The Dynamic of the Equity Premium: Is it Habit Formation or Loss Aversion?

Pierre Monnin

First version: May 2007
Current version: May 2007

This research has been carried out within the NCCR FINRISK project on “Behavioural and Evolutionary Finance”
The Dynamic of the Equity Premium: Is it Habit Formation or Loss Aversion?

Pierre Monnin*

*Author address: Swiss National Bank, Financial Stability, Börsenstrasse 15, P.O. Box, CH.8022 Zurich, Switzerland. Tel: +41 44 631 38 35. Email: pierre.monnin@snb.ch This paper has benefited from the comments and suggestions of Thorsten Hens and of participants to the SOI seminar at the University of Zurich. Scientific support by the National Center of Competence in Research “Financial Valuation and Risk Management” (NCCR Finrisk) is gratefully acknowledged. The opinions expressed herein are those of the author and do not necessarily reflect the views of the Swiss National Bank.
Abstract

The large spread between equity returns and risk free rates observed in most stock markets (the "equity premium puzzle") has been subject of intense debates. Two main families of models claim to solve this puzzle: habit formation models and loss aversion models. The goal of this paper is to assess empirically which of them fits the observed excess returns best. For that, I first show how to express both models in the same form, namely as linear stochastic discount factor models. This form has the main advantage to give explicit and testable constraints for the excess return dynamic. I then compare the theoretical dynamic generated by these models with the observed dynamic. I find that the external habit model and a loss aversion model with a reference level based on past consumption are not likely to correspond to the observed data. In opposition, a loss aversion model with a reference level based on expected consumption and, to some extend, the internal habit model could fit the observed excess return dynamic.

Keywords: Expected stock return; Habit formation; Loss aversion; Consumption-based model.

JEL-Classifications: E21, E44, G12.
1 Introduction

In 1985, Mehra and Prescott introduced the academic and financial profession to what they called the "equity premium puzzle" (Mehra and Prescott 1985). They claimed that the difference between the returns on equities and the risk free return is too large to be compatible with a traditional asset pricing model, based on a representative agent with constant relative risk aversion, who maximizes her (expected) utility. They argued that the coefficient of relative risk aversion (CRRA) estimated with such a model is much higher than the value found in experimental studies. This finding has been replicated several times by other studies. Campbell (1999), for example, estimates that the CRRA must be equal to 316 to match the equity premium observed in the US between 1947 and 1996. In opposition, experiments show that the average CRRA usually lies between 0 and 5, with 10 being considered as an extreme value.

Since then, numerous solutions have been proposed to reconcile the observed equity premium with an asset pricing model based on a representative utility-maximizer agent. It is beyond the scope of this paper to survey the still flourishing literature generated by Mehra and Prescott’s article.¹ The goal of this paper is rather to propose a new way to test two categories of models that have recently caught up much of researchers’ attention on this topic: habit formation models (Abel 1990, Constantinides 1990 and Campbell and Cochrane 1999) and loss aversion models (Benartzi and Thaler 1995, Barberis, Huang and Santos 2001 and De Giorgi, Hens and Mayer 2006). Both models claim to explain the equity premium puzzle by using alternative utility functions. Habit formation models postulate that the utility of the representative agent should be a function of the difference between her consumption and a habit rather than on consumption only. The habit represents a minimum level under which the consumption cannot fall (i.e. a subsistence level). Loss aversion utility functions have two components: the traditional "consumption utility" based on consumption and the "gain-loss

¹See e.g. Mehra and Prescott (2003) for a survey of selected issues on the equity premium puzzle.
utility" which is based on the deviation of consumption from a reference level. Thus a loss-averse agent has an additional source of utility: she gets (loses) some "extra" utility if she consumes more (less) than her reference level. The particularity of loss aversion models is that losses are more "painful" than gains and thus a loss generates a greater decrease in utility than a gain of the same size increased it. The early papers cited above provide theoretical explanations for the equity premium puzzle as well as empirical calibrations, which back their conclusions.

This paper contributes to the literature on the equity premium in two ways: firstly, it shows how to express both habit formation models and loss aversion models in a common framework, namely as linear stochastic discount factor (SDF) asset pricing models. The SDF form has two advantages: the first one is that it allows a transparent comparison of the different excess return dynamic generated by both types of model. Concretely, the SDF form gives two different sets of explicit linear constraints on the dynamic of excess returns, one for habit formation models and one for loss aversion models. The second advantage is that the linear SDF form can be easily estimated using a Vector Autoregressive (VAR) model with multivariate GARCH errors.

The combination of both advantages of the linear SDF form is at the roots of the second innovation of this paper: by comparing the empirical dynamic of excess returns with the theoretical constraints of habit formation models or those of loss aversion models, it is possible to check whether the theoretical dynamic of one model corresponds to the observed dynamic or if it substantially deviates from it. It is even possible to statistically test, e.g. with traditional Wald test, if the null hypothesis of one model is rejected at a significant confidence level. The ultimate goal of this exercise is to determine which model (if any) is the closest to the observed excess returns.\footnote{Note that the methodology presented here can be used to test the validity of any model expressed in the form of a linear SDF model.}

Several previous studies have tried to assess the empirical validity of habit formation or loss aversion models. The usual way to do it is to use GMM estimation to
match the model with the a set of observed unconditional moments and then to check if the estimated coefficients are plausible or if the estimated model generates moments for other variables that are close to reality. These studies give contradictory evidence for habit formation models. On the positive side, GMM estimation generally give parameters in-line with theory (Hyde and Sherif 2005 for UK data, Allais, Cadiou and Déès 2000 for G7 countries and Menzly, Santos and Veronesi 2001 for cross-sectional stock returns). Li (2001) and Li and Zhong (2005) show that habit formation models can explain some part of excess return predictability and outperform power utility model in forecasting excess returns. Other studies are less positive. Tallarini and Zhang (2005) reject a habit formation model with the Efficient Method of Moments at a 1% level (but not at a 0.1% level) and Chapman (2002) cannot reproduce Constantinides’ (1990) results with after-war data. More doubts are cast on habit formation models, when their dynamic properties are examined. Li (2001) finds that returns volatility is positively related to surplus consumption, which is in contradiction with Campbell and Cochrane (1999) model. Duffee (2005) estimate that the price of risk generated by habit formation does not vary enough to justify the evolution of the equity premium. Allais et al. (2000), Lettau and Uhlig (2000) and Pijoan-Mas (2004) show that the consumption implied by habit formation models is much smoother than what is observed in reality.³ Lettau and Uhlig (2000) and Otrok, Ravikumar and Whiteman (2002) find that habit formation models induce counterfactual cyclical behaviour (on labour market or concerning the evolution of the consumption variability and the equity premium, respectively). On the other hand, Semmler and Woermann (2002) find that habit formation models generate a dynamic for Sharpe ratios that fits better the observed dynamic than a model with CRRA does. Empirical evidence on loss aversion models is scarcer. Using GMM estimation, Yogo (2005) finds parameters corresponding to the theory and smaller pricing errors for a model incorporating loss aversion.

