INTERTEMPORAL CONSUMPTION CHOICES, TRANSACTION COSTS
AND LIMITED PARTICIPATION TO FINANCIAL MARKETS:
RECONCILING DATA AND THEORY

Orazio P. Attanasio
Monica Paiella

Intertemporal Consumption Choices, Transaction Costs and Limited Participation to Financial
Markets: Reconciling Data and Theory
Orazio P. Attanasio and Monica Paiella
NBER Working Paper No. 12412
July 2006
JEL No. G11, G12

ABSTRACT

This paper builds a unifying framework that, within the theory of intertemporal consumption choices,
brings together the limited participation -based explanation of the poor empirical performance of the
C-CAPM and the transaction costs-based explanation of incomplete portfolios. Using the
implications of the consumption model and observed household consumption and portfolio choices,
we identify the preference parameters of interest and a lower bound for the costs rationalizing
non-participation in financial markets, in the presence of unobserved heterogeneity in tastes for
consumption and portfolio allocation. Using the US Consumer Expenditure Survey and assuming
isoelastic preferences, we estimate the coefficient of relative risk aversion at 1.7 and a cost bound
of 0.4 percent of non-durable consumption. Our estimate of the preference parameter is theoretically
plausible and the bound sufficiently small to be likely to be exceeded by the actual total (observable
and unobservable) costs of participating to financial markets.

Orazio Attanasio
Department of Economics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM
and NBER
o.attanasio@ucl.ac.uk

Monica Paiella
Research Department
Bank of Italy
Via Nazionale, 91
Roma 00184
ITALY
monica.paiella@bancaditalia.it

1 Introduction[1]

The dynamics of consumption and saving behaviour is obviously related to the demand for assets and, as such, can provide valuable information for equilibrium asset pricing. The pathbreaking contributions of Lucas (1978) and Breeden (1979) made the link between the Euler equation for consumption and equilibrium asset prices explicit and used the first-order conditions of a consumer problem to build what is known as the Consumption Capital Asset Pricing Model (C-CAPM). Unfortunately, despite the formal elegance and the analytical simplicity of the C-CAPM, the empirical performance of the model has been, at best, mixed. Since the early studies by Hansen and Singleton (1982, 1983), it was clear that observed asset returns are inconsistent with the dynamics of consumption choices, at least as observed in aggregate data. Such evidence was re-inforced and confirmed in a large number of other studies. Some studies, such as Hansen and Jagannathan (1991), suggested that one of the reasons for the poor empirical performance of the model is the low level of variability of aggregate consumption growth.

Recently, there have been several attempts at rationalizing this discouraging evidence and several studies have explored the possibility that limited participation in financial markets might explain the disparity between theoretical predictions and empirical evidence. More precisely, since the first order conditions of asset pricing models hold with equality only for those households who own complete portfolios, the models should be tested for this subset of households and not for the whole population. As a consequence, since in practice relatively few households hold shares directly, even abstracting from standard aggregation issues arising from the non-linearity of the marginal rate of substitution, the use of aggregate consumption data in evaluating asset pricing models could be very misleading.

These points have been stressed by Mankiw and Zeldes (1991), Attanasio et al. (2002), Vissing Jorgensen (2002) and Paiella (2004), among others, who propose limited financial market participation as a unified framework for rationalizing the empirical rejection of the C-CAPM. These papers show that accounting for portfolio heterogeneity and in particular for non-participation into financial markets helps reconciling the predictions of the theory with the empirical evidence. Attanasio et al. (2002), for instance, show that focussing

---

[1]Only the authors are responsible for the contents of this paper which do not necessarily reflect the views of the Bank of Italy.

on the consumption of stockholders, not only yields estimates of preference parameters that are in line with the theory, but one does not reject the overidentifying restrictions implied by the model and, relatedly, the moments of the marginal rate of substitution are within the Hansen-Jagannathan bounds.

While these studies have been somewhat successful in reconciling the C-CAPM with the empirical evidence, they take limited participation as given and make no attempt of rationalizing it. Limited participation is in itself a puzzle for the intertemporal consumption model, just like the observed substantial differences in portfolio composition across agents and over the life cycle. Merton (1969) and Samuelson (1969) have illustrated how such behaviour is inconsistent with the maximization of expected lifetime utility, which predicts that rational agents should invest an arbitrarily small amount in all assets with positive expected return, including the risky ones, unless there are non-linearities in the budget constraint.

One possible and obvious way to rationalize incomplete portfolios within the intertemporal consumption model is by invoking non-proportional costs of financial market participation (explicit and non-explicit). As such costs are for the most part unobservable, the plausibility of this explanation depends crucially on the magnitude these costs should have to explain observed data. Should the size of these costs be 'reasonable' one could find this explanation attractive. Should instead the size of the participation costs that rationalize observed data be very large one would probably dismiss this explanation.

One of the first papers to consider this approach was the study by Luttmer (1999) in which, using aggregate data, he provides a lower bound on the transaction costs that would rationalize the model in the face of available data. Paiella (2006), using micro data, provides evidence in support of this hypothesis by bounding from below the costs of participation to some financial markets. Her bounds for the stock market are as small as \$130 per year, which implies that it is likely that the true (unobservable) costs of participation may exceed this level in reality.[2]

This paper brings together the limited participation-based explanation of the poor empirical performance of the C-CAPM and the transaction cost-based explanation of incomplete portfolios to build a powerful test of the theory of the intertemporal allocation of consumption. Using the implications of the consumption model and observed consumption and portfolio choices, we

---

[2]Vissing-Jorgensen (2003) provides additional evidence in favor of the participation cost hypothesis, at least for some consumers.

