Université de Montréal

Intertemporal utility models for asset pricing: reference levels and individual heterogeneity

par

Andrei Semenov

Département de sciences économiques
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
en sciences économiques

Août 2003

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée:

Intertemporal utility models for asset pricing: reference levels and individual heterogeneity

présentée par:

Andrei Semenov

a été évaluée par un jury composé des personnes suivantes:

Président-rapporteur:	Emanuela Cardia
Directeur de recherche:	René Garcia
Codirecteur:	Éric Renault
Membre du jury:	Rui Castro
Examinateur externe - Univ. McGill:	Kris Jacobs
Représentant du doyen de la FES:	Roch Roy

Résumé

La thèse propose de nouveaux modèles d'évaluation des actifs financiers fondés sur la consommation. Ces modèles, soit avec agent représentatif, soit avec consommateurs hétérogènes, permettent de mieux expliquer les primes de risque et le taux sans risque avec des valeurs raisonnables des paramètres de préférence. De plus, ces modèles emboîtent, comme cas particuliers, les modèles les plus connus dans la littérature, ce qui permet des tests de spécification informatifs.

Le premier article introduit la nouvelle fonction d'utilité avec niveau de référence dans un cadre par ailleurs standard d'agent représentatif. Le deuxième article suggère que la séparation de l'aversion pour le risque et la substitution intertemporelle peut être obtenue pas par le remplacement, comme le fait l'utilité récursive, de la consommation future par un équivalent certain de l'utilité future, mais par un niveau de référence exogène qui, d'une manière récursive, évalue la consommation future attendue. Dans le troisième article, un modèle avec agents hétérogènes permet de souligner l'importance de l'asymétrie de la distribution en coupe transversale des consommations individuelles dans la caractérisation des primes de risque. Le quatrième article évalue l'importance de l'hétérogénéité lorsque les agents ont une utilité avec niveau de référence et teste la fonction d'utilité isoélastique dans une économie avec agents hétérogènes.

Dans "A Consumption CAPM with a Reference Level", article conjoint avec René Garcia et Eric Renault, nous étudions un modèle d'utilité espérée dans lequel un agent dérive son utilité à la fois de l'excès relatif de sa consommation par rapport à un niveau de référence et de la valeur absolue de ce niveau de référence. Un des avantages de notre spécification est sa flexibilité. Nous montrons qu'elle peut reproduire la plupart des facteurs d'escompte stochastiques qui ont été proposés dans la littérature empirique sur l'évaluation des actifs financiers. Les tests empiriques du modèle avec la consommation agrégée conduisent à estimer des valeurs économiquement plausibles des paramètres de préférence, contrairement aux deux cas particuliers de la fonction de formation d'habitude et d'utilité non-espérée d'Epstein-Zin. Par ailleurs, nous confirmons l'importance d'inclure le niveau de référence absolu dans la fonction d'utilité.

Dans "Disentangling Risk Aversion and Intertemporal Substitution Through a Reference

Level", article également conjoint avec René Garcia et Eric Renault, nous montrons que si le taux de croissance du niveau de référence dépend du rendement du portefeuille de marché, les conditions de premier ordre pour la fonction d'utilité avec niveau de référence sont équivalentes sur le plan observationnel à ceux qui résultent de la fonction d'utilité récursive d'Epstein-Zin mais conduisent à une interprétation alternative des paramètres de préférence.

Dans "Asset Pricing Puzzles, High-Order Consumption Moments, and Heterogeneous Consumers", nous utilisons une expansion de Taylor de l'utilité marginale pour exprimer l'espérance conditionnelle du rendement de l'actif financier en fonction des moments croisés du rendement avec des moments de la distribution des consommations individuelles. Cette relation permet d'établir si chaque moment augmente ou diminue la prime de risque, sans spécifier une fonction d'utilité particulière. Avec la fonction isoélastique habituelle, nous montrons par calibration et estimation que l'asymétrie de la distribution en coupe transversale des consommations individuelles joue un rôle essentiel dans l'explication de la prime de risque. Nous obtenons également des valeurs économiquement plausibles des paramètres de préférence.

L'objectif de l'article "An Empirical Assessment of a Consumption CAPM with a Reference Level under Incomplete Consumption Insurance" est de tester empiriquement la fonction d'utilité espérée avec niveau de référence proposée dans le premier article sous l'hypothèse d'assurance de consommation incomplète et de participation limitée. Pour ce faire, nous utilisons le modèle d'évaluation des actifs financiers dérivé dans le deuxième article. Nous choisissons assez naturellement comme niveau de référence la consommation agrégée par tête. Lorsque l'asymétrie de la distribution des consommations individuelles et la participation limitée sont prises en compte, on obtient des valeurs économiquement plausibles de tous les paramètres d'intérêt. La fonction d'utilité isoélastique standard et la spécification usuelle de la formation d'habitude sont rejetées statistiquement.

Mots clés: agent représentatif, assurance de consommation incomplète, aversion relative au risque, elasticité de substitution intertemporelle, expansion de Taylor, niveau de référence, participation limitée, prime de risque sur les actions, taux sans risque.

Abstract

The dissertation proposes new consumption-based asset-pricing models. These models, either with a representative agent or with heterogeneous consumers, explain the equity risk premium and the risk-free rate with economically plausible values of the preference parameters. In addition, these models nest, as particular cases, the most well-known models in the literature, allowing for informative specification tests.

The first article introduces a new specification of preferences with a reference level in the representative-agent framework. The second article suggests that the disentangling risk aversion and intertemporal substitution may be obtained not by replacing, as the recursive utility does, the future consumption stream by a certainty equivalent of future utility but by an exogenous reference level which, in a recursive way, assesses the expected future consumption. In the third article, a model with heterogeneous consumers underlines the importance of asymmetry of the cross-sectional distribution of individual consumption in characterizing risk premia. The fourth article studies the importance of consumer heterogeneity when agents have a utility function with a reference level and tests the standard power utility model in the economy with heterogeneous consumers.

In "A Consumption CAPM with a Reference Level", a joint paper with René Garcia and Éric Renault, we study an expected utility model in which an agent derives utility both from consumption relative to an exogenous to the agent reference level and from the absolute value of this reference level. One of the advantages of our specification is its flexibility. We show that it can reproduce most of the stochastic discount factors that have been proposed in the empirical asset pricing literature. The empirical tests of the model with aggregate consumption per capita result in estimating economically plausible values of the parameters of interest, in contrast to the particular cases of the habit formation approach and the Epstein-Zin non-expected utility function. Finally, we confirm the importance of including the absolute value of the reference level of consumption in the utility function.

In "Disentangling Risk Aversion and Intertemporal Substitution Through a Reference Level", also a joint paper with René Garcia and Eric Renault, we show that if the reference level growth rate depends on market portfolio returns, the first-order conditions of the utility specification with a reference level are observationally equivalent to those resulting

from the Epstein-Zin non-expected utility function but lead to an alternative interpretation of the preference parameters.

In "Asset Pricing Puzzles, High-Order Consumption Moments, and Heterogeneous Consumers", we use a Taylor series expansion of the agent's marginal utility function to express the conditional expectation of the asset return in terms of the cross-sectional moments of return with the moments of the distribution of individual consumption. This relationship establishes whether each moment raises or lowers the risk premium without specifying a particular utility function. With the conventional power utility model, we show by calibration and estimation that asymmetry of the cross-sectional distribution of individual consumption plays a key role in explaining the risk premium. We also obtain economically plausible values of the behavioral parameters.

The objective of the paper "An Empirical Assessment of a Consumption CAPM with a Reference Level under Incomplete Consumption Insurance" is to empirically test the expected utility function with a reference level proposed in the first paper under the assumptions of incomplete consumption insurance and limited asset market participation. To convey this test, we use the asset-pricing model derived in the second paper. Assuming the reference level responds gradually to changes in aggregate consumption per capita, we show that when asymmetry of the cross-sectional distribution of individual consumption and limited participation are taken into account, we obtain economically plausible values of all the parameters of interest. The conventional power utility model is rejected statistically.

Keywords: elasticity of intertemporal substitution, equity risque premium, incomplete consumption insurance, limited asset market participation, reference level, relative risk aversion, representative agent, risk-free rate, Taylor series expansion.

Contents

Remerciements

J'adresse mes remerciements à mes directeurs de thèse: René Garcia et Éric Renault. La confiance qu'ils m'ont témoignée, le temps qu'ils m'ont consacré et les précieux conseils qu'ils m'ont donnés tout au long de mes travaux, ainsi que leur générosité et leurs qualités humaines extraordinaires, m'ont largement aidé à mener à bien cette thèse.

Je tiens à remercier le centre interuniversitaire de recherche en économie quantitative (CIREQ) et le centre interuniversitaire de recherche en analyse des organisations (CIRANO) pour leur support financier.

Ma gratitude, également, à tous ceux qui m'ont soutenu, et plus particulièrement ma femme Julie, mon fils Alexandre qui m'ont toujours encouragé, ainsi que mes parents Ninel et Veniamin qui m'ont appris des principes dont je me servirai toute ma vie.

1 A Consumption CAPM with a Reference Level

1.1 Introduction

The canonical consumption-based capital asset pricing model (CCAPM) where a representative agent maximizes his expected time-separable utility over uncertain streams of consumption is the workhorse of financial economists. It allows to understand intuitively the marginal utility trade-offs between different time periods or states of nature given some specification of the agent preferences. After all, consumption is what individuals ultimately care about and it should be reflected in their valuation of assets. However, when per capita consumption enters a power utility function, the model delivers gross inconsistencies with the observed asset returns, whether the empirical assessment is based on calibration or on formal estimation. This resilient empirical misfit has triggered, over the last two decades, a long series of attempts to modify the basic model in order to achieve empirical success.

A useful way to summarize the various directions in which preferences have been enriched is to consider that a state variable needs to be added to the basic model. This variable could be a benchmark level of consumption, as in the rich literature on habit formation. The main idea of the habit formation approach is that an investor derives utility not from the absolute level of consumption but from its level relative to a benchmark which is related to past consumption (Abel (1990, 1996), Campbell and Cochrane (1999), Constantinides (1990), Ferson and Constantinides (1991), Heaton (1995), and Sundaresan (1989)). When this reference level depends on past aggregate consumption, the *catching up* with the Joneses specification of Abel (1990), it captures the idea that the individual wants to maintain his relative status in the economy. The relative social standing is also present in the specification explored by Bakshi and Chen (1996) where absolute or relative wealth besides consumption determines utility.

Recently, Abel (1999) has generalized this specification by making the benchmark level of consumption a function of current as well as recent levels of consumption per capita. It also generalizes Gali's (1994) specification of consumption externalities whereby agents have preferences defined over their own consumption as well as current per capita consumption in the economy. They want to *keep up* with the Joneses. In all these specifications where the state variable is contemporaneous with consumption, it is noteworthy to emphasize that there is a separation between the attitudes towards risk and intertemporal substitution even though the agent maximizes expected time-separable utility. Indeed, this separation is generally associated with the non-expected utility framework of Epstein and Zin (1989) where the agent combines his current consumption with expected future utility in a recursive way.

In the model proposed in this paper, we maintain the representative agent paradigm as well as the time-separable expected utility framework. Our agent derives utility both from the level of consumption relative to a state variable which we call the reference level and from the absolute value of this reference level, that is $U(\frac{C_t}{S_t}, S_t)$ in ratio form or $U(C_t - S_t, S_t)$ in difference form. Such a consumer-investor will use assets to smooth not only fluctuations in the position of consumption with respect to a benchmark level of consumption but also movements in this benchmark level. At the most general level, this benchmark consumption provides a way to extend the intertemporal choice of consumption without uncertainty to risky consumption streams. When no uncertainty prevails, the future levels of reference or benchmark consumption, when seen at time t, coincide with the optimal future consumption values, that is $S_{t+h} = C_{t+h}$, identically for $h \geqslant 0$. In a risky environment, the coincidence prevails only in expectation and the reference level is interpreted as the benchmark consumption the agent has in mind when deciding his risk-taking behavior, $E_t[S_{t+h}] = E_t[C_{t+h}]$, for all $h \geqslant 0$. This formulation leaves open two possibilities. Either the reference level is an expectation of future consumption given past information or it is a function of contemporaneous information, as some macroeconomic variables which belong to the agent's information set at time $(t + h)$ may affect the assessment of the reference level S_{t+h}. We will see that these two modeling avenues relate to different asset pricing models present in the literature and have different implications for the disentangling of the elasticity of intertemporal substitution from the risk aversion coefficient.

The specification is similar to the general formulation in Abel (1990) and was recently used in a saving and growth model by Carroll, Overland, and Weil (2000). This allows to test formally if the absolute level of per capita consumption is important per itself for pricing assets, over and above consumption relative to the reference level as it is specified in the habit formation models of Campbell and Cochrane (1999) and Constantinides (1990). However, contrary to these habit formation specifications, we do not limit the reference level to depend solely on consumption, past or current. Indeed, it can be a function of other variables such as wealth (value of the market portfolio) and we then recover a specification similar to Bakshi and Chen (1996). The novelty of our approach is apparent when the reference level is made a function of both past consumption and the value of the market portfolio since we obtain a stochastic discount factor (SDF) which embeds the usual habit formation approach together with the so-called Kreps-Porteus specification of the recursive framework of Epstein and Zin (1989). When we estimate this new specification with aggregate per capita real consumption and returns on the value-weighted CRSP stock index or size portfolios, we obtain economically plausible values for the parameters, contrary to either the habit formation specification or the Epstein-Zin approach taken separately.

Saying that the future levels of S_t are equal in expectation to the future levels of consumption means that S_t represents the permanent component of consumption. Allowing S_t to depend on variables other than consumption is suggested by the results of Alvarez and Jermann (2002) who show that the size of the permanent component in consumption obtained from consumption data alone is much lower than the size of the permanent component of pricing kernels. Therefore, they recommend that in representative asset pricing models preferences should be such as to magnify the size of the permanent component in consumption. The modeling of our reference level will do just that. When the reference level is made a function of the value of the market portfolio, another permanent component is added to the pricing kernel. In the latter formulation, the consumption wealth ratio will enter the pricing kernel. Lettau and Ludvigson (2001) have emphasized the prominent role played by the log consumption-wealth ratio as a conditioning variable for improving the performance of unconditional specifications.

Another important feature of our approach is to add a moment condition to the set of asset pricing moment conditions. This additional condition relates the growth rate of log consumption to the variables deemed to characterize the growth rate of the log reference level. The estimation of this linear equation delivers an estimated value for the growth rate of the reference level which is used in the asset pricing equations. The estimation of the linear equation and the asset pricing Euler conditions can still be carried out jointly, imposing cross-equation restrictions which improves efficiency of the estimates of preference parameters. Recently, Neely, Roy, and Whiteman (2001) have addressed the issue of near nonidentification of the risk aversion parameter in the intertemporal consumption capital asset pricing model. They conclude that imposing natural identifying restrictions yields stable estimates of the parameters. Hansen and Singleton (1983) have also added information through an equation for predicting consumption growth.

We follow this approach to estimate several generalized versions of the habit formation models. In ratio specifications, habit may depend on one lag of consumption or respond gradually to changes in consumption. In both cases, we find mild support for the presence of the reference level per itself in the utility function. Moreover, the estimates of the time discount parameter are always greater than one. We also test the specification in difference proposed by Campbell and Cochrane (1999). Contrary to the ratio models, we find strong support for the hypothesis that the absolute value of the reference level enters the utility function.

If we assume the reference level growth rate to be a function of the return on the market portfolio, our model of expected utility yields a SDF which is isomorphic in its pricing implications to the Epstein and Zin (1989, 1991) pricing kernel. A striking feature of the comparison between the Epstein and Zin (1989, 1991) non-expected utility model and our expected utility model with a reference level is that the measures of risk aversion differ while the elasticity of intertemporal substitution remains the same in the two models. We explore in detail this difference in the interpretation of the risk aversion parameter in Garcia, Renault, and Semenov (2002). When we estimate this specification of our model which is observationally equivalent to the Epstein and Zin (1991) specification, we obtain a negative point estimate of the elasticity of intertemporal substitution but not significantly different from zero.

Finally, when we allow the reference level growth rate to be determined both by past consumption and by the return on the market portfolio, we obtain a SDF which incorporates habit formation in a Epstein-Zin-like SDF. With this specification, we obtain precise and reasonable estimates of the coefficient of relative risk aversion (RRA) (around 1), of the elasticity of intertemporal substitution (0.86), and of the time discount factor (0.9988).

Loss or disappointment aversion preferences have also received attention to salvage the consumption-based asset pricing model.[1] Recently, Barberis, Huang, and Santos (2001)

[1] See Benartzi and Thaler (1995) for loss aversion and Bekaert, Hodrick, and Marshall (1997), Bonomo

proposed a model of time-varying risk aversion where the investor is loss averse over financial wealth fluctuations. Our reference level model can accommodate specifications in which the investor expresses disappointment aversion whenever his consumption falls under the reference level. Our results show that for consumption above habit, the most plausible assumption is that the representative consumer derives utility from both consumption relative to habit and the absolute level of habit. As consumption declines towards the benchmark level, we cannot reject the assumption that the conventional time- and state-separable utility model well describes the agent's preferences. When we test a model of loss aversion similar to Barberis, Huang, and Santos (2001), we confirm the importance of the absolute value of the reference level in the utility function. Therefore, we are led to conclude that our utility specification not only opens new avenues for modeling the SDF but is robust to existing extensions of the standard CCAPM model.

Our approach can be put fruitfully in relation with the line of research which emphasizes stochastic prices of consumption risks and adds flexibility to the standard CCAPM through the risk aversion specification.[2] By allowing attitudes towards risk to reflect the information set used for consumption and savings choices, risk aversion is no longer fixed, but contingent upon the state of the world.[3] The same individual may be a risk-lover over certain states of the world and risk averse over others, adjusting his tolerance to risk according to the characteristics of the problem that he faces. Such shifts in attitudes could be related to numerous factors. Bakshi and Chen (1996) and St-Amour (1993), for example, allow for wealth-dependent attitudes towards risk, when the equilibrium relative risk aversion is a decreasing function of the individual's wealth, thus implying countercyclicality of risk aversion. In Campbell and Cochrane (1999), Constantinides (1990), and Sundaresan (1989), who introduce time-varying prices of risk through habit formation, relative risk aversion increases as consumption declines towards habit and, therefore, also displays a countercyclical

and Garcia (1993), and Epstein and Zin (2001) for disappointment aversion.

[2]One can include in this line of research Bakshi and Chen (1996), Campbell and Cochrane (1999), Chou, Engle, and Kane (1992), Constantinides (1990), Gordon and St-Amour (1998a, 1998b), Harvey (1991), Mark (1988), McCurdy and Morgan (1991), Melino and Yang (2003), St-Amour (1993), and Sundaresan (1989).

[3]In the standard power utility model, the SDF is just consumption growth raised to the power $-\gamma$ and, thus, one needs a large value of γ to get a volatile pricing kernel. The state dependent risk aversion implies that consumption shocks generate larger unanticipated fluctuations in marginal utility than under fixed preferences and, therefore, one can instead get a volatile SDF from a volatile relative risk aversion.

pattern.

Our model is also related to the nonlinear pricing kernel of Dittmar (2002). He approximates an unknown marginal utility function by a Taylor series expansion. The resulting pricing kernel is a polynomial function in aggregate wealth and human capital. He further shows that a cubic pricing kernel is necessary to describe the data. However, this pricing kernel cannot simultaneously deliver the nonlinearity necessary to price assets and monotonically decrease. Our specification may offer the functional form of the pricing kernel that is missing in Dittmar (2002), which ensures that the pricing kernel is both decreasing and potentially exhibiting a high degree of nonlinearity.

Our model may be given an alternative interpretation. The representative agent can be thought as a portfolio manager whose performance is evaluated in terms of a benchmark as it is the case in practice.[4] The idea of a reference level determining the utility of the investor is related to an older literature. Porter (1974), Fishburn (1977), and Holthausen (1981) present a risk-return model in which risk is associated with outcomes below some specified target level and return is associated with outcomes above the target. A decision maker may display various preference for outcomes above and below the target outcome. They show congruence between that model and a specific form of expected utility function. In investment contexts, decision makers are assumed to derive utility from fluctuations of outcomes relative to some target level of return on investment (Green (1963), Swalm (1966)). Halter and Dean (1971) estimated utility functions for changes in net wealth. Our model can be viewed as a particular extension to a dynamic setting of that risk-return approach, when the reference level is seen as a target.

The rest of the paper is organized as follows. In Section 1.2, we discuss the major features of the model in which a consumer derives utility from consumption relative to some reference consumption level as well as from this level itself. Section 1.3 examines the empirical implications of the proposed utility function under alternative specifications of the reference level generating process and assesses the contribution of the model towards explaining asset returns in US monthly data. Conclusions are presented in Section 1.4.

[4] Gali (1994) also alludes to such an interpretation in his model with consumption externalities.

1.2 An Expected Utility Model with a Reference Consumption Level

We generalize the standard time-separable utility function by assuming that each consumer derives utility from consumption relative to some reference consumption level as well as from this level itself:

$$u_t = \frac{\left(\frac{C_t}{S_t}\right)^{1-\gamma} S_t^{1-\varphi}}{(1-\gamma)(1-\varphi)}, \tag{1}$$

where γ is the curvature parameter for relative consumption, C_t is current consumption, S_t is a time-varying reference consumption level, and the parameter φ controls the curvature of utility over this benchmark level. If $\varphi = \gamma$, we get the standard time-separable power utility function (the reference consumption level plays no role in asset pricing). With $\varphi = 1$, we obtain a preference specification where the ratio of the agent's consumption to the reference consumption level is all that matters. If $\varphi \neq \gamma$ and $\varphi \neq 1$, then the agent takes account of both the ratio of his consumption to the subsistence level and this level itself when choosing how much to consume. Then, when maximizing expected utility over an infinite horizon, the agent assesses:

$$V_t = [(1-\gamma)(1-\varphi)]^{-1} \sum\nolimits_{h=o}^{\infty} \delta^h E_t \left[\left(\frac{C_{t+h}}{S_{t+h}}\right)^{1-\gamma} S_{t+h}^{1-\varphi} \right]. \tag{2}$$

We consider that the reference level S_t is external to the agent and E_t denotes a conditional expectation given the information at time t. At the most general level, this benchmark for consumption provides a way to extend intertemporal choice of consumption without uncertainty to risky consumption streams. When no uncertainty prevails, the future sequence of the reference level at time t, S_{t+h}, $h \geqslant 0$, coincides with the optimal future consumption values:

$$S_{t+h} = C_{t+h} \text{ identically for } h \geqslant 0. \tag{3}$$

In a risky environment, we just generalize condition (3) in terms of conditional expectations:

$$E_t[S_{t+h}] = E_t[C_{t+h}] \text{ for all } h \geqslant 0. \tag{4}$$

Therefore, we can interpret S_{t+h} as the reference level the agent has in mind at time t to decide his risk-taking behavior. In the latter case, some macroeconomic variables which belong to the agent's information set at time $(t+h)$ may affect the assessment of the reference level S_{t+h}. This specification is very general since the reference level remains unspecified. One can see however that several models in the asset pricing literature appear as particular cases of this general utility function. It is the case of the functional form considered by Abel (1990). This specification is obtained with $\gamma = \varphi$ and $S_t = [C_{t-1}^{D}\overline{C}_{t-1}^{1-D}]^{\upsilon}$, where $\overline{C}$ denotes aggregate per capita consumption. For $D = 0$, this is the relative consumption model called "catching up with the Joneses", whereas when $D = 1$, the individual considers as a benchmark his own past consumption. This is the habit formation model. More general models of habit can also be envisioned. Also, one can start with a utility specification in difference instead of ratio and postulate $u_t = \frac{(C_t - S_t)^{1-\gamma} S_t^{\gamma-\varphi} - 1}{1-\gamma}$ and cover the various specifications of the habit formation models in difference used by Campbell and Cochrane (1999) and Constantinides (1990) among others. All asset pricing models with habit formation have imposed the constraint $\varphi = 1$. Recently, Carroll, Overland, and Weil (2000) and Fuhrer (2000) have relaxed this constraint and explored the ensuing implications in a saving and growth model. They found that such a specification could explain that saving and growth were strongly positively correlated. It appears therefore worthwhile to allow φ to be different from 1 and to formally test the constraint $\varphi = 1$ in asset pricing models with habit.

The wealth-induced status model of Bakshi and Chen (1996) is another particular case of this specification. All models proposed in their paper can be formally expressed with a utility function conformable to (1), whether S_t is simply the wealth W_t of the individual or the wealth relative to the wealth of their reference group V_t. However, since our consumer cares about consumption relative to the reference level, our interpretation of this model will be different. Take the case where S_t is W_t. Our model tells us that the individual makes his consumption and portfolio decisions according to his consumption-wealth ratio, and not consumption per se, as well as his level of wealth. In this sense, we incorporate a variable in the model which is important in pricing assets as demonstrated by Lettau and Ludvigson (2001). As we will stress below, the SDF which results from such a specification of the reference level is closely related to the SDF of Epstein and Zin (1989).

Before we discuss in detail the various strategies for modeling the reference level of consumption in relation to the existing asset pricing models and possible extensions of these models, we need to establish how the presence of this level will change the risk premia and the intertemporal consumption trade-offs.

Since the reference level is considered external, the marginal utility of consumption is given by

$$\frac{\partial u_t}{\partial C_t} = C_t^{-\gamma} S_t^{\gamma-\varphi}. \tag{5}$$

Then, when maximizing his expected utility over an infinite horizon, the investor will choose an optimal consumption profile which will satisfy the following Euler equations:

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi} R_{i,t+1}\right] = 1,\ i = 1, ..., I, \tag{6}$$

where I is the number of assets considered and $R_{i,t+1}$ is the gross return of asset i from t to $t+1$. Expectations in (6) are taken conditionally on information available to the individual in period t and R_{it} is the gross return on asset i. The SDF is then

$$M_{t+1} = \delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}. \tag{7}$$

To discuss the implications of asset pricing models, it is common to assume joint conditional lognormality and homoscedasticity of the consumption growth rate and asset returns, since we obtain loglinear relations for asset returns. With our utility model, the risk-free rate will be determined by the following equation:

$$\begin{aligned} r_{f,t+1} &= -log\delta + \gamma E_t\left[\Delta c_{t+1}\right] - \frac{1}{2}\gamma^2\sigma_c^2 \\ &- (\gamma-\varphi) E_t\left[\Delta s_{t+1}\right] - \frac{1}{2}(\gamma-\varphi)^2\sigma_s^2 + \gamma(\gamma-\varphi)\sigma_{cs}, \end{aligned} \tag{8}$$

whereas the risk premium on any asset i will be given by:

$$E_t\left[r_{i,t+1} - r_{f,t+1}\right] = -\frac{1}{2}\sigma_i^2 + \gamma\sigma_{ic} - (\gamma-\varphi)\sigma_{is}, \tag{9}$$

where Δc_{t+1} is the log of the consumption growth rate, Δs_{t+1} is the log of the reference consumption level growth rate, $r_{i,t+1}$ is the log of the simple gross return on asset i, and σ_{xy} denotes generically the unconditional covariance of innovations.

The first three terms on the right-hand side of (8) and the first two terms on the right-hand side of (9) are the same as for a time-separable power utility function of consumption alone. Thus, utility function (1) can help to explain the risk-free rate puzzle if the term $-(\gamma-\varphi)E_t\left[\Delta s_{t+1}\right] - \frac{1}{2}(\gamma-\varphi)^2\sigma_s^2 + \gamma(\gamma-\varphi)\sigma_{cs}$ is negative and the equity premium puzzle if the term $-(\gamma-\varphi)\,\sigma_{is}$ is positive. Therefore, the position of γ with respect to φ and the signs of the covariances between the innovations in the reference level growth rate and the innovations in consumption growth and in asset returns are key in solving the two puzzles.

Another important dimension over which the resolution of the puzzles is discussed is the disentangling of risk aversion and intertemporal substitution. The standard consumption CAPM model with power utility imposes a functional restriction which is not sustainable theoretically nor supported empirically. For our model, we can study this separation by writing the expected return equation, always under the same joint conditional lognormality and homoscedasticity of the consumption growth rate and asset returns:

$$E_t\left[r_{i,t+1}\right] = -log\delta + \gamma E_t\left[\Delta c_{t+1}\right] - \frac{1}{2}\gamma^2\sigma_c^2 - (\gamma-\varphi)\,E_t\left[\Delta s_{t+1}\right]$$
$$-\frac{1}{2}(\gamma-\varphi)^2\sigma_s^2 + \gamma(\gamma-\varphi)\,\sigma_{cs} - \frac{1}{2}\sigma_i^2 + \gamma\sigma_{ic} - (\gamma-\varphi)\,\sigma_{is}. \tag{10}$$

To study intertemporal substitution in a simplified framework, let us assume that all quantities are now deterministic so we can ignore the expectation operators. With the standard power utility function under this assumption, equation (10) reduces to

$$r_{t+1} = -log\delta + \gamma\Delta c_{t+1} - \frac{1}{2}\gamma^2\sigma_c^2, \tag{11}$$

which implies $\sigma = \frac{1}{\gamma} = \frac{\partial\Delta c_{t+1}}{\partial r_{t+1}}$. From (10), it follows that when the agent's preferences are of the form (1), the intertemporal elasticity of consumption is

$$\sigma = \frac{\partial\Delta c_{t+1}}{\partial r_{t+1}} = \frac{1+(\gamma-\varphi)\frac{\partial\Delta s_{t+1}}{\partial r_{t+1}}}{\gamma}, \tag{12}$$

where $\frac{\partial\Delta s_{t+1}}{\partial r_{t+1}}$ can be interpreted as the elasticity of the reference level with respect to the interest rate.[5] The latter equation implies that the elasticity of intertemporal substitution differs from the inverse of the RRA coefficient if $(\gamma-\varphi)\frac{\partial\Delta s_{t+1}}{\partial r_{t+1}} \neq 0$. In our external reference level setting, it will therefore be important to distinguish the specifications where this level depends on past variables, in which case the disentangling will not occur, from

[5]As we can see from equations (9) and (12), since the terms $\frac{\partial\Delta s_{t+1}}{\partial r_{t+1}}$ and σ_{is} have the same sign, if utility model (1) generates an equity premium which is larger than that produced by the basic power utility model, it also generates an elasticity of intertemporal substitution which is less than the inverse of the RRA coefficient.

the ones where it depends on contemporary variables and allows to differentiate σ from the inverse of γ.[6] This justifies to analyze these specifications under two separate subsections under these headings.

Finally, the poor empirical performance of the standard consumption CAPM model has led researchers to explore specifications with a kink in the utility function, preferences changing above and below a certain threshold. Benartzi and Thaler (1995) examine single-period portfolio choice for a loss averse investor, which means that gains and losses will not receive the same weight in terms of utility. Bonomo and Garcia (1993) and Bekaert, Hodrick, and Marshall (1997) have explored the asset pricing implications of a related but different type of preference called disappointment aversion, introduced by Epstein and Zin (1989) in a recursive utility framework.[7] Recently, Barberis, Huang, and Santos (2001) proposed an asset pricing model where the investor is loss averse over financial wealth fluctuations. Given that our model introduces a reference level, it is natural to extend our investigation to specifications where preferences will be different above and below the reference level. When the latter will be related to past consumption variables, we will be able to test generalizations of the habit formation or catching up with the Joneses models. When it will be related to financial wealth, we will provide tests of models similar to the loss aversion model of Barberis, Huang, and Santos (2001). The discussion of these various specifications will be the object of the third subsection.

1.2.1 Reference Level Determined by Past Variables

In this subsection we will model the reference level strictly as a function of the past variables. This will allow us to make the link with the habit formation literature and discuss how this model has the potential to extend it. We will also discuss the persistence of the reference level and an estimation strategy.

[6] Ferson and Constantinides (1991) study an internal habit model, in which the utility is a power function of the difference between the current consumption flow and a fraction of a weighted sum of lagged consumption flows, and prove that habit persistence and/or durability of consumption drive a wedge between the elasticity of consumption with respect to investment returns and the inverse of the RRA coefficient.

[7] Preferences that exhibit disappointment aversion have been axiomatized by Gul (1991) to offer a solution to the so-called Allais paradox.

Modeling of the Reference Level. An approach commonly used in the literature consists in assuming that the reference consumption level, S_{t+1}, is an expectation of consumption C_{t+1} taken conditionally on past consumption levels, that is $S_{t+1} = E\left[C_{t+1}\left|C_t, C_{t-1}, ...\right.\right]$.[8] This is based on the idea that tomorrow's marginal utility of consumption is an increasing function of today's consumption. According to this approach, the time-varying subsistence level, or habit, can be specified either as an internal habit (habit depends on agent's own consumption) (Constantinides (1990), Sundaresan (1989)) or an external habit (the individual's reference consumption level depends on aggregate consumption, which is assumed to be unaffected by any one agent's consumption decisions, rather than on the history of individual's own consumption) (Abel, 1990, 1996, Campbell and Cochrane, 1999).

Let us suppose that $S_t = C_{t-1}^{\alpha}$, as in Abel (1990). As already mentioned, the ratio habit-formation model or the catching up with the Joneses are special cases of (1) when $\varphi = 1$. The utility function is in this case: $u_t = \frac{\left(\frac{C_t}{C_{t-1}^{\alpha}}\right)^{1-\gamma}}{1-\gamma}$, with $\alpha = 0$ giving the standard time-separable model and $\alpha = 1$ the catching up with the Joneses model. In the latter case, only relative consumption matters to the consumer.

Recently, Carroll, Overland, and Weil (2000) and Fuhrer (2000) have argued that one need not impose the constraint that α has to be 0 or 1. For values of α between 0 and 1, both the absolute and relative consumption levels are important to the consumer. The way we have rewritten the utility function lends itself to a different interpretation. A good way to start is to suppose that actual consumption never deviates from the reference level. In this case there is no consumption risk and the consumer needs just decide how to intertemporally substitute consumption over time. The exponent of the reference level is then quite naturally $\rho = 1 - \varphi$, with the elasticity of intertemporal substitution $\sigma = \frac{1}{1-\rho}$. Of course, there is consumption risk and the consumer reacts to it trough the curvature parameter γ which measures risk aversion. Therefore, this specification offers potentially a natural disentangling between risk aversion and intertemporal substitution. We will explore this aspect in the next subsection since this disentangling cannot occur when the reference level depends on past aggregate consumption $\left(\frac{\partial \Delta \overline{C}_t}{\partial r_{t+1}} = 0\right)$.

Since, according to the habit formation approach, the reference consumption level S_t is supposed to depend on past information only, we are allowed to assume that it should be

[8] This assumption is of course compatible with equation (4).

known to an agent at time t. Given this reasoning, we propose to use the following two-stage estimation procedure. In a first stage, we estimate the subsistence level under a particular assumption about the benchmark level formation process. In a second stage, we estimate the Euler equations (6) with the reference consumption level replaced with its estimate obtained in a first stage. Using this approach allows us to estimate a model exploiting the two specifications mentioned above whereby the stock of habit is assumed to be a function of lagged levels of consumption and the parameter which indexes the importance of the reference consumption level is added in the utility function. Given this two-stage estimation procedure, throughout the remainder of this section, we focus not only on the nature of the benchmark level generating process, but also on how this reference level can be estimated.

Persistence of the Reference Level and Estimation Strategy. The persistence of the reference level formation process is an important issue for consumption-based asset pricing models. Alvarez and Jermann (2002) derive a lower bound for the size of the permanent component of asset pricing kernels and find that it is very large. They also show that in the many instances where the pricing kernel is a function of consumption, innovations to consumption must have permanent effects.

As we have seen in the last section, a number of papers have assumed that habit depends on only one lag of consumption. An alternative view is that the subsistence level responds only gradually to changes in consumption (Campbell and Cochrane (1999), Carroll, Overland, and Weil (2000), Constantinides (1990), Fuhrer (2000), Heaton (1995), and Sundaresan (1989)). Carroll, Overland, and Weil (2000), Constantinides (1990), and Fuhrer (2000), for example, assume that the benchmark level evolves according to the adaptive expectations hypothesis, which postulates that the change in expectations, $S_{t+1} - S_t$, is equal to a proportion λ of last period's error in expectations, $C_t - S_t$. That is,

$$S_{t+1} - S_t = \lambda\left(C_t - S_t\right),\ 0 \leqslant \lambda \leqslant 1 \tag{13}$$

or, equivalently,

$$S_{t+1} = \lambda C_t + (1 - \lambda) S_t. \tag{14}$$

In this paper, we consider the unrestricted form of (14):

$$S_{t+1} = a + \lambda C_t + (1-\lambda)S_t, \tag{15}$$

whereas the adaptive expectations hypothesis postulates $a = 0$.

To replace the unobservable expected consumption, S_t, in equation (15) with an observable variable, we repeatedly lag and substitute equation (15) to obtain

$$S_{t+1} = \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i C_{t-i} \tag{16}$$

which means that the habit stock is a weighted average of past consumption flows with the weights $\lambda(1-\lambda)^i$ declining geometrically with time.

Since the subsistence level, S_{t+1}, is assumed to be an expectation of consumption taken conditionally on past consumption levels, we can rewrite (16) as

$$C_{t+1} = \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i C_{t-i} + \varepsilon_{t+1}, \tag{17}$$

where ε_{t+1} is an innovation in C_{t+1}.