The methodology presented in this paper completes and extends the previous studies in three directions. Firstly, it provides of a formal test on the coefficient generating the dynamic of excess returns. This allows comparing the entire dynamic of the different models rather than just their first or second moments. Comparing the observed and the theoretical dynamic rather than unconditional moments is methodologically more robust. Indeed, one can easily imagine a model that has the same first and second moments as its theoretical counterpart but with a totally different dynamic (e.g. see Figure 1). Secondly, the approach used here is based on conditional moments rather than unconditional one. This is closer to reality, where investors act according to the information available at the time when their decision is made (conditional information set). Finally, this paper compares habit formation and loss aversion models in the same methodological framework, which makes the comparison (and the potential ranking of models) easier and coherent.

Concretely, I test two habit formation models (the internal habit model of Constantinides 1990 and the external habit model of Campbell and Cochrane 1999) and two loss aversion models (one in which the reference level is based on expectations as in Benartzi and Thaler 1995 and Barberis et al. 2001, and one in which the reference level is based on past consumption as in Yogo 2005). I use monthly US data on consumption and stock market excess returns over the period 1959-2005. The loss aversion models
tested here differs from their original specification as the losses and gains are expressed in terms of consumption rather than in terms of wealth.

I find that the constant relative risk aversion model and the external habit model are strongly rejected by the data. The observed excess return dynamic does not correspond to the dynamic generated by these models. The constraints that they impose on the dynamic are formally rejected by both Wald and Likelihood ratio tests. The validity of the loss aversion model using past consumption as reference level is also doubtful. Although its restrictions cannot be rejected at a 1% confidence level, the values of the estimated parameters cast some doubt on it. The two remaining models (internal habit formation and loss aversion with expected consumption as reference point) are possible candidates for explaining the excess returns dynamic. They both cannot be rejected at a 1% confidence level and have plausible parameters. The loss aversion model has a slight advantage over the habit formation model as the Wald test does not reject its constraints at a 10% level.\(^4\)

In addition to this ranking of the different models, I find two interesting facts concerning the dynamic of excess returns. Firstly, as already documented by Duffee (2005), the conditional covariance between excess returns and consumption growth seems to be negatively related to excess returns. This is in contradiction with the constant relative risk aversion model and the external habit model. Secondly, consumption growth variance seems to play a role in the excess return dynamic. This factor does not appear in any of the proposed models and thus might constitute a starting point for the development of a more adequate model.

Section 2 presents the asset pricing equation used to test the model as well as the test method. Section 3 shows how to express habit formation and loss aversion models in the form of a linear SDF model. Section 4 gives the empirical results. Section 5 concludes.

\(^4\)Both Wald and Likelihood ratio test reject the internal habit model at a 5% confidence level.
2 Methodology

The methodology used in this paper to assess if a habit formation or loss aversion model are able to match the observed excess return dynamic is based on the tests for exact linear rational expectations models of Hansen and Sargent (1981). This method follows three steps:

1. Work out the theoretical model to derive some hypotheses (or restrictions) on the joint dynamic of the variables.

2. Estimate the empirical joint dynamic.

3. Test if the estimated dynamic is compatible with the theoretical hypothesis made in the first step.

The this section shows how to derive hypotheses on the joint dynamic from any linear SDF model. Section 3 then explains how to express the tradition CRRA model, the habit formation model and the loss aversion model in for of a linear SDF model.

2.1 The C-CAPM no arbitrage equation

The hypotheses on the joint dynamic are directly derived from the following no-arbitrage equation:

\[ E_t (M_{t+1}R_{t+1}) = 0 \]

where \( M_{t+1} \) is the SDF, or pricing kernel, and \( R_{t+1} \) is the excess return of holding a risky asset instead of a risk free bond. This no-arbitrage equation simply states that the discounted expected excess return of risky assets is equal to zero or, in other terms, that \( ex-ante \) there is no systematical discounted excess return.

The SDF \( M_{t+1} \) can take different forms depending on the model used to describe the behavior of asset prices.\(^5\) Each model considered in this paper has a different

---

\(^5\)See Cochrane (2001) for a survey of SDF models.
SDF, but they share a common feature: they are derived from the Consumption-based Capital Asset Pricing Model (C-CAPM). They all start with the same maximization problem, in which a representative agent allocates her resources between savings and consumption in order to maximize the utility of her present and future consumption. The SDF of the C-CAPM is

\[ M_{t+1} = \beta \frac{U'(C_{t+1})}{U'(C_t)} \]  

where \( \beta \) is a subjective discount factor and \( U'(C_t) \) is the marginal utility of consumption at time \( t \) (see Cochrane 2001 for more details). The only difference between the models presented here concerns the form taken by the representative consumer utility function \( U(C_t) \). The SDF associated to the different utility functions are presented in Section 3. In the next section, I will use the general SDF model in equation (1) to explain the methodology and then, when presenting each utility function, I will point out the specific characteristics of each SDF.

Following Söderlind (2006), if we assume that the log SDF \( m_{t+1} \) is normally distributed and that \( m_{t+1} \) and the excess return \( R_{t+1} \) have a bivariate mixture normal (conditional) distribution, we have that

\[ E_t(R_{t+1}) = -Cov_t(m_{t+1}, R_{t+1}) \]  

Economically, this equation specifies that the expected excess return is equal to the expected risk premium for the risky asset. This no-arbitrage condition constitutes the model tested in this paper. The test consists of verifying if the empirical excess return dynamic is compatible with equation (3).

Note that, traditionally in the literature, another linearization of equation (1) is used by assuming log normal returns. Under this assumption, a correction for Jensen’s effect is added to the no arbitrage equation (3), which becomes \( E_t(R_t) + V_t(R_{t+1}) = -Cov_t(m_{t+1}, R_{t+1}) \). Empirically, asset returns distributions are well known to have
fat tails. The mixture of normal distribution used here is more flexible and can take
different shapes that allow for such fat tails. It is therefore more likely to fit the
returns real distribution than the log normal distribution. The cost of this greater
flexibility for the return distribution is to assume that the (conditional) log SDF is
normally distributed. Söderlind (2006) shows that for SDF based on consumption, this
hypothesis is much more likely to be true than for excess returns.

2.2 Specifying the joint dynamic of the variables

Equation (3) suggests that second moments that the joint distribution of \( m_{t+1} \) and
\( R_{t+1} \) are time varying and that the dynamic of \( m_{t+1} \) and \( R_{t+1} \) is a function of their
second moments. These characteristics corresponds to the VAR-GARCH-M model
(Pelloni and Polasek 1999 and Polasek and Lei 2000). This model is an extension of
The VAR-GARCH-M model takes the following form

\[
\begin{align*}
y_t &= \mu + Ay_{t-1} + Bh_{t-1} + De_{t-1} + \varepsilon_t \\
h_t &= \omega + Fh_{t-1} + Ge_{t-1} \\
e_t &= \text{vech} \{ \varepsilon_t \} \\
H_t &= E_{t-1} (\varepsilon_t \varepsilon_t')
\end{align*}
\]

where \( y_t = \left[ R_t, \ldots, R_{t-j}, m_t, \ldots, m_{t-j} \right] \) is a vector of observations, \( A, B, C \)
and \( D \) are matrices of coefficients and \( \mu \) and \( \omega \) are vectors of coefficients. The \textit{vech}
operator converts the lower triangle of a symmetric matrix into a vector. Finally, \( H_t \)
is the expected covariance\textsuperscript{6} matrix at time \( t \). The model is made up of three parts:
equation (5) show that a) the covariances follow a GARCH model and equation (4)

\textsuperscript{6}in this paper, the term covariance includes also the variances of the variables.
show that b) the variables are a linear function of their past values (as in a VAR model) and c) of expected covariances (as in a traditional GARCH-M model).