show how to identify the preference parameters of interest and a lower bound for the costs of participation in financial markets rationalizing participation choices in the presence of unobserved heterogeneity in consumption tastes. The estimation of the parameters of interest is based on the necessary conditions for the optimality of observed behavior of financial market participants and non-participants. The methodology relies on the (empirical) distinction between the consumption path of those households holding a well-diversified portfolio of assets and the consumption path of those holding incomplete portfolios. The former exploit all trading opportunities and their consumption dynamics is consistent with the time series properties of asset prices. The latter do not and by structurally estimating a lower bound to the gains they forego for holding an incomplete portfolio we can bound from below the costs that would rationalize their non-participation. We implement our approach using individual level data to estimate the preference parameters and the cost bound.[3]

Our approach for the cost bound identification builds on and generalizes the works of Luttmer (1999) and Paiella (2006). Luttmer (1999) focuses on the losses for leaving unexploited some trading opportunities and proposes a lower bound on the level of fixed transaction costs reconciling per-capita expenditure and asset returns. Hence, the foregone gains that Luttmer identifies bound from below the cost of trading that would justify not taking advantage of temporary changes in returns not matched by changes in the riskiness of assets. Luttmer's frictions are the costs that a representative agent must pay to trade and modify her consumption path. Luttmer's consumer pays a transaction cost whenever consumption differs from income. Instead, we use individual level data and, by distinguishing between holders and non-holders of risky assets, we focus on the loss for missing out on the equity premium. Our consumers do not pay a cost to save in a safe assets and our frictions are the costs that individual agents must pay in order to participate to the market for risky assets.[4]

---

[3]Earlier work considering the impact of frictions for asset pricing models includes Luttmer (1996). Luttmer shows that small proportional transaction costs on stocks may reconcile aggregate consumption-based discount factors with asset return data.

[4]Luttmer's estimates of the bound to the costs of trading are potentially biased because they are obtained using aggregate expenditure data, which include both the consumption of those who hold financial assets and the consumption of those who do not. For the latter the benefits of trading in financial markets are likely to go beyond those associated to the capturing of excess returns.

Our participation costs can be interpreted in two ways. First of all, they can be thought of as reflecting the costs of information and transaction that would induce households not to invest in some securities; secondly, they can be thought of as the costs of following near-rational decision rules. In the first instance, we would have two types of households: one who pays the fixed cost to invest in asset j and whose consumption, net of the cost, is coherent with the Euler equation for asset j; the other who does not pay the cost because its expected gain from the investment is relatively too low. Households share the same preference parameters, investment opportunities and information sets, but differences in their observable socio-demographic characteristics and possibly in unobservable attributes result in differences in the gains from financial market participation. In the second instance, households behave according to different decision rules. Those who participate follow rational decision processes that can be modelled as solutions to the maximization of the intertemporal choice model. Those who do not participate follow heuristic decision processes. For the latter the gains of fully optimizing, that we estimate by maximizing their utility under the assumption of full rationality, can be expected to be lower than their costs of solving the model for the optimal intertemporal allocation of consumption, which are primarily costs of information, attention, etc.. Differences in socio-demographic characteristics and possibly in unobservable attributes can justify the differences in the costs of behaving according to rational decision processes.

Paiella (2006) focuses on the behaviour of non-participants. While her approach delivers lower bounds for the participation cost that are conceptually similar to those we propose here, these estimates are based on specific assumptions on preference parameters. Instead, we simultaneously estimate the preference parameters and the bounds on participation costs. Moreover, as we use information on both participants and non-participants we need less stringent assumptions about the nature of unobserved heterogeneity.

Using the US Consumer Expenditure Survey and assuming isoelastic preferences with multiplicative preference heterogeneity, we estimate the coefficient of relative risk aversion at 1.7 and a cost bound of 0.4 percent of non-durable consumption. Our estimate of the preference parameter is theoretically plausible and the bound sufficiently small to be likely to be exceeded by the actual total (observable and unobservable) costs of participating to financial markets. This implies that consumption asset pricing models provide an accurate description of the data once limited participation, fixed costs of participation and taste heterogeneity are properly accounted for.

The rest of the paper is organized as follows. In Section 2 we present the framework that we use to identify household preference parameters and the bound to the participation costs, within the type of environment specified by the model of intertemporal choice. Section 3 derives our econometric model based on the conditions for the optimality of consumption of stockholders and of non-stockholders. Section 4 presents the data. In Section 5, we discuss the results from the estimation. Section 6 concludes.

2 The model

Investing in financial assets involves information gathering, decision making, brokerage fees and/or other fixed costs that can create a disincentive to portfolio diversification. Such frictions may end up offseting the positive return paid by the asset. The heterogeneity of portfolio choices can then be explained on the basis of differences in socio-demographic and other, observable and unobservable individual specific characteristics, without the need of assuming heterogeneity in preferences parameters. Such differences would then also be reflected in differences in consumption. The paper tests this hypothesis by jointly estimating the curvature of households' utility function and bounding from below the costs that would rationalize incomplete portfolios for some consumers, but assuming that consumers are homogeneous in terms of the curvature of the utility function. While preferences and participation costs are jointly estimated using observations on both participants and non-participants, intuitively it is clear that the curvature of the utility function is identified by the consumption dynamics of those holding an optimal portfolio of assets vis à vis the dynamics of asset returns. On the other hand, the lower bound on participation cost is identified by the gains that incomplete portfolio holders forego by not-diversifying fully.

Consider an environment where households have rational expectations, intertemporally additively separable preferences over consumption, a strictly increasing and concave per-period utility function, $U(c_{h,t}, \xi_{h,t})$, and a positive subjective discount rate, β. We assume that the instantaneous utility function depends not only on consumption, $c_{h,t}$, but also on an unobservable taste shock, $\xi_{h,t}$. In the empirical specification we will assume that this shock enters multiplicatively. It can therefore be interpreted as representing heterogeneity in discount factors.[5]

---

[5]In addition to the unobserved component, the taste shifter can also have an observed

6

Households have access to two means to substitute consumption over time: a risky asset yielding r_{t+1} and a riskfree asset yielding r_{t+1}^f. Let's assume that in order to invest in the risky asset households must pay a fixed cost. This cost is higher than any cost the riskless asset investment may involve. On the basis of portfolio composition, it is possible to distinguish between two types of households: those who hold both risky and riskless financial assets and those who hold only riskless assets. For the risky asset holders, who have paid the fixed cost, the Euler equation for consumption must hold, i.e.:

$$E_t \left\{ \beta \frac{U'\left(c_{h,t+1}, \xi_{h,t+1}\right)}{U'\left(c_{h,t}, \xi_{h,t}\right)} \left(1 + r_{t+1}\right) \right\} = 1, \qquad (1)$$

where $E_t\{\}$ denotes the expectation conditional on the information available at time t, $U'()$ is the marginal utility of consumption and β is the time discount factor.

