%, 5% and 10% significance level. Regression uses 5107 observations, which includes 1415 uncensored observations. Age cohort dummies, time dummies and region dummies are included in all regressions. Standard errors are in parenthesis.

covariates that are particularly powerful in explaining the decision to invest in stock markets and portfolio allocation of stock market participants.

The estimation of the participation equation and the share of wealth in stocks is reported in Table 2.3. The average participation rate in the sample is 24%. The history of past participation in stock markets has a dramatic effect on the participation rate, it increases the probability of current participation by 44%. Participation rate is increasing and concave in past wealth. It also increases with income. After controlling for age cohorts and time dummies, the effect of age on the decision to participate is still positive. More educated households are more likely to participate in stock markets. The probability of participation in stock markets is substantially greater for households with a white head.

All of the demographic covariates in the outcome equation for the share of wealth invested in stocks are statistically insignificant. Only constant is estimated marginally significant at the level of 0.54.

The evidence that age, education and past participation experience have a significant positive effect on stock market participation may support the existence of information cost to participation. Education generally proxies for the ability to collect and process information. Age and past participation can proxy for the accumulation of the information and experience. Generally, the regression analysis suggests, that participation cost may depend on these demographic characteristics and is likely to decline in education and age and be smaller for households that invested in stock markets before.

Participation in stock markets requires not only financial knowledge, but also time effort. The opportunity cost of participation can be measured as a fixed share of income. Indeed, the leading measure of the participation cost in the literature is the fixed amount of labor income (see Gomes and Michaelides (2003) and the subsequent studies). To relate to this literature I construct participation cost as a share of labor income.

2.2.3 Discussion

Stock market participation is costly in terms of effort and time spent by individuals to collect all necessary information and analyze how stock markets function. Gomes and Michaelides (2003) show that the participation cost have a potential to resolve the discrepancy between theoretical prediction that all investors participate in stock markets if the equity premium is positive, and empirical evidence that only a fraction of individuals participate in stock markets.

Investors face different costs, associated with stock holdings. Part of these costs are the direct transaction costs that have to be paid to the institution that facilitates the interaction between individual investors and stock markets. Brokerage fees can be one example of such costs. On the other hand, in order to enjoy the reward of greater expected return on savings through holding stocks, investor has to obtain a certain amount of knowledge about stock markets. The time and effort spend by individual investors to gain the knowledge represent the opportunity cost of stock market participation. Recent studies show that information barriers is an impediment to stock market participation. Honk, Kubik and Stein (2003) show that social interactions help to promote the participation in stock markets. Guiso and Japelli (2009) argue for the lack of awareness among investors about stock markets as an option to save. According to Alessie, Lusardi and van Rooij (2010), there is a causal effect of the lack of financial literacy on stock market participation.

The awareness and financial illiteracy arguments align with the recent developments in economic and financial policy. Some of the program are designed to regulate financial institutions, however their effect ultimately have to reach the consumers. The goals of one of such programs, Bureau of Consumer Financial Protection, is to set rules to the institutions that provide financial advisory services to promote financial literacy and help consumers to achieve financial independence. As a result, consumers become more informed and more aware of the impact of financial decisions they make. Another program, Improving Access to Mainstream Financial Institutions Act is designed to reach low- and medium-income households through financial education and counseling. This policy is aimed to improve

access to affordable and responsive financial products and services and plays its important role helping low-income households to build and maintain assets.

These economic and financial policies are an important step to help achieving the level of financial independence for individual investors. The economic outcome of these policy programs is yet to be revealed. However, it is possible to evaluate the potential impact of these policies through counterfactual simulations. In the following sections I develop the model of stock market participation, where consumers decisions regarding stock market participation are influenced by participation costs. Participation costs serve as a channel, through which the financial education programs can affect the consumers investment decisions. The model estimates are used to conduct counterfactual simulation exercises to investigate how the participation decision is affected when consumers are provided financial education and consulting.

2.3 MODEL

This section develops the theoretical framework to investigate household portfolio choices.

There are N households indexed by $i = 1, \ldots, N$, each with a lifespan of T periods. In each period, household i observes its accumulated wealth and a current rate of return on the portfolio composed of riskless and (possibly) risky assets. After household's income is received, household observes the amount of cash on hand to be allocated between a single composite consumption good and savings.[7] The household also decides how to reallocate savings between a risky and riskless asset. In each period the household decides whether to invest in stocks by choosing $d_{it} = 1$, or not participate in stock markets and keep all savings in riskless asset by choosing $d_{it} = 0$. If the household decides to invest in stocks, then the share of wealth allocated to stock holdings, represented by α_{it}, is chosen as well.

Let z_{it} be the state vector of household i in period t that is composed of observed and

[7] I do not model household labor supply decisions and assume that the households supply labor in each period.

unobserved characteristics. Let d_{jit} for $j = 0, 1$ be the indicator for participation choices: $d_{1it} = 1$ if household invests in stocks in period t, and $d_{0it} = 1$ if households does not participate in stock markets, where $d_{0it} + d_{1it} = 1$. In every period household i receives a utility payoff $u_j(z_{it})$ that depends on consumption and stock market participation choices. The current utility payoff from choosing j in time t is also affected by a choice-specific shock ε_{jit}, independent over i and t, which is revealed to the household at the beginning of period t. In every period $t = 1, \ldots, T$ household i chooses portfolio composition that solves the following problem:

$$\begin{aligned} &\max_{d_{it}} \mathbb{E}_{it} \sum_{\tau=t}^{T} \sum_{j=0}^{1} \beta^{\tau-t} d_{ji\tau}[u_j(z_{i\tau}) + \varepsilon_{ji\tau}] \\ &\text{subject to:} \quad w_{it} = R^p_{it} w_{it-1} + y_{it} - c_{it} - d_{it}\rho_{it}, \end{aligned} \tag{2.3.1}$$

where the expectation is taken conditional on the state vector z_{it}, $\beta \in (0, 1)$ is the subjective time-discount factor, y is the households income, c is consumption and w is total liquid wealth (stocks, bonds, cash accounts and similar liquid assets). Define α, a share of risky asset in households portfolio as $\alpha = S/w$, where S is the total value of stock holdings. R^s is a real return on risky asset held by household and R^f is return on riskless asset, combination of which gives the return on the portfolio $R^p_t = \alpha_{t-1} R^s_t + (1 - \alpha_{t-1}) R^f_t$. If household decides to invest in stocks, it gives up a fixed per period participation cost ρ. The problem (2.3.1) is maximized by a Markov decision rule.

Let $V(z_{it})$ be the (ex-ante) value function in period t, that corresponds to the optimization problem (2.3.1). The value function $V_j(z_{it})$ conditional on the choice j can be written as $V_j(z_{it}) = v_j(z_{it}) + \varepsilon_{jit}$, where $v_j(z_{it})$ is the conditional value function explained by z_{it} and ε_{jit} is a stochastic component that is not observed by an econometrician. Optimal participation choice involves comparing value functions associated with each choice. The optimal decision d^o_{it} of household i in period t can be expressed as follows:

$$d^o_{it} \equiv \begin{cases} 1 & \text{if} \quad v_1(z_{it}) + \epsilon_{1it} \geqq v_0(z_{it}) + \epsilon_{0it} \\ 0 & \text{otherwise} \end{cases} \tag{2.3.2}$$

From equation (2.3.2), the conditional probability of participation in stock markets can be represented as $p_{it} = p(z_{it}) = \mathbb{E}[d^o_{it}|z_{it}]$. Hotz and Miller (1993) prove the existence of a mapping where the difference in conditional value functions $v_1(z_{it}) - v_0(z_{it})$ can be represented as a function of conditional probabilities. Under assumption that $(\epsilon_{0it}, \epsilon_{1it})$ are identically and independently distributed over (i, t) as Type 1 extreme value random variables, the optimality condition (2.3.2) can be transformed into the following:

$$\ln \frac{p_{it}}{1 - p_{it}} = v_1(z_{it}) - v_0(z_{it}). \tag{2.3.3}$$

The left hand side of (2.3.3) can be estimated directly with the data. The right hand side can be expressed as a function of current and future utility payoffs. The details on the derivation of the right hand side of (2.3.3) follow.

Let $\mathcal{A}^{\tau}_{it}$ denote the set of all possible realizations of the state vector for household i at τ periods after t given the realization of the state vector z_{it} in period t. If the household takes action j at time t, then the state vector z_{it} transitions into z_{it+1} with the probability denoted as $F_j(z_{it+1}|z_{it})$. I use the results established in Hotz and Miller (1993) and Arcidiacono and Miller (2010) to write the conditional value function $v_j(z_{it})$ as follows:

$$\begin{aligned}
v_j(z_{it}) &= u_j(z_{it}) + \beta \sum_{z_{it+1} \in \mathcal{A}^1_{it}} V_{t+1}(z_{it+1}) F_j(z_{it+1}|z_{it}) \\
&= u_j(z_{it}) + \beta \sum_{z_{it+1} \in \mathcal{A}^1_{it}} [v_0(z_{it+1}) + \varphi_0(p_{t+1}(z_{it+1}))] F_j(z_{it+1}|z_{it}) \\
&= u_j(z_{it}) + \sum_{\tau=t+1}^{T} \sum_{z_{i\tau} \in \mathcal{A}^{\tau}_{it}} \beta^{\tau-t} d_{0i\tau} [u_0(z_{i\tau}) + \varphi_0(p_\tau(z_{i\tau}))] F_j(z_{it+1}|z_{it}) \prod_{s=t+2}^{\tau} \sum_{k=0}^{1} d_{kis} F_k(z_{is}|z_{is-1})
\end{aligned}$$

where the first equality establishes that the conditional value function is composed of the utility payoff of action j and the expected future value conditional on following the optimal decision rule from period $t+1$ on; the second line is written using forward substitution of the conditional value function and uses another result of Hotz and Miller (1993) that under assumption that $(\epsilon_{0it}, \epsilon_{1it})$ are identically and independently distributed over (i, t) as Type 1 extreme value random variables the value function can be expressed as a conditional value

function plus a function of the conditional choice probabilities[8]; the third line uses Theorem 1 in Arcidiacono and Miller (2010) to obtain a recursive representation of the conditional value function that is now composed of the current and future utility payoffs, functions of conditional choice probabilities and transition probabilities. The important result from this representation is that in many cases under selected choice alternatives the conditional value function exhibits finite dependence, such that only a small number of future time periods matter beyond the current period. Due to the finite time dependence property the current decisions affect only a limited number of time periods in the future.[9] The model presented in this paper can satisfy the finite dependence property.

In order to take advantage of finite dependence I limit the analysis to only a small number of participation strategies that proves sufficient for estimation of the parameters of interest. Let x_{it} be an L-dimensional vector of exogenous covariates for household i at time t. Income is treated as exogenous as well as forecastable and is a part of the vector of exogenous covariates. Define $\tilde{w}_{it-1} = R^p_{it} w_{it-1}$ as the period t value of household i's accumulated wealth. Then the observed state vector for household i at time t is given by the $(L+2)$-dimensional vector $(\alpha_{it-1}, \tilde{w}_{it-1}, x'_{it})$. The information set z_{it} is composed of the observed state vector, and the (unobserved) individual heterogeneity ν_i and aggregate shock ω_t, so that $z_{it} = (\alpha_{it-1}, \tilde{w}_{it-1}, x'_{it}, \nu_i, \omega_t)'$. Then I define a set of histories as $(2 \times 3 + L)$-dimensional vectors a_{1kit} and a_{0it} as:

$$a_{0it} = (\alpha_{it-1},\ 0\ ,\quad 0\ ,\tilde{w}_{it-1}, \tilde{w}_{it},\quad \tilde{w}_{it+1}, x'_{it}), \tag{2.3.4}$$

$$a_{1kit} = (\alpha_{it-1}, \alpha^*_{it},\quad 0\ ,\tilde{w}_{it-1}, \tilde{w}^k_{it},\quad \tilde{w}_{it+1}, x'_{it}) \tag{2.3.5}$$

where α^*_{it} is the optimal fraction of wealth a household chooses to invest in stocks, conditional on participating in stock markets. Under the strategy (2.3.4) a household chooses not to participate in stock market at date t, and then does not participate in period $t+1$. Likewise, under the strategy (2.3.5) a household chooses to participate in stock market at date t, but

[8] $\varphi(p_{it}) = \xi - \ln(p_{it})$, where $\xi \approx 0.576$ is Euler's constant

[9] Arcidiacono and Miller (2008) describe in detail how finite time dependence property can be applied to a broader class of models.

does not participate in period $t+1$. It is assumed that in period $t+1$ the choice of wealth held in riskless assets is the same under both strategies. Then beyond time period $t+1$ the terms in conditional value function become not consequential from the point of view of the optimal choices. To see this, notice that given the same unobserved characteristics, the observed state vectors for both strategies in period $t+2$, given by $(0, \tilde{w}_{it+1}, x'_{it+2})$, do not differ neither in wealth nor in portfolio allocation. Value functions for both strategies at time $t+2$ that are conditional on identical state vectors, are identical as well. These future value functions will cancel out while taking differences between them in equation (2.3.3), thus inducing finite dependence.