2.3 The joint dynamic implied by the C-CAPM

The next step is to derive the restrictions that the no arbitrage equation (3) imposes on the coefficient of the VAR-GARCH-M model. Let us define \( y_t = \begin{bmatrix} 1 & y_t & h_t & \epsilon_t \end{bmatrix} \)
and rewrite the VAR-GARCH-M with this new variable

\[ y_t = Ay_{t-1} + \epsilon_t \]  

(6)

with

\[ A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \mu & A & B & D \\ \omega & 0 & F & G \\ \omega & 0 & F & G \end{bmatrix} \]  

and \( \epsilon_t = \begin{bmatrix} \epsilon_t \\ \nu_t \end{bmatrix} \)

where \( \nu_t \) is a vector of white noise. We have that \( E_t (u_{t+i}) = 0 \). Rewriting equation (3) in terms of the vector \( y_t \) gives

\[ g'E_t (y_{t+1}) = 0 \]  

(7)

where \( g' \) is a vector that selects the variables appearing in equation (3); namely \( E_t R_{t+1} \) and \( E_t h_{t+1}^{12} = Cov_t (R_{t+1}, m_{t+1}) \).

From equation (6), we can get the expectation about \( y_{t+1} \)

\[ E_t (y_{t+1}) = Ay_t \]  

(8)
Plugging this expectation into equation (7) gives

\[ g' A y_t = 0 \]  \hspace{1cm} (9)

Thus, the coefficients \( A \) of the variables’ dynamic must satisfy the constraint

\[ g' A = 0 \]  \hspace{1cm} (10)

in order for equation (3) to hold. The vector \( g' \) depends on the form of the SDF, which differs for each model studied in this paper. For the general SDF, this constraint imply that the excess return dynamic should be a function of minus the expected covariance between the excess return and the SDF:

\[ R_t = -Cov_{t-1} (R_t, m_t) + \varepsilon_t^R \]  \hspace{1cm} (11)

where \( \varepsilon_t^R \) is a white noise.

The econometric test consists of checking if the estimated unconstrained VAR with coefficient matrix \( A \) is significantly different from the estimated VAR where the above constraints are imposed. This can be done by classical Wald tests or Likelihood Ratio tests (see Greene p. 150).

3 The different utility functions and their SDF

3.1 Power utility

Power utility is the original utility function used by Mehra and Prescott (1985). I use it as a benchmark for habit formation and loss aversion models and to check that the equity premium puzzle is present when conditional moments are used on the specific data set investigated in this paper.
The power utility function is

\[
U(C) = \sum_{i=0}^{\infty} \beta^i \frac{C_{t+i}^{1-\sigma}}{1 - \sigma}
\]

where \(C_t\) is the consumption of the representative agent at time \(t\) and \(\sigma\) is her coefficient of relative risk aversion. The SDF derived from power utility is

\[
m_{t+1} = -\sigma \Delta c_{t+1}
\]

where \(\Delta c_{t+1} = \ln C_{t+1} - \ln C_t\) is the consumption growth rate. With this SDF the no arbitrage equation (3) is

\[
E_t(R_{t+1}) = \sigma \text{cov}_t(\Delta c_{t+1}, R_{t+1})
\]

### 3.2 Habit formation models

The particularity of habit formation models is that the utility is a function of the gap between the consumption and a habit \(H_t\), instead of on consumption level only. The habit is defined as a subsistence level under which the consumption cannot fall. The habit can be a function of past consumptions ("Catching up with the Joneses" functions introduced by Abel 1990) or of past and present consumption ("Keeping up with the Joneses" functions introduced by Gali 1994). There are two main families of habit formation utility functions. In the first one, the utility is a function of the difference between the consumption and the habit (e.g. Constantinides 1990 or Campbell and Cochrane 1999). In the second one, the utility is a function of the ratio between the consumption and the habit (e.g. Abel 1990, 1999, or Gali 1994). This paper focuses on the first kind of habit formation utility functions, which has recently attracted most
of the attention in the literature. The utility function is defined as

\[
U(C) = \sum_{i=0}^{\infty} \beta^i (C_{t+i} - H_{t+i})^{1-\gamma} \frac{1}{1-\gamma}
\]  

where \( \gamma \) is a parameter such that \( 0 < \gamma \leq 1 \).

Meyer and Meyer (2005) demonstrate that habit formation utility function are equivalent to utility function with decreasing relative risk aversion. Thus, they imply time varying relative risk aversion. If the consumption is close to (far from) the habit, the agent has a high (low) relative risk aversion.

Habit formation models make a distinction between internal and external habits. The representative agent has internal habits when she takes into account the impact of her present consumption on the current and future habit level. Internal habit models have been introduced by Constantinides (1990) and Abel (1990). External habits (Campbell and Cochrane 1999) correspond to the case in which agent ignore the impact of their present consumption on the habit level. Chen and Ludvigson (2004) try to determine which alternative best corresponds to the observed data. They conclude their data are better described by internal habit formation than by external habit formation models. Both alternatives are tested in this paper.

**3.2.1 Internal habit model**

The particularity of internal habit utility model is that agents take into account the impact of their present consumption on present and future habit. Thus, using the utility function in equation (15), the marginal utility of present consumption is equal to

\[
U'(C_t) = (C_t - H_t)^{-\gamma} - \sum_{i=0}^{\infty} \beta^i (C_{t+i} - H_{t+i})^{-\gamma} \frac{\partial H_{t+i}}{\partial C_t}
\]  

(16)
The first part of the right hand side of the equation is the utility obtained by consuming one additional unit of good today. The second part is the reduction of present and future utility that this additional unit induces through its effect on the habit. In other terms, consuming more today brings more utility today to the agent but, at the same time, it increases its (present and future) habit, which reduce its (present and future) utility.

As in the original model of Constantinides (1990), I model the habit as a function of past and present consumption. More particularly, the habit is equal to

\[ H_t = C_t^{\theta_0} C_{t-1}^{\theta_1} \cdots C_{t-k}^{\theta_k} \]  

(17)

where the \( \theta \) parameters are smaller than one and chosen such that the consumption does not fall under the habit.

Appendix B shows how to linearize equation (16) to get the log SDF of the internal habit utility function. This log SDF takes the form:

\[ m_{t+1} = \sum_{i=-k+1}^{k+1} \alpha_i \Delta c_{t+i} \]  

(18)

where \( \alpha_i \) are coefficient that mix utility function parameters and linearization parameters.

With internal habit, the SDF is a function of future and past consumption growth rates, whereas it was a function of only the next period growth rate with the simple power utility model. Note also that the coefficients \( \alpha \) are different from the coefficient of relative risk aversion.