Let's turn to those who have chosen not to invest in the risky asset. Let $\{c^h\}_t$, $t = 1, 2, ...T$ be household h observable sequence of consumption choices. Since households choose optimally, conditional on the information available, and at time t they could have chosen any other feasible sequence of consumption bundles, their time t expected *ex-post* utility gain from deviating from $\{c^h\}_t$ must be non-positive. More specifically, we assume that, at time t non-shareholders could have paid a fixed cost of δ units of consumption, invested in the risky asset and adjusted consumption from $(c_{h,t}, c_{h,t+1})$ to $(c_{h,t} + a_{h,t} - \delta c_{h,t}, c_{h,t+1} + b_{h,t+1})$. $a_{h,t}$ and $b_{h,t+1}$ denote a feasible consumption perturbation. Optimality of their observed choices $(c_{h,t}, c_{h,t+1})$ implies that:

$$E_t \left\{ v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, \delta\right) \right\} \leq 0. \qquad (2)$$

where $v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, \delta\right)$ is the *ex-post* utility gain that they could have

---

one. This specification is consistent with the specifications often used in the empirical literature on Euler equations (see Attanasio , 1999). With CRRA utility, the instantaneous utility function would take the form: $U(c, z, \xi) = \frac{(c)^{1-\gamma}}{1-\gamma} \exp\{\theta z + \xi\}$, where the term z is a vector of observable variables, and ξ represents unobserved heterogeneity. In the empirical specification that we use we have not introduced z variables (such as demographic factors) in an unrestrictive fashion. Our utility is expressed in terms of consumption per adult equivalent.

obtained by paying the fixed cost $\delta c_{h,t}$ and perturbing consumption:

$$
\begin{aligned}
v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, \delta\right) \;=\; & \left\{ U\left(c_{h,t} + a_{h,t} - \delta c_{h,t}\right) + \beta U(c_{h,t+1} + b_{h,t+1}\right\} + \\
& - \left\{ U\left(c_{h,t}\right) + \beta U\left(c_{h,t+1}\right)\right\},
\end{aligned}
\tag{3}
$$

where we have suppressed the dependence of the utility function on $\xi_{h,t}$ for notational convenience. Equation (2) says that, net of the cost $\delta c_{h,t}$, the expected utility gain from perturbing the observed consumption path is non-positive. Hence, the investment is not worth. Inequalities like (2) must hold for any t.

The fixed cost δ cannot be observed directly. However, following an approach similar to that proposed by Luttmer (1999) and generalized recently by Pakes et al. (2005) (see also Manski 2003), we can place a lower bound on it. For any given $(a_{h,t}, b_{h,t+1})$, the function $E_t\left\{v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, \delta\right)\right\}$ is continuous and decreasing in δ, as $U_h()$ is continuous and increasing. Hence, for any given $(a_{h,t}, b_{h,t+1})$, there is a unique value $d(a_{h,t}, b_{h,t+1})$ such that (2) is satisfied if $\delta \geq d(a_{h,t}, b_{h,t+1})$. The function $d(a_{h,t}, b_{h,t+1})$ is defined implicitly as the solution to the equation $E_t\left\{v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, d\right)\right\} = 0$. In practice, we are interested in the lower bound d such that (2) is satisfied for any $\delta \geq d$, i.e. if $d = \max d(a_{h,t}, b_{h,t+1})$. As $E_t\left\{v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, \delta\right)\right\}$ is continuous and decreasing in δ, such d solves the equation:

$$
\max_{a_{h,t}, b_{h,t+1}} E_t\left\{v_{h,t+1}\left(a_{h,t}, b_{h,t+1}, d\right)\right\} = 0.
\tag{4}
$$

The parameter d is the Hicks compensating variation for non investing in an asset yielding r_{t+1}. d is a *lower* bound to the foregone gains for holding an incomplete portfolio, which in turn are a lower bound to the cost δ that would rationalize non-participation. The "true" foregone gains for holding a suboptimal portfolio are just a *lower* bound to the participation costs, because the (unobservable) costs δ may be so large that households are never close to deviating from their actual choices. In this instance, by construction, a level of gains that is much smaller than δ will suffice to rationalize observed choices. The bound will be closer to the true cost the more profitable the trading rule $(a_{h,t}, b_{h,t+1})$. Further, d is a lower bound to the foregone gains of incomplete portfolios: the expected utility gains of deviating from observed portfolio choices may be higher than those captured by equation (4) for at least two reasons. First, the framework behind equation (4) measures the expected gains of using an extra instrument to adjust consumption over two

8

periods. Thus, if the conditioning information set of the agent is larger than that of the econometrician, the agent may be actually able to obtain a higher utility gain than the econometrician can estimate. Second, we are approximating the utility from spreading the gains from the investment over the entire lifetime horizon of the utility maximizing agent with the utility from spreading the gains over the two periods when the investment takes place. This set up leaves households' consumption plans unchanged at all other dates and allows to appraise the gains that households forego for non-investing for one period by focusing just on their consumption at two adjacent dates.

Overall d provides the basis for a heuristic test of the cost of participation hypothesis: for the latter to be a plausible explanation of incomplete portfolios, any reasonable cost of participation must be higher than our estimated bound. Although this is not the most powerful test, it is indeed the most reliable. A more powerful test would compare the costs with the true foregone gain -not just with a lower bound, as done here. However, the estimation of the true foregone gain would require a much larger amount of information and/or assumptions.