The uncertainty about the future wealth comes from uncertainty on future returns on stocks. In this case wealth at time $t+1$ depends on the realization of the rates of return on stock holding that is unknown to a household when it makes participation choices at time t. In order to integrate over the uncertain returns on risky portfolio allocations, I discretize all possible realizations of the returns on stock holdings denoting it as k to allow for up to K possible states conditional on investing, generally described as high, moderate and low realizations (H, M, L). Transition probabilities associated with the uncertainty on returns on stock holdings are degenerate if household chooses not to invest in stock market at time t. If household participates in stock market in the current period, then the probability that household's liquid wealth moves to one of the possible K states will depend on the realization of the return on household's portfolio moving over those K states. Transition probabilities are set to be independent from individual investor characteristics, so that $F(z_{kit+1}|z_{it}) = \mathcal{F}^k_{t+1}$.

For strategies (2.3.4) and (2.3.5) the equation (2.3.3) can now be represented as follows:

$$\begin{aligned} \ln\left[\frac{p_{it}}{1-p_{it}}\right] &= u(z_{it}, d_{it}=1) - u(z_{it}, d_{it}=0) \\ &+ \beta\left[\sum_{k=1}^{K}(u(z_{kit+1}, d_{it+1}=0) - u(z_{0it+1}, d_{it+1}=0)) + \ln(p_{kit+1}) - \ln(p_{0it+1})\right]\mathcal{F}^k_{t+1} \end{aligned} \tag{2.3.6}$$

2.4 ESTIMATION

Estimation of the equation (2.3.6) follows a three-step estimation strategy. Followed by specification of participation costs and households' preferences over consumption and investment choices, I derive the equation to be estimated. Next, I discuss estimation of the nuisance parameters. Nuisance parameters include estimates of fixed effects (estimates of unobserved individual heterogeneity), estimates of aggregate shocks as well as estimates of conditional choice probabilities. These parameters need to be estimated before the parameters in the utility and the participation cost are estimated. In particular, fixed effects and aggregate shocks are estimated to be included into the conditioning set of the value function and conditional choice probabilities. The estimates of conditional choice probabilities are then incorporated into the equation (2.3.6) that estimates parameters in the utility and the participation cost.

2.4.1 Preferences

Households derive utility from the consumption good c_{it}, denoted by $u^c_{it} = u(c_{it})$. I specify the utility of consumption as

$$u^c_{it} = c_t - \xi_t c_t^2$$

where ξ_t can be a constant or a function of household's demographic characteristics. Quadratic utility of consumption is one of the leading models in financial industry due to its desirable mean-variance portfolio selection properties. In the current paper, its use is appealing as it renders the estimation problem as linear. One of the potential disadvantages of the quadratic utility is that the negative marginal utility of consumption is not ruled out. Therefore, once the estimation of the utility parameters is achieved, it is important to rule out parameter values that allow marginal utility of consumption to be negative for all possible values of consumption data.

I also allow for non-pecuniary portfolio choice dependent utility cost from seeking advantage of higher equity premium through investing in stocks. The utility cost of adjusting

portfolio composition is denoted as $u_{it}^d = u(d_{it},\, \alpha_{it},\, \alpha_{it-1})$ and specified as:

$$u_{it}^d = \gamma_0 d_{it} + \gamma_1 \alpha_{it} \alpha_{it-1},$$

where γ_0 is the choice specific utility shifter and γ_1 is the parameter of adjusting portfolio allocation between time $t-1$ and t.

Then the utility payoff for household i is defined as $u(z_{it}, d_{it}) = u_{it}^c + u_{it}^d$ and depends on consumption, current and past investment choices, as well as other characteristics in state vector.

2.4.2 Participation Cost

Participation cost ρ_{it} is incorporated into the wealth accumulation budget constraint. In the literature it is common to impose a rather broad meaning on the costs of stock market participation for individuals. It embraces all fixed costs to participation that investors may incur. It may also be interpreted as an opportunity cost of time and effort for stock market participation.

To evaluate the behavior of households under the strategies (2.3.4) and (2.3.5) and compare the conditional value function for identical households taking different strategies, it is important that the participation cost is specified to depend only on the characteristics that are observed for each household regardless if this household participates or not.[10] I parameterize the participation cost as a function of household's labor income as a monetary measure of the opportunity cost, and also as a function of age, education, and past participation choices as proxies for the experience and the ability to absorb and process financial knowledge and information. Specifically, I write participation cost as a linear combination of household's observed characteristics:

$$\rho_t = \delta y_t^l x_t$$

where $x_t = (1,\, edu_t,\, age_t,\, age_t^2,\, d_{t-1})$ and y^l is labor income.

[10]For this reason I cannot use information related to stock buys and sells as well as total accumulated stock holdings because this information is available for participants only.

2.4.3 The Moment Conditions

I proceed to elaborate the equation (2.3.6). The conditional value functions associated with strategies (2.3.4) and (2.3.5) are:

$$\begin{aligned} v_1(z_{it}) &= c_t^1(1-\xi_t c_t^1) + \gamma_0 d_{it} + \gamma_1 \alpha_{it}\alpha_{it-1} \\ &+ \beta\left[\sum_{k=1}^{K} c_{t+1}^1(1-\xi_{t+1}c_{t+1}^1) + \ln(p_{1kit+1}) + \beta V_{t+2}(z_{t+2})\right]\mathcal{F}_{t+1}^k \\ v_0(z_{it}) &= c_t^0(1-\xi_t c_t^0) \\ &+ \beta\left[\sum_{k=1}^{K} c_{t+1}^0(1-\xi_{t+1}c_{t+1}^0) + \ln(p_{0it+1}) + \beta V_{t+2}(z_{t+2})\right]\mathcal{F}_{t+1}^k \end{aligned}$$

While the data on wealth and income are available, total household consumption is not reported in the PSID. To circumvent this problem, I compute the measure of household's consumption from the budget constraint:

$$\begin{aligned} c_{it} &= R_{it}^p w_{it-1} + y_{it} - w_{it} - d_{it}\rho_{it} \\ &= \hat{c}_{it}^p - d_{it}\rho_{it} \end{aligned}$$

where $\hat{c}_{it}^p$ is a function of household's data and the return on portfolio. In addition, $\hat{c}_{it}^f$ will also be used with indication that the only return on portfolio is the risk free rate. Consumption is assumed to be additively separable from participation cost. It is important to note that from the above it does not follow that consumers who invest in stock consume less, however, their wealth is smaller in period t by the amount of participation cost.

The choice of share of wealth invested in stocks, conditional on the participation decision, plays an important role in recovering participation cost. The key feature of the estimation of the participation cost is the intertemporal utility cost and benefits analysis for different participation strategies, where participation cost is tied up with the amount of wealth that households consider investing in stocks. On one hand, participation cost reduces wealth today, while next period a household receives an expected increase in wealth depending on the amount invested in stocks. On the other hand, households may choose not to invest in

stocks today, but instead keep all wealth (including the not forgone participation cost) in risk free assets. The next period, however, the household can only receive a risk free return on the not given up participation cost.

To illustrate the idea, consider two identical households n and i who act under two distinct alternatives (2.3.4) and (2.3.5), respectively. Recall that one of the alternatives is to participate in period t and quit stock markets in period $t+1$. The other alternative is to not participate in periods t and $t+1$. Consider household i whose observed choice is to participate in period t. One of the key assumptions is that the observed wealth on period t for this household is net time t participation cost: $w_{it}|(d_{it}=1) = R^p_{it}w_{it-1} + y_{it} - c_{it} - \rho_{it}$. Wealth in period t of household n will be greater than wealth in period t for household i by the amount of participation cost: $(w_{nt}+\rho_{nt})|(d_{nt}=0) = R^p_{nt}w_{nt-1} + y_{nt} - c_{nt}$. Identical time t state vector implies that time t consumption for both households is the same, therefore utility of consumption for these households is identical in period t. Under these two strategies, however, period $t+1$ consumption is different as it is affected by different wealth in period t. Under the first strategy $c_{it+1} = (\alpha_{it}(R^s_{it+1} - R^f_{t+1}) + R^f_{t+1})w_{it} + y_{it+1} - w_{it+1}$, while under the second strategy $c_{nt+1} = R^f_{t+1}(w_{nt}+\rho_{nt}) + y_{nt+1} - w_{nt+1}$. In the data it is unlikely to see two households with identical continuous state vectors. Therefore, it can also be intuitive to think about the same household that considers making different participation choices in period t. The differences in two participation strategies can be illustrated by looking at transition of wealth from period t to period $t+1$ under these strategies:

Alternative $\quad (d_t = 1, d_{t+1} = 0, ...)$	*Alternative* $\quad (d_t = 0, d_{t+1} = 0, ...)$
$w_t = R^p_t w_{t-1} + y_t - c_t - \rho_t$	$w_t + \rho_t = R^p_t w_{t-1} + y_t - c_t$
$w_{t+1} = (\mathbf{R^s_{t+1}} - \mathbf{R^f_{t+1}})\alpha_\mathbf{t}\mathbf{w_t} + R^f_{t+1}w_t + y_{t+1} - c_{t+1}$	$w_{t+1} = \mathbf{R^f_{t+1}}\rho_\mathbf{t} + R^f_{t+1}w_t + y_{t+1} - c_{t+1}$

It follows that under two distinct strategies, a household chooses between giving up the amount ρ_t in period t and next period get an expected increase in wealth in the amount of $(R^s_{t+1} - R^f_{t+1})\alpha_t w_t$ versus not participating at period t and receive $R^f_{t+1}\rho_t$ next period. At period $t+2$ the state vector z_{t+2} under both strategies becomes identical, therefore the value functions V_{t+2} become identical under two alternatives as well.

After taking differences in contemporaneous utilities, only the utility cost of adjusting one's portfolio will remain: $u(z_{it}, d_{it} = 1) - u(z_{it}, d_{it} = 0) = \gamma_0 d_{it} + \gamma_1 \alpha_{it}\alpha_{it-1}$. The difference in utilities at time $t+1$ is more involved. Denote consumption under participation strategy as c^1_{t+1}, and c^0_{t+1} as consumption for non-participation strategy. Then the difference between utility payoffs is:

$$\begin{aligned}(c^1_{t+1} - c^0_{t+1}) - \xi_t(c^1_{t+1} - c^0_{t+1})(c^1_{t+1} + c^0_{t+1}) = \\ (c^1_{t+1} - c^0_{t+1})\left[1 - \xi_t(c^1_{t+1} + c^0_{t+1})\right] = \\ (\alpha_t \hat{w}_{t+1} - R^f_{t+1}\rho_t)\left[1 - \xi_t(2c^f_{t+1} + \alpha_t \hat{W}_{t+1} + R^f_{t+1}\rho_t)\right]\end{aligned}$$

Therefore the optimality condition (2.3.6) becomes[11]:

$$\begin{aligned}\ln\left(\frac{p_{it}}{1-p_{it}}\right) &= \gamma_0 d_{it} + \gamma_1 \alpha_{it}\alpha_{it-1} \\ &+ \beta \sum_k^K \left[(\alpha_{it}\hat{w}^k_{it+1} - R^f_{t+1}\rho_{it})\left[1 - \xi(2c^f_{it+1} + \alpha_{it}\hat{w}^k_{it+1} + R^f_{t+1}\rho_{it})\right] + \ln\frac{p_{1kt+1}}{p_{0it+1}}\right]\mathcal{F}^k_{t+1}\end{aligned}$$

where $\hat{w}^k_{it+1} = w_{it}(R^{sk}_{it+1} - R^f_{it+1})$ is a function of data, conditional on participation.