Using the habit formation approach to modeling the reference consumption level allows us to determine whether habit persistence in preferences or durability in consumption expenditures is dominant. Habit persistence in preferences implies that today's consumption has a positive effect on tomorrow's marginal utility of consumption: $\frac{\partial^2 u_t}{\partial C_t \partial C_{t-1}} > 0$. For utility function (1), $\frac{\partial^2 u_t}{\partial C_t \partial C_{t-1}} = (\gamma - \varphi)\, C_t^{-\gamma} S_t^{\gamma-\varphi-1} \frac{\partial S_t}{\partial C_{t-1}}$. Hence, if $(\gamma - \varphi)\, \frac{\partial S_t}{\partial C_{t-1}} > 0$, habit persistence dominates durability. If $(\gamma - \varphi)\, \frac{\partial S_t}{\partial C_{t-1}} < 0$, then the effect of durability is dominant.

When Alvarez and Jermann (2002) measure the size of the permanent component of consumption using only consumption data, they find it is well lower than the size of the permanent component of pricing kernels. They suggest that in a representative agent asset pricing framework the specification of preferences should magnify the permanent component in consumption. The reference level in our utility function offers a way to introduce variables which, along with consumption, will contribute to amplify the permanent component of the asset pricing kernel. We will explore these possibilities in the next section where we will most notably look at the link between the reference level and the return on the market portfolio.

1.2.2 Reference Level Determined by Contemporaneous State Variables

A more general approach to modeling the subsistence level formation process is to assume that an agent can take into account not only the information available to him at time t, but also some information available at time $t+1$, when he forms his reference consumption level, S_{t+1}. Abel (1999) and Cochrane (2001), for example, suppose that the agent's benchmark level depends on current period aggregate consumption. Campbell and Cochrane (1999) also make their habit a contemporaneous variable.

According to (6), *the reference consumption level growth rate* is all we need to know about the reference level for asset pricing. We first motivate by an economic argument why the reference level growth rate should be made a function of the return of the market portfolio. We then present a general framework which allows other contemporaneous or past state variables to explain the reference level growth rate. We further show how to nest in this framework both the Epstein and Zin (1989, 1991) pricing kernel and the power utility model of Campbell and Cochrane (1999) with a slow-moving external habit.

Modeling the Growth Rate of the Reference Consumption Level. The SDF defined in (7) implies that the reference level must produce conditional expectations which not only are constrained by (4) but also are consistent with asset prices. Let us consider first the market portfolio pricing condition. If we denote by $R_{M,t+1}$ the gross return on the market portfolio observed at time $(t+1)$, we get

$$E_t \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{S_{t+1}}{S_t} \right)^{\gamma - \varphi} R_{M,t+1} \right] = 1. \tag{18}$$

Condition (18) shows that covariation between the reference level and the market return may compensate for the lack of covariation between consumption and the market return. This extension of the traditional consumption-based asset pricing model may help to solve several asset pricing puzzles features associated with aggregate data. As stressed by Barberis, Huang, and Santos (2001), such an extension has some behavioral foundations since it captures the idea that the degree of loss aversion of the investor depends on his prior investment performance. To make even more explicit this tight relationship between the reference level and investment performance as measured by the market return, we will refer to a loglinearization of conditional moment restrictions (4) and (18) (see Epstein and

Zin (1991) for similar interpretations based on a loglinearization of the Euler equations). Conditional expectations are computed as if the vector

$$(\Delta c_{t+1}, \Delta s_{t+1}, r_{M,t+1}) = \left(log\left(\frac{C_{t+1}}{C_t}\right), log\left(\frac{S_{t+1}}{S_t}\right), logR_{M,t+1}\right) \tag{19}$$

were jointly normal and homoscedastic given the information available at time t. Conditions (4) and (18) at horizon 1 become:

$$E_t[\Delta c_{t+1}] - E_t[\Delta s_{t+1}] = \kappa_1, \tag{20}$$

$$-\gamma E_t[\Delta c_{t+1}] + (\gamma - \varphi)E_t[\Delta s_{t+1}] + E_t[r_{M,t+1}] = \kappa_2 \tag{21}$$

for some constants $\kappa_1$and κ_2. Equivalently, these two restrictions say that both $[\Delta s_{t+1} - \Delta c_{t+1}]$ and $[\Delta s_{t+1} - \frac{1}{\varphi} r_{M,t+1}]$ must be unpredictable at time t. The Epstein and Zin (1989) pricing model is in fact observationally equivalent to the particular case of our CCAPM with reference level where $[\Delta s_{t+1} - \frac{1}{\varphi} r_{M,t+1}]$ is not only unpredictable but constant:

$$\Delta s_{t+1} = \frac{1}{\varphi} r_{M,t+1} + \kappa, \tag{22}$$

for some constant κ. In other words, we consider the particular case where the benchmark growth rate of consumption is loglinearly determined by the current value of the market return, with a slope parameter equal to the elasticity of intertemporal substitution. Note that this is in accordance with the portfolio separation property generally implied by homotheticity of preferences (see Epstein and Zin, 1989), whereby optimal consumption is determined in a second stage, after the portfolio choice has been made.

An interesting generalization is to relate the log of reference level growth to past period consumption growth, as we did in the previous section for habit formation models, and the current period return on the market portfolio in the following way:

$$\Delta s_{t+1} = a_0 + \sum_{i=1}^{n} a_i \times \Delta c_{t+1-i} + b \times r_{M,t+1}.^{9} \tag{23}$$

[9] This assumption also relates our framework to the prospect theory of Kahneman and Tversky (1979) and Tversky and Kahneman (1992). The intuition behind this is that if the level of market portfolio moves up, an agent should think that this increase in his wealth will bring him the additional consumption. It means that the benchmark consumption level, which reflects anticipated consumption, should also move up.

Condition (20) is consistent with a model where consumption growth is equal to the reference level growth rate plus a constant and noise:

$$\Delta c_{t+1} = \kappa_1 + \Delta s_{t+1} + \varepsilon_{t+1}, \tag{24}$$

where ε_{t+1} is an innovation in Δc_{t+1} with $E_t(\varepsilon_{t+1}) = 0$ and $E_t[\Delta s_{t+1}\varepsilon_{t+1}] = 0$. It follows that the log of consumption growth may be described by an affine regression

$$\Delta c_{t+1} = a_0 + \kappa_1 + \sum_{i=1}^{n} a_i \times \Delta c_{t+1-i} + b \times r_{M,t+1} + \varepsilon_{t+1} \tag{25}$$

with $E_t[r_{M,t+1}\varepsilon_{t+1}] = 0$. From (23),

$$\frac{S_{t+1}}{S_t} = A \prod_{i=1}^{n} \left(\frac{C_{t+1-i}}{C_{t-i}}\right)^{a_i} (R_{M,t+1})^b, \tag{26}$$

where $A \equiv exp(a_0)$.

Under the above assumptions, the SDF (7) becomes

$$M_{t+1} = \delta^* \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \prod_{i=1}^{n} \left(\frac{C_{t+1-i}}{C_{t-i}}\right)^{a_i(\gamma-\varphi)} (R_{M,t+1})^{b(\gamma-\varphi)}, \tag{27}$$

where $\delta^* \equiv \delta A^{\gamma-\varphi}$. This specification allows to separate risk aversion from intertemporal substitution, since $\sigma = \frac{1+b(\gamma-\varphi)}{\gamma}$. Therefore, we may rewrite (27) as

$$M_{t+1} = \delta^* \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \prod_{i=1}^{n} \left(\frac{C_{t+1-i}}{C_{t-i}}\right)^{a_i(\gamma-\varphi)} (R_{M,t+1})^{\kappa}, \tag{28}$$

where $\kappa \equiv \sigma\gamma - 1$, so that testing the null hypothesis $H_0 : \kappa = 0$ is equivalent to testing $H_0 : \sigma = \frac{1}{\gamma}$.

This specification of the SDF is interesting for several reasons. First, when $a_i = 0$ $(i = 1, ..., n)$, the SDF in (28) is isomorphic in its pricing implications to the Epstein and Zin (1989, 1991) pricing kernel for a Kreps and Porteus (1978) certainty equivalent. When $b = 0$, the reference level growth depends only on previous period consumption growth, as in the habit formation approach. When neither of these restrictions holds, we have a new asset pricing model which will put together two strands of the literature which evolved in parallel until now.[10] This new framework offers a way to test existing models since they

[10] Recently, Schroder and Skiadas (2002) have shown an isomorphism between competitive equilibrium models with utilities incorporating lineat habit formation and corresponding models without habit formation. In particular, they have offered a solution to problems with utility that combines recursivity with habit formation.

are embedded in the general specification. Let us look in more detail at the comparison between the Epstein and Zin SDF obtained under a non-expected recursive utility model and the SDF under expected utility with a reference level.

Comparison with the Epstein-Zin Stochastic Discount Factor. Under the assumption that $a_i = 0$ $(i = 1, ..., n)$, the SDF in (28) reduces to

$$M_{t+1} = \delta^* \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} (R_{M,t+1})^{\kappa}. \tag{29}$$

When $\gamma = 1/\sigma$ (i.e. $\kappa = 0$), we get the SDF for a standard power utility model. When $\gamma = 0$, the consumption growth rate is irrelevant to the determination of equilibrium asset prices and the market return is sufficient for discounting asset payoffs. In any other case, both the consumption growth rate and the market return are relevant to the determination of equilibrium asset prices.

The Epstein-Zin (1989, 1991) SDF is

$$M_{t+1} = \mu^{\frac{1-\alpha}{\rho}} \left(\frac{C_{t+1}}{C_t}\right)^{\frac{1-\alpha}{\rho}(\rho-1)} (R_{M,t+1})^{\frac{1-\alpha}{\rho}-1}, \tag{30}$$

where ρ is the parameter reflecting intertemporal substitutability (the elasticity of intertemporal substitution is $\psi = 1/(1-\rho)$) and α is the risk aversion parameter. Epstein and Zin (1989) interpret α as a measure of risk aversion for comparative purposes with the degree of risk aversion increasing in α.

The observational equivalence between the SDFs (29) and (30) implies that $\delta^* \equiv \mu^{\frac{1-\alpha}{\rho}}$, $-\gamma \equiv \frac{1-\alpha}{\rho}(\rho - 1)$, and $\sigma\gamma - 1 \equiv \frac{1-\alpha}{\rho} - 1$. The two last identities put together yield $\sigma = \frac{1-\alpha}{\rho\gamma} = 1/(1-\rho) = \psi$, that is the elasticity of intertemporal substitution in model (1) is equivalent to that in the Epstein-Zin non-expected recursive utility specification. In the case of the Epstein-Zin utility function, the elasticity of intertemporal substitution may not be equal to 1, whereas in the case of utility specification (1) any value of σ is allowed.

Since $\sigma = 1/(1-\rho)$, $\gamma = (1-\alpha)/(\sigma - 1)$. It follows that the measure of risk aversion in the Epstein-Zin (1989, 1991) utility specification, α, is equal to $1 - \gamma\sigma + \gamma$. It is easy to see that α is equal to the RRA coefficient, γ, only if $\gamma = 1/\sigma$, what corresponds to the case of the standard power utility model.[11] If γ differs from $1/\sigma$, the parameter α is no longer

[11] In the Epstein-Zin (1989, 1991) preference specification, the parameter α is the RRA coefficient when $1 - \alpha = \rho$ (in this case, we get the SDF for the conventional power utility model, for which $\gamma = 1/\sigma$).

the RRA coefficient and is equal to the RRA coefficient plus the term $1 - \gamma\sigma$.

In our model, risk aversion is defined with respect to the unpredictable discrepancy between actual consumption and the reference level (a quantity independent of the attitude towards risk) and not with respect to the forthcoming level of recursive utility which still mixes attitudes towards risk and intertemporal substitution. Garcia, Renault, and Semenov (2002) develops further the comparison between the two models.

The SDF in (29) yields the following Euler equations:

$$E_t \left[\delta^* \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (R_{M,t+1})^{\kappa} R_{i,t+1} \right] = 1, \; i = 1, ..., I.^{12} \tag{31}$$

A test of the null hypothesis $H_0 : \sigma = 0$ can be carried out by testing the null hypothesis $H_0 : \kappa = -1$ and $\gamma \neq 0$. To examine whether $\sigma = 1/\gamma$, we have to test the null hypothesis $H_0 : \kappa = 0$.

We may rewrite the SDF in (29) as

$$\begin{aligned} M_{t+1} &= \delta^* \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (R_{M,t+1})^{b(\gamma - \varphi)} \\ &= \left(\delta^{*1/\theta} \left(\frac{C_{t+1}}{C_t} \right)^{-\varphi} \right)^{\theta} \left((R_{M,t+1})^{-b\varphi} \right)^{1-\theta}. \end{aligned} \tag{32}$$

Approximating this geometric average with an arithmetic average yields

$$M_{t+1} = \theta \left(\delta^{*1/\theta} \left(\frac{C_{t+1}}{C_t} \right)^{-\varphi} \right) + (1 - \theta) (R_{M,t+1})^{-b\varphi}. \tag{33}$$

After substituting this linear approximation into the Euler equations (31), we obtain

$$1 \approx \theta E_t \left[\delta^{*1/\theta} \left(\frac{C_{t+1}}{C_t} \right)^{-\varphi} (R_{i,t+1}) \right] + (1 - \theta) E_t \left[(R_{M,t+1})^{-b\varphi} (R_{i,t+1}) \right]. \tag{34}$$

It can be viewed from (34) that, as in the Epstein-Zin utility function case, the riskiness of an asset is measured by means of the covariance of its return with the market portfolio return (as in the static CAPM) and the covariance of its return with the consumption growth rate (as in the intertemporal CAPM).

Another usual to illustrate this interpretation is to assume joint lognormality and homoscedasticity of the consumption growth rate and asset returns. Under this assumption,

[12] Since $\kappa = \sigma\gamma - 1$, one may want to estimate σ directly. However, we prefer estimating κ instead of σ because the latter will be unidentified whenever γ is near 0.

we have

$$E_t\left[r_{i,t+1}-r_{f,t+1}\right]=-\frac{1}{2}\sigma_i^2+\gamma\sigma_{ic}-(\gamma-\varphi)\,b\sigma_{iM}=-\frac{1}{2}\sigma_i^2+\varphi\left(\theta\sigma_{ic}+(1-\theta)\,b\sigma_{iM}\right). \quad (35)$$

So, the parameter φ can be thought of as a coefficient measuring the contribution of a weighted combination of asset i's covariance with consumption growth and asset i's covariance with the market return towards the risk premium on asset i.

Alvarez and Jermann (2002) refer to the recursive preferences of Epstein and Zin (1989) and Weil (1989) as a way to increase the size of the permanent component in the pricing kernel. Our utility specification, through the assumed connection between the reference level and the value of the market portfolio, adds similarly a permanent component to the pricing kernel.

Habit Formation in Difference with a Reference Level. Campbell and Cochrane (1999) interpretation of the reference level S_t as an external habit leads them to specify some nonlinear dynamics consistent with the structural restriction $S_t \leq C_t$. Actually they specify the surplus consumption $H_t = \frac{C_t - S_t}{C_t}$ as a conditionally lognormal process. In this setting, we can still introduce our reference level principle by considering that the representative consumer derives utility from consumption relative to his reference level as well as from this level itself

$$u_t = \frac{C_t^{1-\gamma} S_t^{\gamma-\varphi}}{(1-\gamma)(1-\varphi)}, \quad (36)$$

according to (1). However, in order to really nest Campbell and Cochrane (1999) utility model, we must extend this formulation by writing

$$u_t = \frac{(C_t H_t)^{1-\gamma}\, S_t^{\gamma-\varphi}}{(1-\gamma)(1-\varphi)}. \quad (37)$$

Then, testing $H_0 : \gamma = \varphi$ amounts to testing the particular case of Campbell and Cochrane (1999) utility model. At first sight, it may appear a bit artificial to introduce three variables in the definition of the utility functions since any of them is a well-defined function of the two other ones. However, the utility function rewritten in that way (37) helps to better understand the external habit paradigm of Campbell and Cochrane (1999). The statistical model (see (40) below) specifies the joint dynamics of the two lognormal processes (C_t, H_t) while the dynamics of the reference level S_t is only a by-product. However, in the economic model, the optimizing agent considers the product $(C_t H_t)$ as its optimal control variable given the external habit level S_t. Therefore, the resulting SDF is:

$$M_{t+1} = \delta \left(\frac{H_{t+1}}{H_t} \right)^{-\gamma} \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{S_{t+1}}{S_t} \right)^{\gamma - \varphi}. \tag{38}$$

Another way to understand this formula is to realize that the utility function in (37) can also be written :

$$u_t = \frac{(C_t - S_t)^{1-\gamma} S_t^{\gamma - \varphi}}{(1-\gamma)(1-\varphi)}, \tag{39}$$

and the consumer see the reference level S_t as external.

The statistical model proposed by Campbell and Cochrane (1999) is specified in order to make the volatility of the SDF stochastically time-varying with the business cycle pattern. The consumption process is seen as a lognormal random walk while the $\log H_t$ process has the same standardized innovation but evolves as a heteroscedastic AR(1):

$$\Delta c_{t+1} = g + \nu_{t+1}, \ \ \nu_{t+1} \text{ i.i.d. } N(0, \sigma^2), \tag{40}$$

$$h_{t+1} = (1-\phi)\overline{h} + \phi h_t + \lambda(h_t)\nu_{t+1}.$$

We are going to use the sensitivity function $\lambda(h_t)$ proposed by Campbell and Cochrane (1999), that is:

$$\lambda(h_t) = \begin{cases} \frac{1}{\overline{H}}\sqrt{1 - 2(h_t - \overline{h})} - 1 & \text{if } h_t \leq h_{\max} \\ 0 & \text{otherwise,} \end{cases} \tag{41}$$

where $h_{\max} \equiv \overline{h} + \frac{1}{2}(1 - \overline{H}^2)$ and $\overline{H} = \sigma\sqrt{\frac{\gamma}{1-\phi}}$. By choosing this sensitivity function, Campbell and Cochrane (1999) had two objectives in mind. The first was to obtain a constant risk-free rate. This restriction is typically relaxed in our more general setting with $\gamma \neq \varphi$, where the absolute value of the reference level plays an independent role in the utility function. The second one was to ensure that the elasticity of the reference level with respect to consumption is zero in the steady state and is a U-shaped function of h around $h = \overline{h}$. This objective is still achieved in our setting.

Besides allowing for a time-varying risk-free interest rate, our setting with $\gamma \neq \varphi$ disentangles the relative risk aversion coefficient and the elasticity of intertemporal substitution, as already emphasized in the Epstein-Zin-like interpretation of our model in the previous section. This disentangling appears important since in the Campbell and Cochrane (1999) model risk aversion ($\frac{\gamma}{H_t}$) can become very large in the states of the economy where H_t approaches zero, that is when consumption comes very close to the external habit. In our setting, a large risk aversion does not automatically imply a dramatically low level for the elasticity of intertemporal substitution. However, as in Campbell and Cochrane (1999), we make the steady state of the reference level (and in turn the sensitivity function) depend on the preferences only through the risk aversion parameter γ. The issue of a possible additional role of the parameter φ in the steady state of the reference level and the sensitivity function is left for future research.

Campbell and Cochrane calibrated the parameters in order to assess the model implications for asset pricing. Suppose we wanted to estimate this model and test their specification against the more general SDF (38). We should be able to compute the time series of surplus consumption H and to deduce the time series of the growth rate of the reference level S. For the former, given γ and a process for consumption growth, we need the parameter ϕ. In choosing parameters, Campbell and Cochrane match ϕ to the serial correlation of the log price-dividend ratio. But this measure of persistence is tightly related to the persistence of the market portfolio return since one can always write

$$logR_{M,t+1} = \Delta c_{t+1} + log(1 + Q_{t+1}) - logQ_t, \tag{42}$$

where $R_{M,t+1} = \frac{P_{t+1}+C_{t+1}}{P_t}$ and $Q_{t+1} = \frac{P_{t+1}}{C_{t+1}}$ denotes the price-dividend ratio for a claim on aggregate consumption. Therefore, approximately

$$logR_{M,t+1} \cong \Delta c_{t+1} + log(Q_{t+1}) - logQ_t. \tag{43}$$

When Δc_{t+1} is viewed as a white noise (as in Campbell and Cochrane (1999)), the dynamics of the rate of growth of the price dividend ratio is tightly related to the one of the market return. In other words, plugging into the SDF (38) a rate of growth of the reference level that mimics the price dividend ratio dynamics is very similar in spirit to the Epstein and Zin SDF, as revisited in the previous subsection. In this sense, we can claim that our proposed extension of Campbell and Cochrane (1999) nests the Epstein and Zin (1989) asset pricing model expressed with a habit formation model in difference. Therefore, as with the ratio model in Section 1.2.2, we maintain habit formation preferences while disentangling risk aversion from intertemporal substitution in a way observationally equivalent to Epstein and Zin (1989).

1.2.3 Preferences with a Reference Level as a Threshold

Introducing a kink in the utility function has been another way to attempt rescuing the consumption CAPM. Disappointment aversion and loss aversion are two examples of such preferences, the former being defined over intertemporal consumption streams, the latter over wealth. A disappointment averse consumer will put more weight on bad outcomes than on good ones, where bad and good are defined with reference to a certainty equivalent measure of a consumption gamble. Epstein and Zin (1989) integrate these generalized preferences in an intertemporal asset pricing model within a recursive utility framework. Bekaert, Hodrick, and Marshall (1997), Bonomo and Garcia (1993), and Epstein and Zin (2001) explore the asset pricing implications of disappointment aversion. Benartzi and Thaler (1995) also adopt asymmetric preferences over good and bad results, but instead of using an intertemporal asset pricing framework with preferences defined over consumption streams, they start from preferences defined over one-period returns based on Kahneman and Tversky (1979)'s prospect theory of choice. The central idea of prospect theory is that an investor is assumed to derive utility from fluctuations in the value of his financial wealth and to be loss averse over these fluctuations, meaning that he is distinctly more sensitive to reductions in his financial wealth than to increases. Recently, Barberis, Huang, and Santos (2001) have studied asset prices in the context of prospect theory. In their model, investors derive utility both from consumption and changes in the value of their financial wealth. They introduce loss aversion over financial wealth fluctuations and allow the degree of loss aversion to be affected by prior investment performance.

The utility function defined in (1) offers at least two ways to model asymmetric preferences. First, we can use the reference consumption level S_t as a threshold below which outcomes are penalized in terms of utility. In this generalization of habit formation models, investors will have the following utility function:

$$u_t = \begin{cases} \frac{\left(\frac{C_t}{S_t}\right)^{1-\gamma_1} S_t^{1-\varphi_1} - 1}{1-\gamma_1} & \text{if } C_t \geqslant S_t \\ \lambda \frac{\left(\frac{C_t}{S_t}\right)^{1-\gamma_2} S_t^{1-\varphi_2} - 1}{1-\gamma_2} & \text{otherwise,} \end{cases} \tag{44}$$

where λ is a disappointment aversion coefficient. The intuition here is that the investor is likely to be disappointed if his consumption is lower than the reference level and, conversely, satisfied otherwise.

A second type of asymmetry could be built by modeling the threshold level in a fashion similar to Barberis, Huang, and Santos (2001). The reference level S_t in this case could be assimilated to the value of the market portfolio, while the threshold will be given by the position of the return on the market portfolio with respect to the safe asset return. Such a utility specification yields the following SDF:

$$M_{t+1} = \delta \frac{C_{t+1}^{-\gamma_1} V_{t+1}^{\phi_1} I_{[R_{M,t+1} \geqslant z_t R_{f,t+1}]} + \lambda(z_t) C_{t+1}^{-\gamma_2} V_{t+1}^{\phi_2} \left(1 - I_{[R_{M,t+1} \geqslant z_t R_{f,t+1}]}\right)}{C_t^{-\gamma_1} V_t^{\phi_1} I_{[R_{M,t} \geqslant z_t R_{f,t}]} + \lambda C_t^{-\gamma_2} V_t^{\phi_2} \left(1 - I_{[R_{M,t} \geqslant z_t R_{f,t}]}\right)}, \tag{45}$$

where $R_{M,t+1}$ is the return on the market portfolio, $R_{f,t+1}$ is the risk-free interest rate, V_{t+1} is the value of market portfolio, and $I_{[R_{M,t+1} \geqslant z_t R_{f,t+1}]}$ is the indicator function, which takes the value 1 if $R_{M,t+1} \geqslant z_t R_{f,t+1}$ and 0 otherwise. The variable z_t measures the size of prior losses. The larger the prior losses, the more painful subsequent losses will be. Our goal will be to estimate such models by the generalized method of moments to assess the presence of such preferences in the data, as opposed to the calibration exercise of Barberis, Huang, and Santos (2001).

1.3 Empirical Results for Alternative Models of the Reference Level

In this section, we estimate the models described in Section 1.2 using US monthly data. After a brief description of the data construction, we discuss the estimation procedure in light of the identification issues surrounding the preference parameters. We then start by providing the empirical results corresponding to the conventional time- and state-separable preferences. This will provide a basis with which to compare the richer specifications offered by the three types of models with a reference level discussed in Section 1.2. First, we estimate models for the augmented habit formation approach, where the absolute value of the reference level enters in the utility function along with relative consumption. Second, we consider several specifications where the market portfolio return enters in the determination of the reference level growth rate. We estimate models which embed two well-known models: the Campbell and Cochrane (1999) habit formation model and the Epstein and Zin (1989, 1991) non-expected recursive utility model. For the former case, instead of calibrating the model, we estimate and test the model in a GMM framework keeping the original specification of the consumption surplus dynamics. For the latter, we estimate a utility specification where the reference consumption level growth rate is assumed to depend on the previous period consumption growth rates and the return on the market portfolio. As the Epstein-Zin model is seen to be a mixture of the CAPM and the CCAPM models, this generalized specification can be described as a combination of the CAPM with a habit formation CCAPM model. Finally, we estimate models of asymmetric preferences, where the investor draws different utilities above and below the reference level or some function of it. Most notably, we estimate a model of loss aversion very similar in spirit to the model of Barberis, Huang, and Santos (2001), thereby providing a useful complement to their calibration whereas their model was simply calibrated.

1.3.1 Data and Estimation Issues

Consumption and Returns Data. The measure of real aggregate consumption used in this paper is the personal consumption expenditures (in constant 1987 dollars) on nondurables and services (NDS) taken from the United States National Income and Product Accounts.[13] Monthly per capita consumption is obtained by dividing the real aggregate consumption by the total population, including armed forces overseas.[14]

The nominal, monthly risk-free rate of interest is the one-month Treasury bill return from the Center for Research in Security Prices (CRSP) of the University of Chicago. The real risk-free rate is calculated as the nominal risk-free rate, divided by the one-month inflation rate, based on the deflator defined for NDS consumption. As a proxy for the nominal, monthly market return, we take the value-weighted aggregate nominal, monthly return (capital gain plus dividends) on all stocks listed on the NYSE and AMEX, obtained from CRSP. The real, monthly market return is calculated as the nominal market return,

[13] Taken from CITIBASE, mnemonics GMCNQ and GMCSQ.

[14] CITIBASE, mnemonic POP.

divided by the one-month inflation rate.

Estimation and Identification Issues. An iterated GMM approach is used to test Euler equations and estimate model parameters. For each preference specification the Euler equations for the excess market return and the real risk-free interest rate are estimated jointly exploiting two sets of instruments. The first instrument set (INS1) has a constant and the real market return, the real risk-free rate, and the real consumption growth rate lagged once. As our second set (INS2), we use the first set of instruments plus the same variables lagged an additional period. The sampling period is 1959:1 to 1996:12. After allowing for the construction of the lagged variables, the sample period used in the estimation is 1960:3 to 1996:12, for a total of 442 observations.

Under the general assumptions associated with the GMM method, the estimators are asymptotically normal and the main test statistic, the J statistic for overidentifying restrictions, has a chi-square distribution. It is by now well established[15] that asymptotic normality is often a poor approximation for the sampling distributions of the GMM estimators. These distributions can be skewed and have fat tails. Tests of overidentifying restrictions can exhibit important size distortions. Stock and Wright (2000) argue that the problem in GMM with instrumental variables might come from the weak correlation between the instruments and the first order conditions which results in a poor identification of the parameters. In the CCAPM it is the risk aversion parameter which is nearly unidentified. To address these problems, they propose an asymptotic theory for GMM with weak identification. For many of the models we will estimate (CRRA, habit formation and Epstein-Zin preferences), they have derived expressions allowing the computation of new confidence sets for the parameters and the test statistics. Moreover they have shown that the confidence sets based on their theory differ considerably from the conventional confidence sets. This is not a small sample problem in the traditional acceptance of the term. A Monte Carlo study shows that with monthly data, a sample of about 1000 observations is needed for the conventional normal asymptotics to provide a good approximation to the finite sample distributions of the estimators. Given our sample size we will explore with our new specifications.

[15]See Tauchen (1986), Kocherlakota (1990), and Hansen, Heaton, and Yaron (1986) among others.

Recently, Neely, Roy, and Whiteman (2001) address the issue of near nonidentification in the standard CCAPM model. They also argue that the lack of identifiability of the risk aversion parameter is due to the weak correlation between the instruments and the endogenous variables. Lagged values of consumption growth and asset returns are not very useful in predicting either variable. To improve identification, they suggest to add restrictions in the form of a regression. One possibility is to regress returns on consumption growth assuming that the error is uncorrelated with consumption growth. The OLS regression coefficient provides an estimate of the risk aversion parameter γ. Another way to improve identification of γ is to add a regression of consumption growth on asset returns, thus obtaining an estimate of the intertemporal elasticity of substitution $1/\gamma$.[16] Both these regressions lead to more stable estimates of either the coefficient of relative risk aversion or the elasticity of intertemporal substitution.

In the model which embeds habit formation and Epstein-Zin preferences in (28), we will use the regression equation (25), with an orthogonality condition between the market portfolio return and the error term, to estimate the growth rate of the reference level. We therefore provide a more structural justification to the introduction of this additional regression. Indeed, our approach assumes that the agent is using the return on the market portfolio and past consumption growth to assess his target consumption growth rate. Based on this estimated benchmark consumption growth parameters of the utility function will be estimated with the Euler conditions. Of course the estimation of the regression and of the Euler conditions can be carried out simultaneously. Consistently with the results of Neely, Roy, and Whiteman (2001), this approach leads to more sensible and stable estimates than an approach where all the parameters will be estimated using only Euler conditions.

Another way of overcoming the identification problem is to add information by introducing more Euler conditions with other asset returns. We will do that for our joint model of habit formation with Epstein-Zin preferences by introducing size portfolios. However, what counts in the end for identification is that these returns are not too strongly correlated with the market portfolio returns.

[16]They refer to the first regression as a Hansen and Singleton (1982, 1983) normalization and to the second as the Hall (1988) normalization. In both regressions, the simultaneity problem is assumed away.

1.3.2 Results

Time- and State-Separable Preferences. Our benchmark model is the standard time- and state-separable CRRA utility function. Here, we test the Euler equations

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}R_{f,t+1}\right]=1, \tag{46}$$

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(R_{M,t+1}-R_{f,t+1})\right]=0,$$

for the real risk-free interest rate and the excess market return. The parameter estimates and statistical tests for this utility specification are reported in Table I (model M1). For both sets of instruments, the estimate of the RRA coefficient is negative and significantly different from zero at the 5% level. The parameter of time preferences, δ, is always less than 1. According to the J statistic and the conventional asymptotics, the model is not rejected at the 5% significance level when the second set of instruments is used. Nevertheless, the obtained nonsensical negative estimates of the RRA coefficient lead us to conclude that the conventional power utility function is not supported by the data.

The Habit Formation Approach. The external-habit specification of utility function (1) yields the following Euler equations:

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}R_{f,t+1}\right]=1, \tag{47}$$

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}(R_{M,t+1}-R_{f,t+1})\right]=0.$$

In the estimation, we follow a two-stage procedure. In a first stage, we estimate the subsistence level, S_{t+1}, as a function of past consumption levels. In a second stage, we estimate jointly the Euler equations (47) for the excess market return and the real risk-free interest rate with habit replaced with its estimate obtained in a first stage.

To estimate habit as a function of past consumption levels, we follow the adaptive expectations hypothesis whereby the habit, S_{t+1}, is assumed to be an expectation of consumption taken conditionally on information available to the individual at time t, we can rewrite (16) as

$$C_{t+1} = \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i C_{t-i} - \varepsilon_{t+1}, \tag{48}$$

where ε_{t+1} is an innovation in C_{t+1}.

Using the Koyck transformation yields

$$\triangle C_{t+1} = a - \varepsilon_{t+1} + (1-\lambda)\varepsilon_t, \tag{49}$$

where $\triangle C_{t+1} \equiv C_{t+1} - C_t$.

We estimated this model by maximum likelihood and obtained

$$\begin{array}{rcccc} \triangle C_{t+1} = & 1.5620 & + & 0.2362 & \varepsilon_t, \\ & (0.1298) & & (0.0463) & \end{array} \tag{50}$$

which implies $\widehat{a} = 1.5620$ and $\widehat{\lambda} = 0.7638$.

When we estimate and test ARIMA models for NDS consumption, the AIC preferred model is an ARIMA(0,1,2) with a drift term, which is significantly different from zero. It follows that the best model for C_{t+1} is

$$C_{t+1} = \frac{a}{\lambda} + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i C_{t-i} + \varepsilon_{t+1} + \theta\varepsilon_t, \tag{51}$$

which implies that

$$\triangle C_{t+1} = a + \varepsilon_{t+1} + (\theta - (1-\lambda))\,\varepsilon_t - \theta(1-\lambda)\varepsilon_{t-1}.^{17} \tag{52}$$

We also estimated this model by maximum likelihood and found

$$\begin{array}{rcccccc} \triangle C_{t+1} = & 1.5633 & + & 0.3034 & \varepsilon_t & -0.1707 & \varepsilon_{t-1}. \\ & (0.1457) & & (0.0470) & & (0.0470) & \end{array} \tag{53}$$

Therefore, $\widehat{a} = 1.5633$ and $\widehat{\lambda} = 0.7115$.

The corresponding estimation and test results for the Euler equations for the excess market return and the real risk-free interest rate are given in Table I (model M2 and M3, respectively). The results are practically the same for both specifications. The null

[17] If a disturbance term in (48) is allowed to be an AR(p) with $p \geqslant 0$, then C_{t+1} is a random walk with a drift and an AR($p+1$) error structure.

hypothesis $H_0 : \varphi = \gamma$ is rejected at the 5% significance level for both sets of instruments. We find some evidence against the hypothesis that the absolute value of the reference level does not affect utility. For the second instrument set we can reject at the 10% level the null hypothesis $H_0 : \varphi = 1$.[18] The estimate of $\gamma - \varphi$ is negative and significant for both sets of instruments, which implies that durability in consumption expenditures dominates habit persistence. The value of the time preference parameter δ is greater than 1 for any set of instruments. The model is not rejected statistically at the 5% level for both sets of instruments.

Therefore, we can conclude that the data strongly reject the standard expected utility model in favor of the habit specification, that durability dominates habit persistence at the monthly level,[19] and finally that there is mild evidence for both the absolute and the relative importance of the reference level in the utility function.

The Reference Level as a Function of the Return on the Market Portfolio. In Section 1.2.2, we have shown that given a certain specification of the reference level growth rate, we obtained generalized versions of the SDF obtained by Epstein and Zin (1989) in a recursive utility framework when the certainty equivalent of future utility is of the Kreps and Porteus (1978) form. We will start by estimating the constrained version of the model described in Section 1.2.2. We then estimate the generalized version of the model with the SDF in (28). The third subsection reformulates this generalized model with the habit formation model in difference as in (38), of which Campbell and Cochrane (1999) model is a particular case.

The Epstein-Zin Stochastic Discount Factor. Given that (28) is observationally equivalent to the Epstein-Zin (1989, 1991) SDF when $a_i = 0$ $(i = 1, ..., n)$, we estimate the Euler equations

$$E_t \left[\delta^* \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (R_{M,t+1})^{\kappa} R_{f,t+1} \right] = 1, \tag{54}$$

[18] Fuhrer (2000) uses the US quarterly data over the period from 1966:1 to 1995:4 and also rejects the time- and state-separable utility specification and cannot reject the ratio model.

[19] This result is consistent with the empirical evidence in Eichenbaum and Hansen (1990), Eichenbaum, Hansen, and Singleton (1988), Ferson and Constantinides (1991), Gallant and Tauchen (1989), and Heaton (1995).

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(R_{M,t+1})^{\kappa}\left(R_{M,t+1}-R_{f,t+1}\right)\right]=0,$$

jointly with the equation

$$\triangle c_{t+1}=a_0+b\times r_{M,t+1}+\varepsilon_{t+1}, \tag{55}$$

where $\delta^*\equiv\delta\times exp\left(a_0(\gamma-\varphi)\right)$ and $\kappa\equiv b\left(\gamma-\varphi\right)$.

In Table II, in Column EZ, we present the estimation and test results of Euler equations (54) for the excess market return and the real risk-free interest rate together, when the latter equations are estimated jointly with the log consumption growth rate equation (55). The Euler equations are estimated with the instrument set INS2, while the instruments used to test the equation for the log consumption growth rate consists of a constant, the log real market return lagged two periods, and the log consumption growth rate lagged two periods.