Using this SDF in equation (3) yields

\[ E_t (R_{t+1}) = - \sum_{i=1}^{k} \alpha_i Cov_t (\Delta c_{t+i}, R_{t+1}) \]  

(19)
Note that the conditional covariance between $\Delta c_{t+i}$ and $R_{t+1}$ is equal to the covariance between the errors made at time $t$ about $\Delta c_{t+i}$ and $R_{t+1}$ respectively. According to the VAR in equation (6), the error made at time $t$ about the consumption growth rate at time $t+i$ is a linear function of all one-period errors between time $t$ and $t+i$. Thus we have

$$
\varepsilon_{t+1}^R = R_{t+1} - E_t (R_{t+1}) = g'_R \varepsilon_{t+1}
$$

(20)

$$
\varepsilon_{t,t+i}^{\Delta c} = \Delta c_{t+i} - E_t (\Delta c_{t+i}) = g'_{\Delta c} \sum_{j=1}^{i} A^{i-j} \varepsilon_{t+j}
$$

(21)

where $g'_R$ and $g'_{\Delta c}$ are row vector with zero in each column, with the exception of one 1 in the column corresponding to $R_t$ and $\Delta c_t$ in the vector $y_t$, respectively. These vector select the one period error on $R_t$ and $\Delta c_t$ respectively. With this notation and given the fact that the error term are not autocorrelated, we have that

$$
\text{Cov}_t (\Delta c_{t+i}, R_{t+1}) = g'_{R} E_t (\varepsilon_{t+1}^{} \varepsilon_{t+1}^{}) (A')^{i-1} g_{\Delta c}
$$

(22)

It can be shown by developing the matrices and vectors in this expression that it is equivalent to

$$
\text{Cov}_t (\Delta c_{t+i}, R_{t+1}) = a_{i,1} V_t (R_{t+1}) + a_{i,2} \text{Cov}_t (R_{t+1}, \Delta c_{t+1})
$$

(23)

where $a_{i,1}$ and $a_{i,2}$ are coefficients that are a linear combination of the coefficient in the matrix $(A')^{i-1}$ . Plugging this result into equation (19) yields

$$
E_t (R_{t+1}) = a_1 V_t (R_{t+1}) + a_2 \text{Cov}_t (\Delta c_{t+1})
$$

(24)

with $a_1 = \sum_{i=1}^{k} \alpha_i a_{i,1}$ and $a_2 = \sum_{i=1}^{k} \alpha_i a_{i,2}$. 

14
3.2.2 External habit model

In external habit models, the representative agent does not take into account the influence of its consumption on its habit level (future or present). This kind of model has been initiated by Campbell and Cochrane (1999) and I will concentrate on their model. The utility function is the same as in the internal habit model presented in the previous section (cf. equation (15)). With external habit, we have that the marginal utility of consumption is equal to

\[ U'(C_t) = (C_t - H_t)^{-\gamma} \]  \hspace{1cm} (25)

Defining the consumption surplus \( S_t = \frac{C_t - H_t}{C_t} \) and using it to compute the logarithm of the SDF yields

\[ m_{t+1} = -\gamma \Delta c_{t+1} - \gamma \Delta s_{t+1} \]  \hspace{1cm} (26)

where \( \Delta s_t = \ln S_t - \ln S_{t-1} \).

One particularity of Campbell and Cochrane’s model is to define the dynamic for the surplus \( s_t \) as\footnote{In the original model, Campbell and Cochrane use a constant consumption growth to compute \( E_t (\Delta c_{t+1}) \). This generalization is taken from Wachter (2005).}

\[ s_{t+1} = (1 - \omega) \bar{s} + \omega s_t + \lambda(s_t) (\Delta c_{t+1} - E_t (\Delta c_{t+1})) \]  \hspace{1cm} (27)

where \( \bar{s} \) is the steady state surplus level, \( \omega \) is a persistence parameter (0 \( \leq \delta < 1 \)) and \( \lambda(s_t) \) is the sensitivity of the surplus to unexpected changes in consumption, which is defines as

\[ \lambda(s_t) = \begin{cases} \frac{1}{\bar{s}} \sqrt{1 - 2 \left( \frac{s_t - \bar{s}}{\bar{s}} \right) \cdot 1} & \text{if } s_t < s_{\text{max}} \\ 0 & \text{if } s_t \geq s_{\text{max}} \end{cases} \]  \hspace{1cm} (28)
where \( \bar{S} = \exp(\bar{s}) \) and \( s_{\text{max}} = \bar{s} + \frac{1}{2} (1 - S^2) \). With this dynamic, we have that

\[
m_{t+1} = -\gamma \Delta c_{t+1} - \gamma ((1 - \omega) \bar{s} + (\omega - 1) s_t + \lambda(s_t) (\Delta c_{t+1} - E_t(\Delta c_{t+1}))) \tag{29}
\]

Taking the conditional covariance of the log SDF with the expected excess return as equation (3) requires yields

\[
E_t(R_{t+1}) = \gamma (1 + \lambda(s_t)) \text{Cov}_t(R_{t+1}, \Delta c_{t+1}) \tag{30}
\]

With Campbell and Cochrane’s model, the relative risk aversion is equal to \( \gamma (1 + \lambda(s_t)) \) and is thus time varying. Since the sensitivity \( \lambda(s_t) \) is negatively correlated with the surplus, then the closer the agent is from its habit, the higher \( \lambda(s_t) \) is and thus the higher her relative risk aversion.

The constrained model can be estimated by simultaneously estimating the VAR-GARCH-M together with equation (27). The resulting sensitivity \( \lambda(s_t) \) is used for the varying coefficient associated to \( \text{Cov}_t(R_{t+1}, \Delta c_{t+1}) \) in the VAR-GARCH-M model.

### 3.3 Loss aversion functions

Utility functions with loss aversion are based on the Cumulative Prospect Theory developed by Kahneman and Tversky (Kahneman and Tversky 1979 and Tversky and Kahneman 1992). In this framework, utility is a function of the difference between the consumption and a reference level.\(^9\) The particularity of loss aversion functions is that losses have a bigger impact on utility than gains of the same magnitude. In this paper, I study two types of loss aversion utility functions: one in which the reference level is the expected consumption and one in which the reference level is based on past

---

\(^8\)The sensitivity \( \lambda(s_t) \) is set equal to 0 when \( s_t > s_{\text{max}} = \bar{s} + \frac{1}{2} (1 - S^2) \) to insure for a positive term in the square root.

\(^9\)The reference level is different from the habit level, which is defined as a subsistence level. The consumption can go under the reference level in loss aversion model, whereas it has to stay above the subsistence level in habit formation models.
consumption. Note that in the original papers (Benartzi and Thaler 1995, Barberis et al. 2001) use a loss aversion function based on gains and losses in financial wealth. Barberis et al. (2001) justify the use of wealth instead of consumption as a way to capture "feelings unrelated to consumption" (p. 6). To remain close to the C-CAPM framework, in which agent’s utility is only influenced by her consumption, I use a loss aversion function defined on consumption rather than on financial wealth.

Following Köszegi and Rabin (2006), consider the general class of loss aversion function given by

\[ U(C) = \sum_{i=0}^{\infty} \beta^i (\delta V(C_{t+i}) + (1 - \delta) W(V(C_{t+i}) - V(X_{t+i}))) \]  

(31)

with \( \delta \in [0, 1] \). The utility of consumption at each period has two components: the first component \( V(C_{t+i}) \) is the traditional "consumption utility". The second component \( W(V(C_{t+i}) - V(X_{t+i})) \) is the "gain-loss utility", which is the utility derived from the deviation of the consumption \( C_{t+i} \) from a reference level \( X_{t+i} \). With this function, in addition to the utility of its consumption, the agent can get an "extra" utility if she consumes more than her reference level (gain in consumption utility). This extra utility is a function of the difference between the utilities of the consumption and of the reference level. Similarly, if the agent consumes less than her reference level (loss in consumption utility), she will get less utility. Typically with loss aversion, losses are more "painful" than gains and thus a loss generate a greater decrease of utility than the increase of utility associated to a gain of the same size.