Let Θ denote all unknown parameters in the model to be estimated. These parameters include the utility parameters γ_0,γ_1,ξ and the parameters of the participation cost δ. I fix the value of time discount factor β and estimate the remaining parameters conditioning on β. I rearrange and combine terms represented by observables to get the equation:

$$m_{it}(\Theta) = Y_{it} - X_{it}\Theta \tag{2.4.1}$$

where

$$\begin{aligned}Y_{it} &= \ln\left(\frac{p_{it}}{1-p_{it}}\right) - \beta\sum_k^K\left[\alpha_{it}\hat{w}^k_{it+1} + \ln\frac{p_{1kt+1}}{p_{0it+1}}\right]\mathcal{F}^k_{t+1} \\ X_{it} &= (1, \quad \alpha_{it}\alpha_{it-1}, \quad \beta R^f_t y^l_{it} x_{it}, \quad \beta\textstyle\sum_k^K \alpha_t \hat{w}^k_{t+1}(2c^f_{t+1} + \alpha_t \hat{w}^k_{t+1})\mathcal{F}^k_{t+1}, \\ &\qquad 2\beta R^f_t c^f_{t+1} y^l_{it} x_{it}, \quad \beta R^{f2}_t \textstyle\sum_{l=1}^L x^2_{lit}, \quad 2\beta R^{f2}_t \textstyle\sum_{l=1}^{L-1}\sum_{q=l+1}^L x_{lit}x_{qit})\end{aligned}$$

[11]The equation (2.3.6) is derived and estimated in levels of consumption, as typically done in the micro-level consumption models. The effect of consumption growth on current investment decisions is left beyond the scope of this paper.

where L is the dimension of x_{it}. $\Theta = (\Theta_1, \Theta_2)$, where $\Theta_1 = (\gamma_0, \gamma_1, \delta, \xi)$ has dimension $2 + L + H$, and $\Theta_2 = (\xi\delta, \xi\delta^2, \xi(\delta_1\delta_2, ..., \delta_{L-1}\delta_L))$ has dimension $2HL + HL(L-1)/2$, where δ is a vector of parameters $(\delta_1, ..., \delta_L)$ with the same dimension as x_{it} and H is the dimension of ξ. My primary interest is only in Θ_1. In the estimation I do not impose any restrictions on Θ_2 and estimate unrestricted equation (2.4.1) that is linear in parameters.

Once the conditional choice probabilities and the transition probabilities were observed or estimated, I could estimate the unknown structural parameters of interest from a conditional moment condition

$$E\left[m(X_{it}, Y_{it}, \Theta_o)|z_{it}\right] = 0, \tag{2.4.2}$$

where the subscript o denotes the true value of the parameters. The minimum distance estimator is a natural estimator choice in this and similar frameworks.

2.4.4 Individual Effects and Aggregate Shocks

In the literature different methods are proposed to incorporate unobserved heterogeneity in dynamic discrete choice models. Altug and Miller (1998) estimate fixed effects from an auxiliary regression related to the main model. More recent studies of Arcidiacono and Miller (2008) and Kasahara and Shimotsu (2009) propose alternative ways through finite mixture distributions. The approach of Kasahara and Shimotsu is restrictive for my estimation. For identification of unobserved heterogeneity it requires the time dimension of a panel that is beyond the one available in the data at hand. The identification requirements in Arcidiacono and Miller are similar to those as in the study of Kasahara and Shimotsu. I adopt the approach of Altug and Miller and use the earnings equation to estimate individual unobserved effects.

The same earnings equation allows me to estimate time effects as well. In my framework, current wealth accumulation is affected by aggregate shocks on current wages and rates of return. Aggregate shocks on rates of return are captured by the risk-free rate, R_t^f. Aggregate shocks on wages are not directly observed, but can be estimated.

I assume that household's earnings are affected by both time effects and individual effects, hence I can estimate these effects by modeling the household's earnings process. I consider a dynamic earnings equation of the form:

$$\hat{y}_{it} = \phi \hat{y}_{it-1} + x'_{it}\kappa + \omega_t + \nu_i + e_{it} \tag{2.4.3}$$

where $\hat{y}_{it}$ is the log-transformation of household's labor income and x'_{it} is an L-dimensional vector of family characteristics for household i at time t, ω_t is the unobserved time specific effect and ν_i is the unobserved individual specific effect. I follow Arellano and Honore (2001), assuming the predeterminedness condition of $\hat{y}_{it-1}$ and (possibly) x_{it}, precisely $E(e_{it}|\hat{y}_i^{t-1}, x_i'^{t}) = 0$ and $E(\nu_i) = 0$. Subject to a rank condition, $(1 + L + T - 1)$ parameters of the model (2.4.3) are identified with $T \geqslant 3$. With $T = 3$ ϕ, κ's, ω_2, ω_3 are just identified from the $3 + L$ orthogonality conditions:

$$\begin{aligned} E\left[\begin{pmatrix} 1 \\ x'_{it} \\ \hat{y}_{it-2} \end{pmatrix}\left(\Delta\hat{y}_{it} - \phi\Delta\hat{y}_{it-1} - \Delta x'_{it}\kappa - \Delta\omega_t \right)\right] &= 0 \\ E\Big(\hat{y}_{it} - \phi\hat{y}_{it-1} - x'_{it}\kappa - \omega_t\Big) &= 0 \end{aligned} \tag{2.4.4}$$

Estimation of the fixed effects, however, requires a larger T. The extended data set covers $T = 14$ time periods of observations. The vector of family characteristics x'_{it} contains 3 variables: family size, education multiplied by age and age squared. It makes 17 parameters to estimate. The overidentified system of equations (2.4.3) contains 4 moment conditions for each of $T - 2$ time periods making up to 48 orthogonality conditions for each household. The four moment conditions include the unconditional moment condition in levels and the moment conditions in first differences with a constant, income at time $t - 2$ and age squared at time t as instruments. I estimate the system of equations (2.4.4) with the GMM procedure.

Table 2.4 reports the estimates of the earning equation (2.4.3). The estimates on family size, education, and age are consistent with the general expectations. The household income increases with the family size. It responds positively to the level of education and decreases

Table 2.4: Earnings equation [1]

Parameter	Estimate	Std.Err.
ϕ	0.856	0.038
$\kappa's$:		
family size	0.202	0.037
*edu*age*	0.577	0.063
age^2	-0.863	0.088
ω_{1983}	-0.720	0.357
$\Delta\omega_{1985}$	0.078	0.015
$\Delta\omega_{1987}$	-0.063	0.014
$\Delta\omega_{1989}$	-0.025	0.013
$\Delta\omega_{1991}$	-0.088	0.013
$\Delta\omega_{1993}$	0.012	0.013
$\Delta\omega_{1995}$	-0.108	0.016
$\Delta\omega_{1997}$	0.031	0.017
$\Delta\omega_{1999}$	0.084	0.018
$\Delta\omega_{2001}$	-0.025	0.015
$\Delta\omega_{2003}$	-0.071	0.015
$\Delta\omega_{2005}$	0.057	0.017
$\Delta\omega_{2007}$	-0.008	0.019
J test	10.5	
p value	0.99	

[1] Number of time periods T = 14, number of households in the sample 7,744. Instruments include a constant, age of head of household squared at time t, and labor income at time $t-2$.

with the household growing older. The $AR(1)$ coefficient on past earnings is estimated at 0.85 and statistically significant. Most estimates on $\Delta\omega_t$ are also statistically significant. With the parameter estimates of the model (2.4.3), individual fixed effects are trivially estimated.

2.4.5 Conditional Choice Probabilities

I estimate conditional choice probabilities using methodology described in Hotz and Miller (1993) and Altug and Miller (1998). Denote $J[h_N^{-1}(z_{ns}^N - z_{it}^N)]$ as a multivariate kernel function with an appropriately chosen matrix of bandwidths h_N. For $t = 1, ..., T$ I estimate conditional participation rate nonparametrically using the kernel estimator:

$$\hat{p}_{it} = \frac{\sum_{n=1}^{N} d_{nt} J[h_N^{-1}(z_{nt}^N - z_{it}^N)]}{\sum_{n=1}^{N} J[h_N^{-1}(z_{nt}^N - z_{it}^N)]}$$

Define the conditional choice probability $p_j(z_{lkit})$, $j, l = 0, 1$ and $k = 0, ..., K$, as a probability that household i chooses alternative j in period t given that the alternative l was chosen in period $t-1$ and conditioned on realization of the state z_{lkit}. I define the indicator variables as follows:

$$d_{lkit}^{(1)} = \begin{cases} (1 - d_{it-1}) & l = 0, \quad k = 0 \\ d_{it}^k d_{it-1} & \text{otherwise} \end{cases}$$

where d_{it}^k is equal to 1 if a household faces one of the K possible states of realization of returns on stock holdings if the household invests in stock in period $t-1$. Then the estimators for conditional choice probabilities $p_j(z_{lkit})$ for $t = 1, ..., T$ are computed as follows:

$$\hat{p}_j(z_{lkit}) = \frac{\sum_{n=1}^{N} d_{jnt} d_{lknt}^{(1)} J[h_N^{-1}(z_{ns}^N - z_{it}^N)]}{\sum_{n=1}^{N} d_{lknt}^{(1)} J[h_N^{-1}(z_{nt}^N - z_{it}^N)]}$$

The summary of nonparametric estimates of conditional choice probabilities is reported in Table 2.5. The table also reports the number of observations used to calculate the probability associated with each relevant history z_{lkit} as well as sample standard deviations. The estimated conditional choice probabilities are limited to those relevant to the histories used in the equation (2.3.6).

Table 2.5: Sample averages of nonparametric estimates of conditional choice probabilities

		Num.obs.	Sample mean	Sample std.dev.
1.	p_{1it}	10,639	0.256	0.300
2.	$p_0(z_{0it})$	9,412	0.889	0.089
3.	$p_0(z_{1Hit})$	10,590	0.030	0.108
4.	$p_0(z_{1Mit})$	10,588	0.016	0.084
5.	$p_0(z_{1Lit})$	10,608	0.138	0.248

In line 1 of Table 2.5 I report the participation rate p_{1it}. Line 2 reports the probability of not participating in the stock markets in period t conditional on non-participation in period $t-1$. Lines 3-5 report the probability of non-participation in stock markets in period t conditional on participation in the previous time period and receiving high, moderate or low return on the investment in stocks. The average estimated participation rate is 0.25. The results also suggest that the probability of participation decreases if there is a history of non-participation in a previous period to 0.11.

2.5 EMPIRICAL FINDINGS

Table 2.6 contains the estimation results for the utility parameters and parameters of the participation cost. Columns (1)-(3) report estimates for the model with the simple quadratic utility of consumption, columns (5)-(7) report estimates for the model with age augmented quadratic utility of consumption.

Table 2.6: Participation cost and utility parameters.