The obtained point estimate of the RRA coefficient is positive (3.53) but not significantly different from zero at the 5% level. The null hypotheses $H_0:\varphi=\gamma$ and $H_0:\varphi=1$ are both rejected statistically. The null hypothesis $H_0:\kappa=0$ $\left(\sigma=\frac{1}{\gamma}\right)$ is also rejected at the 5% significance level. The point estimate of the elasticity of intertemporal substitution is negative (-1.95) but not significantly different from zero at the 5% level. This negative point estimate for the elasticity of intertemporal substitution is consistent with Stock and Wright (2000) who estimate the Epstein and Zin (1989, 1991) SDF (30) using US monthly data. According to Hansen's test of overidentifying restrictions and the conventional asymptotics, the model is not rejected statistically.

A Generalized Epstein-Zin Model with Habit Formation. We now assume that the reference consumption level growth rate is a function of the previous period consumption growth rate and of the return on the market portfolio. Therefore, we estimate the Euler equations

$$E_t\left[\delta^*\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{C_t}{C_{t-1}}\right)^{\frac{\kappa a_1}{b}}(R_{M,t+1})^{\kappa}R_{f,t+1}\right]=1, \tag{56}$$

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{C_t}{C_{t-1}}\right)^{\frac{\kappa a_1}{b}}(R_{M,t+1})^{\kappa}\left(R_{M,t+1}-R_{f,t+1}\right)\right]=0,$$

jointly with the equation

$$\triangle c_{t+1} = a_0 + a_1 \times \triangle c_t + b \times r_{M,t+1} + \varepsilon_{t+1}, \tag{57}$$

where $\delta^* \equiv \delta \times exp\left(a_0(\gamma - \varphi)\right)$ and $\kappa \equiv b\left(\gamma - \varphi\right)$.

The estimation and test results for equations (56) and (57) are presented in Table II, in Column GRS. As in the case of the Epstein-Zin (1989, 1991) pricing kernel, we use the instrument set INS2 when estimating the Euler equations and the set of instruments which consists of a constant, the log real market return lagged two periods, and the log consumption growth rate lagged two periods when estimating the equation for the log consumption growth rate. This two-period lag is important since ε_{t+1} can be correlated with Δc_t.

The obtained point estimate of the RRA coefficient is in the conventional range (close to 1) and, in contrast to the case of the Epstein-Zin (1989, 1991) SDF, significantly different from 0 at the 5% significance level. The point estimate of κ is significantly different from zero and, therefore, the null hypothesis $H_0 : \sigma = \frac{1}{\gamma}$ is rejected at the 5% level. The estimate of the elasticity of substitution is in the conventional range (0.85) and significantly different from 0. We reject at the 5% level the null hypotheses that the reference consumption level plays no role in asset pricing $(H_0 : \varphi = \gamma)$ and that an agent derives utility solely from the ratio of his consumption to some reference level $(H_0 : \varphi = 1)$. According to Hansen's J statistic, the model is not rejected statistically at the 5% significance level.[20]

As announced, we would like to test our specification under the assumption of weak identification. In conducting such a test, we treat δ^* as strongly identified and $\theta = (\gamma, \sigma, a_0, a_1, b)$ as weakly identified.[21] Under weak identification asymptotics, we compute a 95% confidence interval for γ in which δ^* is concentrated out:

$$\left\{\gamma_0 : TS_{cT}\left(\theta_0, \widehat{\delta^*}\left(\theta_0\right)\right) \leqslant \chi^2_{G-1,0.95}\right\}, \tag{58}$$

[20]When the Euler equations (56) for the excess market portfolio return and the real risk-free interest rate are estimated alone, the point estimate of the RRA coefficient is 5.3307 and that of the elasticity of intertemporal substitution is -1.0047. Both are significantly different from 0. Adding equation (57) allows therefore to obtain more sensible estimates of the preference parameters.

[21]Given θ, the parameter δ^* can be estimated precisely from the Euler equation for the risk-free rate of return, $\widehat{\delta^*} = \left(T^{-1}\sum_{t=1}^{T}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{C_t}{C_{t-1}}\right)^{\frac{(\sigma\gamma-1)a_1}{b}}\left(R_{M,t+1}\right)^{(\sigma\gamma-1)}\left(1+R_{f,t+1}\right)\right)^{-1}$ and, hence, is strongly identified by a constant.

where G is the number of orthogonality conditions, $S_{cT}(\theta)$ is the continuous updating objective function computed using a heteroscedasticity robust weighting matrix, and $\widehat{\delta}^*(\theta) = \left(T^{-1}\sum_{t=1}^{T}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{C_t}{C_{t-1}}\right)^{\frac{(\sigma\gamma-1)a_1}{b}}(R_{M,t+1})^{(\sigma\gamma-1)}R_{f,t+1}\right)^{-1}$.

To assess the plausibility of the GRS SDF under the weak identification assumption, we compute the S-set for γ when θ is in the conventional GMM confidence interval. The result is that our model is not rejected at the 5% significance level under weak identification asymptotics for $\gamma \geqslant 1.985$, a value only slightly greater than that corresponding to the upper bound of the 5% GMM confidence set for γ. This result is consistent with that obtained by Stock and Wright (2000), who also find that the value of γ at which the model is not rejected under the assumption of weak identification is usually higher than that under conventional normal asymptotics. However, in contrast to their result, the value of risk aversion at which our model is not rejected statistically may be recognized as economically plausible.

To check whether our SDF performs better than the Epstein-Zin one, we compute the 95% S-set for γ for the model associated with the Epstein-Zin SDF. For the Epstein-Zin specification of the SDF, $\theta = (\gamma, \sigma, a_0, b)$ and

$$\widehat{\delta}^*(\theta) = \left(T^{-1}\sum_{t=1}^{T}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(1+R_{M,t+1})^{(\sigma\gamma-1)}(1+R_{f,t+1})\right)^{-1}. \tag{59}$$

We find that there is no value of γ in the 5% GMM confidence set for which the model is not rejected at the 5% significance level for θ in the conventional GMM confidence set.[22] Since the upper bound of the 5% confidence set for γ under conventional normal asymptotics is 9.20, it means that even if, as in Stock and Wright (2000), there is some value higher than this upper bound value of the coefficient of risk aversion at which the model is not rejected at the 5% level. Contrary to the case of our SDF, this value appears as less economically plausible.

To check the robustness of the estimates we obtained for the preference parameters and to hopefully improve identification, we estimate the risk-free rate equation in (56) together with (57) and the following set of Euler equations corresponding to the ten decile portfolios formed with all stocks listed on the NYSE and AMEX, obtained from CRSP:

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{C_t}{C_{t-1}}\right)^{\frac{\kappa a_1}{b}}(R_{M,t+1})^{\kappa}(R_{i,t+1}-R_{f,t+1})\right] = 0,\ i = 1, ..., 10. \tag{60}$$

[22] For σ we only consider non-negative values in the 5% GMM confidence interval.

The estimation results are reported in Table II in the columns below the heading GRS deciles. As it can be seen, the preference parameters γ and σ are estimated much more precisely. The point estimate of γ is higher (2.8) than when we used only the market premium (0.98), while the elasticity of intertemporal substitution is slightly lower (0.68 instead of 0.86). The conclusions about the value of φ and the null that $\varphi = 1$ remain unchanged.

The Campbell-Cochrane Stochastic Discount Factor. The Campbell and Cochrane (1999) model occupies a prominent position in the recent literature on empirical asset pricing with consumption. As expressed in (38), the SDF corresponding to their utility specification is

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{H_{t+1}}{H_t} \right)^{-\gamma}, \tag{61}$$

where H_t is the surplus consumption ratio, $H_t \equiv \frac{C_t - S_t}{C_t}$. The success of their calibration exercise relies on specifying the heteroscedastic dynamics specified in (40). However, they do not estimate the coefficient of relative risk aversion and set it equal to 2. In this section we will propose a GMM estimation strategy to see if the model is supported by the data and, if it is the case, for what values of the RRA coefficient. We also want to test the model against the more general model with the absolute value of the reference level in the utility function. Tallarini and Zhang (2000) estimate the Campbell and Cochrane (1999) model by the efficient method of moments using quarterly data. For γ, they find high values between 6 and 8 and reject statistically the model. They also report estimation results obtained with GMM where the initial surplus consumption is estimated jointly with the other structural parameters. Then, the RRA coefficient is smaller than 1 in general. However, they discard these results on the basis that the high persistence parameter estimate coupled with a relatively small sample size render the GMM estimation unreliable.

We start by estimating the Euler equations for the equity premium and the risk-free-rate jointly with the Campbell-Cochrane SDF (61):

$$E_t \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{H_{t+1}}{H_t} \right)^{-\gamma} R_{f,t+1} \right] = 1, \tag{62}$$

$$E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{H_{t+1}}{H_t} \right)^{-\gamma} (R_{M,t+1} - R_{f,t+1}) \right] = 0,$$

together with (40) and (41). To obtain the time series of the unobservable surplus consumption, one needs to set the initial value of surplus consumption and a value of the autoregressive parameter ϕ. The former can be set at the steady state value, which means that we need a starting value of γ. The latter is obtained by Campbell and Cochrane (1999) by the autoregressive parameter of the price-dividend ratio. If we follow this strategy, assuming some starting value of the coefficient of relative risk aversion, γ, say 2, calculating H_t $(t = 0, ..., T)$, estimating the Euler equations and iterating over γ, the estimate of γ approaches 0 and, depending on the starting value, goes into the negative. To better control the estimation of γ, we proceeded by grid search to obtain an initial value that was close to the estimated value by the Euler equations. For the persistence parameter ϕ, we estimated it with the S&P 500 price-dividend ratio and obtained a value of 0.985.

The results presented in Table III (CASE1) are based on a starting value of $\gamma = 0.012$ and on the set of instruments with a constant and the real market return, the real risk-free rate, and the real consumption growth rate lagged one and two periods. The final value is 0.0122 and is significantly different from 0. The parameter δ is estimated very precisely as usual and is less than 1, contrary to the estimate above 1 obtained by Tallarini and Zhang (2000).

Our main interest remains to test the Campbell-Cochrane specification against our model where the absolute value of the reference level enters per itself in the utility function. We therefore apply the same estimation procedure to the following Euler equations:

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{H_{t+1}}{H_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}R_{f,t+1}\right] = 1, \tag{63}$$

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{H_{t+1}}{H_t}\right)^{-\gamma}\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}(R_{M,t+1} - R_{f,t+1})\right] = 0.$$

The estimation and test results are presented in Table III under CASE2. The estimated values for γ and δ are very close to the values obtained with CASE1. The parameter φ which distinguishes between the Campbell-Cochrane specification and our specification has a point estimate of -0.2034. Both null hypotheses $\varphi = \gamma$ and $\varphi = 1$ are rejected, the second very strongly. This means that, similarly to the Epstein-Zin and habit formation specification of the previous section, we reject the hypotheses that the reference consumption level plays no role in asset pricing ($H_0 : \varphi = \gamma$) and that an agent derives utility solely from the difference of his consumption with some reference level $(H_0 : \varphi = 1)$. Interestingly, the value estimated for φ is close to the value of -0.3059 obtained with the Epstein-Zin-habit formation specification. Given the interpretation of the habit mentioned in Section 1.2.2 assimilating it to the value of the market portfolio, this result should not be too surprising at an intuitive level. In a way, the Campbell and Cochrane model is more than just habit formation: it has some flavor of a mixture of Epstein-Zin specification and habit formation.

Asymmetric Preferences. The model proposed recently by Barberis, Huang, and Santos (2001) also figures prominently in the empirical asset pricing literature. In the same spirit as Campbell and Cochrane (1999), they calibrate their model to reproduce several empirical regularities between asset returns and consumption. Before testing a model which is close to their specification, we will first explore to what extent the habit formation model estimated in Section 1.3.2 admits different RRA coefficients around a reference level threshold. We therefore explore the utility specification (44), which yields the following SDF:

$$M_{t+1} = \delta \frac{C_{t+1}^{-\gamma_1} S_{t+1}^{\gamma_1 - \varphi_1} I_{[C_{t+1} \geqslant S_{t+1}]} + \lambda C_{t+1}^{-\gamma_2} S_{t+1}^{\gamma_2 - \varphi_2} I_{[C_{t+1} < S_{t+1}]}}{C_t^{-\gamma_1} S_t^{\gamma_1 - \varphi_1} I_{[C_t \geqslant S_t]} + \lambda C_t^{-\gamma_2} S_t^{\gamma_2 - \varphi_2} I_{[C_t < S_t]}}. \tag{64}$$

We examine whether the investor displays different degrees of risk aversion for outcomes above and below habit assuming that he is not loss averse ($\lambda = 1$) or that he has a more or less marked loss aversion ($\lambda = 1.2$ and $\lambda = 1.5$). In all cases, we maintain the assumption that the value of the parameter φ is the same above and below the threshold ($\varphi_1 = \varphi_2$).

Under these assumptions, we estimate the Euler equations for the excess market return and the real risk-free interest rate:

$$E_t \left[\delta \frac{C_{t+1}^{-\gamma_1} S_{t+1}^{\gamma_1 - \varphi} I_{[C_{t+1} \geqslant S_{t+1}]} + C_{t+1}^{-\gamma_2} S_{t+1}^{\gamma_2 - \varphi} \left(1 - I_{[C_{t+1} \geqslant S_{t+1}]}\right)}{C_t^{-\gamma_1} S_t^{\gamma_1 - \varphi} I_{[C_t \geqslant S_t]} + C_t^{-\gamma_2} S_t^{\gamma_2 - \varphi} \left(1 - I_{[C_t \geqslant S_t]}\right)} R_{f,t+1} \right] = 1, \tag{65}$$

$$E_t \left[\frac{C_{t+1}^{-\gamma_1} S_{t+1}^{\gamma_1 - \varphi} I_{[C_{t+1} \geqslant S_{t+1}]} + C_{t+1}^{-\gamma_2} S_{t+1}^{\gamma_2 - \varphi} \left(1 - I_{[C_{t+1} \geqslant S_{t+1}]}\right)}{C_t^{-\gamma_1} S_t^{\gamma_1 - \varphi} I_{[C_t \geqslant S_t]} + C_t^{-\gamma_2} S_t^{\gamma_2 - \varphi} \left(1 - I_{[C_t \geqslant S_t]}\right)} \left(R_{M,t+1} - R_{f,t+1}\right) \right] = 0.$$

In Table IV, we present comparative results for the cases when the agent is assumed to display the same degree of risk aversion for consumption above and below habit (CASE1) and when this assumption is relaxed (CASE2). The Euler equations for both cases are estimated exploiting the set of instruments which consists of a constant and the return on the market portfolio, the real risk-free rate, and the real consumption growth rate lagged one through four periods.

As the decision maker is assumed to display different degrees of risk aversion for outcomes above and below habit, the point estimate of the RRA coefficient increases when consumption declines towards habit. When the adaptive expectations is maintained, we obtain the RRA coefficient estimates which are significantly different from zero (for outcomes below habit, the RRA coefficient estimate is always significant at the 5% level, whereas for consumption above the subsistence level that estimate is significantly positive at the 10% level only for model M3). However we cannot reject the null hypothesis $\gamma_1 = \gamma_2$. The null hypothesis $H_0 : \varphi = 1$ is rejected at the 5% level for both models M2 and M3. In spite of the empirical evidence of local substitutability of consumption, we cannot reject at the 5% level the null hypothesis that the RRA coefficient significantly differs from the parameter φ only when consumption is above the subsistence level. It follows that the ratio external habit-formation model is rejected by the data given any specification of the habit generating process. For consumption above habit, the most plausible assumption is that the representative consumer derives utility from both consumption relative to habit and the level of habit. As consumption declines towards the benchmark level, we cannot reject the assumption that the conventional time- and state-separable utility model well describes agent's preferences. When we make the consumer loss averse by increasing λ to 1.2 or 1.5, the main effect is to decrease the point estimate of δ to nonsensical values.

We now turn to the second type of asymmetry modeled in Section 1.2.3 and estimate the Euler equations

$$E_t\left[\delta\frac{C_{t+1}^{-\gamma_1}V_{t+1}^{\varphi}I_{[R_{M,t+1}\geqslant z_tR_{f,t+1}]}+\lambda(z_t)C_{t+1}^{-\gamma_2}V_{t+1}^{\varphi}\left(1-I_{[R_{M,t+1}\geqslant z_tR_{f,t+1}]}\right)}{C_t^{-\gamma_1}V_t^{\varphi}I_{[R_{M,t}\geqslant z_tR_{f,t}]}+\lambda C_t^{-\gamma_2}V_t^{\varphi}\left(1-I_{[R_{M,t}\geqslant z_tR_{f,t}]}\right)}R_{f,t+1}\right]=1, \tag{66}$$

$$E_t\left[\frac{C_{t+1}^{-\gamma_1}V_{t+1}^{\varphi}I_{[R_{M,t+1}\geqslant z_tR_{f,t+1}]}+\lambda(z_t)C_{t+1}^{-\gamma_2}V_{t+1}^{\varphi}\left(1-I_{[R_{M,t+1}\geqslant z_tR_{f,t+1}]}\right)}{C_t^{-\gamma_1}V_t^{\varphi}I_{[R_{M,t}\geqslant z_tR_{f,t}]}+\lambda C_t^{-\gamma_2}V_t^{\varphi}\left(1-I_{[R_{M,t}\geqslant z_tR_{f,t}]}\right)}\left(R_{M,t+1}-R_{f,t+1}\right)\right]=0,$$

where, as before, we have maintained the assumption that the parameter φ is the same above and below the threshold. The variable $z_t = \frac{Z_t}{P_t}$ (where Z_t is a value of the stock which happened in the past and P_t is the current stock value) is a benchmark which measures the size of prior losses. It is modeled as in Barberis, Huang, and Santos (2001):

$$z_{t+1} = \eta \left(z_t \frac{\overline{R}}{R_{t+1}} \right) + (1 - \eta)(1). \tag{67}$$

When $\eta = 0$, the benchmark level tracks the stock value one-for-one and moves therefore very fast. When $\eta = 1$, the benchmark level moves sluggishly.[23] Moreover, the agent is loss averse and the degree of loss aversion is a function of z_t, as follows:

$$\lambda(z_t) = \lambda + k(z_t - 1). \tag{68}$$

The larger the prior losses (the larger z_t is), the more painful subsequent losses will be. Barberis, Huang, and Santos (2001) illustrate the effects of loss aversion and prior losses by choosing parameter values, simulating the model and ultimately comparing moments of asset returns with historical figures. Although our model is different from theirs, we have kept the same features regarding loss aversion and the effects of prior losses on this loss aversion. Consistently with the rest of the paper we will however estimate the preference parameters (γ_1, γ_2, φ, and δ) by GMM for various configurations of the parameters λ, κ, and η. Representative results are presented in Table V. Point estimates for the γ are often negative, especially when η is different from 1, but they are not significantly different from zero. It is only when η is equal to 1 that we find positive and significant values for the risk aversion parameters but they are generally quite large. The estimate for φ is also positive, very large and significantly different from zero. The estimate for δ is greater than 1. Results similar to the latter results have been obtained for several values of λ and κ as long as η is equal to 1. The parameter η controls the persistence of z_t and hence also the persistence of the price-dividend ratio. We can see that results are highly sensitive to this parameter. Although our model is different from the original model of Barberis, Huang, and Santos (2001), the results illustrate that when sensible values of the preference parameters are obtained, albeit high, the parameter φ is significantly different from zero and therefore the reference level seems to play a role in the utility function of the agent over and above current consumption.

[23] If the return on the asset is good ($R_{t+1} > \overline{R}$), z_t falls in value, as the benchmark rises less than the stock price. In case of a poor return, z_t rises as the benchmark level falls less than the stock price.

1.4 Concluding Remarks

In this paper, we propose an expected utility model in which a representative agent is assumed to derive utility from both the ratio of consumption to some reference level of consumption and this benchmark level itself. We have explored several specifications for the dynamics of this reference level along the lines of the prominent models in the asset pricing literature. Our main conclusion is that there is ample evidence for the presence of such a reference level in the utility function of the representative agent. Following Alvarez and Jermann (2002), we can rationalize the confirmed presence of this benchmark consumption level by the fact that it adds persistence to the pricing kernel.

Our approach has also led to the generalization of current models in the literature. We succeed in associating habit formation preferences, either in ratio or in difference form, together with so-called Epstein-Zin preferences. In terms of pricing models, we obtained a SDF which is a geometric average of a habit-formation CCAPM and of a CAPM, whereas in Epstein and Zin (1991) it was an average of the standard CCAPM and of the CAPM.

A main feature of our approach is to estimate an additional equation to model the growth rate of this reference level. We have mainly concentrated our analysis on the return of the financial wealth but other possibilities are open here. An obvious one is to make it depend on the return of human capital as in Jagannathan and Wang (1996).

This paper raises a central issue: is this reference level a genuine reflection of individual preferences or is it a feature of aggregation? This is all the more relevant since habit formation has not received yet strong support in micro-data (see Dynan, 2000). Recently, Guvenen (2002) proposed a model where there is limited participation in the stock marker and heterogeneity in the elasticity of intertemporal substitutions. He shows that this model with heterogeneous agents has a reduced-form which is extremely similar to Campbell and Cochrane (1999) framework. The only way to put some light on these issues is to estimate and test the various models proposed in this paper with individual data. This constitutes an exciting agenda for future research.

Appendix: Tables

Table I.
Estimation and Test Results for the Reference Level Model with Past State Variables

The sampling period is from 1960:3 to 1996:12, for a total of 442 observations. An iterated GMM approach is used to test Euler equations and estimate model parameters. Our benchmark model is the standard time- and state-separable CRRA utility function (model M1). In order to estimate habit as a function of past consumption levels, we assume that habit evolves according to the adaptive expectations hypothesis (model M2 (the disturbance term in (48) is an innovation in C_{t+1}) and model M3 (the disturbance term in (48) is allowed to be an AR(1)). In a first stage, we estimate habit, S_{t+1}, as a function of past consumption levels. In a second stage, we estimate the Euler equations (47) for the excess market return and the real risk-free interest rate with habit replaced with its estimate obtained in a first stage. For each preference specification, the Euler equations for the excess market return and the real risk-free interest rate are estimated jointly exploiting two sets of instruments. The first instrument set (INS1) has a constant, the real market return lagged, the real risk-free rate lagged, and the real consumption growth rate lagged. As our second set (INS2), we use the first set of instruments adjusted by the same variables lagged an additional period. The J statistic is Hansen's test of the overidentifying restrictions. The P value is the marginal significance level associated with the J statistic. In Panel A, we report the values of the parameters estimated directly from the Euler equations. The values of the parameters estimated indirectly are presented in Panel B. The standard errors for the parameters estimated indirectly are calculated by using the delta method.

1. Time- and state-separable preferences:
model M1:

	INS1			INS2		
Param.	Estim.	SE	t stat.	Estim.	SE	t stat.
γ	-0.1706	0.0790	-2.1595	-0.2072	0.0647	-3.2025
δ	0.9995	0.0002	4997.500	0.9994	0.0002	4997.0000
J statistic	12.9925			13.6788		
P value	0.0432			0.3217		

2. Habit formation approach:
model M2

	INS1			INS2		
Param.	Estim.	SE	t stat.	Estim.	SE	t stat.
Panel A:						
γ	1.4735	0.5068	2.9075	1.1369	0.2889	3.9353
φ	2.0369	0.6659	3.0589	1.6306	0.3737	4.3634
δ	1.0028	0.0013	771.3846	1.0020	0.0008	1252.5000
J statistic	9.5268			12.3438		
P value	0.0898			0.3384		
Panel B:						
$\gamma - \varphi$	-0.5634	0.1754	-3.2121	-0.4937	0.0999	-4.9419
$1 - \varphi$	-1.0369	0.6659	-1.5571	-0.6306	0.3737	-1.6874

Table I (continued)

model M3

Panel A:						
γ	1.4581	0.5011	2.9098	1.1001	0.2785	3.9501
φ	2.0633	0.6716	3.0722	1.6266	0.3670	4.4322
δ	1.0028	0.0014	716.2857	1.0020	0.0008	1252.5000
J statistic	9.5182			12.4464		
P value	0.0901			0.3310		
Panel B:						
$\gamma - \varphi$	-0.6052	0.1882	-3.2157	-0.5265	0.1047	-5.0287
$1 - \varphi$	-1.0633	0.6716	-1.5832	-0.6266	0.3670	-1.7074

Table II.
Estimation and Test Results for the Epstein-Zin and Generalized Epstein-Zin and Habit Formation Specifications

The sampling period is from 1960:3 to 1996:12, for a total of 442 observations. An iterated GMM approach is used. First, we set $a_i = 0$ $(i = 1, ..., n)$ and estimate the Euler equations (54) for the excess market portfolio return and the risk-free rate jointly with equation (55). This is the case of the Epstein-Zin (1989, 1991) SDF, the results for which are presented in Column EZ. Second, we set $a_i = 0$ $(i = 2, ..., n)$ and estimate the Euler equations (56) for the excess market portfolio return and the risk-free rate jointly with equation (57). For each preference specification, the Euler equations are estimated using as instruments a constant, the real market return, the real risk-free rate, and the real consumption growth rate. All these variables are lagged one and two periods. The set of instruments used to test the equation for the log consumption growth rate consists of a constant, the log real market return lagged two periods, and the log consumption growth rate lagged two periods. The J statistic is Hansen's test of the overidentifying restrictions. The P value is the marginal significance level associated with the J statistic. In Panel A, we report the values of the parameters estimated directly. The values of the parameters estimated indirectly are presented in Panel B. The standard errors for the parameters estimated indirectly are calculated by using the delta method.

	EZ			GRS			GRS	Deciles	
Param.	Estim.	SE	t stat.	Estim.	SE	t stat.	Estim.	SE	t stat.
Panel A:									
a_0	0.0074	0.0008	9.2500	0.0022	0.0003	7.3333	0.0026	0.0001	26.0000
a_1				-0.2273	0.0597	-3.8074	-0.4564	0.0354	-12.8927
b	-0.6309	0.0722	-8.7382	-0.1202	0.0461	-2.6074	0.3257	0.0200	16.2850
γ	3.5275	2.8954	1.2183	0.9847	0.3017	3.2638	2.7965	0.1830	15.2814
κ	-7.8611	0.3514	-22.3708	-0.1552	0.0400	-3.8800	0.8972	0.0416	21.5673
δ^*	1.0424	0.0050	208.48	1.0017	0.0006	1669.5	1.0047	0.0004	2511.75
J statistic	12.1338			12.2398			21.2106		
P value	0.4350			0.3459			1.0000		
Panel B:									
σ	-1.9450	1.6151	-1.2043	0.8580	0.2427	3.5352	0.6784	0.0392	17.3061
φ	-8.9329	2.2761	-3.9247	-0.3059	0.2378	-1.2864	0.0414	0.2409	0.1719
$1-\varphi$	9.9329	2.2761	4.3640	1.3059	0.2378	5.4916	0.9586	0.2409	3.9792
$\gamma-\varphi$	12.4604	1.1194	11.1313	1.2906	0.3251	3.9699	2.7551	0.0844	32.6434
δ	0.9506	0.0059	161.11	0.9988	0.0004	2497.0	0.9975	0.0005	1995.00

Table III.
Estimation and test results for the Campbell-Cochrane SDF specification

The sampling period is from 1960:3 to 1996:12, for a total of 442 observations. An iterated GMM approach is used to test Euler equations and estimate model parameters. The Euler equations for the excess market return and the real risk-free interest rate are estimated jointly with equations (40) and (41) using as instruments a constant, the real market return, the real risk-free rate, and the real consumption growth rate. All these variables are lagged one and two periods. The J statistic is Hansen's test of the overidentifying restrictions. The P value is the marginal significance level associated with the J statistic. In Panel A, we report the values of the parameters estimated directly from the Euler equations. The values of the parameters estimated indirectly are presented in Panel B. The standard errors for the parameters estimated indirectly are calculated by using the delta method.

	CASE1			CASE2		
Param.	Estim.	SE	t stat.	Estim.	SE	t stat.
Panel A:						
γ	0.0122	0.0046	2.6522	0.0115	0.0046	2.5000
φ				-0.2034	0.0830	2.4506
δ	0.9994	0.0002	4997.0	0.9993	0.0002	4996.5
J statistic	15.2328			13.9361		
P value	0.2290			0.2366		
Panel B:						
$\gamma - \varphi$				0.2149	0.0834	2.5767
$1 - \varphi$				1.2034	0.0830	14.4988

Table IV.
Estimation and Test Results for the Reference Level Habit Formation Model with Asymmetric Preferences

The sampling period is 1960:3 to 1996:12, for a total of 442 observations. An iterated GMM approach is used to test Euler equations and estimate model parameters. Our benchmark model is the standard time- and state-separable CRRA utility function (model M1). In order to estimate habit as a function of past consumption levels, we assume that habit evolves according to the adaptive expectations hypothesis (model M2 (the disturbance term in (48) is an innovation in C_{t+1}) and model M3 (the disturbance term in (48) is allowed to be an AR(1)). In a first stage, we estimate habit, S_{t+1}, as a function of past consumption levels. In a second stage, we estimate the Euler equations (65) for the excess market return and the real risk-free interest rate with habit replaced with its estimate obtained in a first stage. We present the comparative analysis of the cases when the agent is assumed to display the same degree of risk aversion for consumption above and below habit (CASE1) and when this assumption is relaxed (CASE2). In both cases, the Euler equations for the excess market return and the real risk-free interest rate are estimated jointly exploiting the set of instruments, which consists of a constant and the return on market portfolio, the real risk-free rate, and the real consumption growth rate lagged one through four periods. The J statistic is Hansen's test of the overidentifying restrictions. The P value is the marginal significance level associated with the J statistic. In Panel A, we report the values of the parameters estimated directly from the Euler equations. The values of the parameters estimated indirectly are presented in Panel B. The standard errors for the parameters estimated indirectly are calculated by using the delta method.

model M2

	CASE1			CASE2		
Param.	Estim.	SE	t stat.	Estim.	SE	t stat.
Panel A:						
γ	0.3052	0.0876	3.4840			
γ_1				0.2180	0.1295	1.6834
γ_2				0.3541	0.1613	2.1953
φ	0.5379	0.1239	4.3414	0.4921	0.1278	3.8505
δ	1.0001	0.0003	3333.6667	1.0000	0.0003	3333.3333
J statistic	27.8911			27.9608		
P value	0.2199			0.1770		
Panel B:						
$\gamma - \varphi$	-0.2327	0.0410	-5.6756			
$1 - \varphi$	0.4621	0.1239	3.7296	0.5079	0.1278	3.9742
$\gamma_1 - \gamma_2$				-0.1361	0.2293	-0.5935
$\gamma_1 - \varphi$				-0.2741	0.1259	-2.1771
$\gamma_2 - \varphi$				-0.1380	0.1181	-1.1685

Table IV (continued)

model M3

Panel A:						
γ	0.3081	0.0862	3.5742			
γ_1				0.2049	0.1278	1.6033
γ_2				0.3622	0.1600	2.2638
φ	0.5660	0.1254	4.5136	0.5072	0.1291	3.9287
δ	1.0001	0.0003	3333.6667	1.0000	0.0003	3333.3333
J statistic	27.7686			27.8292		
P value	0.2247			0.1814		
Panel B:						
$\gamma - \varphi$	-0.2580	0.0444	-5.8108			
$1 - \varphi$	0.4340	0.1254	3.4609	0.4928	0.1291	3.8172
$\gamma_1 - \gamma_2$				-0.1573	0.2281	-0.6896
$\gamma_1 - \varphi$				-0.3023	0.1274	-2.3728
$\gamma_2 - \varphi$				-0.1450	0.1179	-1.2299

Table V.
Estimation and Test Results for the Loss Aversion Model

The sampling period is 1960:3 to 1996:12, for a total of 442 observations. An iterated GMM approach is used to test Euler equations and estimate model parameters. We chose several sets of values for the parameters λ, κ, and η according more or less to the values chosen by Barberis, Huang, and Santos (2001). Given these values we estimate the preference parameters γ_1,γ_2, φ, and δ based on the Euler equations for the excess market return and the real risk-free interest rate. Equations are estimated jointly exploiting the set of instruments, which consists of a constant and the return on market portfolio, the real risk-free rate, and the real consumption growth rate lagged one through two periods. The P value is the marginal significance level associated with the J statistic which is Hansen's test of the overidentifying restrictions.

$\lambda = 2.50$, $\eta = 0.9$

	$\kappa = 1$		$\kappa = 3$		$\kappa = 4$	
Param.	Estim.	t stat.	Estim.	t stat.	Estim.	t stat.
γ_1	-0.94	-0.14	-0.79	-0.12	-0.67	-0.10
γ_2	-0.53	-0.08	-0.36	-0.05	-0.24	-0.04
φ	4.59	0.68	4.26	0.62	4.17	0.61
δ	1.03	84.67	1.02	82.83	1.02	81.24
P value	0.86		0.85		0.84	

$\kappa = 3$, $\eta = 1$

	$\lambda = 1.5$		$\lambda = 2$		$\lambda = 3$	
Param.	Estim.	t stat.	Estim.	t stat.	Estim.	t stat.
γ_1	-1.59	-2.92	-2.34	-4.73	8.69	4.05
γ_2	-1.43	-2.65	-2.07	-4.23	9.17	4.28
φ	-4.62	-7.45	-4.72	-8.43	15.33	6.45
δ	0.98	671.43	0.98	696.86	1.06	160.56
P value	0.77		0.80		0.88	

$\kappa = 3$, $\lambda = 3$

	$\eta = 1$		$\eta = 0.9$		$\eta = 0.8$	
Param.	Estim.	t stat.	Estim.	t stat.	Estim.	t stat.
γ_1	8.69	4.05	-0.94	-0.12	-1.26	-0.17
γ_2	9.17	4.28	-0.43	-0.06	-0.76	-0.10
φ	15.33	6.45	4.45	0.56	5.02	0.67
δ	1.06	160.56	1.02	71.65	1.02	71.27
P value	0.88		0.86		0.88	

References

[1] Abel, A., 1990, Asset Prices under Habit Formation and Catching Up with the Joneses, The American Economic Review 80, Papers and Proceedings, 38-42.

[2] Abel, A., 1996, Risk Premia and Term Premia in General Equilibrium, unpublished paper, University of Pennsylvania.

[3] Abel, A., 1999, Risk Premia and Term Premia in General Equilibrium, Journal of Monetary Economics 43, 3-33.

[4] Alvarez, F. and U. J. Jermann (2002), Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth, unpublished paper, University of Chicago.

[5] Bakshi, G. and Z. Chen, 1996, The Spirit of Capitalism and Stock-Market Prices, American Economic Review 86, 133-157.

[6] Barberis, N., M. Huang, and T. Santos, 2001, Prospect Theory and Asset Prices, Quarterly Journal of Economics 116, 1-53.

[7] Bekaert, G., R. Hodrick, and D. Marshall, 1997, The Implications of First-Order Risk Aversion for Asset Market Risk Premiums, Journal of Monetary Economics 40, 3-39.

[8] Benartzi, S. and R. Thaler, 1995, Myopic Loss Aversion and the Equity Premium Puzzle, Quarterly Journal of Economics 110, 73-92.

[9] Bonomo, M. and R. Garcia, 1993, Disappointment Aversion as a Solution to the Equity Premium and the Risk-Free Rate Puzzles, CRDE Discussion Paper 2793, Université de Montréal.

[10] Campbell, J. and J. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, Journal of Political Economy 107, 205-251.

[11] Carroll, C., J. Overland, and D. Weil, 2000, Saving and Growth with Habit Formation, The American Economic Review 90, 341-355.

[12] Chou, R., R. Engle, and A. Kane, 1992, Measuring Risk Aversion from Excess Returns on a Stock Index, Journal of Econometrics 52, 201-224.

[13] Cochrane, J., 2001, Asset Pricing, Princeton University Press, Princeton, NJ.

[14] Constantinides, G., 1990, Habit Formation: A Resolution of the Equity Premium Puzzle, Journal of Political Economy 98, 519-543.

[15] Dittmar, R. F., 2002, Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns, Journal of Finance 57, 369-403.

[16] Dynan, K., 2000, Habit Formation in Consumer Preferences: Evidence from Panel Data, The American Economic Review 90, 391-406.

[17] Eichenbaum, M. and L. Hansen, 1990, Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data, Journal of Business and Economic Statistics 8, 53-69.

[18] Eichenbaum, M., L. Hansen, and K. Singleton, 1988, A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty, Quarterly Journal of Economics 103, 51-78.

[19] Epstein, L. and S. Zin, 1989, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica 57, 937-969.

[20] Epstein, L. and S. Zin, 1991, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical analysis, Journal of Political Economy 99, 263-288.

[21] Epstein, L. and S. Zin, 2001, The Independence Axiom and Asset Returns, Journal of Empirical Finance 8, 537-572.

[22] Ferson, W. and G. Constantinides, 1991, Habit Persistence and Durability in Aggregate Consumption: Empirical Tests, Journal of Financial Economics 29, 199-240.

[23] Fishburn, P.C., 1977, Mean-Risk Analysis with Risk Associated with Below-Target Returns, American Economic Review 67, 116-126.

[24] Fuhrer, J., 2000, Habit Formation in Consumption and Its Implications for Monetary-Policy Models, The American Economic Review 90, 367-390.

[25] Gali, J., 1994, Keeping up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices, Journal of Money, Credit and Banking 26, 1-8.