3   Empirical specification

The analysis is based on the conventional assumption that utility exhibits constant relative risk aversion. Regarding the trading strategy in case of participation, we assume that after paying the fixed cost for investing in the risky asset, non-shareholders adjust their current savings.[6] Let $x_{h,t}^c(\alpha^c)$ denote the fraction of time t consumption they give up and invest in the risky asset. α^c is a vector of parameters to be estimated. We also assume that they consume all the returns on the investment when they realize it. The ex-post gain of non-shareholders can then be written as:

$$
\begin{aligned}
v_{h,t+1}\left(x_{h,t}^c(\alpha^c),\delta\right) = & \ U\left(c_{h,t}(1 - x_{h,t}^c(\alpha^c) - \delta)\right) + \\
& \ \beta U\left((c_{h,t+1} + x_{h,t}^c(\alpha^c)c_{h,t}(1 + r_{t+1}))\right) \\
& - \left\{U\left(c_{h,t}\right) + \beta U\left(c_{h,t+1}\right)\right\},
\end{aligned}
$$

---

[6]In the appendix, we consider the case where, rather than adjusting only their consumption (and savings), households are also allowed to modify their portfolio and shift resources from the risk free asset to the risky one. Given the limited information we have on portfolio composition we preferred to perform the exercise that we report.

where we have suppressed again the dependence of the utility function on $\xi_{h,t}$ for notational convenience. The estimation of the utility parameter and of the cost bound then relies on two sets of first-order conditions. The first set is the Euler equation in (1) which ensures the optimality of shareholders consumption. The second set consists of the following equations, which must hold for the set of non-shareholders:

$$E_t \left\{ D_1 v_{h,t+1} \left(x_{h,t}^c(\alpha^c), d \right) \right\} = 0, \quad h \in H_{ns}; \tag{5}$$

$$E_t \left\{ v_{h,t+1} \left(x_{h,t}^c(\alpha^c), d \right) \right\} = 0, \quad h \in H_{ns}, \tag{6}$$

where D_1 denotes the derivative with respect to the first argument of $v_{h,t+1}(.)$ and H_{ns} is the set of time t non-shareholders. Equation (5) determines the optimal trading strategy in case of participation, given the cost. Since, in practice, the actual cost, δ, is not observed, nor estimated and only a lower bound to the cost is identified, the optimal portfolio is determined as a function of a cost equal to its estimated bound, d, which is consistent with the rest of the analysis. Equation (6) determines the lower bound d to the participation cost δ, given the optimal investment.

Under the assumption of isoelastic preferences, and re-introducing unobserved heterogeneity, the Euler equation for shareholders in (1) becomes:

$$E_t \left\{ \beta \zeta_{h,t+1} \left(\frac{c_{h,t+1}}{c_{h,t}} \right)^{-\gamma} (1 + r_{t+1}) \right\} = 1, \quad h \in H_s, \tag{7}$$

where γ is the coefficient of relative risk aversion, H_s the set of shareholders and $\zeta_{h,t+1} = \left(\frac{\xi_{h,t+1}}{\xi_{h,t}} \right)$. Equation (6) for non-shareholders can be written as:

$$E_t \left\{ \beta \zeta_{h,t+1} \left(\frac{c_{h,t+1}^{1-\gamma} - \left((c_{h,t+1} + x_{h,t}^c(\alpha^c)c_{h,t}(1 + r_{t+1}) \right)^{1-\gamma}}{\left(c_{h,t}(1 - x_{h,t}^c(\alpha^c) - d) \right)^{1-\gamma} - c_{h,t}^{1-\gamma}} \right) \right\} = 1, \quad h \in H_{ns}, \tag{8}$$

which, as we show in Appendix A, can then be approximated as follows:

$$E_t \left\{ \beta \zeta_{h,t+1} \left(\frac{c_{h,t+1}}{c_{h,t}} \right)^{-\gamma} \frac{(1 + r_{t+1})}{1 + d/x_{h,t}^c(\alpha^c)} \right\} \approx 1. \tag{9}$$

Under the assumption that consumption and the rate of return on stock are jointly lognormal and homoskedastic, we can loglinearize (7) and obtain:

$$\log(1 + r_{t+1}) = \alpha_s + \gamma \Delta \log(c_{h,t+1}) + \varepsilon_{h,t+1}, \quad h \in H_s, \tag{10}$$

10

where α_s is a function of the (conditional) second-order moments of consumption and asset returns and the residual $\varepsilon_{h,t+1}$ includes the expectation errors as well as the transformation of the unobserved heterogeneity term $\zeta_{h,t+1}$. Similarly, we can loglinearize equation (9) and obtain:

$$\log\left(1 + r_{t+1}\right) = \alpha_{ns} + \gamma \Delta \log\left(c_{h,t+1}\right) + \log\left(\frac{x_{h,t}^c(\alpha^c)}{d + x_{h,t}^c(\alpha^c)}\right) + \varepsilon_{h,t+1}, \quad h \in H_{ns}.$$
(11)

Equations (10), and (11) together with (5) allow one to identify and estimate the coefficient of relative risk aversion, γ, and a lower bound, d, to the costs justifying limited financial market participation. They are conditions for optimality that must be satisfied by consumption choices: (10) must hold for shareholders, (5) and (11) for non-shareholders. Notice that (11) is non-linear in the parameters of interest.

The two Euler equations for participants and non-participants can be pooled together to obtain:

$$
\begin{aligned}
\log\left(1 + r_{t+1}\right) &= \alpha_s p_{h,t} + \alpha_{ns}(1 - p_{h,t}) + \gamma \Delta \log\left(c_{h,t+1}\right) + \\
&\quad + (1 - p_{h,t}) \log\left(\frac{x_{h,t}^c(\alpha^c)}{d + x_{h,t}^c(\alpha^c)}\right) + \varepsilon_{h,t+1},
\end{aligned}
$$
(12)

where $p_{h,t}$ is a dummy variable that takes the value of 1 for participants to the stock market.