Parameters	(1)	(2)	(3)	(4)	(5)	(6)
Utility parameters $(c_t - \xi_t c_t^2) + \gamma_0 d_t + \gamma_1 \alpha_t \alpha_{t-1}$						
		$\xi_t = \xi_1$			$\xi_t = \xi_1 + \xi_2 age_t$	
γ_0	31.239* (5.412)	41.486* (3.650)	39.223* (3.289)	29.690* (4.734)	34.671* (2.861)	30.985* (2.045)
γ_1	20.606*** (11.827)	12.106 (9.118)	10.899 (7.249)	19.396*** (9.988)	9.402*** (5.546)	1.525 (3.275)
ξ_1	0.00050* (0.00002)	0.00055* (0.00003)	0.00057* (0.00002)	0.00023** (0.00010)	−0.00030** (0.00014)	−0.00030* (0.00008)
ξ_2				0.000006* (0.000002)	0.000016* (0.000002)	0.000017* (0.000002)
Participation cost $\rho_t = y_t(\delta_1 + \delta_2 edu + \delta_3 age_t + \delta_4 age_t^2 + \delta_5 d_{t-1})$						
δ_1	0.0616* (0.0165)	0.4553* (0.0563)	−0.0411 (0.1129)	0.0553* (0.0144)	0.2964* (0.0484)	0.1659* (0.0658)
δ_2		−0.0194* (0.0031)	−0.0201* (0.0020)		−0.0120* (0.0023)	−0.0122* (0.0016)
δ_3		−0.0011** (0.0005)	0.0229* (0.0048)		−0.0007 (0.0005)	0.0048*** (0.0028)
δ_4			−0.00027* (0.00005)			−0.00006** (0.00003)
δ_5		−0.0179 (0.0139)	−0.0307** (0.0124)		−0.0376* (0.0100)	−0.0477* (0.0075)
Number of parameters	6	20	26	9	34	46
J-test	98	112	125	107	134	138
p-value	5e-7	4e-4	7e-4	5e-7	0.0051	0.087

Participation cost

The estimation results for the participation cost seem to be intuitive and agree with previous findings. First, Gomes and Michaelides (2003) rationalize limited participation by the existence of the participation cost. Second, empirical findings of Bertaut and Starr-McCluer (2002), Guiso and Japelli (2002), Banks and Tanner (2002) show that the participation in stock markets can be explained by such individual characteristics as age and education. The findings in the current study relate to these results from previous work and show that the participation cost is significantly different from zero and varies with household characteristics.

Participation cost is estimated on average about $5.5-6.2\%$ of labor income. Participation cost measured as a share of income is not constant over the life cycle and admits the curvature of the labor income profile. But the results also show that even holding the labor income fixed, participation cost estimated to be significantly affected by age, education and previous participation experience. Table 2.6 shows that participation cost is decreasing in education. It is roughly decreasing with age. When the curvature in age is introduced in the participation cost, it becomes increasing and concave in age, with the turning point around age of 35-40. Finally, participation cost is smaller for households who participated in stock markets in previous periods.

Results in Table 2.6 show that when demographic characteristics in participation cost are not controlled for (columns (1) and (4)), the estimated participation cost admits similar magnitudes. However, when age, education and past participation are controlled for, the estimates in column (2) and (5), and in (3) and (6) differ substantially. On average, participation cost in columns (2) and (3) is greater than in columns (5) and (6) by about 5%. It is illustrated on Figure 2.1. These differences can be explained by the influence of age on participation cost and on the utility. The specification in columns (5) and (6) is different from the one in columns (2) and (3) only through the age dependent parameter ξ_t in the utility of consumption. Once the age is controlled for in the curvature of the utility of consumption, the contribution of age into participation cost becomes smaller.

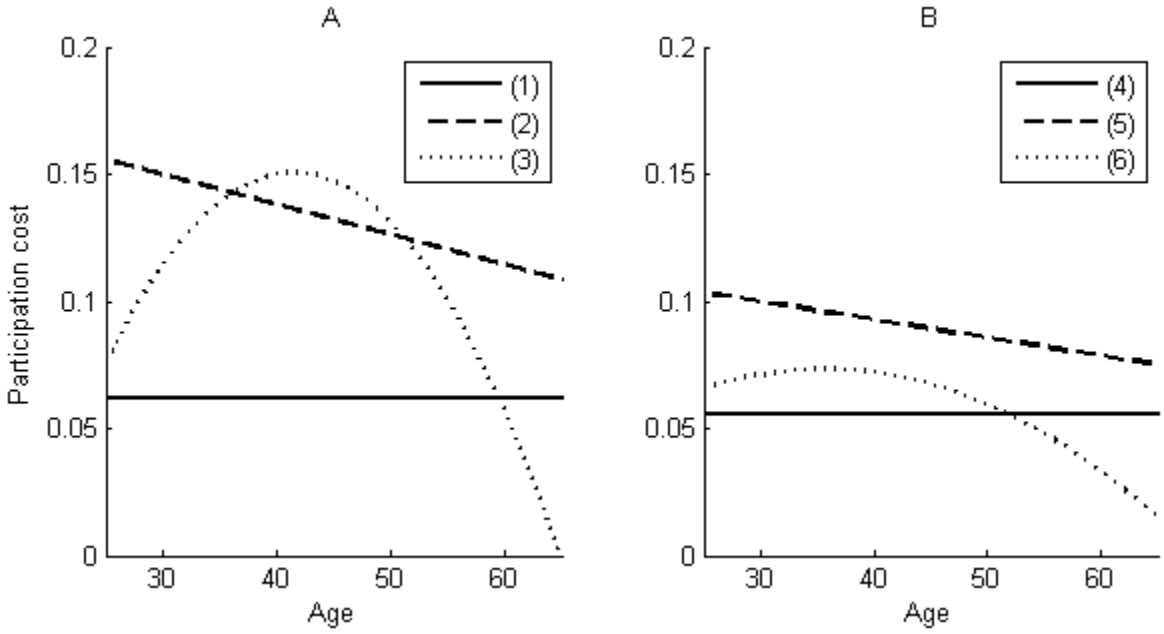

Figure 2.1: Participation cost.

Figure 2.1 shows how participation cost (as a share of labor income) changes over age, with other demographic characteristics fixed at their sample average levels. Panel B shows that participation cost can be as large as 10% of labor income, however the age effect for the line that corresponds to column (5) is estimated imprecisely. The lines corresponding to columns (4) and (6) are constructed from the more precisely estimated coefficients and suggest that the magnitude of participation cost is smaller. On average it is about 5-6%, but decreases to as low as 1% when a household approaches retirement.

Utility parameters

Utility parameters include parameters on the curvature of the utility of consumption and parameters of the utility benefit of stock market participation. The estimates of the ξ_t are significant in all model specifications and indicate that the utility of consumption is concave. The average value of $\xi_t = \xi_1 + \xi_2 age_t$ is similar to the estimates of ξ_t in columns (1)-(3). Even with the negative ξ_1 in columns (5) and (6) ξ_t never takes negative values.

All model specifications in Table 2.6 suggest that there is a positive utility shifter for the households who participate in stock markets. In addition, the parameter on the portfolio adjustment cost is marginally significant and positive. This positive estimate indicates on complimentarity effect of past portfolio allocation on current portfolio allocation. Habits, or rather inertia, in adjusting one's portfolio share is consistent with previous findings in the literature. Namely, Brunnermeier and Nagel (2008) argue that one of the major drivers of households' portfolio allocation is inertia.

The J tests associated with the model specifications reported in Table 2.6 indicate the rejection of overidentifying restrictions implied by the choice of instruments. However, the model specification (6) cannot be rejected at the 5% significance level; parameter estimates from this specification are used in the simulation exercises presented below.

2.6 SIMULATION EXERCISES

I solve the model via backward induction using the estimated parameters of the model. I use the parameter estimates reported in column (6) of Table 2.6. As the estimation uses biannual data, I also solve the model recursively on a biannual basis. To align with the estimation set up that considers households prior to retirement, I simulate households up to the retirement stage. As a termination condition I assume that individuals die at age 67 not leaving a positive bequest.

To solve the model I start with defining a grid on state variables, that is past share of wealth invested in risky stocks and past wealth with earned rate of return on it, as well as income. Then optimal share of wealth invested in risky assets and total wealth can be solved for on the fine grid, with the help of estimated parameters on income process. The interpolation off the grid is achieved with the second order polynomial regression of the value function at each point of the grid on the corresponding state space vector.[12] This regression

[12]Higher order polynomial regression was also tried, however, it produces less accurate predictions.

provides a very accurate approximation to the value functions off the grid with R^2 in range of 0.98-0.99.

Once the solution is obtained, the model is simulated for 5,000 replications.[13] No participation in stock markets is chosen as a starting condition for all simulated individuals. Conditional on the state variables, I compute the conditional choice value function for each combination of the choice variables from the grid, using the interpolation coefficients obtained at the solution as well as the draws of extreme value random variables. Then the maximal conditional choice value function corresponds to the optimal choice of participation decision, portfolio choice, and wealth. If a simulated consumer finds it optimal to participate in stock markets, he receives the rate of return of stock holdings, drawn randomly as low, moderate or high return with corresponding probabilities used in estimation as transition probabilities.

Figure 2.2 compares (1) the average age-specific participation rate computed from the data, (2) participation rate, predicted by probit model, estimated in the section on the data description, and (3) participation rate, predicted by the structural model. Figure 2.2 shows that the baseline model predicts very well the participation rate observed in the data.

2.6.1 Financial Education

The participation cost is an important instrument to evaluate the effectiveness of the economic policies aimed to promote financial literacy among consumers. Recall that the participation cost is estimated to be about 5-6% of labor income. The model of stock market participation allows for a counterfactual simulation exercise aimed to predict the response of consumers' financial behavior to providing financial education and counseling. In order to evaluate the potential impact of the financial education programs intervention, I treat education that affects the participation costs as a proxy to the consumers' ability to process information, including information that is related to participation in financial markets. The simulated response to the policy interventions is interpreted as a lower bound of the possible

[13] Greater number of simulations virtually does not make a difference.

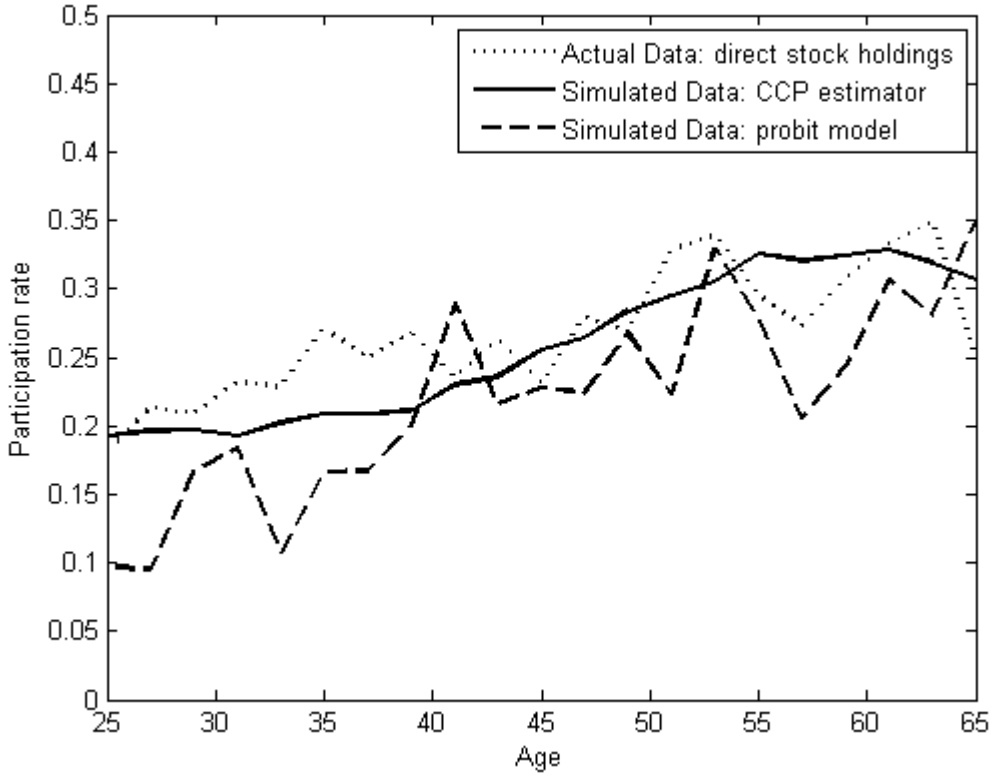

Figure 2.2: Participation rate: Data vs Simulation

response to financial counceling.