In the next steps, I assume that the reference level is similar to the external habit in the sense that agents do not take into account the effect of their consumption on their present and future reference levels (i.e. \( \partial X_{t+i}/\partial C_t = 0 \)). With this assumption,
the marginal utility of consumption is equal to

\[ U' (C_t) = V' (C_t) (\delta + (1 - \delta) W' (V (C_t) - V (X_t))) \]  

(32)

Using once again a first-order Taylor approximation of the logarithm of this expression, we get that (see Appendix A)

\[ u' (C_t) = k + \nu' (C_t) + qw' (V (C_t) - V (X_t)) \]  

(33)

where \( k \) and \( q \) are linearization parameters.

I assume that the consumption utility is the traditional power utility function \((V (C_t) = C_t^{1-\sigma} / (1 - \sigma))\) and that the gain-loss utility takes the form

\[ W (Z) = \begin{cases} 
\frac{Z^{1-\nu}}{1-\nu} & \text{if } Z \geq 0 \\
-\eta \frac{(-Z)^{1-\nu}}{1-\nu} & \text{if } Z < 0 
\end{cases} \]  

(34)

where \( \eta > 1 \) and \( Z \) is the gain or loss of the agent. The parameter \( \lambda \) is the degree of loss aversion. The greater it is, the larger is the decrease in utility associated with a loss. Since \( \lambda > 1 \), the utility gain associated with a gain \( Z \) is smaller than the utility loss associated with a loss of the same size \( Z \). With this specification, we have that the logarithm of the marginal gain-loss utility is

\[ w' (Z) = \begin{cases} 
-\nu \ln Z & \text{if } Z \geq 0 \\
\ln \eta - \nu \ln (-Z) & \text{if } Z < 0 
\end{cases} \]  

(35)

In our case, \( Z_t = V (C_t) - V (X_t) \) and with a first order Taylor approximation of its
logarithm, we get that (see Appendix C)

\[
u'(C_t) = \begin{cases} 
  k^+ + q^+ c_t + p^+ x_t & \text{if } c_t \geq x_t \\
  k^- + q^- c_t + p^- x_t & \text{if } c_t < x_t
\end{cases}
\]

(36)

where \(k_s, q_s\) and \(p_s\) are linear combination of the utility functions parameters and the linearization parameters.

With this linear function, it is possible to compute the log SDF. The SDF take two different forms depending on the initial level of consumption. If the initial level of consumption is below the reference point \(c_t < x_t\) then

\[
m_{t+1} = \begin{cases} 
  \Delta k + q^+ \Delta c_{t+1} + p^+ \Delta x_{t+1} + \Delta q c_t - \Delta p x_t & \text{if } c_{t+1} \geq x_{t+1} \\
  \Delta q c_{t+1} + p^- \Delta x_{t+1} & \text{if } c_{t+1} < x_{t+1}
\end{cases}
\]

(37)

where \(\Delta k = k^+ - k^-\), \(\Delta q = q^+ - q^-\) and \(\Delta p = p^+ - p^-\). If the initial level of consumption is above the reference point \(c_t > x_t\) then

\[
m_{t+1} = \begin{cases} 
  q^+ \Delta c_{t+1} + p^+ \Delta x_{t+1} & \text{if } c_{t+1} \geq x_{t+1} \\
  -\Delta k + q^- \Delta c_{t+1} + p^- \Delta x_{t+1} - \Delta q c_t - \Delta p x_t & \text{if } c_{t+1} < x_{t+1}
\end{cases}
\]

(38)

To compute the covariance between \(R_{t+1}\) and \(m_{t+1}\), we use the following relation (see Appendix for the proof)

\[
\text{Cov}(X, BY) = \left( \frac{E(Y)}{\sqrt{V_1(Y)}} + \Pi \right) \text{Cov}(X, Y)
\]

(39)

where \(X\) and \(Y\) are two variables that are jointly normally distributed and \(B\) is a binary variable that take the value \(B^+\) and \(B^-\) in case of gain or loss, respectively.
Furthermore, \( \pi = \phi (B^+ - B^-) \) and \( \Pi = \frac{\Phi^2 + \phi^2 B^-}{\Phi} + \frac{(1-\Phi)^2 + \phi^2 B^+}{1-\Phi} \) where \( \Phi \) is the probability to observe a loss and \( \phi \) is the probability to observe neither a gain nor a loss. With this relation, it is possible to compute the covariance with the SDF in both situations (with an initial consumption below or above the reference level). The covariance with the constants is equal to zero since that case is equivalent to \( Y = 1 \). Furthermore, as the reference level for the next period is known at time \( t \) (since it is equal to the expected consumption or it is defined on past consumption), its conditional covariance with \( R_{t+1} \) is also equal to 0. Therefore, in both case, we have that

\[
Cov_t (R_{t+1}, m_{t+1}) = \left( \frac{E_t (\Delta c_{t+1})}{V_t^{1/2} (\Delta c_{t+1})} + \Pi_t \right) Cov_t (R_{t+1}, \Delta c_{t+1}) \tag{40}
\]

where \( \pi_t = \phi_t \Delta q \) and \( \Pi_t = \frac{\Phi_t^2 + \phi_t^2 q^-}{\Phi_t} + \frac{(1-\Phi_t)^2 + \phi_t^2 q^+}{1-\Phi_t} \) where \( \Phi_t \) is the probability to observe a loss at time \( t + 1 \) and \( \phi_t \) is the probability to observe neither a gain nor a loss.

Using this equation, we have that

\[
E_t (R_{t+1}) = -\psi_t Cov_t (R_{t+1}, \Delta c_{t+1}) \tag{41}
\]

with \( \psi_t = \frac{E_t (\Delta c_{t+1})}{V_t^{1/2} (\Delta c_{t+1})} + \Pi_t \) \tag{42}

I use two different reference levels to test the model. In the first model (Loss aversion A), the reference level is equal to the expected consumption

\[
X_{t+1} = E_t (C_{t+1}) \tag{43}
\]

The logarithm of the expected consumption can be approximated by a second order
Taylor approximation, which gives

\[ x_{t+1} = c_t + E_t(\Delta c_{t+1}) + \frac{1}{2} V_t(\Delta c_{t+1}) \]  

(44)

With this we can compute the probability \( \Phi_t \) that the consumption level falls below the reference level:

\[ \Pr(c_{t+1} < x_{t+1}) = \Pr\left(\varepsilon_{t+1}^c < \frac{1}{2} V_t(\Delta c_{t+1})\right) = \Phi\left(\frac{1}{2} V_t^{1/2}(\Delta c_{t+1})\right) \]  

(45)

where \( \Phi() \) is the cumulative normal distribution.