[26] Gallant, A. and G. Tauchen, 1989, Seminonparametric Estimation of Conditionally Heterogeneous Processes: Asset Pricing Applications, Econometrica 57, 1091-1120.

[27] Garcia, R., E. Renault, and A. Semenov, 2002, Disentangling Risk Aversion and Intertemporal Substitution Through a Reference Level, CIRANO working paper, 2003s-12.

[28] Gordon, S. and P. St-Amour, 1998a, Estimating a Continuous-Time Asset Pricing Model with State-Dependent Risk Aversion, working paper.

[29] Gordon, S. and P. St-Amour, 1998b, Asset Prices with Contingent Risk Preferences, working paper.

[30] Green, P. E., 1963, Risk Attitudes and Chemical Investment Decisions, Chem. Eng. Progress 59, 35-40.

[31] Gul, F., 1991, A Theory of Disappointment Aversion, Econometrica 59, 667-86.

[32] Guvenen, F., Habit Formation or Cross-Sectional Heterogeneity?, working paper.

[33] Hall, R. E., 1988. Intertemporal Substitution in Consumption, Journal of Political Economy 96, 339-357.

[34] Halter, A. N. and G. W. Dean, 1971, Decisions Under Uncertainty, Cincinnati.

[35] Hansen, L., J. Heaton, and A. Yaron, 1996, Finite Sample Properties of Some Alternative GMM Estimators, Journal of Business and Economic Statistics 14, 262-280.

[36] Hansen, L. P. and K. J. Singleton, 1982, Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, Econometrica 50, 1269-1286.

[37] Hansen, L. P. and K. J. Singleton, 1983, Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns, Journal of Political Economy 91, 249-265.

[38] Harvey, C.R., 1991, The World Price of Covariance Risk, Journal of Finance 46, 111-157.

[39] Heaton, J., 1995, An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications, Econometrica 63, 681-717.

[40] Holthausen, D., 1981, A Risk-Return Model with Risk and Return Measured as Deviation from a Target Return, The American Economic Review 71, 182-188.

[41] Jagannathan, R. and Z. Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, Journal of Finance 51, 3-54.

[42] Kahneman, D. and A. Tversky, 1979, Prospect Theory: An Analysis of Decision Under Risk, Econometrica 47, 263-291.

[43] Kocherlakota, N., 1990, On Tests of Representative Consumer Asset Pricing Models, Journal of Monetary Economics 26, 285-304.

[44] Kreps, D. and E. L. Porteus, 1978, Temporal Resolution of Uncertainty and Dynamic Choice Theory, Econometrica 36, 185-200.

[45] Lettau, M. and S. Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, Journal of Political Economy 109, 1238-1287.

[46] Mark, N., 1988, Time-Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts, Journal of Financial Economics 22, 335-354.

[47] McCurdy, T. and I. Morgan, 1991, Tests for a Systematic Risk Component in Deviations from Uncovered Interest Parity, Review of Economic Studies 58, 587-602.

[48] Melino, A. and A. X. Yang, 2003, State-Dependent Preferences Can Explain the Equity Premium Puzzle, Working Paper, University of Toronto.

[49] Neely, C. J., A. Roy, and C. H. Whiteman, 2001, Risk Aversion Versus Intertemporal Substitution: A Case Study of Identification Failure in the Intertemporal Consumption Capital Asset Pricing Model, Journal of Business and Economic Statistics 19, 395-403.

[50] Porter, R. B., 1974, Semivariance and Stochastic Dominance: A Comparison, The American Economic Review 64, 200-204.

[51] Schroder, M. and C. Skiadas, 2002, An Isomorphism Between Asset Pricing Models With and Without Linear Habit Formation, The Review of Financial Studies 15, 1189-1221.

[52] St-Amour, P., 1993, State-Dependent Risk Aversion, working paper.

[53] Stock, J. and J. Wright, 2000, GMM with Weak Identification, Econometrica 68, 1055-1096.

[54] Sundaresan, S., 1989, Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth, Review of Financial Studies 2, 73-88.

[55] Swalm, R. D., 1966, Utility Theory-Insights into Risk Taking, Harvard Business Review 47, 123-136.

[56] Tallarini, T. D. and H. Zhang, 2000, External Habit and the Cyclicality of Expected Stock Returns, unpublished paper, Graduate School of Industrial Organization, Carnegie Mellon University.

[57] Tauchen, G., 1986, Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data, Journal of Business and Economic Statistics 4, 397-425.

[58] Tversky, A. and D. Kahneman, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty 5, 297-323.

[59] Weil, P., 1989, The Equity Premium Puzzle and the Risk-Free Rate Puzzle, Journal of Monetary Economics 24, 401-21.

2 Disentangling Risk Aversion and Intertemporal Substitution Through a Reference Level

2.1 Introduction

In the standard consumption capital asset pricing model (CCAPM), a representative agent maximizes his time-separable expected utility. The curvature of the utility function captures two aspects of the agent's preferences. As the concavity of the function increases so does his aversion to risk as well as his desire to smooth consumption intertemporally. For a power utility function, it means that the coefficient of relative risk aversion is constrained to be the inverse of the elasticity of intertemporal substitution. This constraint is not supported by empirical observations since agents tend to exhibit an elasticity of intertemporal substitution which is less than the inverse of the relative risk aversion coefficient, as emphasized in Weil (1990). To disentangle the two concepts, Epstein and Zin (1989), often referred to as EZ hereafter, and Weil (1989) have proposed a recursive utility framework that generalizes the dynamic choice model under uncertainty of Kreps and Porteus (1978).

Epstein and Zin (1989) qualify this disentangling by stressing that the risk aversion parameter in their model should not be interpreted independently from the attitude towards intertemporal substitution. However, after reading Epstein and Zin (1991) and a number of papers in the ensuing literature, one realizes that the estimates of the risk aversion parameter are often directly compared with the ones obtained in the standard CCAPM framework as in Hansen and Singleton (1983). We will argue that such a reading of the risk parameter in the Epstein-Zin model could lead to spurious interpretations of allegedly realistic low levels of estimated risk aversion. Moreover, Epstein and Zin's (1991) conclusion that "risk preferences do not differ statistically from the logarithmic specification" could be reinterpreted as an indication that the attitude towards intertemporal substitution does not differ statistically from the logarithmic specification. This reinterpretation is crucially important from an economic point of view since it allows to distinguish myopia in consumption saving decisions from myopia in portfolio allocation (Giovannini and Weil (1989)).

In this paper, we propose a new way to extend preferences to uncertain future consumption flows while maintaining the same unambiguous definition of the elasticity of intertemporal substitution for certain future streams of consumption. We suggest that the requested disentangling may alternatively be obtained not by replacing, as the recursive utility does, the future consumption stream by a certainty equivalent of future utility but by an exogenous reference level which, in a recursive way, assesses the expected future consumption. Therefore, risk aversion is now defined with respect to the unpredictable discrepancy between actual consumption and this reference level (a quantity independent of the attitude towards risk) and not with respect to the forthcoming level of recursive utility which still mixes attitudes towards risk and intertemporal substitution.

In this new framework, preferences are therefore represented by a generalized von Neumann-Morgenstern utility specification whereby satisfaction is derived from consumption relative to an external reference level as well as from this reference level itself. This specification is related to several concepts in the literature. In habit formation models, utility is measured with respect to consumption relative to a time-varying habit or subsistence level either in ratios (Abel (1990, 1996)) or in differences (Constantinides (1990), Sundaresan (1989), and Campbell and Cochrane (1999)), among others. Several variables have also been added to the utility function besides consumption: leisure (Eichenbaum, Hansen, and Singleton (1988)), public expenditures (Aschauer (1985)), durable goods (Startz (1989)), wealth (Bakshi (1996), Smith (2001)). Recently, Carroll, Overland, and Weil (2000), in a growth and saving model, proposed a specification in which the agent can derive utility both from the level of consumption relative to a reference level and from the absolute value of this reference level.

To recover a SDF which is observationally equivalent to the Kreps and Porteus specification in the recursive utility framework of Epstein and Zin (1989), as we do in this paper, we establish a structural link between this reference level and the return on the market portfolio. However, we emphasize that, although observationally equivalent, the two models deliver different measures of risk aversion. As in Barberis, Huang, and Santos (2001), the introduction of a reference level actually changes the measure of risk and in turn the level of risk aversion needed to explain the observed risk premium. [24] In Garcia, Renault, and Semenov (2002), henceforth GRS (2002), we generalize the reference level and make it depend on past consumption as well as on the return on the market portfolio. Therefore,

[24]This result does not depend upon any specific behavioral interpretation of the reference level and may simply be produced by the heterogeneity of agents as in Guvenen (2002).

we embed both habit persistence and the recursive utility Kreps-Porteus model in the same SDF.

Section 2.2 presents the issue of disentangling risk aversion from the elasticity of intertemporal substitution in a recursive utility framework. In Section 2.3, we introduce a new von Neumann-Morgenstern expected utility framework which provides a better separation of the two concepts. In particular, we show in this context that the risk aversion parameter in the recursive utility specification depends on the intertemporal elasticity of substitution. Section 2.4 concludes.

2.2 Disentangling Risk Aversion and Intertemporal Substitution via Recursive Utility

Our focus of interest is the modelling of a preference ordering between stochastic consumption processes $C = (C_t)_{t \geqslant 0}$. Following Duffie and Epstein (1992a), it is quite natural to consider that a utility function U is *risk averse* if, for all processes C in some domain:

$$U[C] \leq U[EC], \tag{69}$$

where EC denotes the deterministic process defined by $[E(C)]_t = E[C_t]$. A more difficult question is to assess the level of risk aversion of a given utility function U in this intertemporal context. Yet, an answer to this question is crucial for contributing to the empirical debate surrounding asset pricing puzzles. For instance, the equity premium puzzle amounts to consider that the level of risk aversion needed to reproduce the observed risk premium on equity is not reasonable. One step in the direction of quantifying risk aversion has been performed by Duffie and Epstein (1992a) through the notion of comparative risk aversion. They define this concept as follows.

Definition 1 *A utility function U^* is said to be more risk averse than U if it rejects any gamble that is rejected by U, that is for any stochastic process C and any deterministic process $\overline{C}$ in some domain: $U[C] \leq U[\overline{C}] \Longrightarrow U^*[C] \leq U^*[\overline{C}]$.*

In other words, if U leads to prefer a deterministic sequence $\overline{C}_t, t \geqslant 0$, to a stochastic consumption process $C_t, t \geqslant 0$, a fortiori U^* will lead to prefer the deterministic path. As acknowledged by Duffie and Epstein (1992a), this definition is not innocuous. To be comparable according to this definition, U^* and U must rank deterministic programs identically.

In particular, one cannot give a sense to the statement "U^* is more risk averse than U" if U^* and U feature different temporal preferences, either for immediate versus late consumption (subjective discounting) or for consumption smoothing (elasticity of intertemporal substitution). This is indeed a fundamental impossibility result about disentangling risk aversion and intertemporal substitution. The only way to escape this general impossibility is to be more specific about the utility model. Epstein and Zin (1989) and Duffie and Epstein (1992 a,b) put forward the recursive utility framework in discrete time and continuous time respectively.

First introduced by Koopmans (1960) in a deterministic setting, the recursive relation

$$V_t = W[C_t, V_{t+1}] \tag{70}$$

specifies the utility index V_t at time t as a function of the consumption C_t in period t and the utility index V_{t+1} of future consumption. The function W has been called an aggregator by Lucas and Stokey (1984). It defines both the rate of time preference and the elasticity of intertemporal substitution. For instance, the time-additive separable (TAS) utility function[25] corresponds to the aggregator

$$W[C, V] = u(C) + \beta V, \tag{71}$$

where β is the subjective discount factor. In the isoelastic case, $u(C) = \frac{C^\rho - 1}{\rho}, \rho \leq 1, \rho = 1 - \frac{1}{\sigma}, \sigma > 0$ is the elasticity of intertemporal substitution. The issue of interest is to extend equation (70) to uncertain consumption streams. Then, the future utility index V_{t+1} appears itself random at time t (we will denote it $\widetilde{V}_{t+1}$ to stress that it is stochastic) and cannot be plugged into (70) without a preliminary treatment.

In other words, we must look for a generalization of (70) which admits the latter equation as a particular case when the future random value of $\widetilde{V}_{t+1}$ is known at time t. The solution proposed by Epstein and Zin (1989) appears to be quite natural in this respect. They consider that the agent first computes the certainty equivalent $m(\widetilde{V}_{t+1}|I_t)$ of the conditional distribution $(\widetilde{V}_{t+1}|I_t)$ of $\widetilde{V}_{t+1}$ given the information at time t and then combines the latter with C_t via the aggregator W :

$$V_t = W[C_t, m(\widetilde{V}_{t+1}|I_t)]. \tag{72}$$

[25]See Becker and Boyd III (1997) for a review of aggregators and their properties.

They refer to Kreps and Porteus (1978) to study (72) under the assumption that m is an expected-utility based certainty equivalent such as

$$m(\widetilde{V}_{t+1}|I_t) = f^{-1}[E[f(\widetilde{V}_{t+1})|I_t]], \tag{73}$$

where they call f a von Neumann-Morgenstern utility index.

This terminology is motivated by the fact that the utility functions defined by (72) and (73) conform with expected utility theory when ranking timeless gambles. To see this, let us consider a lottery on a sequence (C_{t+h}), $h \geqslant 0$, of current and future consumption that is genuinely timeless because the two following conditions are fulfilled. First, randomness is about just one particular future consumption (C_{t+H}) for given H, while the other ones are known at time t. For sake of notational simplicity, let us assume that for any $h \neq H, C_{t+h} = C^*$ given. Second, the uncertainty at time t about C_{t+H} has no temporal features. Basically, the value of C_{t+H} appears to be random at time t but is going to be known no later than time $(t+1)$.

Then, with the aggregator (71), the utility index V_t at time t is given by

$$V_t = u(C^*) + \beta m[(1-\beta)^{-1}u(C^*) + \beta^{H-1}\{u(C_{t+H}) - u(C^*)\}]. \tag{74}$$

We deduce from (74) that it is true that m characterizes the risk aversion preferences for timeless gambles. Typically, for a given level of risk involved in future consumption C_{t+H}, different people will value more or less such a gamble depending upon their level of risk aversion included in m or, equivalently, in the Von Neumann-Morgenstern utility index f.

This risk aversion assessment appears at first sight to be fairly well disentangled from the other features of preferences since the rate of time preference and the elasticity of intertemporal substitution, as described respectively by β and the function u, do not play an important role in this argument. Of course, the risk exposure is not assessed directly in terms of consumption units C_{t+H}, but only through its concave transformation $u(C_{t+H})$. Yet, no genuinely perverse effect results form this concave scaling.

However, if one thinks about more general temporal gambles, it is no longer true that, as commonly believed, m and f will determine the degree of risk taking in portfolio choice problems. We argue that the disentangling of risk aversion and intertemporal substitution is not fully done in the recursive utility framework (72) and (73). To explain this intuitively, we will rely on the analysis of Alvarez and Jermann (2000), who establish a clear distinction between the related concepts of equity premium and cost of consumption uncertainty. The marginal cost of consumption uncertainty is defined, as we did above, from the time t return until maturity of an asset with a single risky payment C_{t+H} at $(t+H)$. However, the consumption equity premium (for an equity with dividends equal to consumption) is defined from the time t shadow price of an asset which pays the full stochastic process of dividends $[C] = [C_{t+h},\ h \geqslant 0]$. In order to control for preferences for the timing of uncertainty resolution, let us maintain the assumption that all uncertainty about this process is revealed at time $(t+1)$. Then the relevant utility index is

$$V_t = u(C_t) + \beta m \left[\sum_{h=1}^{\infty} \beta^{h-1} u(C_{t+h})\right]. \tag{75}$$

Assume for simplicity that $\beta = 1$ and that the stochastic process $[C_{t+h}, h > 0]$ is stationary and ergodic. Then, the higher the elasticity of intertemporal substitution featured by the function u is, the more the individual is able to consider the stochastic process $[C_{t+h}, h > 0]$ as almost equivalent to its smoothed counterpart $[C^*_{t+h}, h > 0]$ defined by

$$C^*_{t+h} = lim_{H=\infty} \left(\frac{1}{H}\sum_{j=1}^{H} C_{t+j}\right). \tag{76}$$

But, by the law of large numbers, the smoothed consumption process is no longer risky. In other words, a high elasticity of intertemporal substitution allows one to think in terms of intertemporal diversification and substantially lowers the level of risk which is significantly borne in a formula like (75). The argument could of course be easily extended to more realistic situations of consumption processes with trends and non zero rate of time preference.

This remark is of course highly relevant when it comes to solving the equity premium puzzle since it implies that m does not provide a meaningful assessment of the individual risk aversion. In other words, one cannot claim to have successfully solved the puzzle when a reasonable level of risk aversion (as described by m or f) is obtained in a representative agent model consistent with (72) and (73). It may only mean that risk aversion has been underestimated through its m (or f) characterization since the agent, with a sufficiently high elasticity of intertemporal substitution, might have perceived that the risk was not so high because of temporal diversification.

This possibility of temporal diversification explains that, as acknowledged by Duffie and Epstein (1992a, 1992b), the significance of the function m for comparative risk aversion arises only for a given elasticity of intertemporal substitution. Our argument goes further though. It would be illusory to rely on a plausible estimate of the risk aversion coefficient in an Epstein-Zin model of asset prices to consider that the equity premium puzzle has been solved. It all depends on the value of the elasticity of intertemporal substitution. The higher it is, the more spurious the inference will be. Following Alvarez and Jermann (2000), this amounts to confuse the equity premium and the cost of consumption uncertainty even though they are clearly distinct, both conceptually and quantitatively. They argue that the steepness of the term structure and the persistence of the shocks are two of the features that make the equity premium different from the marginal cost of consumption uncertainty.

We will propose in the next subsection an expected utility functional form which explicitly takes into account the degree of persistence of the shocks. For this reason, it is better able to disentangle risk aversion from the elasticity of intertemporal substitution. Actually, we derive an asset pricing model which is observationally equivalent to the one of Epstein and Zin (1989) but which modifies the definition of the risk aversion measurement in order to avoid the aforementioned shortcoming of recursive utility, that is the underestimation through m of the true level of risk aversion when the elasticity of intertemporal substitution is high.

Another advantage of our expected utility model is that it only refers to an individual who is neutral with respect to the timing of uncertainty resolution. Indeed, in addition to the temporal aspects of preferences, captured by $\beta$and the function u, and the risk aversion measure given by m, a third aspect of preferences should concern the timing of resolution of uncertainty. Actually, if in the above example we assume now that the risky consumption flow C_{t+H}, $H \geqslant 2$, is going to be revealed only at time $(t+2)$, we realize that the utility index at time t is different from (74). In other words, the definition of (β, u, m) characterizes the conditions under which early or late resolution is preferred. Thus, as recognized by Epstein and Zin (1989), this "latter aspect of preferences seems interwined with both substitutability and risk aversion". While they suspect that this "reflects the inherent inseparability of these three aspects of preference rather than a deficiency" of the framework, the new model proposed in the next section will give more support to the requirement of disentangling preferences for the timing of uncertainty resolution from substitutability and risk aversion. In contrast with Epstein and Zin (1989), this aspect of preferences does no longer seem implied by the comparison of disentangled levels of elasticity of intertemporal substitution and risk aversion. Therefore, one can envision a more general model which would not only disentangle risk aversion from intertemporal substitution, but also describe independently the timing of uncertainty resolution.

2.3 A Consumption CAPM with a Reference Level

In GRS (2002), we develop an intertemporal expected utility model where the representative agent derives utility from consumption measured relatively to a reference level and from this reference level itself:

$$V_t = \lambda(1-a)\sum_{h=o}^{\infty} \delta^h E_t\left[\left(\frac{C_{t+h}}{S_{t+h}}\right)^{1-a} S_{t+h}^{\lambda}\right], \tag{77}$$

where the reference level S_t is considered as external to the agent and E_t denotes a conditional expectation given the information at time t. Depending on the specification of the reference level and on the constraints imposed on the various preference parameters, we show in GRS (2002) that this model produces most of the SDFs that have been used in the empirical asset pricing literature. We will now see how it should be modeled to obtain a SDF which is observationally equivalent to the one derived by Epstein and Zin (1989).

Our argument rests essentially on the fact that the reference level provides a way to extend intertemporal choice of consumption without uncertainty to risky consumption streams. When no uncertainty prevails, the future sequence of the reference level at time t, S_{t+h}, $h \geqslant 0$, coincides with the optimal future consumption values:

$$S_{t+h} = C_{t+h} \text{ identically for } h \geqslant 0. \tag{78}$$

In a risky environment, we just generalize condition (78) in terms of conditional expectations:

$$E_t[S_{t+h}] = E_t[C_{t+h}] \text{ for all } h \geqslant 0. \tag{79}$$

Therefore, we can interpret S_{t+h} as the reference level the agent has in mind at time t to decide his risk-taking behavior. In the spirit of Abel (1990) and Campbell and Cochrane

(1999) models of external habit formation, some macroeconomic variables which belong to the agent's information set at time $(t+h)$ may affect the assessment of the reference level S_{t+h}. In the model of Barberis, Huang, and Santos (2001), when the representative agent's consumption C_{t+h} coincides in equilibrium with the $\overline{C}_{t+h}$ aggregate per capita consumption at time $(t+h)$ (viewed *as*exogenous to the investor), the reference level of consumption will aggregate the gain or loss the agent experiences on his financial investments between $(t+h-1)$ and $(t+h)$. In all these examples, the growth rate $\frac{S_{t+h}}{S_{t+h-1}}$ of benchmark consumption between dates $(t+h-1)$ and $(t+h)$ may include some information contemporaneous with C_{t+h}.[26]

Given condition (78), the parameter λ in (77) can unambiguously be interpreted in terms of intertemporal elasticity of substitution, with $\lambda = 1 - \frac{1}{\xi}$, where ξ denotes the agent's elasticity of intertemporal substitution. Since the reference level is viewed as external by the agent during his optimization, the resulting Euler conditions lead to a generalized CCAPM with the following SDF:

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-a} \left(\frac{S_{t+1}}{S_t} \right)^{a-\frac{1}{\xi}}. \tag{80}$$

Such a SDF implies that the definition of the reference level must produce conditional expectations that are not only constrained by (79), but also consistent with the observed asset prices.

Let us consider first the market portfolio pricing condition. If we denote by $R_{M,t+1}$ the gross return on the market portfolio observed at time $(t+1)$, we get

$$E_t \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-a} \left(\frac{S_{t+1}}{S_t} \right)^{a-\frac{1}{\xi}} R_{M,t+1} \right] = 1. \tag{81}$$

Condition (81) shows that covariation between the reference level and the market return may compensate for the lack of covariation between consumption and the market return. This extension of the traditional consumption-based asset pricing model may help to solve several asset pricing puzzles features associated with aggregate data. As stressed by Barberis, Huang, and Santos (2001), such an extension has some behavioral foundations since

[26]Campbell and Cochrane (1999) also specify that the consumption habit moves in response to current aggregate consumption and not, as in many habit formation models, in proportion to the last period consumption. Since habit is considered as external, the reference level S_{t+h} may even be defined as a function of C_{t+h}.

it captures the idea that the degree of loss aversion of the investor depends on his prior investment performance. To make even more explicit this tight relationship between the reference level and investment performance as measured by the market return, we will refer to a loglinearization of conditional moment restrictions (79) and (81) (see Epstein and Zin (1991) and Campbell (1993) for similar interpretations based on a loglinearization of the Euler equations). Conditional expectations are computed as if the vector

$$(\Delta c_{t+1}, \Delta s_{t+1}, r_{M,t+1}) = \left(log\left(\frac{C_{t+1}}{C_t}\right), log\left(\frac{S_{t+1}}{S_t}\right), logR_{M,t+1} \right) \tag{82}$$

were jointly normal and homoscedastic given the information available at time t. Conditions (79) and (81) at horizon 1 become:

$$E_t[\Delta c_{t+1}] - E_t[\Delta s_{t+1}] = \kappa_1, \tag{83}$$

$$-aE_t[\Delta c_{t+1}] + \left(a - \frac{1}{\xi}\right) E_t[\Delta s_{t+1}] + E_t[r_{M,t+1}] = \kappa_2 \tag{84}$$

for some constants $\kappa_1$and κ_2. Equivalently, these two restrictions say that both $[\Delta s_{t+1} - \Delta c_{t+1}]$ and $[\Delta s_{t+1} - \xi r_{M,t+1}]$ must be unpredictable at time t. We will now see that the Epstein and Zin (1989) pricing model is observationally equivalent to the particular case of our CCAPM with reference level where $[\Delta s_{t+1} - \xi r_{M,t+1}]$ is not only unpredictable but constant:

$$log\frac{S_{t+1}}{S_t} = \xi logR_{M,t+1} + \kappa, \tag{85}$$

for some constant κ. In other words, we consider the particular case where the benchmark growth rate of consumption is log-linearly determined by the currant value of the market return, with a slope parameter equal to elasticity of intertemporal substitution. Note that this is in accordance with the portfolio separation property generally implied by homotheticity of preferences (see Epstein and Zin (1989)), whereby optimal consumption is determined in a second stage, after the portfolio choice has been made.

Given the specification (85) of the reference level, it is clear that the parameters δ and κ cannot be separately identified from this SDF only. We will therefore reparametrize it in the following way:

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t}\right)^{-a} R_{M,t+1}^{a\xi - 1}. \tag{86}$$

At first sight, we obtain a SDF which is observationally equivalent to the one derived by Epstein and Zin (1989) with the TAS aggregator, with some utility function u and a von Neumann-Morgenstern utility index f which are both isoelastic. Yet there are several important differences in the interpretation of the two SDFs.

Let us start with the market return which enters both SDF specifications. In our model, it appears because the investor links the benchmark consumption to the market return. In the recursive utility framework, it appears because the investor cares about the timing of uncertainty resolution. Actually, in (77), the utility index is defined in terms of conditional expectations of future random variables given the information available at time t, and therefore, the investor appears to be neutral with respect to the timing of uncertainty resolution. In this respect, our approach is closer to Bakshi and Chen (1996) who put forward the hypothesis that investors accumulate wealth not only for the sake of consumption but also for wealth-induced social status. Typically, if the reference level S_t were equal to aggregate wealth, a non-zero difference between λ and $(1-a)$ would lead to Model 1 of Bakshi and Chen (1996) where absolute wealth is status. This explains why Bakshi and Chen (1996) also put forward a kind of observational equivalence between their model and Epstein and Zin (1989). However, our approach does not reduce to theirs because they implicitly consider that the rate of growth of aggregate wealth coincides with the market return, which is not true in general. They differ because the share of wealth invested is not constant.[27] On the contrary, Lettau and Ludvigson (2001) have emphasized the prominent role played by the consumption-wealth ratio as a state variable to summarize the relevant conditioning information.

Actually, our model is better understood by reference to the habit formation literature. Our agent derives utility both from the level of consumption relative to the state variable S_t and from the absolute value of this reference level which is similar to a habit. This specification extends to asset pricing applications the one recently used in a saving and growth model by Carroll, Overland, and Weil (2000). In the spirit of the habit forma-

[27]Bakshi and Chen (1996) can be interpreted as a particular case of our model with a unit elasticity of intertemporal substitution. Smith (2001) proposes to extend Bakshi and Chen (1996) by taking into account both the concern about wealth-induced status and the attitude towards the timing of uncertainty resolution. However, it is a simple i.i.d. economy in which the stochastic variation in the invested share of wealth cannot be accommodated.

tion literature, the coefficient a is then interpreted as the risk aversion coefficient. This interpretation immediately raises the following question. Since the SDF (86) is observationally equivalent to the one of Epstein and Zin (1989) with isoelastic functions u and f $\left(u(C) = \left(\frac{1}{\rho}\right)(C^{\rho} - 1) \text{ and } f(V) = \left(\frac{1}{\alpha}\right)(V^{\alpha} - 1)\right)$, it should shed some light on the difficult issue of risk aversion assessment in the context of the recursive utility model of Epstein and Zin (1989). Actually, the exponent of $\frac{C_{t+1}}{C_t}$ in this model (see equation (6.6) p. 958) is

$$\alpha\frac{(\rho - 1)}{\rho} = \frac{\alpha}{1 - \sigma}, \text{ since } \sigma = (1 - \rho)^{-1}. \tag{87}$$

By identification of (86) and (87), we deduce that the quantity $(1 - \alpha)$, instead of being interpreted as a risk aversion parameter, should be seen as

$$(1 - \alpha) = a + 1 - a\sigma. \tag{88}$$

Several comments are in order. First, the ability of the recursive utility model to disentangle risk aversion and intertemporal substitution is questionable. Actually, it is only in the standard expected utility model case, when σ is the inverse of the risk aversion parameter a, that $(1 - \alpha)$ can be interpreted as a risk aversion parameter. Even more problematic is the fact that $(1 - \alpha)$ becomes negative whenever σ is greater than $\frac{1}{a} + 1$. Note that this lack of disentangling manifests itself even without resorting to our interpretation of a as a risk aversion parameter. The natural requirement of a negative exponent for $\frac{C_{t+1}}{C_t}$ in the SDF implies that the alleged risk aversion parameter $(1 - \alpha)$ and $\frac{1}{\sigma}$ should be on the same side of 1.

Second, as soon as σ is greater than $\frac{1}{a}$, the alleged risk aversion parameter $(1 - \alpha)$ underestimates the genuine risk aversion parameter a. Hence, as noted before in Section 2.2, a relatively high level of elasticity of intertemporal substitution may spuriously indicate a moderate risk aversion. If, as documented by Mehra and Prescott (1985), the model can replicate the equity risk premium only for a high level of risk aversion, say $a = 20$, even a moderate elasticity of substitution, say .8, will dramatically lower the perceived risk aversion in the recursive utility model: $(1 - \alpha) = 5$.

Of course, expressing concerns about the recursive utility model does not imply that the alternative model we propose is valid. For tests of its empirical validity, we refer the reader to GRS (2002).[28] However, it is important to stress that, apart from the issues related

[28]We estimate in particular a model where the reference level growth rate is determined both by past

to the interpretation of risk aversion and the attitude towards the timing of uncertainty resolution, the two models are mutually consistent. Indeed, taking (88) into account, we can rewrite the exponent of the market return in the Epstein-Zin SDF as

$$\left(\frac{\alpha}{\rho}\right) - 1 = a\sigma - 1. \tag{89}$$

By identification with (86), we see that our SDF is nothing but a reparametrization of the Epstein-Zin's one with: $\xi = \sigma$ (and $\lambda = \rho$). This correspondence between the parameters of intertemporal substitution of the two models is fully consistent with the interpretation sketched above. It also sheds some interesting light on the issue of myopic portfolio choice. Giovannini and Weil (1989) have stressed that a unit elasticity of intertemporal substitution implies a form of (rational) myopia in consumption and savings decisions but not in portfolio allocation. Actually, a unit elasticity of intertemporal substitution ($\lambda = 0$) reduces our general SDF (80) to

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t}\right)^{-a} \left(\frac{S_{t+1}}{S_t}\right)^{a-1} \tag{90}$$

and, if one admits the log-linearization (85):[29]

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t}\right)^{-a} R_{M,t+1}^{a-1}. \tag{91}$$

This formula can be seen as the explicit solution of equation (B.5) in Giovannini and Weil (1989) for the particular case of conditional log-normality. Except for logarithmic risk preferences ($a = 0$ for us and $\alpha = 0$ for them), the Euler equations for portfolio choice with $\sigma = 1$correspond in general to neither the static CAPM nor the CCAPM. Giovannini and Weil (1989) develop their argument rigorously but it is hard to do in the orthodox approach of the EZ SDF since one gets the spurious feeling that the exponent of consumption growth in the SDF $\left(\frac{\alpha}{1-\sigma}\right)$ is well defined and non zero in the limit cases only if one maintains the equivalence: $\sigma = 1 \iff \alpha = 0$. This is the equivalence between myopic consumption-saving decisions and logarithmic risk preferences. This issue is relevant empirically. For instance,

consumption growth rates (as in habit formation models) and by the return of the market portfolio (as in the Kreps-Porteus specification of the recursive utility model of Epstein and Zin (1989)). The parameters of this specification are economically plausible and estimated with precision.

[29]With or without this log-linearization, this result is fully consistent with the point made by Giovannini and Weil (1989).

Epstein and Zin (1989) conclude that risk preferences do not differ statistically from the logarithmic specification but, by the same token, they find estimates for the elasticity of intertemporal substitution that are not statistically different from 1 for the two first sets of instruments (the ones which, according to the authors, give the most sensible results). Another interesting case is the portfolio applications in Campbell and Viceira (2002). To be able to disentangle myopia in consumption -saving decisions from myopia in portfolio allocation, they set $\sigma = 1$ while allowing any value for the risk aversion parameter. While they are certainly right to do so for economic interpretations, this is strictly speaking not consistent with the EZ parametrization, as it can be seen not only through our interpretation but also through the Appendix B of Giovannini and Weil (1989). By contrast, the exponent a in our SDF (90, 91) is in no way restricted by the condition $\sigma = 1$.

To summarize, the only difference between our approach and the Epstein and Zin (1989) recursive utility model concerns the incorporation of the preferences for intertemporal choice without uncertainty in a risky environment with constant relative risk aversion. While the recursive utility approach replaces the future random utility index by its certainty equivalent, we believe it is preferable to replace upstream the future consumption flows by an external benchmark produced by the first-order conditions for optimal consumption. This benchmark determines the role of the time preference parameters while the risk aversion parameter matters only insofar as uncertainty prevents the agent from meeting his benchmark.

Of course, our argument rests upon some approximations due to a log-linearization of first-order conditions for consumption and neglects volatility predictability. While extensions can be envisioned in this regard, our new insight on risk aversion assessment in the recursive utility model is useful for addressing asset pricing puzzles. In terms of risk premium for individual assets, log-linearization of the pricing equations resulting from our SDF gives

$$E_t[r_{i,t+1}] - r_{f,t+1} = a\sigma_{ic} - (a\sigma - 1)\sigma_{im}, \tag{92}$$

where σ_{ic} and σ_{im} denote the covariances of asset i returns with consumption growth and market returns respectively. This asset pricing model is observationally equivalent to the one of Epstein and Zin (1989) but the interpretation of the coefficients and their orders of magnitude deemed to be reasonable differ. The coefficient of σ_{ic} should be interpreted as a risk aversion parameter, which means, in particular, that it is constrained to be nonnegative. Following the recursive utility parametrization, this would not be the case since $a = (1-\sigma)^{-1}(a^* - 1)$, where $a^* = (1-\alpha)$ is the risk aversion measure. In addition, the coefficient of σ_{im} is $(a\sigma - 1) = (1-\sigma)^{-1}(a^*\sigma - 1)$, which can take very large values for seemingly realistic values of the coefficient a^*.

2.4 Concluding Remarks

In this paper, we have proposed a generalized expected utility framework which disentangles risk aversion and intertemporal substitution in an alternative way to the recursive utility framework proposed by Epstein and Zin (1989). Although observationally equivalent, the two models may lead to significantly different conclusions regarding the well-documented asset pricing puzzles. In particular, a plausible value of the elasticity of intertemporal substitution smaller than one, but not too close to zero, can conceal, within a recursive utility framework, a very large implied value for risk aversion.

One of the advantages of our specification is its flexibility. In GRS (2002), we show that it can reproduce most of the SDFs that have been proposed in the empirical asset pricing literature. In particular, it covers all habit formation approaches and can be seen as a generalization of Campbell and Cochrane (1999). As in the latter, a key assumption is to assume that the reference level is exogenous to the agent. When the growth rate of the reference level is made a function of the return on the market portfolio, we obtain a SDF which is observationally equivalent to the Kreps and Porteus certainty equivalent in the recursive utility framework. Other specifications of the certainty equivalent, in particular disappointment aversion, can also be accommodated in our framework given the right specification of the reference level (see GRS (2002)). The simplicity of our expected utility approach makes it a serious contender for the more involved recursive utility specifications. Moreover, it allows to specify new SDFs which can potentially better explain asset prices. In GRS (2002), we propose a new SDF based on habit persistence and the return on the market portfolio which appears to be supported by the data.

Our generalized expected utility framework basically maintains the assumption of investor neutrality with regard to the timing of uncertainty resolution. Yet, as emphasized by Kreps and Porteus (1979), temporal preference for consumption is only an induced preference. Is earlier resolution of uncertainty better simply because it permits an adaptive choice of the individual activities or do individual preferences for consumption streams include a genuine subjective preference for earlier or later uncertainty resolution? A more general equilibrium model could then justify embedding our Von Neumann-Morgenstern utility with respect to a reference level into a recursive framework with a clear formulation of the timing of outcomes of lotteries and resulting actions taken by the agent. Such a model might provide an answer to the question raised by Epstein and Zin (1989): is there some inherent inseparability of the three aspects of preferences: risk aversion, intertemporal substitution, and concern for the timing of uncertainty resolution? This paper might have provided a first step in finding an answer by a better disentangling of risk aversion from intertemporal substitution without any implication about preference for earlier or later resolution of uncertainty.

References

[1] Abel, A., 1990, Asset Prices under Habit Formation and Catching Up with the Joneses, The American Economic Review 80, Papers and Proceedings, 38-42.

[2] Abel, A., 1996, Risk Premia and Term Premia in General Equilibrium, unpublished paper, University of Pennsylvania.