Providing financial education and counseling may have an effect on the participation decision through the channel of participation cost that is decreasing with education. Recall that the results in column (6) of Table 2.6 suggest that every additional year of education decreases participation cost by 1.2 percentage points. To see how consumers respond to providing financial counseling, I conduct two counterfactual simulation exercises. The first simulation exercise shows how investors respond to an unexpected one unit increase in education that lowers the participation cost. The second exercise evaluates the responde to the same increase in education, however, contrary to the first unexpected increase, this increase is anticipated by an investor. The "financial" education affects participation decision through participation cost, but does not affect the amount of schooling that determines consumers' income process.

The result of this counterfactual simulation is illustrated by Figure 2.3. As expected, simulated consumers respond positively to financial counseling. Their participation rate is greater than that of baseline consumers. The difference between participation of financially advised and baseline consumers grows by up to 4%. The response to the anticipated increase in financial education takes place sooner than the response to an unexpected increase. Consumers who anticipate the decrease in participation costs due to extra financial education show a smoother policy response. Shortly after the policy intervention, both participation rates become equal.

2.7 CONCLUSION

In this paper I make an attempt to rationalize the limited stock market participation at the micro level. I develop and estimate a dynamic structural model of individual stock market participation and portfolio choice; estimation is achieved using longitudinal data on household wealth, stock holdings and income taken from the PSID. The model allows for

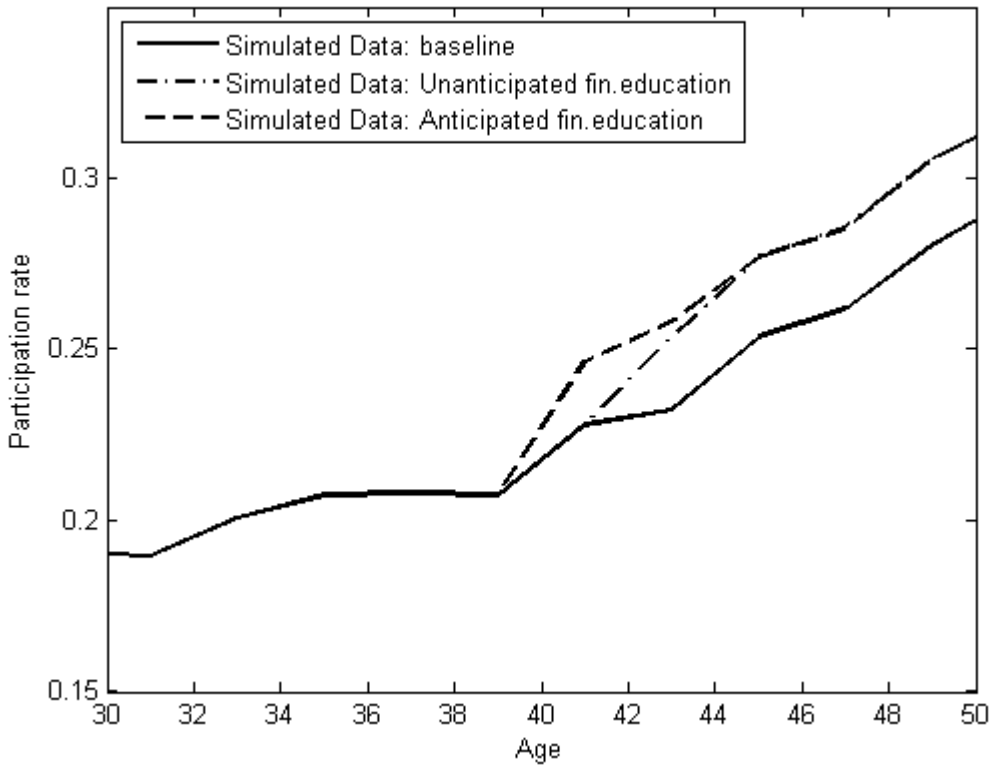

Figure 2.3: Participation rate, financial education program intervention.

endogenous investment choices and intertemporal nonseparabilities in preferences. Decisions regarding stock market participation are influenced by participation costs modelled using a flexible functional form to account for individual heterogeneity. The estimation results support the significant participation cost to stock market participation. Participation cost is estimated on average about $5.5 - 6.2\%$ of labor income and can be explained by such demographic characteristics as consumers' age, education and past stock market participation experience. In addition, I find evidence of inertia in portfolio allocation.

Model estimates are then used to conduct simulation exercises in order to evaluate the possible effectiveness of financial education programs. The model estimates not only make economic sense but also facilitate simulated paths for the decision to participate in stock markets that successfully line up with patterns observed in the data. I investigate how the stock market participation decision can be affected when consumers receive a reasonable amount of financial counseling. In my analysis I allow for up to two weeks of additional financial education per year and show that consumers respond positively to financial education programs in that the financial market participation increases by 4 /

While the results obtained at the estimation of the model seem to be intuitive and agree with previous findings, an important caveats to this analysis need to be mentioned. First, it is likely that the observed income and wealth data are measured with substantial error. In addition to the usual measurement error, the data in wealth supplement of the PSID are known to suffer from systematic underreporting of trades. The "forgotten" trades affect the computed idiosyncratic rates of return. In the current study I address this problem by refereing to the market index as the rate of return. However, the measurement error problem still remains a vital issue and needs a proper treatment. Second, the current analysis focuses on the participation decision and portfolio choice mostly for prime age individuals. However, it also is interesting to look at the transition of the participation rate from prime age, considered in this paper, to the retirement stage. This is an important problem and is deferred for the future research. Finally, the household stock holdings represent only a part of risky assets. Housing, business, other risky accounts such as individual retirement accounts

amount for a bigger part of household financial portfolio. Depending on the structure of all household assets, limited participation in stock markets and smaller shares of liquid wealth invested in stocks directly can be explained by a certain diversification strategy.

With the above limitations to the current analysis this paper is the first analysis of the stock market participation and portfolio composition that allows for heterogeneity in the participation cost and estimates the magnitude of it in micro data. I find a strong evidence that households with experience in stock market participation face smaller participation cost. The results show that the participation cost is increasing and concave with age and decreasing with education. Another novelty of the paper constitutes the counterfactual simulation exercise that shows how education programs can influence the consumers' financial decisions. In general, the results contribute to the broad strand of the literature on life cycle consumption and investment decisions, as well as on idiosyncratic response to consumer financial protection policy initiatives.

BIBLIOGRAPHY

[1] Abel, Andrew B.,1990. "Asset Prices Under Habit Formation and Catching Up With the Jones", American Economic Review, 80, 38-42.

[2] Aguirregabiria, Victor, and Pedro Mira (2007). "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, 1-54.

[3] Ai Chunrong and Xiaohong Chen, 2003. "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," Econometrica, vol. 71(6), p. 1795-1843

[4] Alan, Sule and Martin Browning, 2010. "Estimating Intertemporal Allocation Parameters Using Synthetic Residual Estimation", Review of Economic Studies. vol. 77(4).

[5] Alan, Sule, Orazio Attanasio and Martin Browning, 2009, "Estimating Euler Equations with Noisy Data: Two exact GMM Estimators, Journal of Applied Econometrics, 24, p. 309-324.

[6] Alan, Sule. (2006) "Entry Costs and Stock Market Participation Over the Life Cycle", *Review of Economic Dynamics*, 9(4).

[7] Alessie Rob, Stefan Hochguertel, and Arthur van Soest (2002), "Household Portfolios in the Netherlands", in Household Portfolios (Luigi Guiso, Michael Haliassos, and Tullio Jappelli, eds.). Cambridge, Massachusetts: MIT Press.

[8] Alessie Rob, Stefan Hochguertel, and Arthur van Soest (2004), "Ownership of Stocks and Mutual Funds: A Panel Data Analysis," *The Review of Economics and Statistics*, vol. 86(3), 783-796.

[9] Alessie Rob, Annamaria Lusardi and Maarten van Rooij, (2007). "Financial Literacy and Stock Market Participation," NBER Working Paper.

[10] Altug, Sumru and Robert A. Miller, (1990). "Household Choices in Equilibrium," *Econometrica*, 58, 543570.

[11] Altug, Sumru and Robert A. Miller, (1998). "Effect of Work Experience of Female Wages and Labour Supply," *The Review of Economic Studies*, 65, 4585.

[12] Andersen, Steffen and Kasper Meisner Nielsen (2011), Participation Constraints in the Stock Market: Evidence from Unexpected Inheritance Due to Sudden Death, *The Review of Financial Studies*, forthcoming.

[13] Arcidiacono, Peter and Robert A. Miller (2010), "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity", manuscript, Duke University.

[14] Arellano, Manuel and Bo Honore (2001), "Panel data models: some recent developments," in: J.J. Heckman and E.E. Leamer (ed.), Handbook of Econometrics, edition 1, volume 5, chapter 53, 3229-3296.

[15] Atkeson, Andrew and Ogaki, Masao, 1996. "Wealth-Varying Intertemporal Elasticities of Substitution: Evidence from Panel and Aggregate Data." Journal of Monetary Economics, 38, 507-534.

[16] Attanasio, Orazio and Hamish Low, 2004, "Estimating Euler Equations, Review of Economic Dynamics, 7, p.406-435.

[17] Attanasio Orazio P. and Monica Paiella, (2011), "Intertemporal Consumption Choices, Transaction Costs and Limited Participation in Financial Markets: Reconciling Data and Theory", *Journal of Applied Econometrics*, 26(2), 322343.

[18] Attanasio, Orazio and Guglielmo Weber, 1993. "Consumption Growth, the Interest Rate and Aggregation", The Review of Economic Studies, 60(3): 631-649.

[19] Attanasio, Orazio and Guglielmo Weber, 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey", Journal of Political Economy, vol. 103, p. 1121-1157.

[20] Banks James and Sarah Tanner (2002), "Household Portfolios in the United Kingdom", in Household Portfolios (Luigi Guiso, Michael Haliassos, and Tullio Jappelli, eds.). Cambridge, Massachusetts: MIT Press.

[21] Barsky Robert B., Miles S. Kimball, F. Thomas Juster and Matthew D. Shapiro, 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey," Quarterly Journal of Economics, Vol. 112, 537-579.

[22] Bertaut Carol C. and Martha Starr-McCluer (2002), "Household Portfolios in the United States", in Household Portfolios (Luigi Guiso, Michael Haliassos, and Tullio Jappelli, eds.). Cambridge, Massachusetts: MIT Press.

[23] Browning Martin and M. Dolores Collado, "Habits and Heterogeneity in Demands: A Panel Data Analysis", Journal of Applied Econometrics, 22 (2007), 625-640.

[24] Browning, Martin and Costas Meghir, 1991. "Testing for Separability of Commodity Demands from Male and Female Labor Supply", Econometrica, vol. 59, p. 925-952.

[25] Brunnermeier Markus K. and Stefan Nagel (2008), "Do Wealth Fluctuations Generate Time-varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation", *American Economic Review*, 98(3), 713-736.

[26] Carrasco, Raquel, Jos M. Labeaga and J. David Lopez-Salido, "Consumption and Habits: Evidence from Panel Data", The Economic Journal, 115(January 2005), 144-165.

[27] Carroll, Christopher D., "Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation)", Advances in Macroeconomics, 1(1) (2001), 6.

[28] Chen, Xiaohong and Sydney C.Ludvigson, "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Behavior", Journal of Applied Econometrics. Forthcoming.

[29] Cocco, Joao F., Francisco J. Gomes and Pascal J. Maenhout, (2005) "Consumption and portfolio choice over the life-cycle", *The Review of Financial Studies*,18, 491-533.

[30] Crossley, Thomas and Hamish Low, 2011, Is the EIS Constant?, Journal of the European Economic Association, Forthcoming.

[31] Duflo Esther and Emmanuel Saez, (2003). "The Role Of Information And Social Interactions In Retirement Plan Decisions: Evidence From A Randomized Experiment," *The Quarterly Journal of Economics*, vol. 118(3), 815-842

[32] Dynan Karen E., "Habit Formation in Consumer Preferences. Evidence from Panel Data". American Economic Review, 90 (2000), 391-406.