In the second model (Loss aversion B), the other reference point is defined as a function of past consumption levels such that

\[ X_{t+1} = C_t^\varphi X_t^{1-\varphi} \]  

(46)

With this definition, we can derive the dynamic of the difference \( d_t = c_t - x_t \) between the log consumption and the reference level

\[ d_t = (1 - \varphi) d_{t-1} + \Delta c_t \]  

(47)

With this, the probability \( \Phi_t \) that the consumption level falls below the reference level is

\[ \Pr(c_{t+1} < x_{t+1}) = \Pr\left(\varepsilon_{t+1}^c < -(1 - \varphi) d_t - E_t(\Delta c_{t+1})\right) \]

\[ = \Phi\left(-\frac{(1 - \varphi) d_t + g_{t}^\Delta A y_t}{V_t^{1/2}(\Delta c_{t+1})}\right) \]  

(48)
4 Empirical results

4.1 Data

The data set contains monthly data on U.S. consumption, prices and excess stock returns. The data are provided by Datastream. The excess return is the difference between the S&P500 index monthly return and the secondary market return on 3-month treasury bills. The real consumption growth rate $\Delta c_t$ is computed with the US personal consumption expenditure index and the consumption price index. I examine the period between January 1965 and March 2007 (507 observations for each series).

4.2 Estimation method for the VAR-GARCH-M

As seen in the Section 2.2, the VAR-GARCH-M is an adequate choice to model the joint dynamic. However, it has the disadvantage to be heavily parametrized. To limit the number of coefficients, I include only one lag for each variable of the model, which corresponds to a VAR(1)-GARCH(1,1)-M(1) model. Furthermore, we assume that the consumption growth rate $\Delta c_t$ depends only on their past values and on past excess returns. In other terms, their dynamic do not include a "in-mean" term, which is equivalent to setting the last two rows of the matrix $B$ in equation (4) equal to 0. In addition, the parameters of the variance-covariance dynamic must fulfill some conditions in order to insure a positive definite covariance matrix. For this, I apply the the BEKK specification (Engle and Kroner 1995) to model the multivariate GARCH dynamic. The model can then be estimated by maximizing the log likelihood function.\(^\text{10}\)

I estimate three versions of the VAR-GARCH-M. The first one (Model 1) is the model in equation (4) and (5) with $y'_t = \begin{bmatrix} R_t & \Delta c_t \end{bmatrix}$. This model is used to check if the constraints of the constant relative risk aversion and the internal habit models

\(^{10}\text{See Hamilton (1994) for a more complete introduction to multivariate GARCH models and their estimation with maximum likelihood methods.}\)
are verified by the empirical data. The external habit model (Model 2) is a slightly modified version of Model 1. Firstly, as equation (30) shows, the covariance between the excess returns and the consumption growth enters in the excess return dynamic with a time varying coefficient. To take this into account, one term is added to equation (4)
\[ y_t = \mu + Ay_{t-1} + Bh_{t-1} + De_{t-1} + fw_t + \varepsilon_t \]  
(49)
where \( w'_t = \lambda (s_{t-1}) \begin{bmatrix} h_{t-1}^{12} & e_{t-1}^{12} \end{bmatrix} \) and \( f \) is a vector of coefficients. Since I assumed that \( \Delta c_t \) is independent from covariance term, I set the last row of \( f \) equal to zero. The sensitivity \( \lambda (s_{t-1}) \) is computed with equation (28). The surplus dynamic in equation (27) is estimated simultaneously with the VAR-GARCH-M equations (4) and (5) using likelihood maximization.

Similarly, the third version of the VAR-GARCH-M (Model 3) estimated the loss aversion models. Firstly, I estimate the equation (41) with the following model
\[ y_t = \mu + Ay_{t-1} + Bh_{t-1} + De_{t-1} + vv_t + \varepsilon_t \]  
(50)
where \( v'_t = \left( \pi_t \frac{E_t(\Delta c_{t+1})}{V_t^{1/2}(\Delta c_{t+1})} + \Pi_t \right) \begin{bmatrix} h_{t-1}^{12} & e_{t-1}^{12} \end{bmatrix} \) with \( \pi_t = \phi_t (q_1 - q_2) \) and \( \Pi_t = \frac{\Phi_t^2 + \phi_t^2}{\Phi_t} q_1 + \frac{(1-\Phi_t)^2 + \phi_t^2}{1-\Phi_t} q_2 \). As in Model 2, the second row of the coefficient vector \( v \) is set equal to zero. In the model Loss aversion A (expected consumption as reference level), the probability \( \Phi_t \) and \( \phi_t \) are simply
\[ \Phi_t = \Phi \left( \frac{1}{2} \left( h_t^2 \right)^{1/2} \right) \quad \text{and} \quad \phi_t = \phi \left( \frac{1}{2} \left( h_t^2 \right)^{1/2} \right) \]  
(51)
In the model Loss aversion B (reference level as function of past consumption), the dynamic of the difference between the consumption and the reference level \( d_t \) is
### Table 1: Estimated unconstrained model 1

<table>
<thead>
<tr>
<th></th>
<th>$R_t$</th>
<th>$\Delta c_t$</th>
<th>$V_t (R_{t+1})$</th>
<th>$V_t (\Delta c_{t+1})$</th>
<th>$Cov_t (R_{t+1}, \Delta c_{t+1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>-0.0064</td>
<td>0.0029**</td>
<td>7.56e-5**</td>
<td>1.50e-5*</td>
<td>0.98e-5*</td>
</tr>
<tr>
<td>$R_{t-1}$</td>
<td>-0.0275</td>
<td>0.0251**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\Delta c_{t-1}$</td>
<td>0.4062</td>
<td>-0.2063**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (R_t)$</td>
<td>3.8866</td>
<td>.</td>
<td>0.8681**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (\Delta c_t)$</td>
<td>258.5174*</td>
<td>.</td>
<td>0.2908**</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>$Cov_{t-1} (R_t, c_t)$</td>
<td>-114.9734**</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.5024**</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1})^2$</td>
<td>-1.1415</td>
<td>.</td>
<td>0.0918**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$(\varepsilon_{\Delta c}^2)$</td>
<td>15.9545</td>
<td>.</td>
<td>.</td>
<td>0.2596**</td>
<td>.</td>
</tr>
<tr>
<td>$(\varepsilon_{R_{t-1} \Delta c_{t-1}})$</td>
<td>2.7790</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.1544**</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>2823.608</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

* (***) denotes that the coefficient is significant at the 5% (1%) confidence level.

estimated simultaneously with equation (47). The probability $\Phi_t$ and $\phi_t$ are then

$$
\Phi_t = \Phi \left( -\frac{(1 - \eta) d_t + g_{\Delta c} A y_t}{(h_t^2)^{1/2}} \right) \quad \text{and} \quad \phi_t = \phi \left( -\frac{(1 - \eta) d_t + g_{\Delta c} A y_t}{(h_t^2)^{1/2}} \right)
$$

(52)

### 4.3 Estimation of the joint dynamic

The estimation of the unconstrained Model 1 is displayed in Table 1. Figure 2 shows the observed excess returns and the expected excess return given by the joint dynamic in Model 1. Most of the coefficients of the excess return dynamic are not significant with the exception of the coefficients associated to past consumption growth variance and to past covariance between consumption and excess returns. Note that the latter is significantly negative. The parameters of the consumption growth rate dynamic, of both conditional variances and of the conditional covariance are all significant. The excess return variance has the strongest inertia of the three second moments, followed by the covariance. Figure 3 shows the estimated second moments of excess returns and consumption growth rate.