Given the parameters of the investment rule, equation (12) can be estimated by standard GMM methods. Any instrument that is uncorrelated with the expectational errors and the unobserved heterogeneity terms will be a valid instrument.[7]

One difficulty with equation (12) is the presence of the term $\log\left(\frac{x_{h,t}^c(\alpha^c)}{d + x_{h,t}^c(\alpha^c)}\right)$, which implies a non-linearity in parameters. In principle, one could further linearize (11) by applying a first-order Taylor expansion to the only term that is non-linear in parameters. In particular, $\log\left(\frac{x_{h,t}^c(\alpha^c)}{d + x_{h,t}^c(\alpha^c)}\right)$ could be approximated by $-\frac{d}{x_{h,t}^c(\alpha^c)}$. However, as we expect this ratio to be in the order

---

[7]The residual terms of these equations will also include the deviation between the conditional second moments in the intercept term and their unconditional value. We will therefore require that these deviations are orthogonal to the instrument used. See Attanasio and Low (2003) for a discussion of these issues.

of 0.1, the approximation would be a poor one. We therefore prefer to apply non-linear GMM techniques to estimate our parameters.

Equation (11) differs from the standard Euler equation for the term $\log\left(\frac{x_{h,t}^c(\alpha^c)}{d+x_{h,t}^c(\alpha^c)}\right)$, which captures the difference in consumption growth between the relatively steep consumption path of shareholders and the flatter one of non-shareholders. If returns are high, the optimal investment in case of participation $(x_{h,t}^c(\alpha^c))$ would be large, unless non-shareholders' consumption is correlated to the return on the risky asset due, for example, to some correlation between individual income and the stock market. In this instance, in order to justify non-shareholding or, equivalently, significant differences between shareholders' and non-shareholders' consumption, costs must be high too. Furthermore, the more risk averse, the smaller the impact of costs: the more risk averse, the smaller the investment in risky assets and, consequently, the smaller the covariance of shareholders' expenditure with asset returns and the closer such covariance to that of non-shareholders' expenditure with returns.

As to the estimation of the investment rules based on equations (5), since the data used for the analysis consist of repeated cross-sections, and not of long individual consumption series, we cannot estimate individual optimal rules. However, we can estimate the trading rules by summing over the set of households who do not invest in the asset considered at t and taking unconditional expectations, which yields:

$$E\left\{\sum_{h\in H_{ns_t}} D_1 v_{h,t+1}\left(x_{h,t}^c(\alpha^c), d\right)\right\} = 0, \quad h \in H_{ns}. \tag{13}$$

The trading rule $x_{h,t}^c(\alpha^c)$ is assumed to be linear in a set of forecasting variables $z_{h,t}$ that help to select the most profitable level of investment which is then linear in consumption and wealth. In particular, in what follows, we assume $x_{h,t}^c(\alpha^c) = \alpha^{c\prime} z_{h,t}$. This set up allows capturing in the estimate the predictability of the components of asset returns that are correlated with consumption growth and the set of forecasting variables $z_{h,t}$. As we discuss below, the choice of the variables that determine the trading rule is somewhat arbitrary. The lower bound on the cost structure is then a function of the variables used in the trading rule.

4 Data

The estimation of the preference parameters and of the financial participation cost bound is based on data from the US Consumer Expenditure Survey, which is run on a continuous basis by the Bureau of Labor Statistics. The CEX is a representative sample of the US population. It is a rotating panel in which interviews occur continuously throughout the year, each consumer unit being interviewed every three months over a twelve-month period, apart from attrition. As households complete their participation, new ones are introduced into the panel on a regular basis and, as a whole, about 4500 households are interviewed each quarter, more or less evenly spread over the three months.

At the time of the last interview, households provide information on their asset holdings at that date and on the dollar difference with respect to the amounts held twelve months earlier. The asset categories in the CEX are: 1. checking, brokerage and other accounts; 2. saving accounts; 3. US saving bonds; 4. stocks, bonds, mutual funds and other securities. As a measure of risky asset holdings, we take the amounts held in stocks, bonds, mutual funds and other securities and US saving bonds. As a measure of riskless asset holdings, we take the amounts held in checking and saving accounts. In order to avoid problems arising from the simultaneity of expenditure growth between t and $t+1$ and portfolio composition at $t+1$, the asset holding status must be defined at the beginning of period t. For this purpose, for each asset category, we subtract from the stocks held at the time of the last interview the amount of savings (the dollar change) carried out over the previous twelve months. Hence, for each household we can define only one observation on the expected utility gain, $E_t \left\{ v_{h,t+1} \left(x_{h,t}^c(\alpha^c), d \right) \right\}$.

The consumption measure that we use is deseasonalized, real monthly per-adult equivalent expenditure on non-durable goods and services. Each quarterly interview collects household monthly expenditure data on a variety of goods and services for the previous three months. However, since the information on asset holdings is annual, we use only two observations on consumption and denote as c_t and as c_{t+1} household per-adult equivalent consumption based on the expenditure reported for the first and last month of the year covered by the survey.

The data used for the analysis cover the period from 1982 to 2001, first

quarter.[8] Since interviews occur every month, t runs for a total of 208 periods (months). From the initial sample we exclude those households that do not participate to all interviews, those living in rural areas or in university housing and those whose head was under 21 or over 75 years old. We also exclude those with incomplete income responses, those whose financial supplement contains invalid blanks either in the stocks of assets or in the dollar changes occurred with respect to the previous year and those whose stocks of checking and saving accounts and/or of shares and bonds are non-positive (15 percent of the sample). Finally, we drop those households whose monthly consumption falls in the 1 percent tails of the distribution or whose consumption growth over the year falls in the 5 percent tails. Overall, the sample used consists of 24, 016 households. The fraction of non-shareholders has fallen from almost 65 percent in the first half of the eighties to less than 60 percent towards the end of the past decade. Table 1 reports some descriptive statistics. Stockholders are slightly older than non-stockholders, they are significantly more educated, their consumption is higher and substantially more correlated with stock returns.