[3] Alvarez, F. and U.J. Jermann, 2000, Using Asset Prices to Measure the Cost of Business Cycles, NBER Working Paper 7978.

[4] Aschauer, D., 1985, Fiscal Policy and Aggregate Demand, American Economic Review 75, 117-127.

[5] Bakshi, G. and Z. Chen, 1996, The Spirit of Capitalism and Stock-Market Prices, American Economic Review 86, 133-157.

[6] Barberis, N., M. Huang, and T. Santos, 2001, Prospect Theory and Asset Prices, Quarterly Journal of Economics 116, 1-53.

[7] Becker, R. and J. Boyd III, 1997, Capital Theory, Equilibrium Analysis and Recursive Utility, Blackwell Publishers.

[8] Campbell, J., 1993, Intertemporal Asset Pricing without Consumption Data, The American Economic Review 83, 487-512.

[9] Campbell, J. and J. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, Journal of Political Economy 107, 205-251.

[10] Campbell, J. and L. Viceira, 2002, Strategic Asset Allocation, Portfolio Choice for Long-Term Investors, Clarendon Lectures in Economics, Oxford University Press.

[11] Carroll, C., J. Overland, and D. Weil, 2000, Saving and Growth with Habit Formation, The American Economic Review 90, 341-355.

[12] Constantinides, G., 1990, Habit Formation: A Resolution of the Equity Premium Puzzle, Journal of Political Economy 98, 519-543.

[13] Duffie, D. and L.G. Epstein, 1992a, Stochastic Differential Utility, Econometrica 60, 353-394.

[14] Duffie, D. and L.G. Epstein, 1992b, Asset Pricing with Stochastic Differential Utility, The Review of Financial Studies 5, 411-436.

[15] Eichenbaum, M., L. Hansen, and K. Singleton, 1988, A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty, Quarterly Journal of Economics 103, 51-78.

[16] Epstein, L. and S. Zin, 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica 57, 937-969.

[17] Epstein, L. and S. Zin, 1991, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns I: An Empirical Analysis, Journal of Political Economy 99, 263-286.

[18] Garcia, R., E. Renault, and A. Semenov, 2002, A Consumption Capital Asset Pricing Model with A Reference Level, working paper, CIRANO and Université de Montréal.

[19] Giovannini, A. and P. Weil, 1989, Risk Aversion and Intertemporal Substitution in the Capital Asset Pricing Model, NBER Working paper No. 2824, National Bureau of Economic Research, Cambridge, MA.

[20] Guvenen, F., 2002, Habit Formation or Cross-Sectional Heterogeneity?, working paper, University of Rochester.

[21] Hansen, L. and K. Singleton, 1983, Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns, Journal of Political Economy 91, 249-265.

[22] Koopmans, T., 1960, Stationary Ordinal Utility and Impatience, Econometrica 28, 287-309.

[23] Kreps, D. and E. Porteus, 1978, Temporal Resolution of Uncertainty and Dynamic Choice Theory, Econometrica 46, 185-200.

[24] Kreps, D. and E. Porteus, 1979, Temporal von Neumann-Morgenstern and Induced Preferences, Journal of Economic Theory 20, 81-109.

[25] Lettau, M. and S. Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, Journal of Political Economy 109, 1238-1287.

[26] Lucas, R. and N. Stokey, 1984, Optimal Growth with Many Consumers, Journal of Economic Theory 32, 139-171.

[27] Mehra, R. and E. Prescott, 1985, The Equity Premium: A Puzzle, Journal of Monetary Economics 15, 145-162.

[28] Smith, W.T., 2001, How Does the Spirit of Capitalism Affect Stock Market Prices, The Review of Financial Studies 14, 1215-1232.

[29] Startz, R., 1989, The Stochastic Behavior of Durable and Non-Durable Consumption, Review of Economics and Statistics 71, 356-363.

[30] Sundaresan, S., 1989, Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth, Review of Financial Studies 2, 73-88.

[31] Weil, P., 1989, The Equity Premium Puzzle and the Risk-Free Rate Puzzle, Journal of Monetary Economics 24, 401-421.

[32] Weil, P., 1990, Non-Expected Utility in Macroeconomics, Quarterly Journal of Economics 105, 29-42.

3 Asset Pricing Puzzles, High-Order Consumption Moments, and Heterogeneous Consumers

3.1 Introduction

Numerous studies over the past two decades have focused on a representative-agent consumption based asset pricing model that treats asset prices as being determined by the consumption and savings decisions of a single representative agent assumed to have time-separable power utility. A number of empirical investigations show that the representative-agent model is apparently inconsistent with the data on consumption and asset returns in many respects. Thus, a reasonably parametrized representative-agent model generates an average equity premium which is too low (the equity premium puzzle) and a risk-free rate which is too high (the risk-free rate puzzle) compared to the observed values. One plausible response to the equity premium and risk-free rate puzzles is to argue that the poor empirical performance of the representative-agent model is due to the fact that the model abstracts from limited participation of consumers in asset markets and from incomplete consumption insurance.

Full consumption insurance implies that heterogeneous consumers can use financial markets to diversify away any idiosyncratic differences in their consumption streams. According to the first-order condition of a single investor's intertemporal consumption and portfolio choice problem, under complete consumption insurance, intertemporal marginal rates of substitution are identical across individuals. Moreover, the first-order condition for excess returns implies that agents are able to equalize their marginal utilities as well. If all investors in the economy have the same utility function (it is common to assume an agent to be risk-averse and, therefore, his preferences to be adequately presented by some increasing and strictly concave utility function), with complete consumption insurance they are able to equalize, state by state, their consumption. It follows that under the assumption of full consumption insurance, aggregate consumption per capita can be used in place of individual consumption and, hence, the pricing implications of a heterogeneous-consumer model are similar to those of the representative-consumer economy. With incomplete consumption insurance, individuals are not able to self-insure against uninsurable risks and, therefore, are heterogeneous. Intuitively, incomplete consumption insurance seems to be a very plausible assumption given the lack of certain types of insurance, such as insurance against the idiosyncratic shocks to the households' income, for example. Limited capital market participation can be viewed as another special form of incomplete markets when one group of people is prevented from trading in capital markets and simply consumes its labor income.

Bertaut (1998), Blume and Zeldes (1993), Haliassos and Bertaut (1995), and Mankiw and Zeldes (1991) observe that only about 30-40% of US individuals hold stocks either directly or through defined contribution pension funds.[30] Using data on food consumption from the Panel Study of Income Dynamics (PSID), Mankiw and Zeldes (1991) find that the consumption of stockholders is more volatile and more highly correlated with the stock market premium as compared with that of nonstockholders. They estimate Euler equations of per capita consumption for stockholders and nonstockholders separately and find that the estimate of risk aversion decreases as the threshold value in the definition of stockholders is raised. Brav and Géczy (1995), Brav, Constantinides, and Géczy (2002), hereafter BCG, and Vissing-Jorgensen (1998) use total consumption of nondurables and services reconstructed from the Consumer Expenditure Survey (CEX) and also find evidence for an estimate of risk aversion which is decreasing in asset holdings. This suggests limited asset market participation as a plausible explanation for the empirical failures of the representative-agent model. Nevertheless, the empirical results in BCG (2002), Mankiw and Zeldes (1991), and Vissing-Jorgensen (1998) show that even when the wealthiest group of assetholders is considered, relatively high risk aversion is needed to explain the excess market portfolio return. This provides some evidence that although limited asset market participation helps to explain the size of the market premium, it is insufficient to resolve the equity premium puzzle when taken alone.

The potential for the incomplete market model to explain the equilibrium behavior of stock and bond returns, both in terms of the level of equilibrium rates and the discrepancy between equity and bond returns, was first suggested by Mehra and Prescott (1985). Weil (1992) studies a two-period model in which consumers face, in addition to aggregate div-

[30]According to the Current Population Reports, only about 20% of US households hold publicly traded stocks and/or mutual fund shares (about 20% of the US population owned such assets in 1984, 21.8% in 1988, and 20.7% in 1991).

idend risk, idiosyncratic and undiversifiable labor income risk. He shows that decreasing absolute risk aversion and decreasing absolute prudence are sufficient to guarantee that the model predicts a smaller bond return and a larger equity premium than a representative agent model calibrated on the basis of aggregate data solely.

In the infinite horizon setting, individuals are able to make risk-free loans to one another and borrow to buffer any short-lived jump in their consumption. This reasoning suggests that in the infinite horizon economy, the additional demand for savings induced by the market incompleteness will generally be smaller than that in a two-period model. Hence, the absence of market completeness may have little impact on interest rates. Aiyagari and Gertler (1991), Bewley (1982), Heaton and D. Lucas (1992, 1995, 1996), Huggett (1993), Lucas (1994), Mankiw (1986), and Telmer (1993) confirm this intuition.[31]

In contrast to earlier work which assumes that the idiosyncratic income shocks are transitory and homoscedastic, Constantinides and Duffie (1996), henceforth CD, model the time-series process of each consumer's ratio of labor income to aggregate income as nonstationary and heteroscedastic. Given the joint process of arbitrage-free asset prices, dividends, and aggregate income satisfying a certain joint restriction, CD (1996) show that in the equilibrium of an economy with heterogeneity in the form of uninsurable, persistent, and heteroscedastic labor income shocks, the pricing kernel is a function not only of per capita consumption growth, but also of the cross-sectional variance of the logarithmic individual consumption growth rate. One of the key features of the CD (1996) model is that idiosyncratic shocks to labor income must be persistent. However, Heaton and D. Lucas (1996) and Storesletten, Telmer, and Yaron (1997) use data from the PSID and show that the conclusion whether labor income shocks are persistent or not depends on auxiliary modelling assumptions. BCG (2002) test empirically the CD (1996) pricing kernel using the CEX database and find that this SDF fails to explain the equity premium.[32]

[31]In the Bewley (1982), Lucas (1994), Mankiw (1986), and Telmer (1993) models, consumers face uninsurable income risk and borrowing or short-selling constraints, whereas Aiyagari and Gertler (1991) and Heaton and D. Lucas (1992, 1995, 1996) calibrate an economy in which consumers face uninsurable income risk and transaction or borrowing costs. Aiyagari and Gertler (1991) and Heaton and D. Lucas (1992, 1995, 1996) show that the pricing implications of an incomplete market model do not differ substantially from those of a representative-consumer model, unless the ratio of the net supply of bonds to aggregate income is restricted to an unrealistically low level.

[32]Balduzzi and Yao (2000) derive a SDF which differs from the CD (1996) pricing kernel in that the second

Jacobs and Wang (2001) investigate a pricing kernel linear in the cross-sectional mean and variance of consumption growth without specifying implicitly any utility function. They find that although both these pricing factors are almost always significantly estimated, the sign estimated for the cross-sectional variance is sensitive to the chosen measure of consumption and set of households.

Jacobs (1999) investigates the importance of consumer heterogeneity by testing Euler equations involving household-level consumption data from the PSID and does not find evidence that the equity premium and risk-free rate puzzles result from aggregation problems. His result is that the Euler equation involving the risk-free asset is strongly rejected by the data, while equilibrium restrictions pertaining to the risky asset are not rejected for certain instrument sets. When the Euler equations for the risk-free asset and for the risky asset are estimated jointly, the model is strongly rejected statistically. This result is similar to that obtained by Hansen and Singleton (1982, 1983, 1984) within the representative agent framework.

BCG (2002) find empirical evidence for the importance of the skewness of the cross-sectional distribution of the individual consumption growth rate, combined with the mean and variance, in explaining the equity premium. Specifically, their calibration result is that the SDF, given by a third-order Taylor series expansion to the equal-weighted average of the household's intertemporal marginal rate of substitution (IMRS), explains the premium of the market portfolio return over the risk-free rate with low and economically plausible (between two and four) value of the RRA coefficient. This result is at odds with that in Cogley (2002). Cogley (2002) uses a Taylor series expansion to the individual's IMRS and develops an equilibrium factor model in which the pricing factors for the equity premium are the cross-moments of the excess market portfolio return with the first three moments of the cross-sectional distribution of log consumption growth. He finds that this model is not able to explain the observed mean equity premium with economically plausible value of risk aversion even when the model includes the first three cross-sectional moments of log

pricing factor is the difference of the cross-sectional variance of log consumption and not the cross-sectional variance of the log consumption growth rate. Although this SDF specification allows to explain the equity premium with a value of the relative risk aversion (RRA) coefficient which is substantially lower than that obtained using the conventional power utility model, the value of risk aversion needed to explain the equity premium remains rather high (larger than 9).

consumption growth.[33]

In this paper, we use a different approach to assess the importance of incomplete consumption insurance for explaining the equity premium and the risk-free rate of return. Following Mankiw (1986) and Dittmar (2002), we use a Taylor series expansion to an unknown marginal utility function. A distinctive feature of our approach is that we take a Taylor series expansion to the individual's marginal utility of consumption around the conditional expectation of consumption and not around the unconditional expectation as in Mankiw (1986) and Dittmar (2002). Using a Taylor series expansion to the individual's marginal utility of consumption around the conditional expectation of consumption allows us to derive an approximate equilibrium model for expected returns in which the priced risk factors are the cross-moments of return with the moments of the cross-sectional distribution of individual consumption. The attractiveness of this approach comes from the possibility of avoiding an ad hoc specification of preferences and considering a general class of utility functions when addressing the question of the sign of the effect of a particular moment of the cross-sectional distribution of individual consumption on the expected excess market portfolio return and risk-free interest rate, while an ad hoc specification of the utility function is necessary when taking a Taylor series expansion to the agent's IMRS or to the mean of the individual's IMRS.[34]

We show that if preferences exhibit decreasing and convex absolute prudence, then the cross-sectional mean and skewness of individual consumption help to explain the equity premium if their cross-moments with the excess market portfolio return are positive, while the cross-sectional variance and kurtosis of individual consumption always lower the equity premium explained by the model.

In our empirical investigation, we use several approaches to assess the plausibility of the approximate equilibrium model for reproducing different features of expected returns.

[33]With low (below 5) values of the RRA coefficient, the model can explain only about one-fourth of the observed mean equity premium.

[34]All we need to know to answer the question whether considering a particular moment of the cross-sectional distribution of individual consumption generates a smaller or, on the contrary, larger predicted asset return, is the sign of its cross-moment with return and the sign of the corresponding derivative of the utility function. Considering some special form of preferences is, however, necessary when assessing the size of that effect.

First, we perform a Hansen-Jagannathan volatility bound analysis. In this part, we assess the plausibility of the SDF for the market premium by studying the mean and standard deviation of the pricing kernel for different values of the risk aversion coefficient. Second, a calibration exercise is done. The purpose of this part is to test whether the observed mean equity premium can be explained with an economically plausible value of the RRA coefficient. The third part provides an empirical investigation of the conditional version of the Euler equations for the equity premium and the risk-free rate using a non-linear generalized method of moments (GMM) estimation approach. In this part, we exploit information about time-series properties of consumption and asset returns. In each of the three parts, the empirical analysis is performed under the assumption of limited asset market participation.[35]

The rest of the paper proceeds as follows. In Section 3.2, we derive an approximate equilibrium model for the expected equity premium and risk-free rate using a general class of utility functions. Section 3.3 describes the data and presents the empirical results under the CRRA preferences. Section 3.4 concludes.

3.2 An Approximate Equilibrium Asset Pricing Model

Consider an economy in which an agent maximizes expected lifetime discounted utility:

$$E_t\left[\sum_{j=0}^{\infty}\delta^j u\left(C_{k,t+j}\right)\right]. \tag{93}$$

In (93), δ is the time discount factor, $C_{k,t}$ is the individual k's consumption in period t, $u(\cdot)$ is a single-period von Neumann-Morgenstern utility function, and $E_t[\cdot]$ denotes an expectation which is conditional on the period-t information set, Ω_t, that is common to all agents.[36]

Let us consider a set of agents, $k = 1, ..., K$, that participate in asset markets. In equilibrium, the investor k's optimal consumption profile must satisfy the following first-

[35] To examine the asset pricing implications of the hypothesis of limited participation of consumers in asset markets, we consider different sets of households defined as assetholders according to a criterion of asset holdings above a certain threshold value ranging from \$2 to \$20000.

[36] We assume $u(\cdot)$ to be increasing, strictly concave, and differentiable.

order condition:

$$\delta E_t \left[u' \left(C_{k,t+1} \right) R_{i,t+1} \right] = u' \left(C_{k,t} \right), \; k = 1, ..., K, \; i = 1, ..., I. \tag{94}$$

The right-hand side of (94) is the marginal utility cost of decreasing consumption by $dC_{k,t}$ in period t. The left-hand side is the increase in expected utility in period $t+1$ which results from investing $dC_{k,t}$ in asset i in period t and consuming the proceeds in period $t+1$. $R_{i,t+1}$ is the simple gross return on asset i and I is the number of traded securities.

For the excess return on asset i over some reference asset j, equation (94) can be rewritten as

$$E_t \left[u' \left(C_{k,t+1} \right) \left(R_{i,t+1} - R_{j,t+1} \right) \right] = 0, \; k = 1, ..., K, \; i = 1, ..., I. \tag{95}$$

Assume that $u(\cdot)$ is $N+1$ times differentiable and take a N-order Taylor series expansion to the individual k's marginal utility around the conditional expectation of consumption, $h_{t+1} \equiv E_t \left[C_{k,t+1} \right]$, assumed to be the same for all investors:

$$u' \left(C_{k,t} \right) = \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_t \right) \left(C_{k,t} - h_t \right)^n, \; k = 1, ..., K.^{37} \tag{96}$$

Substituting (96) into (94) yields

$$\delta \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_{t+1} \right) \times E_t \left[\left(C_{k,t+1} - h_{t+1} \right)^n R_{i,t+1} \right] = \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_t \right) \left(C_{k,t} - h_t \right)^n, \tag{97}$$

$k = 1, ..., K, \; i = 1, ..., I.$

We can rearrange (97) to explicitly determine expected asset returns:

$$\begin{aligned} E_t \left[R_{i,t+1} \right] = {} & \delta^{-1} \sum_{n=0}^{N} \frac{1}{n!} \frac{u^{(n+1)} \left(h_t \right)}{u' \left(h_{t+1} \right)} \left(C_{k,t} - h_t \right)^n \\ & - \sum_{n=1}^{N} \frac{1}{n!} \frac{u^{(n+1)} \left(h_{t+1} \right)}{u' \left(h_{t+1} \right)} \times E_t \left[\left(C_{k,t+1} - h_{t+1} \right)^n R_{i,t+1} \right], \end{aligned} \tag{98}$$

$k = 1, ..., K, \; i = 1, ..., I.$

These equations can now be summed over investors and then divided by the number of investors in the economy to yield the following approximate relationship between expected

[37] Here, and throughout the paper, $u^{(n)}(\cdot)$ denotes the nth derivative of $u(\cdot)$. Mankiw (1986) limits his analysis to a second-order Taylor approximation to the agent's marginal utility of consumption ($N = 2$).

asset returns and priced risk factors:

$$E_t\left[R_{i,t+1}\right] = \delta^{-1}\sum_{n=0}^{N}\frac{1}{n!}\frac{u^{(n+1)}\left(h_t\right)}{u'\left(h_{t+1}\right)}Z_{n,t} - \sum_{n=1}^{N}\frac{1}{n!}\frac{u^{(n+1)}\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z_{n,t+1}R_{i,t+1}\right], \qquad (99)$$

$i = 1, ..., I$, where $Z_{n,t} \equiv \frac{1}{K}\sum_{k=1}^{K}\left(C_{k,t} - h_t\right)^n$.[38]

This is an approximate equilibrium asset pricing model in which the priced risk factors are the cross-moments of return with the moments of the cross-sectional distribution of individual consumption.

The multifactor pricing model (99) can be seen as an attractive alternative to the multifactor models based on the Arbitrage Pricing Theory (APT). The first attractive feature of model (99) is that, in contrast to the APT which does not provide the identification of the risk factors, the set of factors and the form of the pricing kernel obtain endogenously from the first-order condition of a single investor's intertemporal consumption and portfolio choice problem. That allows to avoid some serious problems arising from an ad hoc specification of a factor structure.[39] First, choosing factors without regard to economic theory may lead to overfitting the data. The second potential danger is the lack of power of tests which ignore the theoretical restrictions implied by a structural equilibrium model. Another attractive feature of model (99) is that the signs of the risk factor coefficients are driven by preference assumptions, while they are unrestricted in the multifactor models based on the APT. The problem with both the multifactor pricing model (99) and the multifactor models based on the APT approach is the unknown number of risk factors. In the case of model (99), this problem translates into deciding at which point to truncate the Taylor series expansion. This issue is explored in Section 3.3.3.

As previously mentioned, the conclusion about the role of consumer heterogeneity in explaining asset returns depends on the assumed degree of shock persistence. A virtue of our approach is that it dos not need to make any assumption about shock persistence. That differs our approach from those by Aiyagari and Gertler (1991), Bewley (1982), Heaton and

[38] A major problem with testing equation (98) directly is the observation error in reported individual consumption. Averaging over investors seems to mitigate the measurement error effect. However, it is quite plausible that the observation error in individual consumption makes it difficult to precisely estimate the cross-moments of return with the high-order moments of the cross-sectional distribution of individual consumption. An issue of the measurement error effect will be addressed in Section 3.2.3.

[39] See Campbell, Lo, and MacKinlay (1997).

D. Lucas (1992, 1995, 1996), Huggett (1993), Lucas (1994), Mankiw (1986), Telmer (1993), and CD (1996).

3.2.1 Uninsurable Background Risk and the Equity Premium Puzzle

The intuition that relaxing the assumption of complete consumption insurance has the potential for explaining the equity premium puzzle is due to recognizing the fact that in the real world, consumers face, in addition to the risk associated with the portfolio choice, multiple uninsurable and idiosyncratic risks such as loss of employment or divorce, for example.

Mankiw (1986), for instance, argues that if aggregate shocks to consumption are not dispersed equally across all consumers, then the level of the equity premium is in part attributable to the distribution of aggregate shocks among the population. Specifically, he takes a second-order Taylor series expansion to the agent's marginal utility around the unconditional expectation of consumption which is assumed to be the same for all individuals and shows that the expected excess return on the market portfolio over the return on the risk-free asset depends on the cross-moment of the equity premium with the cross-sectional variance of individual consumption.

If risks are substitutes, then the presence of an exogenous risk should reduce the demand for any other independent risk.[40] Nevertheless, the presence of one undesirable risk can make another undesirable risk desirable. This is the case of complementarity in independent risks.[41] Whether risks are substitutes or complements may depend on their nature. It is possible that the effect of adding one risk to another one is mixed and it is rather a question of which effect, substitutability or complementarity in independent risks, is dominating.

Weil (1992) demonstrates that if consumers exhibit decreasing absolute risk aversion and decreasing absolute prudence (i.e., the absolute level of precautionary savings declines as wealth rises), then neglecting the existence of the undiversifiable labor income risk leads to an underprediction of the magnitude of the equity premium.[42] Gollier (2001) shows

[40]See Gollier and Pratt (1996), Kimball (1993), Pratt and Zeckhauser (1987), and Samuelson (1963).

[41]See Ross (1999).

[42]If consumers' tastes exhibit decreasing absolute risk aversion and decreasing absolute prudence, then the nonavailability of insurance against an additional idiosyncratic and undiversifiable labor income risk makes consumers more unwilling to bear aggregate dividend risk and the equilibrium return premium on equity

that if absolute risk aversion is decreasing and convex and/or absolute risk aversion and absolute prudence are decreasing, the presence of the uninsurable background risk in wealth raises the aversion of a decision maker to any other independent risk. If at least one of these sufficient conditions is satisfied, such preferences preserve substitutability in the uninsurable background risk in wealth and the portfolio risk and, therefore, can help in solving the equity premium puzzle. Since decreasing and convex absolute risk aversion and decreasing absolute prudence are widely recognized as realistic assumptions, these results are overwhelmingly in favor of substitutability in the uninsurable background risk in wealth and the portfolio risk. It follows that the presence of the uninsurable background risk in wealth should reduce the agent's optimal demand for any risky asset and, therefore, should increase the expected equity premium.

For the excess return on the market portfolio over the risk-free rate, $RP_{t+1} \equiv R_{M,t+1} - R_{f,t+1}$, equation (99) reduces to

$$E_t\left[RP_{t+1}\right] = -\sum_{n=1}^{N} \frac{1}{n!} \frac{u^{(n+1)}\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z_{n,t+1} RP_{t+1}\right]. \tag{100}$$

Another way to represent this equation is to rewrite it in terms of the deviations of individual consumption from per capita consumption and the cross-moments of excess return with per capita consumption. In particular, when a second-order Taylor approximation to marginal utility ($N = 2$) is taken, (100) can be rewritten as

$$E_t\left[RP_{t+1}\right] = -\frac{u''\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[\left(C_{t+1} - h_{t+1}\right) RP_{t+1}\right] - \frac{1}{2}\frac{u'''\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times$$
$$E_t\left[\left(C_{t+1} - h_{t+1}\right)^2 RP_{t+1}\right] - \frac{1}{2}\frac{u'''\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[\frac{\sum_{k=1}^{K}\left(C_{k,t+1} - C_{t+1}\right)^2}{K} RP_{t+1}\right] \tag{101}$$

where C_{t+1} denotes aggregate consumption per capita.[43] The first two terms on the right-hand side of (101) show how relative asset yields depend on the second and third cross-moments of excess return with per capita consumption, while the third term reflects influence of consumer heterogeneity on the expected excess return. For a higher-order Taylor approximation, a similar equation in terms of higher-order cross-moments can be derived.

In the complete consumption insurance framework, $C_{k,t+1} = C_{t+1}$, $k = 1, ..., K$, and,

rises relative to the full-insurance case.

[43]See Mankiw (1986).

consequently, equation (100) reduces to that in the representative-agent framework:

$$E_t\left[RP_{t+1}\right] = -\sum_{n=1}^{N} \frac{1}{n!} \frac{u^{(n+1)}\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z^a_{n,t+1} RP_{t+1}\right], \qquad (102)$$

where $Z^a_{n,t+1} \equiv \left(C_{t+1} - h_{t+1}\right)^n$ and $Z^a_{1,t+1} = Z_{1,t+1}$ at all t.

The cross-moments of the excess market portfolio return with the moments of the cross-sectional distribution of individual consumption can be calculated from data on individual consumption expenditures and the excess return on the market portfolio. It follows that to determine which effect (substitutability or complementarity in the portfolio risk and the background risk in wealth) is generated by each of the moments of the cross-sectional distribution of individual consumption, it suffices to sign the first five derivatives of $u(\cdot)$. As is conventional in the literature, we assume that the marginal utility of consumption is positive ($u'(\cdot) > 0$) and decreasing ($u''(\cdot) < 0$). We also assume that an agent is prudent ($u'''(\cdot) > 0$).[44] We now turn to the signs of the fourth and fifth derivatives of $u(\cdot)$. Assume that absolute prudence, $AP(\cdot)$, is decreasing.[45]

Proposition 2 *Absolute prudence is decreasing (DAP) if and only if* $u''''(\cdot) < -AP(\cdot)\, u'''(\cdot)$. *The condition* $u''''(\cdot) < 0$ *is necessary for DAP.*

Proof. DAP implies that

$$AP'(\cdot) = -\frac{u''''(\cdot)\, u''(\cdot) - \left(u'''(\cdot)\right)^2}{\left(u''(\cdot)\right)^2} < 0. \qquad (103)$$

In order to prove that the condition $u''''(\cdot) < 0$ is necessary for DAP suppose, in contrast, that $u''''(\cdot) \geqslant 0$. When $u''''(\cdot) \geqslant 0$, $u''''(\cdot)\, u''(\cdot) \leqslant 0$ and, therefore, $AP'(\cdot) > 0$, what contradicts the assumption that absolute prudence is decreasing.

Inequality (103) means that $u''''(\cdot)\, u''(\cdot) - \left(u'''(\cdot)\right)^2 > 0$ is the necessary and sufficient condition for DAP. We can rewrite this condition as

$$u''''(\cdot) < \frac{\left(u'''(\cdot)\right)^2}{u''(\cdot)} = -AP(\cdot)\, u'''(\cdot). \qquad (104)$$

[44] Kimball (1990) defines "prudence" as a measure of the sensitivity of the optimal choice of a decision variable to risk (of the intensity of the precautionary saving motive in the context of the consumption-saving decision under uncertainty). A precautionary saving motive is positive when $-u'(\cdot)$ is concave ($u'''(\cdot) > 0$) just as an individual is risk averse when $u(\cdot)$ is concave.

[45] $AP(\cdot) = -\frac{u'''(\cdot)}{u''(\cdot)} > 0$. Intuitively, the willingness to save is an increasing function of the expected marginal utility of future wealth. Since marginal utility is decreasing in wealth, the absolute level of precautionary savings must also be expected to decline as wealth rises.

Since an agent is assumed to be prudent, the term on the right-hand side of (104) is negative. ■

A natural assumption is that, likewise absolute risk aversion, absolute prudence is convex (the absolute level of precautionary savings is decreasing in wealth at a decreasing rate).

Proposition 3 *Absolute prudence is convex (CAP) if and only if* $u'''''(\cdot) > -2AP'(\cdot) \times u'''(\cdot) - AP(\cdot)\, u''''(\cdot)$. *If preferences exhibit prudence and decreasing absolute prudence, then* $u'''''(\cdot) > 0$ *is the necessary condition for CAP.*

Proof. Absolute prudence is convex if the following condition is satisfied:

$$AP''(\cdot) = -\frac{A-B}{C} > 0, \tag{105}$$

where $A \equiv (u''(\cdot))^2 (u'''''(\cdot)\, u''(\cdot) - u'''(\cdot)\, u''''(\cdot))$, $B \equiv 2u''(\cdot)\, u'''(\cdot)\, (u''''(\cdot)\, u''(\cdot) - (u'''(\cdot))^2)$, and $C \equiv (u''(\cdot))^4$.

To prove that $u'''''(\cdot) > 0$ is necessary for CAP under prudence and DAP, assume that $u'''''(\cdot) \leqslant 0$. An agent is prudent $(AP(\cdot) > 0)$ if and only if $u'''(\cdot) > 0$. By Proposition 1, we know that the necessary condition for DAP is that $u''''(\cdot) < 0$. Then, under prudence and DAP, $A > 0$. Since $u''''(\cdot)\, u''(\cdot) - (u'''(\cdot))^2 > 0$ is the necessary and sufficient condition for DAP, prudence and DAP also imply that $B < 0$. In consequence, $AP''(\cdot) < 0$, what contradicts the initial assumption that absolute prudence is convex.

It follows from (105) that the necessary and sufficient condition for CAP is $A - B < 0$. This condition can be written as follows:

$$u''''' > \frac{2u'''(\cdot)\left(u''''(\cdot)\, u''(\cdot) - (u'''(\cdot))^2\right)}{(u''(\cdot))^2} + \frac{u'''(\cdot)\, u''''(\cdot)}{u''(\cdot)} \tag{106}$$

or, equivalently,

$$u''''' > -2AP'(\cdot)\, u'''(\cdot) - AP(\cdot)\, u''''(\cdot)\,. \tag{107}$$

Under prudence and DAP, the term $-2AP'(\cdot)\, u'''(\cdot) - AP(\cdot)\, u''''(\cdot)$ is positive.[46] ■

So, we obtain that under DAP and CAP, $u''''(\cdot) < 0$ (the necessary condition for DAP) and $u'''''(\cdot) > 0$ (this condition is necessary for CAP). Combined with the conditions $u'(\cdot) >$

[46] If an agent exhibits prudence, then $AP(\cdot) > 0$ and $u'''(\cdot) > 0$. The condition $u''''(\cdot) < 0$ is necessary for DAP.

0, $u''(\cdot) < 0$, and $u'''(\cdot) > 0$, it follows that the cross-sectional mean and skewness of individual consumption help to explain the equity premium if their cross-moments with the excess market portfolio return are positive. Since the cross-moments of the equity premium with the cross-sectional variance and kurtosis of individual consumption are always positive, taking them into account lowers the equity premium explained by model (100).

3.2.2 Uninsurable Background Risk and the Risk-Free Rate

According to (99), the equilibrium rate of return on the risk-free asset is

$$E_t\left[R_{f,t+1}\right] = \delta^{-1}\sum_{n=0}^{N}\frac{1}{n!}\frac{u^{(n+1)}\left(h_t\right)}{u'\left(h_{t+1}\right)}Z_{n,t} - \sum_{n=1}^{N}\frac{1}{n!}\frac{u^{(n+1)}\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z_{n,t+1}R_{f,t+1}\right] \qquad (108)$$

or, equivalently,

$$E_t\left[R_{f,t+1}\right] = \left(\delta\frac{u'\left(h_{t+1}\right)}{u'\left(h_t\right)}\right)^{-1} + \left(\delta^{-1}\frac{u''\left(h_t\right)}{u'\left(h_{t+1}\right)}Z_{1,t} - \frac{u''\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z_{1,t+1}R_{f,t+1}\right]\right)$$
$$+\sum_{n=2}^{N}\frac{1}{n!}\left(\delta^{-1}\frac{u^{(n+1)}\left(h_t\right)}{u'\left(h_{t+1}\right)}Z_{n,t} - \frac{u^{(n+1)}\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z_{n,t+1}R_{f,t+1}\right]\right).^{47} \qquad (109)$$

When there is complete consumption insurance, the expected risk-free rate is

$$E_t\left[R_{f,t+1}\right] = \left(\delta\frac{u'\left(h_{t+1}\right)}{u'\left(h_t\right)}\right)^{-1} + \left(\delta^{-1}\frac{u''\left(h_t\right)}{u'\left(h_{t+1}\right)}Z^a_{1,t} - \frac{u''\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z^a_{1,t+1}R_{f,t+1}\right]\right)$$
$$+\sum_{n=2}^{N}\frac{1}{n!}\left(\delta^{-1}\frac{u^{(n+1)}\left(h_t\right)}{u'\left(h_{t+1}\right)}Z^a_{n,t} - \frac{u^{(n+1)}\left(h_{t+1}\right)}{u'\left(h_{t+1}\right)} \times E_t\left[Z^a_{n,t+1}R_{f,t+1}\right]\right) \qquad (110)$$

with $Z^a_{1,t} = Z_{1,t}$ at all t.

The expected return on the risk-free asset in equations (109) and (110) is expressed as a sum of three terms. The first term, $\left(\delta\frac{u'(h_{t+1})}{u'(h_t)}\right)^{-1}$, characterizes the effect of preference for the present. Since the agent's utility function is concave, the investor has preferences for smoothing his consumption over time. In order to make the agent not to smooth his consumption, the risk-free rate must be larger than $\left(\delta\frac{u'(h_{t+1})}{u'(h_t)}\right)^{-1}$ (the consumption smoothing effect). This effect is reflected by the second term on the right-hand side of equations (109) and (110). The size of the consumption smoothing effect depends on the

[47] Likewise (100), this equation can also be rewritten in terms of the cross-moments of return with per capita consumption and the cross-moments of return with the moments of the cross-sectional distribution of individual consumption.

degree of concavity of the agent's utility function.[48] When the agent is prudent and, hence, wants to save more in order to self-insure against uninsurable risks, the risk-free rate must be lower than $\left(\delta \frac{u'(h_{t+1})}{u'(h_t)}\right)^{-1}$ to sustain the equilibrium (the precautionary saving effect). The precautionary saving effect is represented by the third term on the right-hand side of equations (109) and (110). Since $Z^a_{1,t} = Z_{1,t}$ at all t, the first two terms on the right-hand side of equations (109) and (110) are the same and, therefore, taking into account consumer heterogeneity has the potential for explaining the risk-free rate puzzle if the third term on the right-hand side of (109) is less than that in (110).

3.2.3 Measurement Error Issue

A well documented potential problem with using household level data is the large measurement error in reported individual consumption.[49] The widely used solution to mitigate the impact of measurement error consists in averaging over the level of consumption or consumption growth. Since measurement error is not observable, the choice of the optimal method remains somewhat arbitrary and depends on what type of measurement error is assumed.[50]

We assume that the observation error in the consumption level is additive. Since individual consumption is assumed to be misreported by some stochastic dollar amount $\epsilon_{k,t}$, the observed consumption level is $C_{k,t} = C^*_{k,t} + \epsilon_{k,t}$, where $C^*_{k,t}$ is the true level of the agent k's consumption in period t. We further assume that for all k and at all t, $\epsilon_{k,t} \sim D\left(0, \sigma^2_{\epsilon,t}\right)$ and $\epsilon_{k,t}$ is independent of the true consumption level.

By the law of large numbers, when $K \longrightarrow \infty$, $\frac{1}{K}\sum_{k=1}^{K} C_{k,t} \xrightarrow{P} E\left[C_{k,t}\right] = E\left[C^*_{k,t}\right]$ and, hence, averaging over the level of consumption should mitigate the additive idiosyncratic measurement error effect. It follows that when $K \longrightarrow \infty$, $Z^a_{n,t} \longrightarrow Z^{a*}_{n,t}$ for all n at all t and $Z_{1,t} \longrightarrow Z^*_{1,t}$ at all t.[51]

[48]The more concave the agent's utility function, the higher the risk-free rate needed to compensate the agent for not smoothing his consumption over time.

[49]See Runkle (1991) and Zeldes (1989).

[50]Additive measurement error suggests averaging over the level of consumption, while in the case of multiplicative measurement error, averaging over consumption growth may be preferable.