An important observation about the plausibility of the different models can already be made by looking at the joint dynamic coefficients. Note that the past conditional
covariance between excess returns and consumption growth rates enters the equation with a significant negative coefficient and that the product of past residuals coefficient is not significative. Since the conditional covariance is a positive function of both terms, it suggests that the conditional covariance is inversely related to expected excess returns, as documented in Duffee (2005). The relation is the opposite when unconditional moments are used. A negative correlation between the conditional covariance and expected excess returns is in contradiction with models based on power utility or external habit. Such correlation is however possible for some particular configurations of the external habit coefficients or for loss aversion utility. In the loss aversion case, this can be explained by the fact that loss averse agents can be risk lover when they expect losses in consumption. Concretely, since losses causes negative utility, loss averse agents are ready to take more risk to avoid them. Risk loving agents (or at least "occasional" risk loving agents) can then explain the observed negative correlation between the expected return and the conditional covariance.
Figure 3: Estimated conditional second moments

Excess returns volatility

Consumption growth rate volatility

Covariance between excess returns and consumption growth rate
### Table 2: Estimated constant relative risk aversion model

<table>
<thead>
<tr>
<th></th>
<th>$R_t$</th>
<th>$\Delta c_t$</th>
<th>$V_t (R_{t+1})$</th>
<th>$V_t (\Delta c_{t+1})$</th>
<th>$\text{Cov}<em>t (R</em>{t+1}, \Delta c_{t+1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>.</td>
<td>0.0029**</td>
<td>9.28e-5**</td>
<td>1.30e-5*</td>
<td>1.02e-5*</td>
</tr>
<tr>
<td>$R_{t-1}$</td>
<td>.</td>
<td>0.0229**</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>$\Delta c_{t-1}$</td>
<td>.</td>
<td>-0.2050**</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>$\sigma$</td>
<td>41.9390**</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>$V_{t-1} (R_t)$</td>
<td>.</td>
<td>.</td>
<td>0.8829**</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>$V_{t-1} (\Delta c_t)$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.3620**</td>
<td></td>
</tr>
<tr>
<td>$\text{Cov}_{t-1} (R_t, c_t)$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.5654**</td>
</tr>
<tr>
<td>$(\epsilon_{R_{t-1}})^2$</td>
<td>.</td>
<td>.</td>
<td>0.0650**</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>$(\epsilon_{\Delta c_{t-1}})^2$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.2492**</td>
<td></td>
</tr>
<tr>
<td>$\epsilon_{R_{t-1}} \epsilon_{\Delta c_{t-1}}$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.1273**</td>
</tr>
</tbody>
</table>

Log likelihood 2812.971

* (**) denotes that the coefficient is significant at the 5% (1%) confidence level.

### 4.4 Constant relative risk aversion model

The estimated coefficients of the constant relative risk aversion model are presented in Table 2. Not surprisingly, the coefficient of the consumption growth and of the second moments are very similar to those of the unconstrained model.

With the VAR-GARCH-M approach, the estimated coefficient of relative aversion is approximately equal to 42. This value is in-line with other previous studies and is well above the values that are traditionally considered as admissible for the coefficient of relative aversion. However, some authors claim that such a value might still be compatible with a representative agent with constant relative risk aversion. Hens and Woehrmann (2006), for example, show that when one applies mental accounting to the constant relative risk aversion model (as it is normally done in experiment), then an estimated CRRA higher than 10 is compatible with experimental results. The methodology proposed here offers another way to test the validity of the constant relative risk aversion model by comparing the theoretical and the observed dynamic rather than by assessing the plausibility of the CRRA value. The unconstrained and the constrained expected excess returns are presented in Figure 4. A first visual comparison suggests that both dynamic are quite different. In other terms, the factors included
in the unconstrained model and not in the constant relative risk aversion model seem to have an impact on the dynamic of expected excess returns. Whether this impact is statistically significant is formally tested in section 4.9. A few statistic can already give us a hint about it. The rank correlation\(^{11}\) between the expected excess returns given by the unrestricted model and the one given by the constant relative risk aversion model is -0.05, which is negative but not significant. More worrying, when we restrict the sample to the months when the excess return is more extreme (inferior to the first quartile or superior to the third quartile), the rank correlation becomes significant and equal to -0.13. This means that the CRRA model tends to gives expected returns of the wrong direction, when the excess returns (positive or negative) are large in historical comparison.

\(^{11}\)The rank correlation (Spearman correlation) is similar to the traditional correlation. It measure the link between two variables. The difference between traditional and rank correlations is that the former indicates the strength of the linear link between two variables whereas the latter gives the strength of the link without assuming any particular form for this link.
### Table 3: Estimated internal habit model

<table>
<thead>
<tr>
<th></th>
<th>( R_t )</th>
<th>( \Delta c_t )</th>
<th>( V_t (R_{t+1}) )</th>
<th>( V_t (\Delta c_{t+1}) )</th>
<th>( Cov_t (R_{t+1}, \Delta c_{t+1}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c )</td>
<td>.</td>
<td>0.0029**</td>
<td>6.96e-5**</td>
<td>1.34e-5*</td>
<td>1.03e-5*</td>
</tr>
<tr>
<td>( R_{t-1} )</td>
<td>.</td>
<td>0.0238**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( \Delta c_{t-1} )</td>
<td>.</td>
<td>-0.2008**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>2.2691**</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-0.0887</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( V_{t-1} (R_t) )</td>
<td>.</td>
<td>.</td>
<td>0.8890**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( V_{t-1} (\Delta c_t) )</td>
<td>.</td>
<td>.</td>
<td>0.3512**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( Cov_{t-1} (R_t, c_t) )</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.5587**</td>
</tr>
<tr>
<td>( (\varepsilon_{t-1})^2 )</td>
<td>.</td>
<td>.</td>
<td>0.0740**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( (\varepsilon_{t-1})^2 )</td>
<td>.</td>
<td>.</td>
<td>0.2490**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>( R_{t-1} \Delta c_{t-1} )</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.1358**</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>2816.191</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

* (***) denotes that the coefficient is significant at the 5% (1%) confidence level.

### 4.5 Internal habit formation

The estimated coefficients of the internal habit formation model are presented in Table 3. The coefficients of the consumption growth rate dynamic and of the second moments are similar to those of the unconstrained model. In opposition to the CRRA model, the coefficient associated with the covariance \( (\alpha_2) \) is non significant (and negative). In this model, the excess return variance has a significant positive impact. The higher the excess return variance is, the higher the expected return. This positive effect between excess return and its variance is well known.

Figure 5 compares the expected excess returns of the unrestricted model with the excess return of the internal habit formation model. With this model, the rank correlation between them is significant and equal to 0.25. This implies that the excess returns generated with the internal habit formation model go in the same direction as the observed excess returns. Since the constant relative risk aversion model is a constrained version of the internal habit model (with \( \alpha_1 = 0 \)), we can test if the internal habit formation model significantly improve the constant relative risk aversion model with a Wald test. The null hypothesis (\( \alpha_1 = 0 \)) is rejected at a 1% confidence level.
All these facts hints that the internal habit formation model improves the traditional constant relative risk aversion model.