5   Results

As discussed above, the basis for our estimation is constituted by equations (12) and (13). These two equations are orthogonality conditions that we will exploit to obtain GMM estimates of the structural parameters. It is worth stressing that, although we have several thousands individual observations, consistency in estimation is achieved by having a large number of time periods. As discussed in Chamberlain (1984), large T asymptotic is necessary in such a situation if one is not willing to assume the presence of complete markets that make aggregate shocks identical for all consumers. Therefore, the fact that we have 208 time periods is crucial. In the estimation, we will also recognize the presence of aggregate shocks by allowing for arbitrary

---

[8]Around 1985-86 and 1995-1996, the sample design and the household identification numbers were changed and after the first quarter of 1986 and of 1996 no track is kept of those who had entered the survey in 1985 and in 1995, respectively. As a consequence of this and of the fact that the information on financial asset holdings is collected during the last interview, those households who have their first interview in the third and fourth quarter of 1985 or of 1995 had to be excluded from the sample. Thus, the sample used consists of households who have their first interview between January 1982 and June 1985, between January 1986 and June 1995 and between January 1996 and June 2000.

correlations among the residuals of individuals observed in the same time period, as described in the appendix. Given the data structure, which includes annual consumption growth observed at a monthly frequency, residuals for individuals observed in adjacient mnoths will be correlated. This correlation will decline only for individuals that are further than 11 months apart. As discussed in Appendix B, in the compuation of standard errors we take this structure into account. Finally, we also introduce cohort dummies in equations (12) and (13) to take into account possible differencs in tastes (and in particular discount factors) across cohorts.

While in principle it is possible to estimate the parameters of these models considering equations (12) and (13) simultaneously, we used a recursive approach. Given an initial guess for d and γ, we estimate the parameters of the trading rule by maximizing the utility gain in (13). Given the parameters of the trading rule, we then estimate the parameters in (12) by non-linear GMM. This second steps gives us new estimates for d and γ. We iterate this procedure until convergence.

The variables that enter the trading rule are somewhat arbitrary. It should be remembered, however, that an incorrect specification of the trading rule has only implications for the tightness of the bound. We specify the trading rule as a function of three variables: the risk free rate, the price/earning ratio and the term premium. The last variable is lagged three periods, while the other two are lagged two periods.

For the GMM procedure, we need instruments that are uncorrelated with the unobserved heterogeneity in tastes and that are lagged two or more periods given the structure of our residuals, arising from the overlapping of the observations on consumption growth.[9] In addition to the variables that enter the trading rule, we include a polynomial in the age of the household head, cohort dummies, a time trend and other aggregate (lagged) price variables, such the return on the risky asset, the risk premium and the price dividend ratio. As a measure of risky asset return, we take the return on the S&P500 CI, as riskless return we take the return on 3-month Treasury bills, the risk premium measure is given by the ratio of the yield of BAA bonds to that of AAA bonds, and the term premium measure is computed as ratio of the yield of 10-year government bonds to 3-month Treasury bills. The choice of the instruments is based on a regression aimed at identifying which of the available exogenous variables contributes the most to the prediction of the

---

[9]Interviews occur every month of the year.

15

return on the S&P500.

Table 2 reports the results of the estimation of the parameters of interest. The upper panel of the table reports the point estimates of the intercepts and of the slope of equation (12), and of the cost bound, given the investment rule. The bottom panel displays the coefficients of non-stockholders' investment rule in case of participation, given risk aversion and the cost.

The estimates in the upper panel indicate a coefficient of risk aversion of 1.7. While this coefficient is not estimated extremely precisely, the point estimate indicates a theoretically plausible magnitude. The estimate of the non-linear term imply a point etimate for the cost bound of 0.4 with a 95 percent confidence interval for the cost bound ranging from 0.1 percent to 1 percent of non-durable consumption. The cost bound is sufficiently small to suggest that it is likely that the actual total (observable and unobservable) costs of participation exceed it in reality. In fact, if we take average per-adult equivalent non-stockholders monthly non-durable consumption from Table 1 and multiply it by 12 and by 2.5, which is the mean of the per-adult equivalent scale, and then multiply this by the estimated bound, we obtain a dollar estimate of the cost bound of approximately \$72 per year.

The investment rule coefficients are precisely estimated based on equation (13). They imply that, given a cost of 0.4 percent of non-durable consumption and a risk aversion of 1.7, non-shareholders would maximize their gains from participation by investing in the risky asset 4.2 percent of their current consumption, on average. For costs higher than 0.4 percent they are better off by non-investing at all.

6   Concluding Remarks

This paper considers the poor empirical performance of the consumption-based capital asset pricing model and builds a unifying framework that, within the theory of intertemporal consumption choice, brings together limited participation and fixed participation costs to appraise their joint role in explaining the disparity between the standard model predictions and the empirical evidence. Allowing explicitly for heterogeneity in tastes, for differences in the consumption paths of shareholders and non-shareholders and for financial market participation frictions, we show how to identify the preference parameters of interest and a bound to the costs rationalizing incomplete portfolios. Our approach yields Euler equation-based estimates of relative

risk aversion around 1.7, which is a theoretically plausible value for the curvature of the utility function. The bound to the costs needed to reconcile the model with observed behavior turns out to be around 0.4 percent of non-durable consumption. Costs higher than this bound would offset the gains of investing optimally in stocks for a large fraction of the population. Our estimate is sufficiently low to make the participation cost-based explanation of incomplete portfolios a reasonable explanation, because it is likely that the true total costs of participation exceed this bound. Overall, our results suggest that the intertemporal consumption model provides a suitable description of household behavior once fixed costs of participation and unobservable heterogeneity in tastes are properly accounted for.

References

[1] Attanasio, O.P. , 2000, Consumpion, in Handbook of Macroecnomics.

[2] Attanasio, O., Banks, J. and S. Tanner, 2002, "Asset Holding and Consumption Volatility," Journal of Political Economy, 110, 771-792.