[51]The sign * means that a value is calculated using true levels of consumption.

It may be shown that

$$Z_{2,t} \xrightarrow{P} E\left[C_{k,t} - h_t\right]^2 = E\left[C^*_{k,t} - h_t\right]^2 + \sigma^2_{\epsilon,t}, \tag{111}$$

$$Z_{3,t} \xrightarrow{P} E\left[C_{k,t} - h_t\right]^3 = E\left[C^*_{k,t} - h_t\right]^3 + E\left[\epsilon_{k,t}\right]^3 \tag{112}$$

and

$$Z_{4,t} \xrightarrow{P} E\left[C_{k,t} - h_t\right]^4 = E\left[C^*_{k,t} - h_t\right]^4 + 6\sigma^2_{\epsilon,t} E\left[C^*_{k,t} - h_t\right]^2 + E\left[\epsilon_{k,t}\right]^4. \tag{113}$$

Therefore, when the number of households in a sample is large, equations (102) and (110) yield asymptotically unbiased estimates of the coefficient of risk aversion and δ. The same is also true for equations (100) and (109) with $N = 1$. Under the assumption of consumer heterogeneity, a Taylor series expansion of order higher than 1 can lead to biased estimates of both the risk aversion parameter and the time discount factor δ. Observe that with measurement error of the type assumed here, little may be said about the signs and magnitudes of the biases in the estimates of the coefficient of risk aversion and δ.[52] However, as we can see from (111)-(113), it seems to be plausible that the magnitude of the bias in the estimates of the moments of the cross-sectional distribution of individual consumption and, therefore, the cross-moments of return with the moments of the cross-sectional distribution of individual consumption increases as the order of a Taylor series expansion rises.

3.3 Empirical Results

In Section 3.2, we have shown that the hypothesis of incomplete consumption insurance has the potential for explaining both the excess market portfolio return and the rate of return on the risk-free asset. In this Section, we assess the quantitative importance of the hypotheses of incomplete consumption insurance and limited asset market participation in explaining the equity premium and the risk-free rate.

A class of utility functions widely used in the literature is the set of utility functions exhibiting an harmonic absolute risk aversion (HARA). HARA utility functions take the

[52] In Section 3.3.5, we perform an empirical analysis of the effect of additive measurement error in the consumption level on the estimates of the preference parameters in the context of the hypothesis of incomplete consumption insurance.

following form:

$$u(C_t) = a\left(b + \frac{cC_t}{\gamma}\right)^{1-\gamma}, \tag{114}$$

where a, b, and c are constants, $b + \frac{cC_t}{\gamma} > 0$, and $a\frac{c(1-\gamma)}{\gamma} > 0$.[53]

There are three special cases of HARA utility functions.[54] Constant relative risk aversion (CRRA) utility functions can be obtained by selecting $b = 0$:

$$u(C_t) = a\left(\frac{cC_t}{\gamma}\right)^{1-\gamma}. \tag{115}$$

In the special case when $a = \frac{(c/\gamma)^{\gamma-1}}{(1-\gamma)}$, we get $u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$, where γ is the RRA coefficient, $\gamma \neq 1$.[55]

If $\gamma \longrightarrow \infty$, we obtain constant absolute risk aversion (CARA) utility functions:

$$u(C_t) = -exp\left(-\frac{c}{b}C_t\right). \tag{116}$$

With $\gamma = -1$, we get quadratic utility functions:

$$u(C_t) = a\left(b + \frac{cC_t}{\gamma}\right)^2. \tag{117}$$

It is easy to check that when the first five derivatives of the HARA utility function exist, $u'(\cdot) > 0$ and $u''(\cdot) < 0$ imply $u'''(\cdot) > 0$, $u''''(\cdot) < 0$, and $u'''''(\cdot) > 0$. Given the results obtained in Section 3.2, it follows that any HARA class utility function may be used when addressing the question of the effect of a particular cross-moment of return with the moment of the cross-sectional distribution of individual consumption on the expected equity premium and risk-free rate. For the equations derived in Section 3.2 to be scale-invariant, we need $b = 0$ when using a HARA class preference specification. That corresponds to the CRRA preferences.

Assuming the CRRA homogeneous preferences, $u(C_{k,t}) = \frac{C_{k,t}^{1-\gamma}-1}{1-\gamma}$, $k = 1, ..., K$, we can rewrite (100) and (109) as

$$E_t[RP_{t+1}] = -E_t\left[\left(\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)RP_{t+1}\right] \tag{118}$$

[53]The last inequality is necessary to insure that $u'(\cdot) > 0$.

[54]See Golier (2001).

[55]A logarithmic utility specification, $u(C_t) = log(C_t)$, corresponds to the case when $\gamma = 1$.

and

$$E_t\left[R_{f,t+1}\right] = \delta^{-1}\left(\frac{h_{t+1}}{h_t}\right)^{\gamma}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t}}{h_t^n}\right)$$
$$-E_t\left[\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}R_{f,t+1}\right].^{56} \tag{119}$$

If consumption insurance is complete, then $Z_{n,t+1} = Z_{n,t+1}^a$ and, hence, equations (118) and (119) can be rewritten as

$$E_t\left[RP_{t+1}\right] = -E_t\left[\left(\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}^a}{h_{t+1}^n}\right)RP_{t+1}\right] \tag{120}$$

and

$$E_t\left[R_{f,t+1}\right] = \delta^{-1}\left(\frac{h_{t+1}}{h_t}\right)^{\gamma}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t}^a}{h_t^n}\right)$$
$$-E_t\left[\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}^a}{h_{t+1}^n}R_{f,t+1}\right], \tag{121}$$

respectively.

We first focus on the mean and standard deviation of the SDF derived in Section 3.2. In this part, we perform the Hansen-Jagannathan volatility bound analysis to explore the potential for this pricing kernel to explain the market premium. The second part is a model calibration. Here, we look for the values of the risk aversion coefficient γ and the time discount factor δ which allow to fit the observed mean equity premium and risk-free rate. In the third part, we exploit the time series properties of consumption and asset returns and use a non-linear GMM estimation approach to test the conditional Euler equations for the excess market portfolio return and the risk-free rate.

3.3.1 Description of the Data

The Consumption Data. The consumption data used in our analysis are taken from the CEX. As opposed to the PSID, which offers only food consumption data on an annual basis,

[56]Under the CRRA preferences, $\frac{Z_{n,t+1}}{h_{t+1}^n} = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{C_{k,t+1}-h_{t+1}}{h_{t+1}}\right)^n$. If h_{t+1} is the reference level of consumption, we can define $\frac{C_{k,t+1}-h_{t+1}}{h_{t+1}}$ as a surplus consumption ratio, as in Campbell and Cochrane (1999).

the CEX contains highly detailed data on monthly consumption expenditures.[57] The CEX attempts to account for an estimated 70% of total household consumption expenditures. Since the CEX is designed with the purpose of collecting consumption data, measurement error in consumption is likely to be smaller for the CEX consumption data compared to the PSID consumption data.

The CEX data available cover the period from 1979:10 to 1996:2. It is a collection of data on approximately 5000 households per quarter in the US. Each household in the sample is interviewed every three months over five consecutive quarters.[58] As households complete their participation, they are dropped and new households move into the sample. Thus, each quarter about 20% of the consumer units are new. The second through fifth interviews use uniform questionnaires to collect demographic and family characteristics as well as data on monthly consumption expenditures for the previous three months made by households in the survey.[59] Various income information is collected in the second and fifth interviews as well as information on the employment of each household member.

The measure of consumption used in this empirical investigation is consumption of nondurables and services (NDS). For each household, we calculate monthly consumption expenditures for all the disaggregate consumption categories offered by the CEX. Then, we deflate obtained values in 1982-84 dollars with the CPI's (not seasonally adjusted, urban consumers) for appropriate consumption categories.[60] Aggregating the household's monthly consumption across these categories is made according to the National Income and Product Account definitions of consumption aggregates. In order to transform our consumption data to a per capita basis, we normalize the consumption of each household by dividing it by the number of family members in the household.

The Returns Data. The measures of the nominal market return are the value-weighted and equal-weighted returns (capital gain plus dividends) on all stocks listed on the NYSE and AMEX obtained from the CRSP. The real monthly market return is calculated as the

[57] Food consumption is likely to be one of the most stable consumption components. Furthermore, as is pointed out by Carroll (1994), 95% of the measured food consumption in the PSID is noise due to the absence of interview training.

[58] The first interview is practice and is not included in the published data set.

[59] Demographic variables are based upon heads of households.

[60] The CPI series are obtained from the Bureau of Labor Statistics through CITIBASE.

nominal market return divided by the 1-month inflation rate based on the deflator defined for NDS consumption. The nominal monthly risk-free rate of interest is the 1-month Treasury bill return from CRSP. The real monthly risk-free interest rate is calculated as the nominal risk-free rate divided by the 1-month inflation rate. Market premium is calculated as the difference between the real market return and the real risk-free rate of interest.

Asset Holders. For the consumer units completing their participation in the first through third quarters of 1986, the Bureau of Labor Statistics (BLS) has changed, beginning the first quarter of 1986, the consumer unit identification numbers so that the identification numbers for the same household in 1985 (when this household has been interviewed for the first time) and in 1986 (when it has completed its participation) are not the same. To match consumer units between the 1985 and 1986 data tapes, we use household characteristics which allow us to identify consumer units uniquely. As a result, we manage to match 47.0% of households between the 1985 and 1986 data tapes. The detailed description of the procedure used to match consumers units is given in the Appendix A.

In the fifth (final) interview, the household is asked to report end-of-period estimated market value of all stocks, bonds, mutual funds, and other such securities (market value of all securities) held by the consumer unit on the last day of the previous month as well as the difference in this estimated market value compared with the value of all securities held a year ago last month. Using these two values, we calculate asset holdings at the beginning of a 12-month recall period. The consumer unit is considered as an assetholder if the household's asset holdings at the beginning of a 12-month recall period exceed a certain threshold.

To assess the quantitative importance of limited participation of households in asset markets, we consider four sets of households. The first set (SET1) consists of all consumer units irrespectively of the reported market value of all securities. To take into consideration that only a fraction of households participates in asset markets, we use three sets of households defined as assetholders: the first one (SET2) consists of consumer units whose asset holdings are equal to or exceed $2 in 1999 dollars, the two others consist of all households with reported total assets equal to or exceeding $10000 (SET3) and $20000 (SET4).[61]

[61]Over the period 1991-1996 about 18% of households, for whom the market value of all securities held

Per capita consumption of a set of households is calculated as the equal-weighted average of normalized consumption expenditures of the households in the set. Obtained per capita consumption is seasonally adjusted by using the X-11 seasonal adjustment program.[62] We seasonally adjust the normalized consumption of each household by using the additive adjustments obtained from per capita consumption.

Data Selection Criteria. We drop from the sample nonurban households, households residing in student housing, households with incomplete income responses, and households who do not have a fifth interview. Following BCG (2002), in any given month, we drop from the sample households that report in that month as zero either their food consumption or their NDS consumption, or their total consumption, as well as households with missing information on the above items. Additionally, we keep in the sample only households whose head is between 19 and 75 years of age.

3.3.2 Estimation of the Conditional Expectation of Consumption

Since the conditional expectation of consumption, h_{t+1}, is supposed to be the same for all investors, we may assume it to be equal to the conditional expectation of aggregate consumption per capita, $h_{t+1} \equiv E_t\left[C_{t+1}\right]$. Assuming, as in Campbell and Cochrane (1999), a random walk model of consumption,

$$\Delta c_{t+1} = g + \eta_{t+1}, \tag{122}$$

where $\Delta c_{t+1} \equiv log\frac{C_{t+1}}{C_t}$ and $\eta_{t+1} \sim N\left(0, \sigma_\eta^2\right)$, we get

$$h_{t+1} = exp\left(g + \frac{\sigma_\eta^2}{2}\right) C_t. \tag{123}$$

Table I presents the usual ML parameter estimates and tests of model (122).

a year ago last month is not missing, reported asset holdings of $1 at the beginning of a 12-month recall period. That occurs when the household reported owning securities without precising their value (see Vissing-Jorgensen (1998)). Following Vissing-Jorgensen (1998), we classify these households as nonassetholders.

[62]Ferson and Harvey (1992) point out that since the X-11 program uses past and future information in the time-averaging it performs, this type of seasonal adjustment may induce spurious correlation between the error terms of a model and lagged values of the variables and, hence, may cause improper rejections of the model based on tests of overidentifying restrictions. As alternatives to using the X-11 program, Brav and Géczy (1995) propose to use a simpler linear filter (Davidson and MacKinnon (1993)) or the Ferson-Harvey (1992) method of incorporating forms of seasonal habit persistence directly in the Euler equation.

3.3.3 Required Order of a Taylor Series Expansion

One approach to determine the order at which the expansion should be truncated is to allow data to motivate the point of truncation.[63] This approach consists in repeating the estimation of the model for increasing values of N and truncating the expansion at the point when further increasing in N does not significantly affect the estimation results. As Dittmar (2002) points out, there are at least two difficulties with allowing data to determine the required order of a Taylor series expansion. The first one is the possibility of overfitting the data. Another problem is that when a high-order expansion is taken, preference theory no longer guides in determining the signs of the priced risk factors. To avoid the last problem, Dittmar (2002) proposes to let preference arguments determine the point of truncation. He shows that increasing marginal utility, risk aversion, decreasing absolute risk aversion, and decreasing absolute prudence imply the fourth derivative of utility functions to be negative. Since preference assumptions do not guide in determining the signs of the higher-order derivatives, Dittmar (2002) assumes that the Taylor series expansion terms of order higher than three do not matter for asset pricing and truncates a Taylor series expansion after the cubic term.[64] His point of view is that the advantage coming from signing the Taylor series expansion terms outweighs a loss of power due to omitting the terms of order four and higher.

In this paper, we let both preference theory and the data guide the truncation. The restriction of decreasing absolute prudence allows us to sign the fifth derivative of utility functions and, therefore, pursue the expansion further than it is usually done. Following Dittmar (2002), we should truncate a Taylor series expansion after the cross-moment of return with the cross-sectional kurtosis of individual consumption. The question here is whether the cross-moments of return with the first four moments of the cross-sectional distribution of individual consumption can be estimated precisely given the previously mentioned potentially severe effect of measurement error on the estimates of the high-order moments of the cross-sectional distribution of individual consumption. To answer this question, we estimate the cross-moments $E\left[Z_{n,t+1}RP_{t+1}\right]$ and $E_t\left[Z_{n,t+1}RP_{t+1}\right]$ for n ranging from 1 to

[63] See Bansal, Hsieh, and Viswanathan (1993).

[64] BCG (2002) also limit their analysis to a third-order approximation when using a Taylor series expansion to the equal-weighted average of the household's IMRS.

4 using an iterated GMM approach.[65] We find that the precision of estimation decreases as n rises so that for all the sets of households classified as assetholders, the null hypotheses $E[Z_{n,t+1}RP_{t+1}] = 0$ and $E_t[Z_{n,t+1}RP_{t+1}] = 0$ are not rejected statistically at the 5% level for $n = 4$. This result confirms the conjecture that the observation error in reported individual consumption can make it difficult to precisely estimate the high-order cross-sectional moments of individual consumption and, therefore, their cross-moments with the excess market portfolio return.

We find that the cross-moments of the equity premium with the cross-sectional variance and skewness of individual consumption are both positive. It follows that the variance of the cross-sectional distribution of individual consumption represents the effect of complementarity in the portfolio risk and the background risk in wealth, while the cross-sectional skewness of individual consumption represents the effect of substitutability. Since both the unconditional and conditional cross-moments of the equity premium with the cross-sectional skewness of individual consumption are still estimated precisely, we limit our empirical investigation to the first three moments of the cross-sectional distribution of individual consumption.[66]

3.3.4 The Hansen-Jagannathan Volatility Bound Analysis

In this Section, we assess the plausibility of the pricing kernel derived in Section 3.2 by studying its mean and standard deviation for different values of the risk aversion coefficient.

Assuming that a candidate SDF $M_t^*(m)$ may be formed as a linear combination of asset returns, $M_t^*(m) = m + (R_t - E[R_t])' \lambda_m$, Hansen and Jagannathan (1991) show that a lower bound on the volatility of any SDF M_t, that has unconditional mean m and satisfies the first-order condition $\iota = E[M_t R_t]$, is given by

$$\sigma(M_t^*(m)) = \left((\iota - mE[R_t])' \Sigma^{-1} (\iota - mE[R_t])\right)^{1/2}, \tag{124}$$

where ι is the vector of ones, R_t is the vector of time-t asset gross returns, and Σ is the unconditional variance-covariance matrix of asset returns.

[65] When the conditional cross-moments are estimated, the set of instruments consists of a constant and the term inside brackets lagged one period.

[66] Given the same problem with estimating high-order moments due to measurement error, Cogley (2002) also stops at a third-order polynomial when taking a Taylor series expansion to the individual's IMRS.

When an unconditionally risk-free asset (or, more generally, an unconditional zero-beta asset[67]) exists, the first-order condition for the risk-free interest rate, $1 = E[M_t R_{f,t}]$, implies that the unconditional expectation of the SDF is the reciprocal of the expected gross return on this asset, $m = \frac{1}{E[R_{f,t}]}$. However, the restriction of nonsingularity of the second-moment matrix of asset returns, Σ, implies that there is no unconditionally risk-free asset or combination of assets[68] and, hence, m must be treated as an unknown parameter.

If excess returns are used, condition (124) becomes

$$\sigma\left(M_t^*(m)\right) = \left(m^2 E\left[R_t^e\right]' \widetilde{\Sigma}^{-1} E\left[R_t^e\right]\right)^{1/2}. \tag{125}$$

Here, R_t^e is the vector of excess returns[69] and $\widetilde{\Sigma}$ denotes the variance-covariance matrix of excess returns. Working with excess returns, we are allowed to assume that there is an unconditionally risk-free asset.[70] When such an asset exists, m is no longer an unknown parameter and may be calculated from data on the risk-free interest rate.

If there is a single excess return, the lower bound on the volatility of the SDF is given by

$$\sigma\left(M_t^*(m)\right) = m \frac{E\left[R_{i,t}^e\right]}{\sigma\left(R_{i,t}^e\right)}. \tag{126}$$

It means that the Sharpe ratio of any excess return $R_{i,t}^e$ must obey

$$\frac{E\left[R_{i,t}^e\right]}{\sigma\left(R_{i,t}^e\right)} \leqslant \frac{\sigma\left(M_t^*(m)\right)}{m} = \frac{\sigma\left(M_t^*(m)\right)}{E\left[M_t^*(m)\right]}. \tag{127}$$

The Hansen-Jagannathan or Sharpe ratio inequality (127) implies that for the SDF to be consistent with a given set of asset return data, it must lie above a ray from the origin with slope equal to the Sharpe ratio of the single risky excess return $R_{i,t}^e$, $\frac{E[R_{i,t}^e]}{\sigma(R_{i,t}^e)}$.

The only excess return used in this empirical research is the excess return on the market portfolio over the risk-free rate. Over the period from 1979:11 to 1996:2, the mean market premium is 0.73% per month with a standard deviation of 4.24%. The dotted line in Figure

[67] An asset which unconditional covariance with the SDF is zero.

[68] At least, its identity is not known beforehand.

[69] $R_{i,t}^e \equiv R_{i,t} - R_{j,t}$ is the excess return on asset i over some reference asset j.

[70] This asset is usually used as a reference one.

1 indicates the lower volatility bound for the pricing kernels implied by the Sharpe ratio inequality (127) and the average monthly excess return on the US stock market.

Our benchmark model is the conventional power utility specification under the assumption of complete consumption insurance. Under complete consumption insurance, the SDF in the Euler equation for the market premium is $\widetilde{M}_{t+1} \equiv \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}$. The black "boxes" in Figure 1 represent mean-standard deviation pairs implied by the SDF $\widetilde{M}_{t+1}$ for γ ranging from 1 to 13 with increments of 1 for the set of all households in the sample (SET1). The white "boxes" denote mean-standard deviation points for the set of households whose asset holdings are equal to or exceed \$2 (SET2). The "triangles" and the "crosses" are used for the groups of households who reported total assets equal to or exceeding \$10000 (SET3) and \$20000 (SET4), respectively.

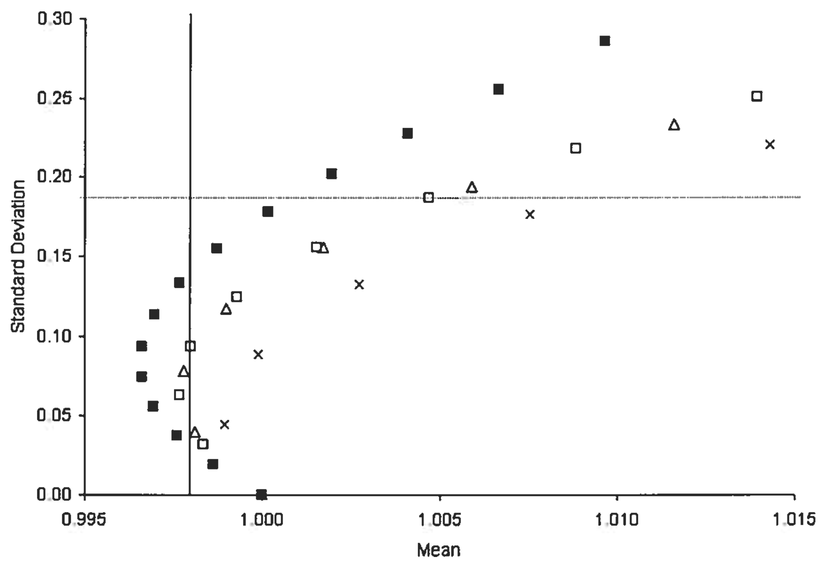

Figure 1: Standard deviation-mean diagram for $\widetilde{M}_{t+1} = \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}$. The dotted line indicates the lower volatility bound for SDFs implied by the Sharpe ratio inequality (64) and the average monthly excess return on the US stock market. The solid vertical line at $\frac{1}{E[R_{F,t+1}]} = 0.9979$ indicates $E\left[\widetilde{M}_{t+1}\right]$ for δ equal to 1. The CEX data, 1979:10 to 1996:2.

To explain a Sharpe ratio of 0.17, the complete consumption insurance model needs a risk aversion coefficient $\gamma \geqslant 10$ for SET1, $\gamma \geqslant 6$ for SET2, and $\gamma \geqslant 5$ for SET3 and SET4.

Although, as expected, the coefficient of risk aversion, γ, at which the mean-standard deviation points implied by $\widetilde{M}_{t+1}$ enter the feasible region, decreases in asset holdings, it is only slightly different across the sets of consumer units defined as assetholders.

Given that over the period from 1979:11 to 1996:2 the sample mean real risk-free rate is 0.21% per month, the unconditional mean of the SDF is $m = \frac{1}{E[R_{f,t+1}]} = 0.9979$.

From

$$m = E\left[M_{t+1}\right] = E\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\right] = E\left[\delta\widetilde{M}_{t+1}\right] = \frac{1}{E\left[R_{f,t+1}\right]}, \tag{128}$$

it follows that $E\left[\widetilde{M}_{t+1}\right] = \frac{1}{\delta E[R_{f,t+1}]}$ and $\delta = \frac{1}{E[R_{f,t+1}]\times E[\widetilde{M}_{t+1}]}$.

The solid vertical line at $\frac{1}{E[R_{f,t+1}]} = 0.9979$ in Figure 1 indicates the mean of $\widetilde{M}_{t+1}$ for δ set to 1. By looking at Figure 1, we can see that when the mean-standard deviation points are in the admissible region for SDFs, the mean of $\widetilde{M}_{t+1}$ is larger than the reciprocal of the mean real risk-free interest rate, what implies δ less than 1. However, for economically plausible (less than 3) values of the risk aversion coefficient, $\widetilde{M}_{t+1}$ has, in most cases, a mean which is less than $\frac{1}{E[R_{f,t+1}]}$ and, then, the Euler equation for the risk-free rate can be satisfied only with δ larger than 1 (the risk-free rate puzzle).[71] At the same time, for the values of the RRA coefficient in the conventional range, the SDF has a standard deviation which is much less than that required by the Sharpe ratio. This is the equity premium puzzle.

Now, let us assume that consumption insurance is complete and take a Taylor series expansion to the representative-agent's marginal utility of consumption around the conditional expectation of aggregate consumption per capita. That implies the following pricing kernel:

$$\widetilde{M}_{t+1} \equiv 1 + \sum_{n=1}^{N} \frac{1}{n!}(-1)^n \left(\prod_{l=0}^{n-1}(\gamma + l)\right)\frac{Z^a_{n,t+1}}{h^n_{t+1}}. \tag{129}$$

The Hansen-Jagannathan volatility bound analysis results for $\widetilde{M}_{t+1}$ given by (129) for a third-order ($N = 3$) Taylor series expansion to the agent's marginal utility of consumption are presented in Figure 2. These results show that even when a third-order Taylor series expansion is taken and only a fraction of consumers is assumed to participate in asset

[71] For SET1, for example, the value of δ larger than 1 is required for fitting the return on the risk-free asset for any value of γ between 2 and 7.

markets, relatively high risk aversion is needed to make the mean-standard deviation points enter the feasible region.

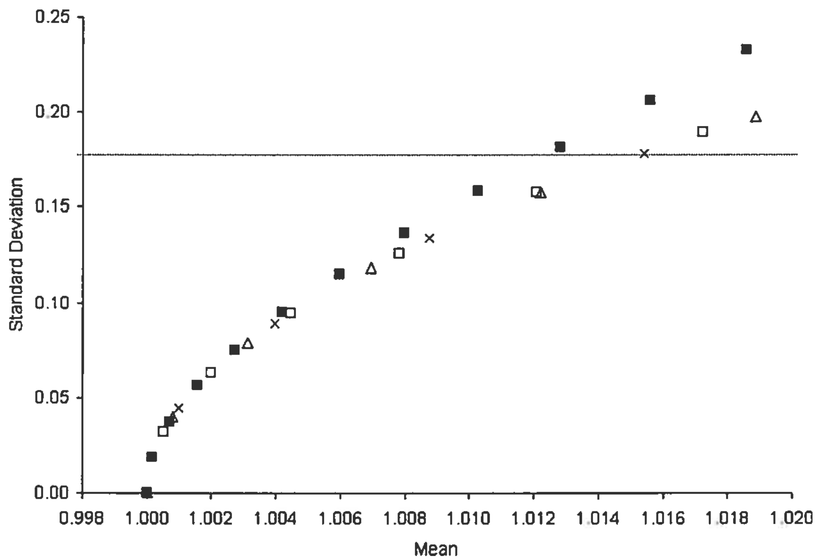

Figure 2: Standard deviation-mean diagram for $\widetilde{M}_{t+1} = 1 + \sum_{n=1}^{3} \frac{1}{n!} (-1)^n \cdot \left(\prod_{l=0}^{n-1} (\gamma + l)\right) \frac{Z^a_{n,t+1}}{h^n_{t+1}}$. The dotted line indicates the lower volatility bound for SDFs implied by the Sharpe ratio inequality (64) and the average monthly excess return on the US stock market. The CEX data, 1979:10 to 1996:2.

If we assume that there are uninsurable shocks to consumption, the SDF in the Euler equation for the equity premium is

$$\widetilde{M}_{t+1} \equiv 1 + \sum_{n=1}^{N} \frac{1}{n!} (-1)^n \left(\prod_{l=0}^{n-1} (\gamma + l)\right) \frac{Z_{n,t+1}}{h^n_{t+1}}. \tag{130}$$

The results for the SDF specification (130) corresponding to a third-order ($N = 3$) Taylor series expansion to the agent's marginal utility of consumption are shown in Figure 3. The first mean-standard variation point above the horizontal axis corresponds to the RRA coefficient of 0.1, successive points have relative risk aversion of 0.2, 0.3, and so on. Given the results in Figure 3, we may conclude that taking into account consumer heterogeneity substantially affects the mean and standard deviation of the SDF. The Hansen-Jagannathan analysis shows that the mean-standard deviation points enter the feasible region at plausible values of γ (less than 1), unlike the case when a representative agent within each group of households is assumed.[72] For a given value of γ, volatility of the SDF is only slightly affected by the size of asset holdings, so that the value of risk aversion allowing to explain the market premium does not significantly differ across sets of households.

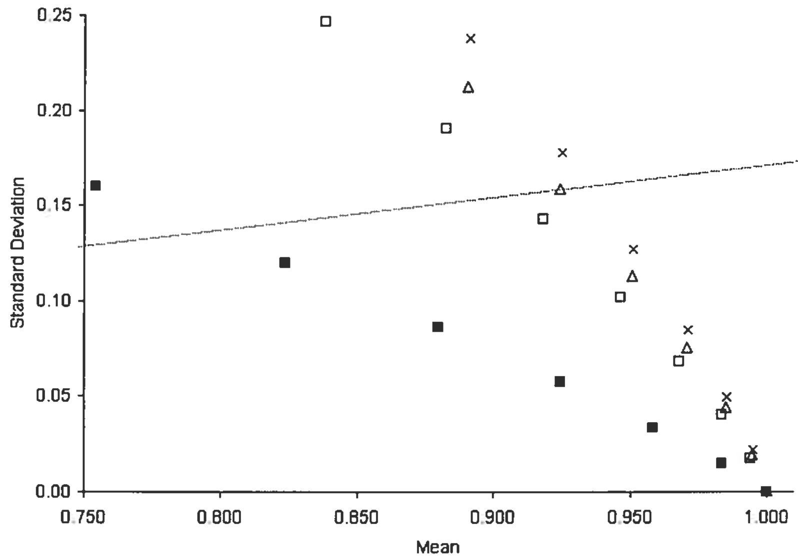

Figure 3: Standard deviation-mean diagram for $\widetilde{M}_{t+1} = 1 + \sum_{n=1}^{3} \frac{1}{n!} (-1)^n \cdot \left(\prod_{l=0}^{n-1} (\gamma + l)\right) \frac{Z_{n,t+1}}{h_{t+1}^n}$. The dotted line indicates the lower volatility bound for SDFs implied by the Sharpe ratio inequality (64) and the average monthly excess return on the US stock market. The CEX data, 1979:10 to 1996:2.

3.3.5 Model Calibration

In this Section, we address the question of whether the size of the adverse effect of the independent background risk in wealth on the attitude towards the portfolio risk is such that it makes it possible to explain the observed mean excess return on the market portfolio and risk-free rate with an economically plausible value of the RRA coefficient. The results are presented in Table II.

As in the preceding Section, our benchmark case is complete consumption insurance. To

[72]Although both the volatility bound and the mean-standard deviation points implied by the SDF are estimated with error, it seems, however, to be improbable that this error is so large to substantially affect the results and to make the points enter the feasible region at implausibly high values of γ.

assess the contribution of the hypothesis of complete consumption insurance, we calculate the unexplained mean equity premium as

$$v_1 = \frac{1}{T}\sum_{t=0}^{T-1}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} RP_{t+1} \tag{131}$$

for the values of the risk aversion coefficient γ increasing from 0 with increments of 0.1.[73] When the RRA coefficient is set to zero, the unexplained mean premium is equal to the sample mean of the excess market portfolio return.[74] In Table II, we report the values of γ for which the unexplained mean premium of the value-weighted market portfolio becomes negative. As we can see in Table II, even when limited asset market participation is taken into consideration, the complete consumption insurance model is able to fit the observed mean equity premium only if an individual is assumed to be implausibly risk averse.

When a Taylor series expansion to aggregate consumption per capita is taken, we calculate the statistic v_2 as

$$v_2 = \frac{1}{T}\sum_{t=0}^{T-1}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z^a_{n,t+1}}{h^n_{t+1}}\right) RP_{t+1}. \tag{132}$$

The results in Table II show that a Taylor series expansion of any order fails to explain the mean equity premium with an economically plausible value of the RRA coefficient. When a first-order Taylor series expansion is taken, the mean premium of the value-weighted market portfolio can be explained with the RRA coefficient ranging from 22.4 to 113 for different sets of households. In the case of a third-order Taylor series expansion, the mean equity premium can be explained with the values of the RRA coefficient which are slightly lower than those in the first-order Taylor series expansion case (between 23.2 and 63), but, nevertheless, remain too large to be recognized as economically plausible. As the representative-agent's marginal utility of consumption is expanded as a Taylor series up to terms capturing the third cross-moment of the excess market portfolio return with aggregate consumption, the statistic v_2 increases as the RRA coefficient rises, so that there is no positive value of γ allowing to fit the observed mean premium of the value-weighted market portfolio.

[73] See BCG (2002).

[74] Over the period from 1979:11 to 1996:2, the mean premium of the value-weighted market portfolio is 0.73% per month.

Under the assumption of limited asset market participation, we find some evidence that the risk aversion coefficient decreases as the threshold value in the definition of assetholders is raised. When the threshold value is quite large ($10000 in 1999 dollars or larger), one can explain the mean equity premium with γ which is lower than that under the standard power utility model. However, even after limited participation is taken into consideration, the model fails to explain the mean excess market portfolio return with an economically plausible value of risk aversion.

Assuming uninsurable shocks to consumption, we calculate the unexplained mean equity premium as

$$v_3 = \frac{1}{T}\sum_{t=0}^{T-1}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)RP_{t+1}. \tag{133}$$

In contrast to the complete consumption insurance case, taking into account consumer heterogeneity allows to fit the mean excess market portfolio return with an economically plausible (between 1 and 2) value of the RRA coefficient when the agent's marginal utility of consumption is expanded as a Taylor series up to cubic terms. Under the hypothesis of limited asset market participation, empirical evidence that the RRA coefficient decreases as the threshold value rises is weak. There is no positive value of the RRA coefficient allowing to explain the mean equity premium when the agent's marginal utility of consumption is expanded as a Taylor series up to terms capturing the cross-sectional variance of individual consumption.[75] Given the values of the risk aversion parameter γ which allow to explain the observed mean premium of the value-weighted market portfolio, we estimate the time discount factor δ needed to fit the observed mean risk-free rate as

$$\delta = \frac{\frac{1}{T}\sum_{t=0}^{T-1}\left[\left(\frac{h_{t+1}}{h_t}\right)^{\gamma}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t}}{h_t^n}\right)\right]}{\frac{1}{T}\sum_{t=0}^{T-1}\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)R_{f,t+1}\right]}. \tag{134}$$

[75] Although BCG (2002) take a Taylor series expansion to the equal-weighted sum of the household's IMRS and not that to the agent's marginal utility of consumption, as in our work, their results are similar to ours. Specifically, they find that when the SDF is expressed in terms of the cross-sectional mean and variance of the household consumption growth rate, the average unexplained premium increases as the RRA coefficient rises. When a Taylor series expansion captures the cross-sectional skewness, in addition to the mean and variance, the average unexplained excess return on the market portfolio is less than that for the SDF given by the equal-weighted sum of the household's IMRS.

To test whether the obtained results are susceptible to additive measurement error in the consumption level, we assume that observation error is normally distributed with zero mean, $\epsilon_{k,t} \sim N\left(0, \sigma_{\epsilon,t}^2\right)$, and independent of true consumption. We further assume that the cross-sectional variance of measurement error is 20% of the cross-sectional variance of the household consumption observed in the data, $\sigma_{\epsilon,t}^2 = 0.2\frac{1}{K}\sum_{k=1}^{K}\left(C_{k,t} - C_t\right)^2$. The row "allowing for observation error" in Table II presents the results obtained when $\widetilde{C}_{k,t} = C_{k,t} + \epsilon_{k,t}$ is used in the calibration. These results illustrate that in a small sample framework with the measurement error of the type analyzed here, the estimate of γ will be biased upward. Empirical evidence is that, in contrast to the estimate of γ, the estimate of δ is quite sensitive to idiosyncratic observation error in the consumption level.

3.3.6 GMM Results

The Hansen-Jagannathan volatility bound analysis and the calibration results show that the complete consumption insurance model does not perform well when the standard power utility model is used as well as when a Taylor series expansion to the representative-agent's marginal utility of consumption around the conditional expectation of aggregate consumption per capita is taken. However, taking into account asymmetry of the cross-sectional distribution of individual consumption allows to explain both the equity premium and the return on the risk-free asset with economically plausible values of the RRA coefficient and the time discount factor. Given this result, in this Section, we limit our analysis to the incomplete consumption insurance case.

An iterated GMM approach is used to test Euler equations and estimate model parameters. We estimate the Euler equation for the excess value-weighted market return as

$$E_t\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)RP_{t+1}\right]=0 \tag{135}$$

and the Euler equation for the gross return on the real risk-free interest rate as

$$\delta E_t\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)R_{f,t+1}\right]= \\ \left(\frac{h_{t+1}}{h_t}\right)^{\gamma}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t}}{h_t^n}\right) \tag{136}$$

jointly exploiting three sets of instruments. The first instrument set (INSTR1) consists of a constant, the real value-weighted and equal-weighted market returns, the real risk-free rate, and the real consumption growth rate lagged one period. The second set of instruments (INSTR2) is the first set extended with the same variables lagged an additional period. The third set (INSTR3) has a constant, the real value-weighted and equal-weighted market returns, the real risk-free rate, and the real consumption growth rate lagged one, two, and three periods.