[33] Eisenhauer Joseph G. and Luigi Ventura, 2003. "Survey measures of risk aversion and prudence," Applied Economics, vol. 35(13), 1477-1484

[34] Fuhrer, Jeffrey C., "Habit Formation in Consumption and Its Implications for Monetary-Policy Models". American Economic Review, 90 (2000), 367-390.

[35] Gomes Francisco J. and Alexander Michaelides, (2003) "Portfolio choice with internal habit formation: a life-cycle model with uninsurable labor income risk", *Review of Economic Dynamics*, 6, 729-766.

[36] Grinblatt Mark , Juhani Linnainmaa, and Matti Keloharju (2010), "IQ and Stock Market Participation", Working paper.

[37] Guiso Luigi and Tullio Jappelli (2002), "Household Portfolios in Italy", in Household Portfolios (Luigi Guiso, Michael Haliassos, and Tullio Jappelli, eds.). Cambridge, Massachusetts: MIT Press.

[38] Guiso Luigi and Tullio Jappelli (2005), "Awareness and Stock Market Participation", *Review of Finance*, 9(4), 537-567.

[39] Guiso, Luigi and Paiella, Monica, 2004. "The Role of Risk Aversion in Predicting Individual Behaviours," CEPR Discussion Paper.

[40] Haliassos, Michael and Alexander Michaelides, (2003), "Portfolio Choice and Liquidity Constraints". *International Economic Review*, 44(1), 143-177.

[41] Hansen, Lars Peter, John C. Heaton and Amir Yaron, "Finite Sample Properties of some Alternative GMM Estimators," Journal of Business and Economic Statistics, 14 (1996), 262-280.

[42] Hausman Jerry A., Whitney K. Newey, Hidehiko Ichimura and James L. Powell, "Identification and Estimation of Polynomial Errors-in-variables Models", Journal of Econometrics, 50 (1991), 273-295.

[43] Hong, Han and Elie Tamer, "A Simple Estimator for Nonlinear Error in Variable Models", Journal of Econometrics, 117(2003), 1-19.

[44] Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, (2004), "Social interaction and stock market participation", *Journal of Finance* 59, 137-163.

[45] Horowitz, Joel and and Charles F. Manski, 2000. "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data," Journal of the American Statistical Association, 95, 7784.

[46] Hotz, V. Joseph, Kydland, Finn E. and Sedlacek, Guilherme L, (1988). "Intertemporal Preferences and Labor Supply," *Econometrica*, 56(2), 335-60.

[47] Hotz, Joseph V. and Robert A. Miller (1993) "Conditional choice probabilities and the estimation of dynamic models". *Review of Economic Studies*, 60(3), 497530.

[48] Imbens, Guido W. and Charles F. Manski, 2004 "Confidence Intervals for Partially Identified Parameters," Econometrica, 72(6), 1845-1857.

[49] Kasahara Hiroyuki and Katsumi Shimotsu (2009) "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, 77(1), 135-175.

[50] Meghir, Costas and Guglielmo Weber, "Intertemporal Nonseparability or Borrowing Restrictions? A Disaggregate Analysis Using a U.S. Consumption Panel," Econometrica, 64(5) (1996), 1151-81.

[51] Munoz Sonia (2006) "Habit Formation and Persistence in Individual Asset Portfolio Holdings: The Case of Italy", IMF Working Paper No. 06/29.

[52] Newey Whitney K. and James L.Powell (2003). "Instrumental Variable Estimation of Nonparametric Models", Econometrica, vol.71(5), 1565-1578.

[53] Newey Whitney K. and Richard J. Smith, "Asymptotic Bias and Equivalence of GMM and GEL Estimators", MIMEO MIT Department of Economics, 2000.

[54] Muellbauer, John, "Habits, Rationality and Myopia in the Life Cycle Consumption Function," Annales d' Economie et de Statistique, 9 (1988), 4770.

[55] Naik, Narayan Y. and Michael J. Moore. 1996. "Habit Formation and Intertemporal Substitution in Individual Food Consumption." Review of Economics and Statistics, 78(2): 321-328.

[56] Paiella Monica (2007), "The Forgone Gains of Incomplete Portfolios", *Review of Financial Studies*, 20, 16231646.

[57] Pissarides, Christopher A. and Guglielmo Weber, 1989, "An Expenditure-Based Estimate of Britain's Black Economy", Journal of Public Economics, vol. 39, p. 17-32.

[58] Runkle, David E. 1991. "Liquidity Constraints and the Permanent Income Hypothesis", Journal of Monetary Economics, 27(1): 73-98.

[59] Schennach, Susanne, M., "Estimation of Nonlinear Models with Measurement Error." Econometrica, 72 (2004), 33-75.

[60] Shapiro, Matthew D., "The Permanent Income Hypothesis and the Real Interest Rate". Economics Letters, 14 (1984), 93-100.

[61] Smith, William T. and Qiang Zhang, "Asset Pricing with Multiplicative Habit and Power-expo Preferences". Economics Letters, 94 (3) (2007), 319-325.

[62] Tauchen George, 1986. "Finite state Markov-chain approximations to univariate and vector autoregressions", Economics Letters, 20, 177181.

[63] Ventura, Eva, "A Note on Measurement Error and Euler Equations". Economics Letters, 45 (1994), 305-308.

[64] Vissing-Jorgensen, Annette, 2002. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." Journal of Political Economy, 110(4), 825- 853.

[65] Vissing-Jorgensen, Annette.(2002) "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures", NBER Working Paper.

APPENDIX A

DERIVATION OF THE MOMENT CONDITION

In order to obtain an expression in terms of observed consumption, we consider Equation (3.1) piece by piece and express observed consumption in terms of true consumption and measurement error as stated above. Let $\mathbb{C}_{it} = (c_{it-1}, c_{it}, c_{it+1}, c_{it+2}, r_{it+1})$, let $\mathbb{C}^*_{it}$ be the counterpart of the vector $\mathbb{C}_{it}$ where true consumption is replaced by observed consumption, and let $\Im_{it} = (\eta_{it-1}, \eta_{it}, \eta_{it+1}, \eta_{it+2})$ be the corresponding vector of measurement errors. We start with the first term.

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{C}_{it},\Im_{it}|\mathbb{Z}_{it}}\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right]= \\
&\mathbb{E}_{\mathbb{C}_{it},\Im_{it}|\mathbb{Z}_{it}}\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g_{it+1}}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}\frac{1}{v_{it+1}}\left(\frac{v_{it+1}}{v^{\alpha}_{it}}\right)^{1-\gamma}\right]= \\
&\mathbb{E}_{\mathbb{C}_{it}|\mathbb{Z}_{it}}\left[\beta(1+r_{it+1})\frac{\varphi_{it+1}}{g_{it+1}}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}\right]\mathbb{E}_{\Im_{it}|\mathbb{Z}_{it}}\left[\frac{1}{v_{it+1}}\left(\frac{v_{it+1}}{v^{\alpha}_{it}}\right)^{1-\gamma}\right]= \\
&\mathbb{E}_{\mathbb{C}_{it}|\mathbb{Z}_{it}}\left[\beta(1+r_{i,t+1})\frac{\varphi_{it+1}}{g_{it+1}}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}\right]\mathcal{A}_1,
\end{aligned}
$$

Under assumption 1.3.4, it can be shown that $\mathcal{A}_1 = \exp\{\sigma^2\left(\alpha^2(1-\gamma)^2 + \gamma^2 - \alpha\gamma(1-\gamma)\right)\}$. Hence

$$
\mathbb{E}_{\mathbb{C}_{it}|\mathbb{Z}_{it}}\left[\beta(1+r_{i,t+1})\frac{\varphi_{it+1}}{g_{it+1}}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}\right]=\mathbb{E}_{\mathbb{C}_{it},\Im_{it}|\mathbb{Z}_{it}}\left[\beta\mathcal{A}_1^{-1}(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right].
$$

The second and the third terms are transformed in the same way to get

$$\begin{aligned}
\mathbb{E}_{\mathbb{C}_{it}|\mathbb{Z}_{it}}\left[\beta(1+r_{i,t+1})\frac{\varphi_{it+1}\varphi_{it+2}}{g_{it+1}}\left(\frac{g_{it+1}g_{it+2}}{(g_{it}g_{it+1})^{\alpha}}\right)^{1-\gamma}\right] &= \mathbb{E}_{\mathbb{C}_{it},\Im_{it}|\mathbb{Z}_{it}}\left[\alpha\beta^2\mathcal{A}_2^{-1}(1+r_{it+1})\frac{\varphi_{it+1}\varphi_{it+2}}{g^*_{it+1}}\left(\frac{g^*_{it+1}g^*_{it+2}}{(g^*_{it}g^*_{it+1})^{\alpha}}\right)^{1-\gamma}\right] \\
\mathbb{E}_{\mathbb{C}_{it}|\mathbb{Z}_{it}}\left[\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right] &= \mathbb{E}_{\mathbb{C}_{it},\Im_{it}|\mathbb{Z}_{it}}\left[\alpha\beta\mathcal{A}_3^{-1}\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right],
\end{aligned}$$

Again, under Assumption 1.3.4 we find that

$$\begin{aligned}
\mathcal{A}_2 &= \exp\{\sigma^2\left(\alpha^2(1-\gamma)^2+\gamma^2+(1-\gamma)(1+\alpha)\right)\}, \\
\mathcal{A}_3 &= \exp\{\sigma^2\left((1+\alpha+\alpha^2)(1-\gamma)^2\right)\}.
\end{aligned}$$

The moment condition (Equation 1.3.1) for (unobserved) true consumption is therefore transformed into a moment condition for observed consumption:

$$\begin{aligned}
\mathbb{E}_{\mathbb{C}_{it},\Im_{it}|\mathbb{Z}^*_{it}}\Bigg[\beta\mathcal{A}_1^{-1}(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\left(1-\alpha\beta\frac{\mathcal{A}_1}{\mathcal{A}_2}\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)- \\
\left(1-\alpha\beta\mathcal{A}_3^{-1}\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)\Bigg] = 0
\end{aligned} \tag{A.0.1}$$

where $\mathbb{Z}^*{}_{it}$ is a q-dimensional observable subset of $\mathbb{Z}_{it}$ that can include current and past interest rates as well as observable consumption growth up to time $t-2$.

APPENDIX B

PROOF OF THEOREM ??

Let $\tilde{\theta}$ be an alternative vector of parameters that satisfy equation (1.3.3) and define $\Gamma(x^*_{i,t+2}) = \rho(x^*_{it+2}, \theta_0) - \rho(x^*_{it+2}, \tilde{\theta})$ so that

$$E\left[\Gamma(x^*_{i,t+2})|z^*_{it}\right] = 0.$$

Then by Assumption 1.3.5.1, for at least one t we have that

$$\Gamma(x^*_{i,t+2}) = 0. \tag{B.0.1}$$

Setting $\delta = 0$, $\gamma = 1$, and $\alpha\beta = \kappa_1 = \kappa_2/\kappa_3$, then equation (B.0.1) is solved trivially. Assumption 1.3.5.3 eliminates this trivial solution. Differentiating equation (B.0.1) by $g^*_{i,t+2}$ and noting that $g^*_{i,t} > 0$ for i and t obtains

$$\begin{aligned}&\alpha_0(1-\gamma_0)\beta_0^2(1+r_{it+1})\varphi_{it+1}(\delta_0)\varphi_{it+2}(\delta_0)\left(\frac{g^*_{it+1}}{g^{*\alpha_0}_{it}}\right)^{1-\gamma_0}\left(\frac{g^*_{it+2}}{g^{*\alpha_0}_{it+1}}\right)^{1-\gamma_0} = \\ &\tilde{\alpha}(1-\tilde{\gamma})\tilde{\beta}^2(1+r_{it+1})\varphi_{it+1}(\tilde{\delta})\varphi_{it+2}(\tilde{\delta})\left(\frac{g^*_{it+1}}{g^{*\tilde{\alpha}}_{it}}\right)^{1-\tilde{\gamma}}\left(\frac{g^*_{it+2}}{g^{*\tilde{\alpha}}_{it+1}}\right)^{1-\tilde{\gamma}}.\end{aligned} \tag{B.0.2}$$