[3] Attanasio, O. and Low, H. (2003): "Estimating Euler Equations," Review of Economic Dynamics.

[4] Hansen, L. and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies", Journal of Political Economy, 99, 225-62.

[5] Hansen, L. and K. Singleton, 1982, "Generalized Instrumental Variable Estimation of Nonlinear Rational Expectation Models," Econometrica, 50, 1269-1286.

[6] Luttmer, E., 1995, "Asset Pricing in Economies with Frictions," Econometrica, 64, 1439-1467.

[7] Luttmer, E., 1999, "What Level of Fixed Costs Can Reconcile Asset Returns and Consumption Choices?," Journal of Political Economy, 107, 969-997.

[8] Mankiw, N. and S. Zeldes, 1991, "The Consumption of Stockholders and Non-Stockholders," Journal of Financial Economics, 17, 211-219.

[9] Manski, C., 2003, Partial Identification of Probability Distributions, New York: Springer-Verlag.

[10] Mehra, R. and E. Prescott, 1985, "The Equity Premium Puzzle," Journal of Monetary Economics, 15, 145-161.

[11] Merton, R., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case," Review of Economics and Statistics, 51, 247-257.

[12] Paiella, M., 2006, The Foregone Gains of Incomplete Portfolios, CSEF Working Paper No. 156.

[13] Paiella, M., 2004, "Heterogeneity in Financial market Participation: Appraising its Implications for the C-CAPM," Review of Finance, 8, 1-36.

[14] Pakes, A. Porter, J., Ho, K. and J. Ishii (2005): "Moment Inequalities and Their Application," Mimeo

[15] Samuleson, P., 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming", Review of Economics and Statistics, 51, 39-46.

[16] Vissing-Jørgensen, A., 2002, "Limited Asset Market Participation, Intertemporal Substitution and Risk Aversion," Journal of Political Economy, 110, 825-853.

[17] Vissing-Jørgensen, A., 2003, "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," NBER Macroeconomics Annual.

Appendix A: An Alternative Investment Strategy and the Approximation of the First-Order Condition for Non-Stockholders

Let's assume that after paying the fixed cost for investing in the risky asset, non-shareholders may adjust both their savings and their wealth allocation. Let $x_{h,t}^c(\alpha^c)$ denote the fraction of time t consumption they give up and invest in the risky asset and $x_{h,t}^w(\alpha^w)$ is the fraction of their wealth $W_{h,t}$, invested at the riskless rate, that they move into the risky asset. α^c and α^w denote the vectors of parameters. We also assume that they consume the return on the investment when they realize it. The *ex-post* gain of non-shareholders is then given by:

$$
\begin{aligned}
v_{h,t+1}\left(x_{h,t}^c(\alpha^c), x_{h,t}^w(\alpha^w), \delta\right) = \ & U\left(c_{h,t}(1 - x_{h,t}^c(\alpha^c) - \delta)\right) + \\
& \beta U\left((c_{h,t+1} + x_{h,t}^c(\alpha^c)c_{h,t}(1 + r_{t+1}) + \right. \\
& \left. \left. x_{h,t}^w(\alpha^w)W_{h,t}(r_{t+1} - r_{t+1}^f)\right)\right\} \\
& - \left\{U\left(c_{h,t}\right) + \beta U\left(c_{h,t+1}\right)\right\},
\end{aligned}
\tag{14}
$$

where we have suppressed the dependency of the utility function on $\xi_{h,t}$ for notational convenience. Under the assumption of isoelastic preferences and multiplicative heterogeneity, the first-order condition for the optimality of non-shareholders consumption can be written as:

$$
E_t\left\{\beta\zeta_{h,t+1} \right.
$$

$$
\left. \frac{c_{h,t+1}^{1-\gamma} - \left(\left(c_{h,t+1} + x_{h,t}^c(\alpha^c)c_{h,t}(1 + r_{t+1}) + x_{h,t}^w(\alpha^w)W_{h,t}(r_{t+1} - r_{t+1}^f)\right)\right)^{1-\gamma}}{\left(c_{h,t}(1 - x_{h,t}^c(\alpha^c) - d)\right)^{1-\gamma} - c_{h,t}^{1-\gamma}}\right\}
$$

$$
= 1.
\tag{15}
$$

Equation (15) can be approximated as follows. If we multiply and divide the numerator of the term on the left-hand-side of (15) by $c_{h,t+1}^{1-\gamma}$ and the denominator by $c_{h,t+1}^{1-\gamma}$, we can re-write the ratio as:

$$
\frac{c_{h,t+1}^{1-\gamma}}{c_{h,t}^{1-\gamma}}\left(\frac{1 - \left(1 + x_{h,t}^c(\alpha^c)\frac{c_{h,t}}{c_{h,t+1}}\left(1 + r_{t+1}\right) + x_{h,t}^w(\alpha^w)\frac{W_{h,t}}{c_{h,t+1}}\left(r_{t+1} - r_{t+1}^f\right)\right)^{1-\gamma}}{\left(1 - x_{h,t}^c(\alpha^c) - d\right)^{1-\gamma} - 1}\right).
\tag{16}
$$

Taking first-order Taylor expansions[10] around 1 of the polynomials raised to the $(1-\gamma)$ at the numerator and denominator, we can approximate (16) by:

$$\frac{c_{h,t+1}^{1-\gamma}}{c_{h,t}^{1-\gamma}} \left(\frac{(1-\gamma)\left(x_{h,t}^c(\alpha^c)\frac{c_{h,t}}{c_{h,t+1}}(1+r_{t+1}) + x_{h,t}^w(\alpha^w)\frac{W_{h,t}}{c_{h,t+1}}(r_{t+1}-r_{t+1}^f)\right)}{(1-\gamma)\left(x_{h,t}^c(\alpha^c)+d\right)} \right).$$

(17)