To study the role of the first three moments of the cross-sectional distribution of individual consumption in explaining the equity premium and risk-free rate puzzles, we truncate the series expansion after the cubic term and estimate the Euler equations with $N = 3$. The model is able to fit the excess return on the market portfolio and the risk-free rate with an economically plausible (between 1 and 2) and statistically significant value of risk aversion for any set of households whatever the set of instruments (see Table III). The estimate of the RRA coefficient is decreasing in asset holdings, as anticipated. According to Hansen's test of the overidentifying restrictions, the model is not rejected statistically at the 5% level.

3.4 Concluding Remarks

Empirical evidence suggests that the complete consumption insurance model fails to fit the observed equity premium with an economically plausible value of risk aversion when the standard power utility model is used as well as when a Taylor series expansion to the representative-agent's marginal utility of consumption around the conditional expectation of aggregate consumption per capita is taken. This result is robust to the threshold value in the definition of assetholders and the used analysis method.

When aggregate shocks are assumed to be uninsurable, the impact of incomplete consumption insurance on the expected equity premium is mixed. Thus, we find that the cross-moments of the excess market portfolio return with the cross-sectional variance and skewness of individual consumption are both positive. It follows that the cross-sectional variance of individual consumption represents the effect of complementarity in the portfolio risk and the background risk in wealth, while the cross-sectional skewness of individual consumption represents the effect of substitutability. The empirical results show that the effect of substitutability dominates.[76]

Another important result is that both the equity premium and the risk-free rate may be explained with economically plausible values of the RRA coefficient and the time discount factor when asymmetry of the cross-sectional distribution of individual consumption is taken into account. This result is robust to the threshold value in the definition of assetholders.

Appendix A: Matching Consumer Units between the 1985 and 1986 Data Tapes

In the CEX, each household is interviewed every three months over five consecutive quarters. The initial interview collects demographic and family characteristics and is not placed on the tape. Each quarter, consumer units that have completed their final interview in the previous quarter (about one-fifth of the sample) are replaced by new households introduced for the first time. The households remained on the tape complete their participation. For the consumer units completing their participation in the first through third quarters of 1986, BLS has changed beginning the first quarter of 1986 the consumer unit identification numbers (NEWID). As a result, the NEWIDs for the same household in 1985 (when this household has been interviewed for the first time) and in 1986 (when he has completed his participation) are no longer the same.

To match consumer units between the 1985 and 1986 data tapes, we use the following household characteristics:

AGE_REF - age of reference person
COMP1 - number of males age 16 and over in family
COMP2 - number of females age 16 and over in family
COMP3 - number of males age 2 through 15 in family
COMP4 - number of females age 2 through 15 in family
COMP5 - number of members under age 2 in family
BLS_URBN - area of residence (urban/rural)
BUILDING - description of building

[76]Taking into account that the CRRA preferences exhibit decreasing and convex absolute risk aversion and decreasing absolute prudence, this result confirms the findings in Weil (1992) and Gollier (2001).

EDUC_REF - education of reference person

MARITAL1 - marital status of reference person

ORIGIN1 - origin or ancestry of reference person

POPSIZE - population size of the primary sampling unit

REF_RACE - race of reference person

REGION - region (for urban areas only)

SEX_REF - sex of reference person.

The values of the variables SEX_REF, ORIGIN1, and REF_RACE must be the same for the same household on both the 1985 and 1986 data tapes. Moreover, the CEX is constructed so that the values of the variables BLS_URBN, REGION, and BUILDING are also the same for the same consumer unit over all interviews.

As a rule, the variable POPSIZE also has the same value for the same consumer unit over all interviews. However, on the 1985 and 1986 data tapes, this variable is coded differently:

	1985	1986
1	More than 4 million	More than 4 million
2	1.25 million - 4 million	1.20 million - 4 million
3	0.385 - 1.249 million	0.33 - 1.19 million
4	75 - 384.9 thousand	75 - 329.9 thousand
5	Less than 75 thousand	Less than 75 thousand

It follows that in the case of these two years, the same consumer unit has the same code for POPSIZE only if in 1985 this code is 1, 2, or 5. If a consumer unit has in 1985 the code 3, then it can have in 1986 the code 2 or 3. Households having in 1985 the code 4 can have the code 3 or 4 in 1986. It is valid for all households living in the Northeast, the Midwest, and the South region. For consumer units residing in the West, the variable POPSIZE is suppressed by BLS on the 1986 data tape.

It is more difficult to deal with the variables MARITAL1, EDUC_REF, COMP1, COMP2, COMP3, COMP4, and COMP5 which can take different values over interviews. For these variables, we determine the set of all possible values that they can take in 1986 given their values in 1985.

The variable MARITAL1 is coded as follows:

1 Married

2 Widowed

3 Divorced

4 Separated

5 Never married

If a reference person is married today, then tomorrow he may be either married or widowed, or divorced, or separated. If he is widowed or divorced, he may either keep this status or be married. In the case, when a reference person is separated, he remains to be separated or else becomes to be widowed or divorced. A never married person can either keep this marital status or be married.

Let $COMP1_{h,t}$ denote the number of males age 16 and over in family h in period t, $COMP2_{h,t}$ - the number of females age 16 and over, $COMP3_{h,t}$ - the number of males age 2 through 15, $COMP4_{h,t}$ - the number of females age 2 through 15, $COMP5_{h,t}$ - the number of members under age 2 in family, $SUM\_T_{h,s} = COMP3_{h,s} + COMP4_{h,s} + COMP5_{h,s}$ - the total number of children age 15 and under in family in any period s ($s > t$, $t = 2, 3, 4$), $SUM\_M_{h,s} = COMP1_{h,s} + COMP3_{h,s}$ and $SUM\_F_{h,s} = COMP2_{h,s} + COMP4_{h,s}$ - the total number of males age 2 and over and the total number of females age 2 and over, respectively.

We apply the following restrictions:

$$COMP5_{h,t} \leqslant SUM\_T_{h,s} \tag{A.1}$$

and

$$COMP3_{h,t} \leqslant SUM\_M_{h,s} \leqslant COMP1_{h,t} + COMP3_{h,t} + COMP5_{h,t}. \tag{A.2}$$

It follows that

$$0 \leqslant SUM\_M_{h,s} - COMP3_{h,t} \leqslant COMP1_{h,t} + COMP5_{h,t}. \tag{A.3}$$

Similarly, for $SUM\_F_{h,s}$, we get

$$COMP4_{h,t} \leqslant SUM\_F_{h,s} \leqslant COMP2_{h,t} + COMP4_{h,t} + COMP5_{h,t} \tag{A.4}$$

and

$$0 \leqslant SUM\_F_{h,s} - COMP4_{h,t} \leqslant COMP2_{h,t} + COMP5_{h,t}. \tag{A.5}$$

Finally,

$$\begin{aligned} SUM\_M_{h,s} + SUM\_F_{h,s} \leqslant \; & COMP1_{h,t} + COMP2_{h,t} + COMP3_{h,t} \\ & + COMP4_{h,t} + COMP5_{h,t}. \end{aligned} \tag{A.6}$$

In the CEX, the variable EDUC_REF is coded as follows:

1 Elementary (1-8 years)

2 High school (1-4 years), less than High school graduate

3 High school graduate (4 years)

4 College (1-4 years), less than College graduate

5 College graduate (4 years)

6 More than 4 years of college

7 Never attended school

Changing the code from 7 to 0 enables us to introduce the following restriction:

$$EDUC\_REF_{h,s} - EDUC\_REF_{h,t} \leqslant 1. \tag{A.7}$$

We apply the following restriction to the value of the variable AGE_REF:

$$AGE\_REF_{h,t} \leqslant AGE\_REF_{h,s}. \tag{A.8}$$

We define three groups of households. The first group consists of consumer units that should have the following stream of interviews: the second interview in the second quarter of 1985, the third - in the third quarter of 1985, the fourth - in the fourth quarter of 1985, and the fifth interview in the first quarter of 1986. For this group, we construct two data sets. In the 1985 data set, we include all households who completed at least one of the interviews in 1985. The 1986 data set consists of households who completed their fifth interview in the first quarter of 1986. As a second group are considered consumer units whose second interview should happened in the third quarter of 1985, the third - in the fourth quarter of 1985, the fourth - in the first quarter of 1986, and the fifth interview should happened in the second quarter of 1986. For this group, the 1985 data set includes all consumer units who completed at least one interview in 1985 and the 1986 data set consists of households who completed at least one of the interviews in 1986. In the third group of households, we include consumer units that should have the second interview in the fourth quarter of 1985, the third - in the first quarter of 1986, the fourth - in the second quarter of 1986, and the fifth interview in the third quarter of 1986. For this group of consumer units, the 1985 data set consists of all households who completed their second interview in the fourth quarter of 1985 and the 1986 data set consists of consumer units who completed in 1986 at least one of the interviews.

So, for each group of consumer units there are two data sets (one for 1985 and another one for 1986). To find the same household in both data sets, each consumer unit included in the 1986 data set is compared with each household included in the 1985 data set for the same group of households. As a result, for the first group of consumer units, there are 521 households (in 1019 included in the 1986 data set, 51.1%), for which we find households with the same characteristics in the corresponding 1985 data set. For the second group, this number is 756 (in 1612 consumer units, 49.6%). For the third group, we match 807 households (in 1800, 44.8%). For all the three groups, the number of matched households is 2084 (in 4431 included in the 1986 data sets, 47.0%).[77]

[77] In order to see how well this procedure works, we test it using the 1986 and 1987 data tapes (in both these years, the same household has the same NEWID) for the first group of consumer units constructed in the same way as in the case of using the 1985 and 1986 data tapes. We have 1465 households who completed their fifth interview in the first quarter of 1987 and 1787 consumer units included in the 1986 data set. Only 1344 in 1465 households have at least one interview happened in 1986. As a result, we match 1281 households (95.3% of the real number). In 96% of the cases, there is only one household in the 1986 data set corresponding to the household included in the 1987 data set. In only 4% of the cases, there are more than 1 (between 2 and 4) corresponding households.

Appendix B: Tables

Table I.

Parameter Estimates and Tests of Model $\Delta c_{t+1} = g + \eta_{t+1}$

The sampling period is from 1979:10 to 1996:2. Four sets of households are considered. The first set (SET1) consists of all consumer units irrespectively of the reported market value of all securities. We also use three sets of households classified as assetholders: SET2 consists of households who reported asset holdings equal to or exceeding \$2 in 1999 dollars, the two other sets consist of households who reported total assets equal to or exceeding \$10000 (SET3) and \$20000 (SET4). The model is estimated by ML. Standard errors in parentheses.

Parameters	SET1	SET2	SET3	SET4
g	0.0016	0.0022	0.0027	0.0021
	(0.0014)	(0.0023)	(0.0029)	(0.0032)
σ_η^2	0.0004	0.0010	0.0016	0.0020

Table II.

Values of the RRA Coefficient Allowing to Fit the Equity Premium

Four sets of households are considered. The first set (SET1) consists of all consumer units irrespectively of the reported market value of all securities. We also use three sets of households classified as assetholders: SET2 consists of households who reported asset holdings equal to or exceeding $2 in 1999 dollars, the two other sets consist of households who reported total assets equal to or exceeding $10000 (SET3) and $20000 (SET4). Panel A provides the results under the assumption of complete consumption insurance. In panel B, we report the results under incomplete consumption insurance. Under the assumption of incomplete consumption insurance, we present the results for both the raw consumption data (the row "raw consumption data") and under the assumption that 20% of the observed cross-sectional variance of individual consumption is noise (the row "allowing for observation error"). The sign "-" means that there is no positive value of the RRA coefficient allowing to explain the observed mean equity premium (the unexplained mean equity premium remains positive for all considered values of risk aversion and increases as the RRA coefficient rises).

Model	Parameters	SET1	SET2	SET3	SET4
Panel A: Complete consumption insurance					
1. Standard power utility model	γ	36.00	51.00	30.00	36.00
2. First-order Taylor series expansion	γ	113.00	76.00	27.00	22.40
3. Second-order Taylor series expansion		-	-	-	-
4. Third-order Taylor series expansion	γ	36.00	63.00	23.40	23.20
Panel B: Incomplete consumption insurance					
1. First-order Taylor series expansion	γ	113.00	76.00	27.00	22.40
2. Second-order Taylor series expansion		-	-	-	-
3. Third-order Taylor series expansion:					
- raw consumption data	γ	1.06	1.55	1.38	1.41
	δ	0.9840	1.3098	1.0073	0.9169
- allowing for observation error	γ	1.13	1.75	1.45	1.51
	δ	0.9847	0.8895	1.0095	0.9141

Table III.

GMM Results under Incomplete Consumption Insurance. Third-Order Taylor Series Expansion

The sampling period is from 1979:10 to 1996:2. Four sets of households are considered. The first set (SET1) consists of all consumer units irrespectively of the reported market value of all securities. We also use three sets of households classified as assetholders: SET2 consists of households who reported asset holdings equal to or exceeding \$2 in 1999 dollars, the two other sets consist of households who reported total assets equal to or exceeding \$10000 (SET3) and \$20000 (SET4). The agent's marginal utility of consumption is expanded as a Taylor series up to cubic terms ($N = 3$). The Euler equations for the excess value-weighted and equal-weighted market returns are estimated jointly with the Euler equation for the real risk-free interest rate using an iterated GMM approach (standard errors in parentheses). Three sets of instruments are exploited. The first instrument set (INSTR1) consists of a constant, the real value-weighted and equal-weighted market returns, the real risk-free rate, and the real consumption growth rate lagged one period. The second set of instruments (INSTR2) is the first set extended with the same variables lagged an additional period. The third set (INSTR3) has a constant, the real value-weighted and equal-weighted market returns, the real risk-free rate, and the real consumption growth rate lagged one, two, and three periods. The J statistic is Hansen's test of the overidentifying restrictions. The P value is the marginal significance level associated with the J statistic.

Parameters	SET1	SET2	SET3	SET4
		INSTR1		
γ	1.2126	1.1282	0.0715	0.0848
	(0.0505)	(0.0780)	(0.0189)	(0.0223)
δ	0.9901	1.0113	0.9978	0.9977
	(0.2349)	(0.0223)	(0.0005)	(0.0005)
J statistic	4.9871	7.1099	7.6519	7.7094
P value	0.7590	0.5248	0.4682	0.4624
		INSTR2		
γ	1.1495	1.4911	1.3999	1.4609
	(0.0296)	(0.0574)	(0.0554)	(0.0584)
δ	0.9878	0.9854	1.0138	0.9733
	(0.0517)	(0.0971)	(0.0595)	(0.1049)
J statistic	6.5296	8.9943	8.2458	7.6449
P value	0.9813	0.9137	0.9412	0.9587
		INSTR3		
γ	1.0611	0.0128	0.0077	0.0091
	(0.0206)	(0.0016)	(0.0014)	(0.0011)
δ	0.9977	0.9988	0.9982	0.9979
	(0.0231)	(0.0001)	(0.0001)	(0.0001)
J statistic	8.9334	9.4825	9.5840	9.6172
P value	0.9977	0.9964	0.9961	0.9960

References

[1] Aiyagari, R.S. and M. Gertler, 1991, Asset Returns with Transactions Costs and Uninsured Individual Risk, Journal of Monetary Economics 27, 311-331.

[2] Balduzzi, P. and T. Yao, 2000, Does Heterogeneity Matter for Asset Pricing?, working paper.

[3] Bansal, R., D.A. Hsieh, and S. Viswanathan, 1993, A New Approach to International Arbitrage Pricing, Journal of Finance 48, 1719-1747.

[4] Bertaut, C.C., 1998, Stockholding Behavior of U.S. Households: Evidence from the 1983-1989 Survey of Consumer Finances, Review of Economics and Statistics 80, 263-275.

[5] Bewley, T.F., 1982, Thoughts on Tests of the Intertemporal Asset Pricing Model, working paper, Evanston, Ill.: Northwestern University.

[6] Blume, M.E. and S.P. Zeldes, 1993, The Structure of Stock Ownership in the U.S., working paper, Philadelphia, University of Pennsylvania.

[7] Brav, A., G.M. Constantinides, and C.C. Géczy, 2002, Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence, Journal of Political Economy, forthcoming.

[8] Brav, A. and C.C. Géczy, 1995, An Empirical Resurrection of the Simple Consumption CAPM with Power Utility, working paper, University of Chicago.

[9] Campbell, J. and J. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, Journal of Political Economy 107, 205-251.

[10] Campbell, J.Y., A.W. Lo, and A.C. MacKinlay, 1997, The Econometrics of Financial Markets, Princeton University Press.

[11] Carroll, C.D., 1994, How Does Future Income Affect Current Consumption, Quarterly Journal of Economics 109, 111-147.

[12] Cogley, T., 2002, Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey, Journal of Monetary Economics 49, 309-334.

[13] Constantinides, G.M. and D. Duffie, 1996, Asset Pricing with Heterogeneous Consumers, Journal of Political Economy 104, 219-240.

[14] Davidson, R. and J.G. MacKinnon, 1993, Estimation and Inference in Econometrics, Oxford University Press.

[15] Dittmar, R.F., 2002, Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns, Journal of Finance 57, 369-403.

[16] Ferson, W.E. and C.R. Harvey, 1992, Seasonality and Consumption-Based Asset Pricing, Journal of Finance 47, 511-552.

[17] Gollier, C., 2001, The Economics of Risk and Time, The MIT Press, Cambridge, Massachusetts, London, England.

[18] Gollier, C. and J.W. Pratt, 1996, Risk Vulnerability and the Tempering Effect of Background Risk, Econometrica 64, 1109-1124.

[19] Haliassos, M. and C.C. Bertaut, 1995, Why Do So Few Hold Stocks?, The Economic Journal 105, 1110-1129.

[20] Hansen, L.P. and R. Jagannathan, 1991, Implications of Security Market Data for Models of Dynamic Economies, Journal of Political Economy 99, 225-262.

[21] Hansen, L.P. and K. Singleton, 1982, Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, Econometrica 50, 1269-1286.

[22] Hansen, L.P. and K. Singleton, 1983, Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns, Journal of Political Economy 91, 249-265.

[23] Hansen, L.P. and K. Singleton, 1984, Errata, Econometrica 52, 267-268.

[24] Heaton, J. and D.J. Lucas, 1992, The Effects of Incomplete Insurance Markets and Trading Costs in a Consumption-Based Asset Pricing Model, Journal of Economic Dynamics and Control 16, 601-620.

[25] Heaton, J. and D.J. Lucas, 1995, The Importance of Investor Heterogeneity and Financial Market Imperfections for the Behavior of Asset Prices, Carnegie-Rochester Ser. Public Pol.

[26] Heaton, J. and D.J. Lucas, 1996, Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing, Journal of Political Economy 104, 443-487.

[27] Huggett, M., 1993, The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies, Journal of Economic Dynamics and Control 17, 953-969.

[28] Jacobs, K., 1999, Incomplete Markets and Security Prices: Do Asset-Pricing Puzzles Result from Aggregation Problems?, Journal of Finance 54, 123-163.

[29] Jacobs, K. and K.Q. Wang, 2001, Idiosyncratic Consumption Risk and the Cross-Section of Asset Returns, working paper.

[30] Kimball, M.S., 1990, Precautionary Savings in the Small and in the Large, Econometrica 58, 53-73.

[31] Kimball, M.S., 1993, Standard Risk Aversion, Econometrica 61, 589-611.

[32] Lucas, D.J., 1994, Asset Pricing with Undiversifiable Income Risk and Short Sales Constraints: Deepening the Equity Premium Puzzle, Journal of Monetary Economics 34, 325-341.

[33] Mankiw, N.G., 1986, The Equity Premium and the Concentration of Aggregate Shocks, Journal of Financial Economics 17, 211-219.

[34] Mankiw, N.G. and S.P. Zeldes, 1991, The Consumption of Stockholders and Nonstockholders, Journal of Financial Economics 29, 97-112.

[35] Mehra, R. and E.C. Prescott, 1985, The Equity Premium: A Puzzle, Journal of Monetary Economics 15, 145-162.

[36] Pratt, J.W. and R.J. Zeckhauser, 1987, Proper Risk Aversion, Econometrica 55, 143-154.

[37] Ross, S.A., 1999, Adding Risks: Samuelson's Fallacy of Large Numbers Revisited, Journal of Financial and Quantitative Analysis 34, 323-340.

[38] Runkle, D.E., 1991, Liquidity Constraints and the Permanent-Income Hypothesis: Evidence from Panel Data, Journal of Monetary Economics 27, 73-98.

[39] Samuelson, P.A., 1963, Risk and Uncertainty: the Fallacy of the Law of Large Numbers, Scientia 98, 108-113.

[40] Storesletten, K., C.I. Telmer, and A. Yaron, 1997, Persistent Idiosyncratic Shocks and Incomplete Markets, unpublished manuscript.

[41] Telmer, C.I., 1993, Asset-Pricing Puzzles and Incomplete Markets, Journal of Finance 48, 1803-1832.

[42] Vissing-Jorgensen, A., 1998, Limited Stock Market Participation, working paper.

[43] Weil, P., 1992, Equilibrium Asset Prices with Undiversifiable Labor Income Risk, Journal of Economic Dynamics and Control 16, 769-790.

[44] Zeldes, S., 1989, Consumption and Liquidity Constraints: An Empirical Investigation, Journal of Political Economy 97, 305-346.

4 An Empirical Assessment of a Consumption CAPM with a Reference Level under Incomplete Consumption Insurance

4.1 Introduction

It is common in consumption-based asset pricing to assume the existence of a representative consumer and to use aggregate consumption per capita in place of the consumption of any particular agent. Empirical tests of the representative-agent model reject the model in several ways. Thus, Mehra and Prescott (1985) show that the representative-agent model is not able to explain the observed average excess return on the US stock market unless risk aversion is assumed to be implausibly high (the equity premium puzzle). The large estimate of risk aversion implies another puzzle: if investors are extremely risk-averse, then the observed average growth rate of per capita consumption is consistent with the low short-term real interest rate only if the representative agent has a negative rate of time preference. This is the risk-free rate puzzle (Weil (1989)). Ferson and Constantinides (1991), Hansen and Jagannathan (1991), and Hansen and Singleton (1982, 1983) test the conditional Euler equations for an assumed representative agent and find that the overidentifying restrictions strongly reject the model when the Euler equations for the equity premium and the risk-free rate are estimated jointly.

Since Mehra and Prescott's original investigation, several generalizations of essential features of the representative-agent model have been suggested to mitigate its poor empirical performance. Thus, Brav, Constantinides, and Géczy (2002), Constantinides and Duffie (1996), Mankiw (1986), Mehra and Prescott (1985), Semenov (2002), and Weil (1992) suggest that deviations from complete consumption insurance have the potential to explain the equity premium and risk-free rate puzzles. In particular, Semenov (2002) develop an approximate equilibrium multifactor model for expected returns in which the priced risk factors are cross-moments of return with the moments of individual consumption and find that the model can explain both the equity premium and the risk-free rate with economically plausible (less than 2) values of risk aversion and the time discount factor when the agent's marginal utility of consumption is expanded as a Taylor series up to terms capturing the skewness of the distribution of individual consumption around its conditional expectation.

Another possible explanation of empirical rejections of the representative-agent model is excessive rigidity of the conventional time- and state-separable utility function which constrains the elasticity of intertemporal substitution to be the reciprocal of the RRA coefficient. Constantinides (1990) studies an internal habit model in which the utility is a power of the difference between the current consumption flow and a fraction of a weighted sum of lagged consumption flows and proves that habit persistence and/or durability of consumption drives a wedge between the elasticity of consumption with respect to investment returns and the inverse of the RRA coefficient. Epstein and Zin (1989, 1991), hereafter EZ, assume that the agent's lifetime utility depends on both current consumption and a certainty equivalent of a random future utility through an intertemporal constant elasticity of substitution utility function. For the certainty equivalent, EZ (1989, 1991) consider a constant relative risk aversion expected utility specification. This generalized specification of intertemporal utility allows a separation of risk aversion (reflected in the certainty equivalent function) from intertemporal substitution (encoded in the aggregator function).

Garcia, Renault, Semenov (2002), henceforth GRS, propose another way to disentangle intertemporal substitution and risk aversion. They assume an agent to derive utility from both the ratio of his consumption to some benchmark level of consumption and this level itself. They show that if the external reference level matters for a decision maker and the reference consumption level growth rate is correlated with the market portfolio return, this expected utility model has the ability to explain both the equity premium and the risk-free rate as well as to separate the elasticity of intertemporal substitution from the inverse of the RRA coefficient. An important result is that if the reference consumption level growth rate is assumed to be a function of the market portfolio return alone, this utility specification yields a SDF which is isomorphic in its pricing implications to the pricing kernel corresponding to the EZ (1989, 1991) non-expected utility specification. The comparison between the EZ (1989, 1991) non-expected utility model and the GRS (2002) expected utility model with a reference level shows that the elasticity of intertemporal substitution remains the same in the two models, while the measure of risk aversion in the GRS (2002) utility specification differs from that in the EZ (1989, 1991) utility model. An attractive feature of the GRS (2002) preference specification is that, in contrast to the Constantinides (1990) internal habit model, consumption is not required to be always above the reference level for marginal utility to be positive. GRS (2002) test this utility function under the assumption of complete consumption insurance and obtain the point estimate of the elasticity of intertemporal substitution that is in the conventional range and statistically different from the inverse of the RRA coefficient. Besides, their empirical result is that the SDF corresponding to the preference specification with a reference level outperforms the EZ (1989, 1991) pricing kernel.

The goal of this paper is to examine the asset pricing implications of the preference specification with a reference level under the assumptions of incomplete consumption insurance and limited participation of consumers in the asset markets using the approximate equilibrium multifactor model for expected asset returns developed in Semenov (2002). The common to all agents contemporaneous macroeconomic factors posited to affect the reference level are assumed to be adequately proxied by the level of aggregate consumption per capita. Assuming further the substistence level to response gradually to changes in aggregate consumption per capita, we use the following two-stage procedure to estimate the parameters of interest. In the first step, we estimate sensitivity of the reference level to changes in aggregate consumption per capita. The second step is to use the iterated generalized method of moments (GMM) approach to estimate the conditional Euler equations for the equity premium and the risk-free rate of return implied by the Semenov (2002) approximate equilibrium multifactor model for expected asset returns and the GRS (2002) preference specification using the estimate of the speed of adjustment parameter obtained in the first step.

The remaining of this paper is organized as follows. In Section 4.2, we briefly review the major features of the approximate equilibrium multifactor model for expected asset returns developed in Semenov (2002). Section 4.3 details the preference specification with a reference level responding gradually to changes in aggregate consumption per capita. Section 4.4 describes the data, estimation and testing methodology and presents estimation results. Section 4.5 concludes.

4.2 The Equilibrium Multifactor Pricing Model

Consider the intertemporal consumption and portfolio choice problem of a single representative investor who can trade freely in asset i and who maximizes expected lifetime discounted utility

$$MaxE_t\left[\sum_{j=0}^{\infty}\delta^j u\left(C_{k,t+j}\right)\right] \tag{137}$$

subject to his budget constraint

$$C_{k,t+1} - R_{i,t+1}(W_{k,t} - C_{k,t}) = 0, \tag{138}$$

where δ is the subjective discount factor, $C_{k,t+j}$ is the investor's consumption in period $t+j, u\left(C_{k,t+j}\right)$ is the one-period utility of consumption at $t+j$, $W_{k,t}$ is the investor's welfare in period t, $R_{i,t+1}$ is the simple gross return on asset i, and $E_t\left[\cdot\right]$ denotes the mathematical expectation conditioned on the period-t information set, Ω_t, that is common to all agents.[78]

The first-order condition describing the investor's optimal consumption and portfolio plan is

$$\delta E_t\left[u'\left(C_{k,t+1}\right)R_{i,t+1}\right] = u'\left(C_{k,t}\right),\ k=1,...,K,\ i=1,...,I. \tag{139}$$

The right side of (139) is the loss in utility if the investor buys another unit of the asset, the left side is the increase in discounted, expected utility he obtains from the extra payoff at time $t+1$. Hence, in the optimum the investor equates the marginal loss and the marginal gain from holding of his portfolio.

Assuming $u\left(\cdot\right)$ to be $N+1$ times differentiable, Semenov (2002) uses a N-order Taylor expansion to the individual k's marginal utility around the conditional expectation of consumption, $h_{t+1} \equiv E_t\left[C_{k,t+1}\right]$:[79]

$$u'\left(C_{k,t}\right) = \sum_{n=0}^{N}\frac{1}{n!}u^{(n+1)}\left(h_t\right)\left(C_{k,t}-h_t\right)^n,\ k=1,...,K. \tag{140}$$

Substituting for $u'\left(C_{k,t}\right)$ from (140), we obtain

$$\delta\sum_{n=0}^{N}\frac{1}{n!}u^{(n+1)}\left(h_{t+1}\right)E_t\left[\left(C_{k,t+1}-h_{t+1}\right)^n R_{i,t+1}\right] = \sum_{n=0}^{N}\frac{1}{n!}u^{(n+1)}\left(h_t\right)\left(C_{k,t}-h_t\right)^n, \tag{141}$$

$k=1,...,K,\ i=1,...,I.$

[78] $u\left(\cdot\right)$ is assumed to be increasing, strictly concave, and differentiable.

[79] Here, and throughout the paper, $u^{(n)}\left(\cdot\right)$ denotes the nth derivative of $u\left(\cdot\right)$.

By summing these equations over individuals and dividing by the number of individuals in the population, we obtain the following set of equations:

$$\delta \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_{t+1}\right) E_t \left[Z_{n,t+1} R_{i,t+1}\right] = \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_t\right) Z_{n,t}, \tag{142}$$

$i = 1, ..., I$, where $Z_{n,t} \equiv \frac{1}{K} \sum_{k=1}^{K} \left(C_{k,t} - h_t\right)^n$.

Equation (142) can be rewritten as

$$E_t \left[R_{i,t+1}\right] = \delta^{-1} \sum_{n=0}^{N} \frac{1}{n!} \frac{u^{(n+1)} \left(h_t\right)}{u' \left(h_{t+1}\right)} Z_{n,t} - \sum_{n=1}^{N} \frac{1}{n!} \frac{u^{(n+1)} \left(h_{t+1}\right)}{u' \left(h_{t+1}\right)} \times E_t \left[Z_{n,t+1} R_{i,t+1}\right]. \tag{143}$$

This is the approximate equilibrium multifactor model for expected asset returns.[80]

For the expected excess return on the market portfolio over the risk-free rate, $RP_{t+1} \equiv R_{M,t+1} - R_{f,t+1}$, equation (142) reduces to

$$\sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_{t+1}\right) E_t \left[Z_{n,t+1} RP_{t+1}\right] = 0. \tag{144}$$

The Euler equation for the expected equilibrium risk-free rate is

$$\delta \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_{t+1}\right) E_t \left[Z_{n,t+1} R_{f,t+1}\right] = \sum_{n=0}^{N} \frac{1}{n!} u^{(n+1)} \left(h_t\right) Z_{n,t}. \tag{145}$$

4.3 Preferences

Following GRS (2002), consider a single investor whose period t utility function is given by

$$u \left(C_{k,t}, S_{k,t}\right) = \frac{\left(\frac{C_{k,t}}{S_{k,t}}\right)^{1-\gamma} S_{k,t}^{1-\varphi} - 1}{1-\gamma}. \tag{146}$$

Here, $S_{k,t}$ is the agent's k time-varying subsistence or reference consumption level in period t, γ is the Arrow-Pratt measure of relative risk aversion, and the parameter φ controls the curvature of utility over the reference level of consumption.

Utility function (146) nests some preference specifications which can be obtained given different values of the curvature parameter φ. Thus, if $\varphi = \gamma$, the reference consumption level plays no role in asset pricing and we get the standard time-separable power utility model. When $\varphi < \gamma$, an increase in the reference level raises the marginal utility of the agent's own consumption. Gali (1994) refers to this type of externalities as positive consumption

[80] See Semenov (2002).

externalities. Alternatively, when $\varphi > \gamma$, an increase in the benchmark level lowers the marginal utility of consumption. These are negative consumption externalities.[81] With $\varphi = 1$, we obtain the ratio preference specification when the agent derives utility from consumption relative to the benchmark level. If $\varphi \neq \gamma$ and $\varphi \neq 1$, then the agent takes into account both the ratio of his consumption to the subsistence level and this level itself when choosing how much to consume.

Assume the time-varying substistence level to be unaffected by any one agent's consumption decisions. GRS (2002) show that if the reference consumption level is exogenous to an individual consumer, this utility specification not only has the potential to explain the equity premium and risk-free rate puzzles but also allows to disentangle intertemporal substitution and risk aversion.[82] In particular, they show that in this model the intertemporal

[81] In the special case, when the reference level is proxied by past consumption levels, positive consumption externalities are usually referred to as habit persistence in preferences, while negative consumption externalities correspond to durability in consumption expenditures (see, for example, Eichenbaum and Hansen (1990), Eichenbaum, Hansen, and Singleton (1988), Ferson and Constantinides (1991), Gallant and Tauchen (1989), and Heaton (1995)).

[82] Since the reference consumption level is external, the SDF is then

$$M_{t+1} \equiv \delta \frac{u'\left(C_{k,t+1}, S_{k,t+1}\right)}{u'\left(C_{k,t}, S_{k,t}\right)} = \delta \left(\frac{C_{k,t+1}}{C_{k,t}}\right)^{-\gamma} \left(\frac{S_{k,t+1}}{S_{k,t}}\right)^{\gamma-\varphi}. \tag{147}$$

Under the assumption that there is an representative agent, we can rewrite (147) as

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}. \tag{148}$$

Assuming joint conditional lognormality and homoskedasticity of the consumption growth rate and asset returns, GRS (2002) obtain

$$r_{f,t+1} = -log\delta + \gamma E_t\left[\Delta c_{t+1}\right] - \frac{1}{2}\gamma^2\sigma_c^2 - (\gamma-\varphi) E_t\left[\Delta s_{t+1}\right] - \frac{1}{2}(\gamma-\varphi)^2\sigma_s^2 + \gamma(\gamma-\varphi)\sigma_{cs} \tag{149}$$

and

$$E_t\left[r_{i,t+1} - r_{f,t+1}\right] = -\frac{1}{2}\sigma_i^2 + \gamma\sigma_{ic} - (\gamma-\varphi)\sigma_{is}, \tag{150}$$

where Δc_{t+1} is the log of the consumption growth rate, Δs_{t+1} is the log of the reference consumption level growth rate, $r_{i,t+1}$ is the log of the simple gross return on asset i, and σ_{xy} denotes the unconditional covariance of innovations. The first three terms on the right-hand side of (149) and the first two terms on the right-hand side of (150) are the same as for a time-separable power utility function of consumption alone. Thus, utility function (146) has the ability to explain the equity premium puzzle if the term $-(\gamma-\varphi)\sigma_{is}$ is positive and the risk-free rate puzzle if the term $-(\gamma-\varphi)E_t\left[\Delta s_{t+1}\right] - \frac{1}{2}(\gamma-\varphi)^2\sigma_s^2 + \gamma(\gamma-\varphi)\sigma_{cs}$ is negative.

elasticity of consumption is

$$\sigma_k = \frac{\partial \triangle c_{k,t+1}}{\partial r_{i,t+1}} = \frac{1 + (\gamma - \varphi)\frac{\partial \triangle s_{k,t+1}}{\partial r_{i,t+1}}}{\gamma}, \tag{151}$$

where $\frac{\partial \triangle s_{k,t+1}}{\partial r_{i,t+1}}$ can be interpreted as the elasticity of the subsistence level with respect to investment returns. Equation (151) implies that the elasticity of intertemporal substitution differs from the inverse of the RRA coefficient if $(\gamma - \varphi)\frac{\partial \triangle s_{k,t+1}}{\partial r_{i,t+1}} \neq 0$.[83][84]

Since the reference level of consumption is not observable, this model is of little use without specifying a way to measure the factors that are posited to affect subsistence requirements. In GRS (2002), it is shown that, given different assumptions about the reference level generating process, the pricing kernel corresponding to preference specification (146) nests several the most often used in asset pricing SDFs. Thus, when the benchmark level of consumption is assumed to be determined by past consumption levels only, the model generalizes the usual external habit formation specifications. One reasonable approach is to assert that the agent's reference level could be affected not only by past consumption, but also by some contemporaneous macro- and microeconomic factors such as business cycles, inflation, age of reference person, his education, marital status, etc. GRS (2002) demonstrate that if we assume that the return on the market portfolio is a valid proxy for the common to all agents macroeconomic factors and the benchmark consumption level does not depend on past consumption, preference specification (146) yields a SDF which is observationally equivalent to the pricing kernel corresponding to the EZ (1989, 1991) non-expected recursive utility function.[85]

---

[83]Since $\frac{\partial \triangle s_{t+1}}{\partial r_{i,t+1}}$ and σ_{is} have the same sign, if utility specification (146) contributes towards a solution of the equity premium puzzle, it also yields the elasticity of intertemporal substitution which is less than the inverse of the risk aversion coefficient (see equations (150) and (151)).

[84]Another example of the utility specification allowing to separate the elasticity of intertemporal substitution from risk aversion in the expected utility framework is the Ferson-Constantinides (1991) internal habit model, in which the utility is a power function of the difference between the current consumption flow and a fraction of a weighted sum of lagged consumption flows. However, this model is restrictive in that consumption must always be above habit for marginal utility to be positive, what is not required in model (146).