Differentiating equation (B.0.2) with respect to $g^*_{i,t+2}$ and again noting that $g^*_{i,t} > 0$ for i and t gives

$$\begin{aligned}&\alpha_0(1-\gamma_0)^2\beta_0^2(1+r_{it+1})\varphi_{it+1}(\delta_0)\varphi_{it+2}(\delta_0)\left(\frac{g^*_{it+1}}{g^{*\alpha_0}_{it}}\right)^{1-\gamma_0}\left(\frac{g^*_{it+2}}{g^{*\alpha_0}_{it+1}}\right)^{1-\gamma_0} = \\ &\tilde{\alpha}(1-\tilde{\gamma})^2\tilde{\beta}^2(1+r_{it+1})\varphi_{it+1}(\tilde{\delta})\varphi_{it+2}(\tilde{\delta})\left(\frac{g^*_{it+1}}{g^{*\tilde{\alpha}}_{it}}\right)^{1-\tilde{\gamma}}\left(\frac{g^*_{it+2}}{g^{*\tilde{\alpha}}_{it+1}}\right)^{1-\tilde{\gamma}}.\end{aligned} \tag{B.0.3}$$

Therefore, from equations (B.0.2) and (B.0.3) we have that $\gamma_0 = \tilde{\gamma}$. Differentiating equation (B.0.2) with respect to $g^*_{i,t}$ obtains

$$\alpha_0^2(1-\gamma_0)^2\beta_0^2(1+r_{it+1})\varphi_{it+1}(\delta_0)\varphi_{it+2}(\delta_0)\left(\frac{g^*_{it+1}}{g^{*\alpha_0}_{it}}\right)^{1-\gamma_0}\left(\frac{g^*_{it+2}}{g^{*\alpha_0}_{it+1}}\right)^{1-\gamma_0} =$$
$$\tilde{\alpha}^2(1-\tilde{\gamma})^2\tilde{\beta}^2(1+r_{it+1})\varphi_{it+1}(\tilde{\delta})\varphi_{it+2}(\tilde{\delta})\left(\frac{g^*_{it+1}}{g^{*\tilde{\alpha}}_{it}}\right)^{1-\tilde{\gamma}}\left(\frac{g^*_{it+2}}{g^{*\tilde{\alpha}}_{it+1}}\right)^{1-\tilde{\gamma}}. \tag{B.0.4}$$

From equations (B.0.2) and (B.0.4) and $\gamma_0 = \tilde{\gamma}$ conclude that $\alpha_0 = \tilde{\alpha}$. Substituting what we have so far into equation (B.0.2) obtains

$$\begin{aligned}
\beta_0^2\varphi_{it+1}(\delta_0)\varphi_{it+2}(\delta_0) &= \tilde{\beta}^2\varphi_{it+1}(\tilde{\delta})\varphi_{it+2}(\tilde{\delta}) \Rightarrow \\
2\ln(\beta_0) + \Delta_2 w_{i,t+2}\delta_0 &= 2\ln(\tilde{\beta}) + \Delta_2 w_{i,t+2}\tilde{\delta} \Rightarrow \\
Q_{i,t+2}D_0 &= Q_{i,t+2}\tilde{D},
\end{aligned} \tag{B.0.5}$$

where $Q_{i,t+2} = (1, \Delta_2 w_{i,t+2})$ and $D = (2\ln(\beta), \delta')'$. Then from equation (B.0.5) we have that

$$E[Q'_{i,t+2}Q_{i,t+2}]D_0 = E[Q'_{i,t+2}Q_{i,t+2}]\tilde{D},$$

which by Assumption 1.3.5.2 obtains $D_0 = D_1$, that is, $\delta_0 = \tilde{\delta}$ and $\beta_0 = \tilde{\beta}$. Finally substituting what we have so far into equation (B.0.1) obtains

$$X_1\kappa_{0,1} + X_2\kappa_{0,2} + X_3\kappa_{0,3} = X_1\tilde{\kappa}_1 + X_2\tilde{\kappa}_2 + X_3\tilde{\kappa}_3, \tag{B.0.6}$$

where $X_1 = 1$, $X_2 = \beta_0(1+r_{it+1})\frac{\varphi_{it+1}}{g^*_{it+1}}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma_0}$, and $X_3 = \alpha_0\beta_0\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma_0}$. Full rank of $E[X'X]$, where $X = (X_1, X_2, X_3)$, comes from variation in $r_{i,t+1}$ and the functional difference between $g^{*(1-\gamma_0)}_{it+1}$ and $g^{*(-\gamma_0)}_{it+1}$. This and equation (B.0.6) obtains $\kappa_{0,1} = \tilde{\kappa}_1$, $\kappa_{0,2} = \tilde{\kappa}_2$, and $\kappa_{0,3} = \tilde{\kappa}_3$.

APPENDIX C

CONSTRUCTION OF THE SIMULATION EXERCISE

The generic household maximizes remaining lifetime discounted expected utility with respect to consumption stream $\{c_s, s = t, \cdots, T\}$ subject to intertemporal budget constraint:

$$\begin{aligned} & E_t \sum_{s=t}^{T} \beta^{s-t} \phi_s \frac{\tilde{c}_s^{1-\gamma} - 1}{1-\gamma} \\ & \text{s.t.} \qquad a_{t+1} = (1 + r_{t+1})(a_t + y_t - c_t) \end{aligned} \tag{C.0.1}$$

where $\tilde{c}_t = c_t / c_{t-1}^{\alpha}$ and expectation is taken conditional on all information for the household at time t, c_t denotes real consumption of a single homogenous nondurable good, a_t denotes assets held by the household, y_t household's labor income and r_t the real net interest rate faced by the household at time t. In simulation we also augment utility with household-specific "taste shifter" ϕ_t, defined similar to estimation part of the paper. It is assumed that the length of life T is finite and known with certainty. It is also assumed that the household has no bequest motive so that $a_T = 0$. The time discount factor β and habit formation parameter α are positive, and $\gamma > 1$.

Labor income y_t is decomposed into the following components: permanent p_t and transitory u_t so that we have:

$$y_t = p_t u_t$$

$$p_{t+1} = p_t v_{t+1}$$

where the transitory income shock u_t and permanent income shock v_t are *iid* lognormal with unit mean and constant variances $(e^{\sigma_u^2} - 1)$ and $(e^{\sigma_v^2} - 1)$, respectively. We assume that innovations to income are independent over time and across households, hence there are no aggregate income shocks.

The interest rate is generated by a stationary $AR(1)$ process:

$$r_{t+1} = (1 - \rho)\mu_r + \rho r_t + \upsilon_{t+1}$$

where μ_r is the unconditional mean, $\rho \in (0, 1)$ is the autoregressive coefficient and υ_{t+1} is *iid* normal with zero mean and variance σ_υ^2.

We redefine the budget constraint in (C.0.1) in terms of cash on hand $x_t = a_t + y_t$:

$$x_{t+1} = (1 + r_{t+1})(x_t - c_t) + y_{t+1} \tag{C.0.2}$$

Then we rewrite all relevant variables in terms of their ratios to the permanent income: $\hat{z}_t = z_t/p_t$, where $z_t = x_t, c_t, y_t$. Then the budget constraint (C.0.2) takes the form:

$$\hat{x}_{t+1} = (1 + r_{t+1})\frac{(\hat{x}_t - \hat{c}_t)}{v_{t+1}} + u_{t+1}$$

Represent consumption growth rate $\hat{g}_t = \frac{\hat{c}_t}{\hat{c}_{t-1}}$. Then the Euler condition is given as:

$$E_t\left[\beta(1+r_{t+1})\frac{\varphi_{t+1}}{\hat{g}_{t+1}v_{t+1}}\left(\frac{\hat{g}_{t+1}}{\hat{g}_t^\alpha}\frac{v_{t+1}}{v_t^\alpha}\right)^{1-\gamma}\left(1-\alpha\beta\varphi_{t+2}\left(\frac{\hat{g}_{t+2}}{\hat{g}_{t+1}^\alpha}\frac{v_{t+2}}{v_{t+1}^\alpha}\right)^{1-\gamma}\right)-\left(1-\alpha\beta\varphi_{t+1}\left(\frac{\hat{g}_{t+1}}{\hat{g}_t^\alpha}\frac{v_{t+1}}{v_t^\alpha}\right)^{1-\gamma}\right)\right]=0 \tag{C.0.3}$$

where $\varphi_t = \exp(\delta'\Delta w_t)$ and w_t are household-related characteristics other than consumption. In particular, we use past income and current age squared as such characteristics.

In Table C we summarize the values of parameters we use in this simulation exercise. The "deep" parameters of the model are taken as follows: utility curvature parameter equals to 5, habit formation parameter equals to 0.85, and time discount factor equals to 0.95. The values of parameters governing the stochastic environment are largely borrowed from the related literature. Given the shocks to income process and the interest rate, we solve for optimal consumption rule recursively, starting from terminal condition $\hat{c}_T = \hat{x}_T$. The consumption rule at time t is a function of three endogenous state variables $\hat{c}_{t-1}, \hat{x}_t, r_t$ and two exogenous

state variables u_t and v_t. The solution is obtained over a fine grid on continuous state variables $\hat{c}_{t-1}$ and $\hat{x}_t$. Every time period the household is allowed to borrow the amount it can pay back with certainty by the end of its life, which determines $\hat{x}_{min}$ for the grid on $\hat{x}$ for each time period t. Following Tauchen (1986) we approximate the interest rate by first order discrete Markov process. The algorithm produces $\hat{c}_t$ as the solution for the Euler equation (C.0.3) for each quintuple $(\hat{c}_{t-1}, \hat{x}_t, r_t, v_t, u_t)$.

As the state space is relatively large for this simulation problem, we limit grid points to 20 for $\hat{c}_{t-1}$ and $\hat{x}_t$, to 5 for r_t and to 3 for v_t and u_t. After obtaining the solution for $\hat{c}_t$ we use a polynomial regression to approximate $\hat{c}(\hat{c}_{t-1}, \hat{x}_t)$ for each (discretized) choice of r_t, v_t and u_t. Life-time consumption paths are simulated using starting values on $\hat{c}_0$ and $\hat{x}_1$ and generated random draws for interest rates and shocks related to income process. We simulate 40-period consumption paths for 4000 ex ante identical individuals and remove first 8 and last 8 years to obtain 24-period time dimension of the simulated panel data.