After simplifying and collecting terms, we can re-write the first-order condition for non-shareholders in (15) as follows:

$$E_t\left\{ \beta\zeta_{h,t+1}\left(\frac{c_{h,t+1}}{c_{h,t}}\right)^{-\gamma}\frac{(1+r_{t+1})+\frac{x_{h,t}^w(\alpha^w)W_{h,t}}{x_{h,t}^c(\alpha^c)c_{h,t}}(r_{t+1}-r_{t+1}^f)}{1+d/x_{h,t}^c(\alpha^c)} \right\} \approx 1. \quad (18)$$

$x_{h,t}^w(\alpha^w)=0$ corresponds to the case we focus upon, where non-stockholders' consumption must satisfy:

$$E_t\left\{ \beta\left(\frac{c_{h,t+1}}{c_{h,t}}\right)^{-\gamma}\frac{(1+r_{t+1})}{1+d/x_{h,t}^c(\alpha^c)} \right\} \approx 1. \quad (19)$$

Appendix B: The Variance-Covariance Matrix of the Errors

The error structure of the main equation we estimate is complicated by several factors. First, we deal with annual changes in consumption observed at a monthly frequency. In a time series context this would induce MA(12) residuals. Second, individuals observed over the same time periods or, given the time frame just mentioned, over adjacent months, will be affected by similar aggregate shocks. This implies correlation in the cross sectional dimension of the residuals.

While the instrumenting strategy we used takes into account this complex structure and guarantees that we obtain consistent estimates, in computing the standard errors, we need to take it into consideration explicitly. The residuals of equation (12) are expectational errors for an Euler equation and can be expressed as the sum of 12 monthly innovations:

---

[10]Second- and higher-order terms can be ignored because they are small for reasonable values of the parameters and of the variables.

$$\varepsilon_t^h = v_t^h + v_{t-1}^h... + v_{t-11}^h. \tag{20}$$

Each of the monthly innovations can be expressed as the sum of two errors, one representing aggregate shocks and one purely idiosyncratic ones. In other words, we express v_t^h as the sum of its cross-sectional mean and deviations from the same and assume that these deviations are independent across consumers:

$$v_t^h = \eta_t + u_t^h. \tag{21}$$

Let:

$$Var(\eta_t) = \sigma_\eta^2, \tag{22}$$

and

$$Var(u_t^h) = \sigma_{u,t}^2, \tag{23}$$

i.e. it is time-varying. Then:

$$Var(\varepsilon_t^h) = 12\,\sigma_\eta^2 + \sum_{j=0}^{11} \sigma_{u,t-j}^2, \tag{24}$$

$$Cov(\varepsilon_t^h \varepsilon_t^k) = 12\,\sigma_\eta^2, \tag{25}$$

and

$$\begin{aligned} Cov(\varepsilon_t^h \varepsilon_{t-j}^k) &= Cov(u_t^h u_{t-j}^k) \neq 0 \text{ if } 0 < |t-j| \leq 11 \\ &= 0 \text{ if } |t-j| \geq 11. \end{aligned} \tag{26}$$

Our estimate of the elements of the variance-covariance matrix is based on the internal product of the GMM residuals; hence it is heteroskedasticity robust.

Table 1 – Descriptive Statistics

	Shareholders	Non-shareholders	All
Age	48	47	47
Education: Less than high school	0.07	0.18	0.14
High school diploma	0.44	0.51	0.49
University degree	0.49	0.30	0.38
Gender (male=1)	0.74	0.66	0.69
Race: White	0.93	0.87	0.89
Marital status (married=1)	0.74	0.61	0.66
Household with children	0.43	0.41	0.42
Per-adult equivalent monthly expenditure	$646	$536	$580
	(353)	(306)	(329)
Consumption growth (Δlog c) in the cross-section	0.06	0.04	0.05
	(0.31)	(0.31)	(0.31)
Corr(consumption growth, risky return)	0.0998	0.0069	0.0744
Risky return: return of the S&P500CI	0.144		
	(0.143)		
Riskless return: return on 3-month T-bills	0.025		
	(0.018)		
Risk premium (BAA/AAA)	1.009		
	(0.004)		
Term premium (10yr gov. bonds/1yr gov. bonds)	1.011		
	(0.008)		
Price/Earnings ratio	20.019		
	(9.640)		
Price/Dividend ratio	41.308		
	(19.384)		
N. obs.	9,329	14,687	24,016

Note: Shareholders hold both risky (stocks and bonds) and riskless (checking and saving accounts) assets; non-shareholders hold only riskless assets. Standard errors in parentheses. Expenditure is on non-durable and services, it is deseasonalized and in dollars of 2002. The returns on the risky and riskless asset are real returns. The Price/Earnings and Price/Dividend ratios are taken from Shiller's homepage.

Table 2 – GMM estimates

LHS: $\log(1+r_{t+1})$	Coefficients
$\Delta\log(c_{h,t+1})$	1.722
	(1.258)
$\exp(d)$: $\log(x^c_{h,t}/(\exp(d)+x^c_{h,t}))$	-5.589
	(0.504)
Non-shareholder dummy	0.304
	(0.451)
Cohort: 1920-1929	-0.030
	(0.027)
Cohort: 1930-1939	-0.015
	(0.019)
Cohort: 1940-1949	-0.020
	(0.019)
Cohort: 1950-1959	-0.024
	(0.023)
Cohort: 1960+	-0.001
	(0.039)
Constant	-0.049
	(0.261)
Implied γ	~ 1.7
Implied d *(cost bound)*	**0.4%**
Nobs	24,016
Trading rule ($f(z)$)	
r^f_{t-2}	-0.390
	(0.008)
$(P/E)_{t-2}$	-0.102
	(0.048)
$r^{gb10}_{t-3}/ r^{gb1}_{t-3}$	0.444
	(0.008)
Average consumption share to be invested	*0.030*
	(0.010)

Note: The set of instruments include the risky and the risk free rates, the price/earning and the price/dividend ratios, the risk and the term premiums, a third-order polynomial in the age of the household head, cohort dummies, a time trend and a constant. The standard errors in parentheses are heteroskedasticity robust and allow for the clusters and the correlation over time.