[85]Empirical evidence in GRS (2002) is that when the representative agent's reference consumption level is assumed to depend on both the market portfolio return and lagged aggregate consumption per capita, we are able to fit empirical data on asset returns with economically plausible and statistically significant

In this paper, we assume that the common to all agents contemporaneous macroeconomic factors posited to affect the reference level may be adequately proxied by the level of aggregate consumption per capita. Assume further that the substistence level responses gradually to changes in aggregate consumption per capita and the dynamics of $\{logS_{k,t+1}\}$ are given by the following equation:

$$logS_{k,t+1} = a_{k,t+1} + (1 - \lambda_k)\, logS_{k,t} + \lambda_k logC_{t+1},\ 0 \leqslant \lambda_k \leqslant 1, \tag{152}$$

where $a_{k,t+1}$ is the rate of reference level growth caused by the increase in the standard of living.[86]

If we repeatedly lag and substitute equation (152), we can write $logS_{k,t+1}$ as a weighted sum of current and past measured consumption:

$$logS_{k,t+1} = \frac{a_{k,t+1}}{\lambda_k} + \lambda_k \sum_{i=0}^{\infty} (1 - \lambda_k)^i\, logC_{t+1-i},\ 0 < \lambda_k \leqslant 1. \tag{153}$$

Let us assume the reference consumption level to be the same for all agents, $S_{k,t+1} = S_{t+1}$ for all k ($\lambda_k = \lambda$ and $a_{k,t+1} = a_{t+1}$ for all k). Assume further that the rate of growth of the reference consumption level with the passage of time, a_{t+1}, is correlated with the return on the market portfolio, $a_{t+1} = a + b \times r_{M,t+1}$, where $r_{M,t+1}$ is the continuously compounded market portfolio return.

When $\lambda = 0$, (152) implies $logS_{t+1} = a + b \times r_{M,t+1} + logS_t$. Consequently, $\triangle s_{t+1} = a + b \times r_{M,t+1}$ and, therefore, $\frac{\partial \triangle s_{t+1}}{\partial r_{M,t+1}} = b$. From (151), we obtain that all investors have the same elasticity of intertemporal substitution $\sigma = \frac{1+(\gamma-\varphi)b}{\gamma}$. When $\varphi = \gamma$, we get the conventional power utility model for which $\sigma = \frac{1}{\gamma}$ whatever the value of b.

values of risk aversion and the elasticity of intertemporal substitution (the null hypothesis the elasticity of intertemporal substitution equals the inverse of the RRA coefficient is rejected statistically at the 5% significance level), in opposite to the Epstein-Zin (1989, 1991) SDF which yields a negative estimate of elasticity of substitution.

[86]The higher the value of λ_k, the more rapid the adjustment process. If $\lambda_k = 0$, then $S_{k,t+1} = exp\,(a_{k,t+1}) \cdot S_{k,t}$ at all t (the reference level grows simply with the passage of time). At the other extreme, if $\lambda_k = 1$, there is full adjustment in one period, $S_{k,t+1} = exp\,(a_{k,t+1}) \cdot C_{t+1}$. This case corresponds to the formulation of the benchmark level in Gali (1994) according to which the reference level of consumption only depends on the contemporaneous per capita consumption level in the economy. A similar approach to make the reference level of consumption grow with the passage of time is used in Abel (1999). Specifically, Abel (1999) states $S_t \equiv C_t^{\alpha_0} C_{t-1}^{\alpha_1} \left(G^t\right)^{\alpha_2}$, where G is a constant, $G \geqslant 1$.

Under the assumptions above, when there is an representative agent, the SDF for model (146) is

$$M_{t+1} = \delta^* \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} (1 + R_{M,t+1})^{\sigma\gamma-1}, \tag{154}$$

where $\delta^* \equiv \delta \times exp\,(a\,(\gamma - \varphi))$. This SDF is observationally equivalent to the EZ (1989, 1991) pricing kernel.[87]

If $\lambda > 0$, according to (153)

$$\triangle s_{t+1} = \frac{b}{\lambda}\,(r_{M,t+1} - r_{M,t}) + \lambda\triangle c_{t+1} + \lambda\sum_{i=1}^{\infty}(1-\lambda)^i\,\triangle c_{t+1-i}. \tag{155}$$

It implies $\frac{\partial\triangle s_{t+1}}{\partial r_{M,t+1}} = \frac{b}{\lambda} + \lambda\sigma$ and, therefore, $\sigma = \frac{1+(\gamma-\varphi)\left(\frac{b}{\lambda}+\lambda\sigma\right)}{\gamma}$. This relationship can be rearranged so that we obtain $\sigma = \frac{1+(\gamma-\varphi)\frac{b}{\lambda}}{(1-\lambda)\gamma+\lambda\varphi}$. When $\varphi = \gamma$, our utility specification reduces to the standard time- and state-separable power utility function with $\sigma = \frac{1}{\gamma}$ for any values of λ $(0 < \lambda \leqslant 1)$ and b.

In this paper, we will consider only the case when the rate of reference level growth caused by the increase in the standard of living is constant over time, $b = 0$. We further hypothesize that C_{t+1} is related to S_{t+1} by the relationship $C_{t+1} = S_{t+1} \times \varepsilon_{t+1}$, where a disturbance term ε_{t+1} is independent of S_{t+1} and lognormally distributed, $log\varepsilon_{t+1} \sim N\left(0, \sigma_\varepsilon^2\right)$.[88]

Under these assumptions

$$logC_{t+1} = \frac{a}{\lambda} + \lambda\sum_{i=0}^{\infty}(1-\lambda)^i\,logC_{t+1-i} + log\varepsilon_{t+1},\; 0 < \lambda \leqslant 1. \tag{156}$$

[87] See GRS (2002) for the detailed comparative analysis of these two SDFs.

[88] A disturbance term ε_{t+1} is assumed to represent a contemporaneous shock to realized aggregate consumption. If the shock is positive, $log\varepsilon_{t+1} > 0$ $(\varepsilon_{t+1} > 1)$, consumption is above the benchmark level. However, when the shock is negative, $log\varepsilon_{t+1} < 0$ $(\varepsilon_{t+1} < 1)$, consumption is presumed to be below the reference level. The only case, when consumption coincides with the benchmark level is the absence of any shock. Under the assumptions $a_{k,t+1} = a$ and $\lambda_k = \lambda$ for all k, substituting $logC_{t+1} = logS_{t+1} + log\varepsilon_{t+1}$ into (152) yields

$$\triangle s_{t+1} = \frac{a}{1-\lambda} + \frac{\lambda}{1-\lambda}log\varepsilon_{t+1},\; 0 \leqslant \lambda < 1.$$

If $log\varepsilon_{t+1}$ are independent, then $\triangle s_{t+1}$ are IID normal variates with mean $\frac{a}{1-\lambda}$ and variance $\left(\frac{\lambda}{1-\lambda}\right)^2\sigma_\varepsilon^2$. As λ approaches 1, $var(\triangle s_{t+1})$ approaches $var(\triangle c_{t+1})$.

Equation (156) may be rewritten as

$$logC_{t+1} = \frac{a}{\lambda(1-\lambda)} + \lambda \sum_{i=1}^{\infty} (1-\lambda)^{i-1} logC_{t+1-i} + \frac{log\varepsilon_{t+1}}{(1-\lambda)},\ 0 < \lambda < 1. \tag{157}$$

Using the Koyck transformation yields

$$\Delta c_{t+1} = \frac{a}{1-\lambda} - \eta_{t+1} + (1-\lambda)\eta_t,\ 0 < \lambda < 1, \tag{158}$$

where $\eta_{t+1} \equiv -\frac{log\varepsilon_{t+1}}{(1-\lambda)}$. This is an MA(1) model in which the coefficient of η_t characterizes persistence in the reference consumption level process. This model can be estimated from the time series of aggregate consumption per capita.

In opposite to Cecchetti, Lam, and Mark (1990, 1993) and Kandel and Stambaugh (1991), who assume a persistent discrete-state Markov process for expected aggregate consumption growth z_t, Campbell (1999) assumes z_t to follow an AR(1) process with mean g and persistence ψ:

$$\begin{aligned} \Delta c_{t+1} &= z_t + v_{t+1}, \\ z_{t+1} &= (1-\psi) g + \psi z_t + u_{t+1},\ \psi > 0, \end{aligned} \tag{159}$$

$$\begin{bmatrix} v_{t+1} \\ u_{t+1} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_v^2 & \rho\sigma_v\sigma_u \\ \rho\sigma_v\sigma_u & \sigma_u^2 \end{bmatrix}\right). \tag{160}$$

Wachter (2002) shows that conditional on consumption data, system (159) has the same likelihood function as the following ARMA(1,1) process:

$$\Delta c_{t+1} = (1-\psi) g + \psi \Delta c_t + \eta_{t+1} + \theta\eta_t, \tag{161}$$

where

$$\left((\theta+\psi)^2 + 1 - \psi^2\right)\sigma_\eta^2 = \sigma_v^2 \left(1-\psi^2\right) + \sigma_u^2 \tag{162}$$

and

$$\theta\sigma_\eta^2 = \rho\sigma_v\sigma_u - \psi\sigma_v^2. \tag{163}$$

If $\theta = 0$ ($u_{t+1} = \psi v_{t+1}$), we obtain a linear version of the Mehra-Prescott (1985) model in which consumption growth follows an AR(1) process. Setting $\theta = 0$ and $\psi = 0$ results in the random walk model of consumption (Campbell and Cochrane (1999)). Equation (158) can be obtained from (161) when ψ is set to 0 with the free parameter θ and, hence, is less restrictive than the model assumed by Campbell and Cochrane (1999).

The more general case is to assume that not only the current period shock but also some previous period shocks may affect the level of current consumption: $C_{t+1} = S_{t+1} \times \prod_{i=0}^{\infty} \varepsilon_{t+1-i}^{\alpha_i}$, $\alpha_0 = 1$. Under this assumption,

$$\Delta c_{t+1} = \frac{a}{1-\lambda} - \eta_{t+1} - \sum_{i=1}^{\infty} (\alpha_i - \alpha_{i-1}(1-\lambda))\, \eta_{t+1-i},\ 0 < \lambda < 1. \tag{164}$$

Another way to take into account some persistence in shocks is to assume an AR(p) model for $log\varepsilon_{t+1}$. If we assume, for example, that $log\varepsilon_{t+1}$ follows an AR(1) process, $log\varepsilon_{t+1} = b \times log\varepsilon_t + u_{t+1}$, we get $C_{t+1} = S_{t+1} \times \prod_{i=0}^{\infty} exp(b^i \times u_{t+1-i})$ and, hence,

$$\Delta c_{t+1} = \frac{a}{1-\lambda} - \xi_{t+1} - \sum_{i=1}^{\infty} b^{i-1}(b-(1-\lambda))\, \xi_{t+1-i},\ 0 < \lambda < 1, \tag{165}$$

where $\xi_{t+1} \equiv -\frac{u_{t+1}}{1-\lambda}$.

As we saw above, when $\varphi = \gamma$ (the standard power utility model), $\sigma = \frac{1}{\gamma}$ for any values of λ and b. When $\varphi \neq \gamma$ and $b = 0$, the elasticity of intertemporal substitution is the reciprocal of the arithmetic average of the RRA coefficient γ and the parameter φ for any value of λ $(0 \leqslant \lambda \leqslant 1)$:

$$\sigma = \frac{1}{(1-\lambda)\gamma + \lambda\varphi}. \tag{166}$$

When $\lambda = 0$, $S_{t+1} = exp\,(a)\, S_t$ and, hence, the SDF for model (146),

$$M_{t+1} = \delta^* \left(\frac{C_{k,t+1}}{C_{k,t}}\right)^{-\gamma},\ \delta^* \equiv \delta \times exp\,(a\,(\gamma - \varphi))\,, \tag{167}$$

is observationally equivalent to that for the power utility model. So, it is not astonishing that with $\lambda = 0$, both models yield the same elasticity of intertemporal substitution, $\sigma = \frac{1}{\gamma}$. In another extreme case, when there is full adjustment in one period ($\lambda = 1$), $\sigma = \frac{1}{\varphi}$. It follows that under the assumption that the reference consumption level fully adjusts in one period, utility function (146) allows us not only to directly estimate the parameter of elasticity of intertemporal substitution, σ, but also completely disentangle risk aversion and elasticity of substitution.[89]

[89] There are two different parameters γ and φ which govern risk aversion and the elasticity of intertemporal substitution, respectively.

4.4 Empirical Analysis

Empirical evidence in Semenov (2002) is that the approximate equilibrium multifactor model for expected asset returns is able to explain both the equity premium and the return on the risk-free asset with economically plausible values of the RRA coefficient and the time discount factor when all individuals are assumed to have the CRRA homogeneous preferences and the agent's marginal utility of consumption is expanded as a Taylor series up to cubic terms.

An undesirable property of the CRRA utility specification is that the elasticity of intertemporal substitution is constrained to be the reciprocal of the RRA coefficient. An attractive feature of the expected utility model with a reference level of consumption is that it allows to disentangle risk aversion and intertemporal substitution. GRS (2002) test this utility function under the assumption of market completeness and find that this specification of preferences allows to obtain the point estimate of the elasticity of intertemporal substitution that is in the conventional range and statistically different from the inverse of the RRA coefficient.

In this section, we assume incomplete consumption insurance and limited participation and test the expected utility function with a reference level of consumption using the approximate equilibrium multifactor model for expected asset returns.

4.4.1 Description of the Data

The Consumption Data. The consumption data are taken from the CEX. As opposed to the PSID which offers only food consumption data on an annual basis, the CEX contains highly detailed data on monthly consumption expenditures.[90] The CEX attempts to account for an estimated 70% of total household consumption expenditures. Since the CEX is designed with the purpose of collecting consumption data, measurement error in consumption is likely to be smaller for CEX consumption data compared to the PSID consumption data.

The CEX data available cover the period from 1979:10 to 1996:2. It is a collection of data

[90]Food consumption is likely to be one of the most stable consumption components. Furthermore, as Carroll (1994) points out, 95% of measured in the PSID food consumption is noise due to the absence of interview training.

on approximately 5000 households per quarter in the United States. Each household in the sample is interviewed every three months over five consecutive quarters.[91] As households complete their participation, they are dropped and new households move into the sample. Thus, each quarter about 20% of the consumer units are new. The second through fifth interviews use uniform questionnaires to collect demographic and family characteristics as well as data on monthly consumption expenditures for the previous three months made by households in the survey.[92] Various income information is collected in the second and fifth interviews as well as information on the employment of each household member.

The measure of consumption used in this empirical investigation is consumption of nondurables and services (NDS). For each household, we calculate monthly consumption expenditures for all the disaggregate consumption categories offered by the CEX. Then, we deflate obtained values in 1982-84 dollars with the CPI's (not seasonally adjusted, urban consumers) for appropriate consumption categories.[93] Aggregating the household's monthly consumption across these categories is made according to the National Income and Product Account definitions of consumption aggregates. In order to transform my consumption data to a per capita basis, we normalize the consumption of each household by dividing it by the number of family members in the household.

The Returns Data. The measures of the nominal market return are the value-weighted and equal-weighted returns (capital gain plus dividends) on all stocks listed on the NYSE and AMEX obtained from the CRSP. The real, monthly market return is calculated as the nominal market return divided by the 1-month inflation rate based on the deflator defined for NDS consumption. The nominal, monthly risk-free rate of interest is the 1-month Treasury bill return from CRSP. The real, monthly risk-free interest rate is calculated as the nominal risk-free rate divided by the 1-month inflation rate. The market premium is calculated as the difference between the real market return and the real risk-free rate of interest.

---

[91]The first interview is practice and is not included in the published data set.

[92]Demographic variables are based upon heads of households.

[93]The CPI's series are obtained from the BLS through CITIBASE.

Asset Holders. For the consumer units completing their participation in the first through third quarters of 1986, BLS has changed, beginning the first quarter of 1986, the consumer unit identification numbers so that the identification numbers for the same household in 1985 (when this household has been interviewed for the first time) and in 1986 (when he has completed his interviews) are not the same. To match the consumer units between the 1985 and 1986 data tapes, we use the household characteristics which allow to identify consumer units uniquely. As a result, we manage to match 47.0% of households between the 1985 and 1986 data tapes. The detailed description of the procedure used to match the consumer units between the 1985 and 1986 data tapes is given in Semenov (2002).

In the fifth (final) interview, the household is asked to report end-of-period estimated market value of all stocks, bonds, mutual funds, and other such securities (market value of all securities) held by the consumer unit on the last day of the previous month as well as the difference in the estimated market value of all securities compared with the value of all securities held a year ago last month. Using these two values, we calculate asset holdings at the beginning of a 12-month recall period. The consumer unit is considered as an assetholder if the household's asset holdings at the beginning of a 12-month recall period exceed a certain threshold. To assess the quantitative importance of limited participation of households in the asset markets, we consider four sets of households. The first set (SET1) consists of all consumer units independently of the reported market value of all securities. To take into consideration that only a part of households participates in the asset markets, we use three sets of households defined as assetholders. The first one (SET2) consists of the consumer units whose asset holdings are equal to or exceed $2 in 1999 dollars, the two others consist of the households reported total assets equal to or exceeding $10000 (SET3) and $20000 (SET4).[94]

Per capita consumption of a set of households is calculated as the equal-weighted average of normalized consumption expenditures of the households in the set. Obtained per capita consumption is seasonally adjusted by using the X-11 seasonal adjustment program.[95]

[94]Over the period 1991-1996 about 18% of households, for which the market value of all securities held a year ago last month is not missing, reported asset holdings of $1 at the beginning of a 12-month recall period. That occurs when the household reported owning securities without precising their value (see Vissing-Jorgensen (1998)). Following Vissing-Jorgensen (1998), we classify these households as nonassetholders.

[95]Ferson and Harvey (1992) point out that since the X-11 program uses of past and future information in

We seasonally adjust the normalized consumption of each household by using the additive adjustments obtained from per capita consumption.

Data Selection Criteria. Following Vissing-Jorgensen (1998), we drop from the sample the bottom and the top percent of consumption growth observations for each month (under the assumption that these extreme values reflect reporting or coding errors). In addition, we drop nonurban households, households residing in student housing, households with incomplete income responses, and households who do not have a fifth interview. Following Brav, Constantinides, and Géczy (2002), in any given month, we drop from the sample households that report in that month as zero either their food consumption or their consumption of nondurables and services, or their total consumption, as well as households with missing information on the above items. Additionally, we keep in the sample only the households whose head is between 19 and 75 years of age.

4.4.2 The Estimation Methodology

When all investors have homogeneous preferences of the form (146), the Euler equations for the equity premium (144) and the risk-free rate (145) can be written as

$$E_t\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}RP_{t+1}\right]=0 \tag{168}$$

and

$$\delta E_t\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h_{t+1}^n}\right)\left(\frac{S_{t+1}}{S_t}\right)^{\gamma-\varphi}R_{f,t+1}\right]=$$
$$\left(\frac{h_{t+1}}{h_t}\right)^{\gamma}\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t}}{h_t^n}\right), \tag{169}$$

respectively.

Assuming the dynamics of the log reference level to be given by equation (152), we use the following two-stage procedure to estimate the parameters of interest. The first step is

the time-averaging it performs, this type of seasonal adjustment may induce spurious correlation between the error terms of a model and lagged values of the variables and, hence, may cause improper rejections of the model based on tests of overidentifying restrictions. As alternatives to using X-11 program, Brav and Géczy (1995) propose to use a simpler linear filter (Davidson and MacKinnon (1993)) or the Ferson-Harvey (1992) method of incorporating forms of seasonal habit persistence directly in the Euler equation.

maximum likelihood (ML) estimation of the following regression model:

$$\Delta c_{t+1} = g - \eta_{t+1} + \theta \eta_t. \tag{170}$$

Given the estimates of g and θ obtained from (170), we are able to estimate the parameters a and λ of the behavioral model (158). The coefficient of η_t yields an estimate of $(1-\lambda)$ and, hence, of λ. The constant term g, when multiplied by $(1-\lambda)$, yields an estimate of a.

The second step is to use the iterated GMM approach to estimate the Euler equations for the premium of the real value-weighted and equal-weighted market portfolio returns over the risk-free rate as

$$E_t\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h^n}\right)\left(\prod_{i=0}^{I}\left(\frac{C_{t+1-i}}{C_{t-i}}\right)^{\lambda(1-\lambda)^i}\right)^{\gamma-\varphi}RP_{t+1}\right]=0 \tag{171}$$

and for the real risk-free rate as

$$\delta E_t\left[\left(1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t+1}}{h^n}\right)\left(\prod_{i=0}^{I}\left(\frac{C_{t+1-i}}{C_{t-i}}\right)^{\lambda(1-\lambda)^i}\right)^{\gamma-\varphi}R_{f,t+1}\right]=$$
$$1+\sum_{n=1}^{N}\frac{1}{n!}(-1)^n\left(\prod_{l=0}^{n-1}(\gamma+l)\right)\frac{Z_{n,t}}{h^n} \tag{172}$$

with λ replaced by its estimate obtained from (170).[96]

The sample period is from 1979:10 to 1996:2. As in Semenov (2002), we expand the agent's marginal utility of consumption as a Taylor series up to cubic terms ($N=3$). The Euler equations for the excess value-weighted and equal-weighted market returns (171)

[96]It can be seen that when Abel's (1999) specification of consumption externalities is used, the parameters δ, G, and α_2 are not identifiable from (169). All we are able to identify is the parameter $\delta^* \equiv \delta G^{\alpha_2(\gamma-\varphi)}$. This leads to another problem. Given that $G \geqslant 1$ and $0 \leqslant \alpha_2 \leqslant 1$, the parameter δ cannot be estimated consistently when $G \neq 1$ and $\alpha_2(\gamma-\varphi) \neq 0$. The estimate of δ is upward biased by the factor of $G^{\alpha_2(\gamma-\varphi)}$ when $G \neq 1$, $\alpha_2 \neq 0$, and $\gamma-\varphi > 0$ and downward biased when $G \neq 1$, $\alpha_2 \neq 0$, and $\gamma-\varphi < 0$. An attractive feature of our specification of the benchmark consumption level is that when $\lambda > 0$, the term $exp(a(\gamma-\varphi))$ vanishes from the Euler equations as the ratio of benchmark levels in two successive periods is taken and, therefore, unlike Abel's (1999) specification of the reference consumption level, an unbiased estimate of δ can be obtained. Moreover, unlike Abel's (1999) specification, the presented in this paper specification of consumption externalities allows to estimate the growth rate of the benchmark level reflecting the increase in the standard of living (the parameter a in equation (170)).

and the Euler equation for the real risk-free interest rate (172) are estimated jointly using an iterated GMM approach. We exploit two sets of instruments. The first instrument set (INSTR1) consists of a constant, the real value-weighted and equal-weighted market returns, the real risk-free rate, and the real consumption growth rate lagged one period. The second set of instruments (INSTR2) is the first set extended with the same variables lagged an additional period.

4.4.3 Estimation Results

As in Semenov (2002), we assume the conditional expectation of consumption to be equal to the conditional expectation of aggregate consumption per capita, $h_{t+1} = E_t\left[C_{t+1}\right]$, and estimate the following random walk model of consumption:

$$\Delta c_{t+1} = g + \eta_{t+1}, \tag{173}$$

where $\Delta c_{t+1} \equiv log\frac{C_{t+1}}{C_t}$ and $\eta_{t+1} \sim N\left(0, \sigma_\eta^2\right)$. It follows that $h_{t+1} = exp\left(g + \frac{\sigma_\eta^2}{2}\right) C_t$. Table I presents the usual ML estimates for model (173).

The results of the ML estimation of (170) are reported in Table II. Neither the null hypothesis $H_0 : g = 0$ nor $H_0 : a = 0$ is rejected at the 5% level. The point estimates of λ are significantly different from both 0 and 1 at the 5% level for all the sets of households. The greatest value of λ is obtained for SET1, what means that the reference consumption level of non-assetholders adapts to changes in aggregate consumption per capita quicker than that of assetholders.

Given that the weight of consumption lagged ten periods becomes so small that one can neglect further past values of consumption, we estimate the Euler equations (171) and (172) with $I = 10$. The results of estimation and testing the Euler equations (171) and (172) with $N = 3$ are reported in Table III. Using the set of instruments INSTR1, we obtain the estimates of the RRA coefficient which are in the conventional range and significantly different from 0 for all the sets of households.[97] The point estimate of the elasticity of intertemporal substitution, σ, is positive only when a part of consumers is assumed to participate in the asset markets. However, only for SET2, σ is significantly positive at the

[97] The point estimates of γ are significantly different from 0 at the 5% significance level for SET1, SET2, and SET4. For SET3, the point estimate of risk aversion is significantly positive at the 10% level.

5% level. The standard power utility model $\left(H_0 : \gamma - \varphi = 0 \; (H_0 : \sigma = \frac{1}{\gamma})\right)$ is rejected for the households whose asset holdings are less than \$10000. The ratio model $(H_0 : \varphi = 1)$ is rejected at the 5% significance level for all the households reported market value of all securities of less than \$20000. According to Hansen's J statistic, the model is not rejected statistically.

As the instrument set INSTR2 is used, we obtain the point estimates of γ which are significantly positive for all the sets. A little evidence of predictable variation in consumption growth in the face of predictable asset returns suggests that the elasticity of intertemporal substitution, σ, is small.[98] The results in Table III show that the point estimate of σ is small and significantly positive when only the households reported total assets equal to or exceeding \$10000 are classified as assetholders. Both the standard power utility specification and the ratio model are rejected statistically at the 5% significance level for all the sets of consumers. For SET1 and SET2, the point estimate of $\gamma - \varphi$ is significantly positive, what suggests that for the households reported total assets less than \$10000, consumption externalities are positive, while they are negative for the consumer units whose asset holdings are equal to or exceed \$10000 (for SET3 and SET4, the point estimate of $\gamma - \varphi$ is significantly negative). According to Hansen's test of the overidentifying restrictions, the model is not rejected statistically at the 5% level.

4.5 Concluding Remarks

The empirical results provide some evidence that the reference consumption level responses only gradually to changes in contemporaneous aggregate consumption per capita. The null hypotheses that the reference level only grows with the passage of time and that there is full adjustment in one period are both rejected statistically at the 5% significance level. This result is robust to the threshold value in the definition of assetholders. The rejection of the null hypothesis $\lambda = 1$ allows to conclude that Gali's specification of consumption externalities is not supported by the data. Empirical evidence is also presented that for nonassetholders, the reference level adapts to changes in aggregate consumption per capita more quickly than that for assetholders.

Another important result is that given partial adjustment of the reference level to

[98]See Campbell, Lo, and MacKinlay (1997) and Campbell and Mankiw (1990).

changes in aggregate consumption, we are able to disentangle risk aversion and intertemporal substitution. The obtained estimate of the elasticity of intertemporal substitution is in the conventional range and significantly different from the inverse of the RRA coefficient at the 5% level when the households reported the market value of all securities equal to or exceeding $10000 are classified as assetholders. Both the standard time-separable power utility model and the ratio preference specification are rejected statistically.

Appendix: Tables

Table I.

Parameter Estimates for $\Delta c_{t+1} = g + \eta_{t+1}$

The sampling period is from 1979:10 to 1996:2. Four sets of households are considered. The first set (SET1) consists of all consumer units with any reported market value of all securities. We also use three sets of households classified as assetholders: SET2 consists of the households whose asset holdings are equal to or exceed \$2 in 1999 dollars, the two others consist of the households reported total assets equal to or exceeding \$10000 (SET3) and \$20000 (SET4). The model is estimated by ML (standard errors in parentheses).

Parameters	SET1	SET2	SET3	SET4
g	0.0016	0.0022	0.0027	0.0021
	(0.0014)	(0.0023)	(0.0029)	(0.0032)
σ_η^2	0.0004	0.0010	0.0016	0.0020

Table II.

Parameter Estimates for $\Delta c_{t+1} = g - \eta_{t+1} + \theta\eta_t$

The sampling period is from 1979:10 to 1996:2. Four sets of households are considered. The first set (SET1) consists of all consumer units with any reported market value of all securities. We also use three sets of households classified as assetholders: SET2 consists of the households whose asset holdings are equal to or exceed \$2 in 1999 dollars, the two others consist of the households reported total assets equal to or exceeding \$10000 (SET3) and \$20000 (SET4). The model is estimated by ML (standard errors in parentheses). In Panel A, we report the values of the parameters estimated directly. The values of the parameters estimated indirectly are presented in Panel B. The standard errors for the parameters estimated indirectly are calculated by using the delta method.

Parameters	SET1	SET2	SET3	SET4
Panel A:				
g	0.0015	0.0020	0.0020	0.0018
	(0.0011)	(0.0014)	(0.0018)	(0.0021)
$\theta = 1 - \lambda$	0.2050	0.3486	0.3347	0.3401
	(0.0741)	(0.0736)	(0.0714)	(0.0698)
Panel B:				
a	0.0003	0.0007	0.0007	0.0006
	(0.0003)	(0.0005)	(0.0006)	(0.0007)
λ	0.7950	0.6514	0.6653	0.6599
	(0.0741)	(0.0736)	(0.0714)	(0.0698)

Table III.

Parameter Estimates and Test Statistics for the Utility Specification with a Reference Level under Incomplete Consumption Insurance

The proxy for the reference consumption level growth rate is constructed using 10 lags of consumption growth. This has the effect of reducing the length of the sample by 10 months, so that the sampling period used in the estimation is from 1980:8 to 1996:2 rather than from 1979:10 to 1996:2. Four sets of households are considered. The first set (SET1) consists of all consumer units with any reported market value of all securities. We also use three sets of households classified as assetholders: SET2 consists of the households whose asset holdings are equal to or exceed $2 in 1999 dollars, the two others consist of the households reported total assets equal to or exceeding $10000 (SET3) and $20000 (SET4). The agent's marginal utility of consumption is expanded as a Taylor series up to cubic terms ($N = 3$). The Euler equations for the excess value-weighted and equal-weighted market returns and for the real risk-free interest rate are estimated jointly using an iterated GMM approach (standard errors in parentheses). Two sets of instruments are exploited. The first instrument set (INSTR1) consists of a constant, the real value-weighted and equal-weighted market returns, the real risk-free rate, and the real consumption growth rate lagged one period. The second set of instruments (INSTR2) is the first set extended with the same variables lagged an additional period. The J statistic is Hansen's test of the overidentifying restrictions. The P value is the marginal significance level associated with the J statistic. In Panel A, we report the values of the parameters estimated directly from the Euler equations. The values of the parameters estimated indirectly are presented in Panel B. The standard errors for the parameters estimated indirectly are calculated by using the delta method.

Parameters	SET1	SET2	SET3	SET4
		INSTR1		
Panel A:				
γ	1.0912	1.2337	0.0238	1.1598
	(0.0201)	(0.0575)	(0.0125)	(0.0504)
φ	-46.6953	17.3727	0.2992	9.3589
	(3.0307)	(4.4998)	(0.2925)	(6.0644)
δ	0.9374	0.8701	0.9974	0.8976
	(0.0853)	(0.0509)	(0.0008)	(0.0954)
J statistic	6.3931	7.3585	8.0852	6.9422
P value	0.8950	0.8330	0.7784	0.8614
Panel B:				
σ	-0.0271	0.0851	4.8301	0.1522
	(0.0031)	(0.0230)	(4.5033)	(0.0937)
$\varphi - 1$	-47.6953	16.3727	-0.7008	8.3589
	(3.0307)	(4.4998)	(0.2925)	(6.0644)
$\gamma - \varphi$	47.7866	-16.1390	-0.2755	-8.1991
	(3.0313)	(4.4766)	(0.3003)	(6.0600)

Table III (continued)

Parameters	SET1	SET2	SET3	SET4
		INSTR2		
Panel A:				
γ	1.0887	1.0224	0.2528	0.3829
	(0.0130)	(0.0239)	(0.0538)	(0.0617)
φ	-33.6584	-21.2708	7.2052	12.5850
	(2.2501)	(2.6538)	(1.5513)	(2.1250)
δ	1.0737	0.9627	1.0218	0.9975
	(0.0324)	(0.0174)	(0.0054)	(0.0139)
J statistic	8.0897	8.7905	8.7992	8.6911
P value	0.9990	0.9980	0.9980	0.9982
Panel B:				
σ	-0.0377	-0.0741	0.2050	0.1186
	(0.0045)	(0.0131)	(0.0482)	(0.0231)
$\varphi - 1$	-34.6584	-22.2708	6.2052	11.5850
	(2.2501)	(2.6538)	(1.5513)	(2.1250)
$\gamma - \varphi$	34.7471	22.2932	-6.9524	-12.2021
	(2.2489)	(2.6565)	(1.5447)	(2.1108)

References

[1] Abel, A.B., 1999, Risk Premia and Term Premia in General Equilibrium, Journal of Monetary Economics 43, 3-33.

[2] Brav, A., G.M. Constantinides, and C.C. Géczy, 2002, Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence, Journal of Political Economy, forthcoming.

[3] Brav, A. and C.C. Géczy, 1995, An Empirical Resurrection of the Simple Consumption CAPM with Power Utility, working paper, University of Chicago.

[4] Campbell, J.Y., 1999, Asset Prices, Consumption, and the Business Cycle, in: J.B. Taylor and M. Woodford, eds., Handbook of Macroeconomics, Vol. 1, North-Holland, Amsterdam.

[5] Campbell, J.Y. and J. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, Journal of Political Economy 107, 205-251.

[6] Campbell, J.Y., A.W. Lo, and A.C. MacKinlay, 1997, The Econometrics of Financial Markets, Princeton University Press.

[7] Campbell, J.Y. and N.G. Mankiw, 1990, Permanent Income, Current Income, and Consumption, Journal of Business and Economic Statistics 8, 265-278.

[8] Carroll, C.D., 1994, How Does Future Income Affect Current Consumption, Quarterly Journal of Economics 109, 111-147.

[9] Cecchetti, S.G., P.-S. Lam, and N.C. Mark, 1990, Mean Reversion in Equilibrium Asset Prices, American Economic Review 80, 398-418.

[10] Cecchetti, S.G., P.-S. Lam, and N.C. Mark, 1993, The Equity Premium and the Risk-Free Rate: Matching the Moments, Journal of Monetary Economics 31, 21-45.

[11] Constantinides, G.M., 1990, Habit Formation: A Resolution of the Equity Premium Puzzle, Journal of Political Economy 98, 519-543.

[12] Constantinides, G.M. and D. Duffie, 1996, Asset Pricing with Heterogeneous Consumers, Journal of Political Economy 104, 219-240.

[13] Davidson, R. and J.G. MacKinnon, 1993, Estimation and Inference in Econometrics, Oxford University Press.

[14] Eichenbaum, M. and L. Hansen, 1990, Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data, Journal of Business and Economic Statistics 8, 53-69.

[15] Eichenbaum, M., L. Hansen, and K. Singleton, 1988, A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty, Quarterly Journal of Economics 103, 51-78.

[16] Epstein, L.G. and S.E. Zin, 1989, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica 57, 937-969.

[17] Epstein, L.G. and S.E. Zin, 1991, Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical analysis, Journal of Political Economy 99, 263-288.

[18] Ferson, W.E. and G.M. Constantinides, 1991, Habit Persistence and Durability in Aggregate Consumption: Empirical Tests, Journal of Financial Economics 29, 199-240.

[19] Ferson, W.E. and C.R. Harvey, 1992, Seasonality and Consumption-Based Asset Pricing, Journal of Finance, June 1992.

[20] Gali, J., 1994, Keeping up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices, Journal of Money, Credit and Banking 26, 1-8.

[21] Gallant, A. and G. Tauchen, 1989, Seminonparametric Estimation of Conditionally Heterogeneous Processes: Asset Pricing Applications, Econometrica 57, 1091-1120.

[22] Garcia, R., É. Renault, and A. Semenov, 2002, A Consumption Capital Asset Pricing Model with A Reference Level, working paper, CIRANO and Université de Montréal.

[23] Hansen, L.P. and R. Jagannathan, 1991, Implications of Security Market Data for Models of Dynamic Economies, Journal of Political Economy 99, 225-262.

[24] Hansen, L.P. and K. Singleton, 1982, Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, Econometrica 50, 1269-1286.

[25] Hansen, L.P. and K. Singleton, 1983, Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns, Journal of Political Economy 91, 249-265.

[26] Heaton, J., 1995, An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications, Econometrica 63, 681-717.

[27] Kandel, S. and R.F. Stambaugh, 1991, Asset Returns and Intertemporal Preferences, Journal of Monetary Economics 27, 39-71.

[28] Mankiw, N.G., 1986, The Equity Premium and the Concentration of Aggregate Shocks, Journal of Financial Economics 17, 211-219.

[29] Mehra, R. and E.C. Prescott, 1985, The Equity Premium: A Puzzle, Journal of Monetary Economics 15, 145-162.

[30] Semenov, A., 2002, Asset Pricing Puzzles, High-Order Consumption Moments, and Heterogeneous Consumers, working paper, Université de Montréal.

[31] Vissing-Jorgensen, A., 1998, Limited Stock Market Participation, working paper.

[32] Wachter, J. A., 2002, Habit Formation and Returns on Bonds and Stocks, working paper, New York University.

[33] Weil, P., 1989, The Equity Premium Puzzle and the Risk-Free Rate Puzzle, Journal of Monetary Economics 24, 401-421.

[34] Weil, P., 1992, Equilibrium Asset Prices with Undiversifiable Labor Income Risk, Journal of Economic Dynamics and Control 16, 769-790.