Table C1: Parameter values used in Monte Carlo simulation

Parameter	
Utility curvature parameter, γ	5.0
Time discount factor, β	0.95
Habit formation parameter, α	0.85
Mean of the interest rate, μ^r	0.05
Autoregressive coefficient of r_{it}, ρ	0.60
Standard deviation of interest rate shocks, σ_υ	0.025
Standard deviation of permanent income innovation, σ_v	0.015
Standard deviation of transitory income innovation, σ_u	0.100
Variance of measurement errors, σ_η^2	0.06 (20% noise)

APPENDIX D

INTERTEMPORAL ELASTICITY OF SUBSTITUTION

In this section we calculate individual-specific intertemporal elasticities of substitution . Individual-specific intertemporal elasticity of substitution can be found from:

$$\frac{1}{IES_{it}} = \left| \frac{\partial \ln \frac{MU_{it}}{MU_{it+1}}}{\partial \ln \frac{c_{it+1}}{c_{it}}} \right|$$

where

$$\begin{aligned} \frac{MU_{it}}{MU_{it+1}} &= \frac{\frac{\phi_{it}}{c_{it}} \left(\frac{c_{it}}{c_{it-1}^{\alpha}} \right)^{1-\gamma} - \alpha\beta \frac{\phi_{it+1}}{c_{it}} \left(\frac{c_{it+1}}{c_{it}^{\alpha}} \right)^{1-\gamma}}{\frac{\phi_{it+1}}{c_{it+1}} \left(\frac{c_{it+1}}{c_{it}^{\alpha}} \right)^{1-\gamma} - \alpha\beta \frac{\phi_{it+2}}{c_{it+1}} \left(\frac{c_{it+2}}{c_{it+1}^{\alpha}} \right)^{1-\gamma}} \\ &= \frac{\left(1 - \alpha\beta\varphi_{it+1} \left(\frac{g_{it+1}}{g_{it}^{\alpha}} \right)^{1-\gamma} \right)}{\frac{\varphi_{it+1}}{g_{it+1}} \left(\frac{g_{it+1}}{g_{it}^{\alpha}} \right)^{1-\gamma} \left(1 - \alpha\beta\varphi_{it+2} \left(\frac{g_{it+2}}{g_{it+1}^{\alpha}} \right)^{1-\gamma} \right)} \end{aligned} \tag{D.0.1}$$

Taking logs of (D.0.1) and partial derivatives with respect to $\ln g_{it+1} = \ln(c_{it+1}\, c_{it})$ we obtain:

$$\frac{1}{IES_{it}} = \gamma - \frac{\alpha\beta(1-\gamma)\varphi_{it+1} \left(\frac{g_{it+1}}{g_{it}^{\alpha}} \right)^{1-\gamma}}{1 - \alpha\beta\varphi_{it+1} \left(\frac{g_{it+1}}{g_{it}^{\alpha}} \right)^{1-\gamma}} - \frac{\alpha^2\beta(1-\gamma)\varphi_{it+2} \left(\frac{g_{it+2}}{g_{it+1}^{\alpha}} \right)^{1-\gamma}}{1 - \alpha\beta\varphi_{it+2} \left(\frac{g_{it+2}}{g_{it+1}^{\alpha}} \right)^{1-\gamma}} \tag{D.0.2}$$

Since $\alpha \geq 0$ and $\gamma \geq 1$, we obtain the following bounds

$$\frac{1}{IES_{it}} \geq \gamma, \qquad IES_{it} \leq \frac{1}{\gamma}. \tag{D.0.3}$$

These inequalities are strict for $\alpha > 0$ and $\gamma > 1$. In order to derive bounds for the (inverse) IES, we must take into account measurement errors in observed consumption. To do so we first rewrite equation (D.0.2) as follows

$$\begin{aligned}
\tfrac{1}{IES_{it}} = & \ \gamma - \alpha\beta(1-\gamma)\varphi_{it+1}\left(\tfrac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\textstyle\sum_{j=0}^{\infty}\left(\alpha\beta\varphi_{it+1}\left(\tfrac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j} \\
& -\alpha^2\beta(1-\gamma)\varphi_{it+2}\left(\tfrac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}\textstyle\sum_{j=0}^{\infty}\left(\alpha\beta\varphi_{it+2}\left(\tfrac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}\right)^{j} \\
= & \ \gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty}\left(\alpha\beta\varphi_{it+1}\left(\tfrac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j} \\
& -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty}\left(\alpha\beta\varphi_{it+2}\left(\tfrac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}\right)^{j},
\end{aligned}$$

which is a valid representation since the assumption of positive marginal utilities implies that each term in the infinite sum is between 0 and 1. For the same reason, the dominated convergence theorem applies and we find that

$$\begin{aligned}
E\left[\tfrac{1}{IES_{it}}|z_{it}\right] = & \ \gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty}E\left[\left(\alpha\beta\varphi_{it+1}\left(\tfrac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right] \\
& -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty}E\left[\left(\alpha\beta\varphi_{it+2}\left(\tfrac{g_{it+2}}{g_{it+1}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right],
\end{aligned}$$

Next, for each j we have that

$$\begin{aligned}
E\left[\left(\alpha\beta\varphi_{it+1}\left(\tfrac{g_{it+1}^{*}}{g_{it}^{*\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right] &= E\left[\left(\alpha\beta\varphi_{it+1}\left(\tfrac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}\left(\left(\tfrac{v_{it+1}}{v_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right] \\
&= E\left[\left(\alpha\beta\varphi_{it+1}\left(\tfrac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right]E\left[\left(\left(\tfrac{v_{it+1}}{v_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right] \qquad \text{(D.0.4)}
\end{aligned}$$

Because $j \geq 1$. Jensen's inequality implies that

$$E\left[\left(\left(\frac{v_{it+1}}{v_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right] \geq \left(E\left[\left(\frac{v_{it+1}}{v_{it}^{\alpha}}\right)^{1-\gamma}|z_{it}\right]\right)^{j} = \mathcal{A}_3^{j}. \qquad \text{(D.0.5)}$$

Notice that $\mathcal{A}_3$ is exactly the quantity defined in the derivation of the moment condition in Appendix A. Therefore,

$$E\left[\left(\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}^{*}}{g_{it}^{*\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right] \geq E\left[\left(\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g_{it}^{\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right]\mathcal{A}_3^{j} \qquad \text{(D.0.6)}$$

Combining equations (D.0.3), (D.0.4) and (D.0.6), we find that

$$\begin{aligned}\gamma \leq E\left[\tfrac{1}{IES_{it}}|z_{it}\right] \leq \quad & \gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j |z_{it}\right] \\ & -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)^j |z_{it}\right],\end{aligned}$$

with strict inequalities if $\alpha > 0$ and $\gamma > 1$. Again, by Jensen's inequality we have that $(E[1/IES_{it}|z_{it}])^{-1} \leq E[IES_{it}|z_{it}]$. The corresponding bound for $E[IES_{it}|z_{it}]$ is given by

$$\begin{aligned}\tfrac{1}{\gamma} \geq E\left[IES_{it}|z_{it}\right] \geq \quad & \Bigg(\gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j |z_{it}\right] \\ & -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)^j |z_{it}\right]\Bigg)^{-1}.\end{aligned}$$

Under the assumption that measurement errors are distributed log-normal as in Section 1.3, the inverse IES is point identified. To see this, note that under this assumption,

$$E\left[\left(\left(\frac{v_{it+1}}{v_{it}^{\alpha}}\right)^{1-\gamma}\right)^j |z_{it}\right] = \mathcal{A}_3^{j^2},$$

where $A_3 = \exp\{\sigma^2\left((1+\alpha+\alpha^2)(1-\gamma)^2\right)\}$. Then straightforward calculations give

$$\begin{aligned}E\left[\tfrac{1}{IES_{it}}|z_{it}\right] = \quad & \gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j |z_{it}\right] \\ & -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)^j |z_{it}\right].\end{aligned}$$

The parametric distributional assumption for measurement errors does not entail point identification of the IES, but does reduce the bound as follows

$$\begin{aligned}\tfrac{1}{\gamma} \geq E\left[IES_{it}|z_{it}\right] \geq \quad & \Bigg(\gamma - (1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+1}\left(\frac{g^*_{it+1}}{g^{*\alpha}_{it}}\right)^{1-\gamma}\right)^j |z_{it}\right] \\ & -\alpha(1-\gamma)\textstyle\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+2}\left(\frac{g^*_{it+2}}{g^{*\alpha}_{it+1}}\right)^{1-\gamma}\right)^j |z_{it}\right]\Bigg)^{-1}.\end{aligned}$$

This bound is typically more narrow that that in equation (8.14) since $\mathcal{A}_3 > 1$ so that $\mathcal{A}_3^{-j^2} < \mathcal{A}_3^{-j}$.

APPENDIX E

RELATIVE RISK AVERSION

In this section we calculate individual-specific relative risk aversion parameters. These coefficients correspond to curvature and are closely related to the elasticities of the marginal utility of consumption with respect to consumption. Individual-specific relative risk aversion is defined as:

$$RRA_{it} = -c_{it}\frac{\Lambda^{cc}_{it}}{\Lambda^{c}_{it}}$$

where $\Lambda^{c}_{it} = MU_{it}$ and $\Lambda^{cc}_{it} = \frac{\partial MU_{it}}{\partial c_{it}}$. Consequently, the risk aversion parameters implied by our model are given by:

$$RRA_{it} = \frac{\gamma - \alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma} - \alpha^2\beta(1-\gamma)\varphi_{it+1}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}}{1 - \alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}} \tag{E.0.1}$$

If the observed consumption c^* is contaminated with measurement errors, we must take these into account when calculating individual-specific RRAs. We first rewrite equation (E.0.1) as follows

$$RRA_{it} = \gamma - (1+\alpha)(1-\gamma)\sum_{j=1}^{\infty}\left(\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}}{g^{\alpha}_{it}}\right)^{1-\gamma}\right)^{j}.$$

The same arguments as those used in the previous section also validate this expression. Therefore, calculations similar to those in the previous section lead us to the inequalities

$$\gamma \leq E\left[RRA_{it}|z_{it}\right] \leq \gamma - (1+\alpha)(1-\gamma)\sum_{j=1}^{\infty} E\left[\left(\mathcal{A}_3^{-1}\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}^{*}}{g_{it}^{*\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right],$$

with strict inequality if $\alpha > 0$ and $\gamma > 1$. Furthermore, if the log-normal assumption on measurement errors is maintained, then we find that

$$E\left[RRA_{it}|z_{it}\right] = \gamma - (1+\alpha)(1-\gamma)\sum_{j=1}^{\infty} E\left[\mathcal{A}_3^{-j^2}\left(\alpha\beta\varphi_{it+1}\left(\frac{g_{it+1}^{*}}{g_{it}^{*\alpha}}\right)^{1-\gamma}\right)^{j}|z_{it}\right],$$

where, again, $A_3 = \exp\{\sigma^2\left((1+\alpha+\alpha^2)(1-\gamma)^2\right)\}$.

Table E1: Estimation of the Euler equation with habit formation (Detailed version of Table 1.3)

Parameters	Nonparametric ME				Log-normal ME			No ME
	Internal habit			External habit				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
γ	5.279	5.089	4.950	6.024	6.899	3.522	3.029	3.126
	(0.057)	(0.107)	(0.099)	(0.071)	(2.460)	(0.662)	(0.366)	(0.368)
β	0.987	0.973	0.929	0.983	0.955	0.897	0.956	0.801
	(0.068)	(0.013)	(0.043)	(0.025)	(0.023)	(0.049)	(0.030)	(0.170)
α	0.831	0.801	0.807	0.800	0.780	0.573	0.799	0.585
	(0.016)	(0.028)	(0.028)	(0.011)	(0.092)	(0.128)	(0.115)	(0.252)
a				0.473				
				(0.014)				
σ^2					0.042	0.072	0.036	
					(0.012)	(0.008)	(0.002)	
% noise					35	45	32	
Demographics:								
family size	0.014	0.516	1.471	-0.365	0.078	0.025	-2.131	2.332
	(0.144)	(0.243)	(0.178)	(0.087)	(0.545)	(0.538)	(0.731)	(1.488)
age squared	1.567	0.304	0.882	2.018	0.099	0.395	2.122	-0.964
	(0.206)	(0.338)	(0.191)	(0.129)	(0.544)	(0.540)	(0.634)	(2.592)
Nonparametric ME:								
κ_1	0.016	0.176	0.028	0.037				
	(0.005)	(0.053)	(0.011)	(0.015)				
κ_2	0.744	0.699	0.871	0.799				
	(0.385)	(0.719)	(0.650)	(1.658)				
κ_3	0.763	0.377	0.958	0.540				
	(0.328)	(1.138)	(0.600)	(1.258)				
Some savings			yes				yes	
Aggr. shocks:								
t+1		-1.116				-0.244		
		(0.445)				(0.179)		
t+2		-0.375				-0.154		
		(0.413)				(0.152)		
t+3		-0.021				0.123		
		(0.476)				(0.177)		
t+4		0.987				0.161		
		(0.451)				(0.168)		
t+5		0.932				-0.026		
		(0.458)				(0.160)		
t+6		-1.169				-0.176		
		(0.453)				(0.165)		
t+7		-0.176				-0.093		
		(0.441)				(0.150)		
t+8		0.454				0.037		
		(0.455)				(0.166)		
t+9		-0.255				-0.145		
		(0.491)				(0.162)		
$Q(\theta)$	0.008	0.006	0.014	0.009	0.009	0.007	0.017	0.181
J statistic	29.1	23.0	25.6	31.4	33.5	23.2	30.6	619.0
d.f.	42	33	42	41	44	35	44	45
p value	0.933	0.902	0.978	0.858	0.873	0.936	0.936	0.000
N	3,402	3,402	1,754	3,402	3,402	3,402	1,754	3,402