### 4.6 External habit formation

As explained in Section 4.2, the unconstrained model for the external habit model is slightly different from Model 1, in order to take the term $\lambda (s_{t-1}) \text{Cov}_t (R, \Delta c_t)$ into account. The results of both the unconstrained and the constrained models for external habit formation model are presented in Table 4. The coefficients of the consumption growth dynamic and of the second moments dynamics are very similar to Model 1 in both models. Contrary to Model 1, in the excess return equation, the coefficient associated to the past covariance is not significant anymore. The coefficient of the consumption growth variance remains significant. All other coefficients are not significant. The constrained model indicates that the estimated parameter $\gamma$ of the habit utility function is 0.61, which is consistent with its theoretical value. Figure 6 displays the expected excess returns of the unconstrained and of the constrained model. The rank correlation between the two series is statistically significant and equal to -0.15. This implies that the excess returns generated by the external habit formation models move
Table 4: Estimated external habit model

<table>
<thead>
<tr>
<th></th>
<th>$R_t$</th>
<th>$\Delta c_t$</th>
<th>$V_t (R_{t+1})$</th>
<th>$V_t (\Delta c_{t+1})$</th>
<th>$Cov_t (R_{t+1}, \Delta c_{t+1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>-0.0177</td>
<td>0.0029**</td>
<td>5.71e-5**</td>
<td>1.85e-5**</td>
<td>1.40e-5**</td>
</tr>
<tr>
<td>$R_{t-1}$</td>
<td>-0.0525</td>
<td>0.0261**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\Delta c_{t-1}$</td>
<td>0.5299</td>
<td>-0.2161**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (R_t)$</td>
<td>2.3552</td>
<td>.</td>
<td>0.8615**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (\Delta c_t)$</td>
<td>348.8674*</td>
<td>.</td>
<td>.</td>
<td>0.1815**</td>
<td>.</td>
</tr>
<tr>
<td>$Cov_{t-1} (R_t, c_t)$</td>
<td>-1290.8180</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.3955**</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1} R)^2$</td>
<td>-1.3288</td>
<td>.</td>
<td>0.1110**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1} \Delta c)^2$</td>
<td>16.2469</td>
<td>.</td>
<td>0.2673**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\varepsilon_{t-1} \varepsilon_{t-1}$</td>
<td>82.2660</td>
<td>.</td>
<td>.</td>
<td>0.1722**</td>
<td>.</td>
</tr>
<tr>
<td>$\lambda (s_{t-1})$</td>
<td>0.0164</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\lambda (s_{t-1}) Cov_{t-1} (R_t, \Delta c_t)$</td>
<td>1830.7340</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\lambda (s_{t-1}) \varepsilon_{t-1} R \varepsilon_{t-1} \Delta c$</td>
<td>-115.7720</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>2832.693</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c$</td>
<td></td>
<td>0.0029**</td>
<td>8.42e-5**</td>
<td>1.24e-5**</td>
<td>0.77e-5**</td>
</tr>
<tr>
<td>$R_{t-1}$</td>
<td>.</td>
<td>0.0241**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\Delta c_{t-1}$</td>
<td>.</td>
<td>-0.2049**</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (R_t)$</td>
<td>.</td>
<td>.</td>
<td>0.8782**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (\Delta c_t)$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.3812**</td>
<td>.</td>
</tr>
<tr>
<td>$Cov_{t-1} (R_t, c_t)$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.5785**</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1} R)^2$</td>
<td>.</td>
<td>.</td>
<td>0.0772**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1} \Delta c)^2$</td>
<td>.</td>
<td>.</td>
<td>0.2483**</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\varepsilon_{t-1} \varepsilon_{t-1}$</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0.1385**</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.6142</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>2815.027</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

* (***) denotes that the coefficient is significant at the 5% (1%) confidence level.

In the opposite direction of the observed excess returns. This is a first negative sign for the external habit model.

In addition to the results presented in Table 4 and in Figure 6, the estimation of the external habit also provides us with an estimation of the surplus dynamic given by equation (27). The estimated parameters that are relevant for the surplus dynamic are

$$\omega = 0.9924 \text{ and } \bar{S} = 59.36\%$$
I find an average surplus equal to 59% of the total consumption. Figure 7 gives the estimated dynamic for the surplus generated by these parameters has well as the dynamic of the sensitivity and its estimated distribution. The estimation show that the surplus was the lowest between 1981 and 1983 with a minimum of 55.8% of total consumption in June 1982. The first half of the nineties is also a period in which the surplus was relatively low. Since then it has evolved below its mean but close to it. These results suggest that investors were the most risk averse in 1981-1983 an in the first half of the nineties. Accordingly, during these period they required higher returns to compensate for the risk. Note that, I observe no significant correlation between the coefficient of risk aversion and the covariance between consumption growth and excess returns. Thus, according to this estimation, the risk aversion is independent of the risk associated with the stock market.

The estimated persistence factor $\omega$ estimated here is slightly higher than in Campbell and Cochrane (1999) (0.9924 vs. 0.9856). The estimated average surplus is about 10 times bigger than the one of Campbell and Cochrane: 59.36% vs. 5.7%! One consequence of this differences is that the surplus estimated here evolves much further...
In each panel, the solid black line corresponds to the theoretical maximum surplus and the dashed grey line to the average surplus. The upper right panel shows the sensitivity corresponding to the estimated surplus. The lower right panel displays the function linking the sensitivity (vertical axis) to the surplus (horizontal axis).
away from the habit than in Campbell and Cochrane. A direct consequence of this gap is that the sensitivity is relatively stable and low. This implies that the price of risk generated by external habit formation model is also stable and low. This is in-line with Duffee (2005), who argues that price of risk in external habit formation model does not vary enough to justify the large swings in the equity premium and speaks against Campbell and Cochrane’s external habit formation model.

### 4.7 Loss aversion model A (reference point based on expected consumption)

The estimation of the loss aversion model with a reference point based on expected consumption is presented in Table 5. As for previous models, the coefficients for consumption growth and for conditional second moments are similar to Model 1. In the unconstrained model, all coefficients are non significant.

Figure 8 presents the expected excess returns given by the unconstrained model and by the restricted model. We observe a rank correlation of 0.06, which is statistically
Table 5: Estimated loss aversion model A

<table>
<thead>
<tr>
<th></th>
<th>Unconstrained model</th>
<th>3</th>
<th>Constrained model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R_t$</td>
<td>$\Delta c_t$</td>
<td>$V_t (R_{t+1})$</td>
</tr>
<tr>
<td>$c$</td>
<td>-0.0048</td>
<td>0.0028**</td>
<td>6.90e-5**</td>
</tr>
<tr>
<td>$R_{t-1}$</td>
<td>-0.0136</td>
<td>0.0269**</td>
<td>.</td>
</tr>
<tr>
<td>$\Delta c_{t-1}$</td>
<td>0.3605</td>
<td>-0.1849**</td>
<td>.</td>
</tr>
<tr>
<td>$V_{t-1} (R_t)$</td>
<td>3.8066</td>
<td>.</td>
<td>0.8620**</td>
</tr>
<tr>
<td>$V_{t-1} (\Delta c_t)$</td>
<td>272.9046</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$Cov_{t-1} (R_t, \Delta c_t)$</td>
<td>-128.9434</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1}^R)^2$</td>
<td>-1.3869</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$(\varepsilon_{t-1}^{\Delta c})^2$</td>
<td>9.6495</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\varepsilon_{t-1}^R \Delta c_{t-1}$</td>
<td>9.6022</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.0002</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\eta_t Cov_{t-1} (R_t, \Delta c_t)$</td>
<td>2.7163</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$\eta_t \varepsilon_{t-1}^R \varepsilon_{t-1}^{\Delta c}$</td>
<td>-0.6421</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>2825.084</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

|                      |                      |              | Constrained model  |
|                      |                      |              | 3            |
| $c$                  |                      | 0.0028**     | 8.97e-5**     | 1.39e-5**               | 1.09e-5**               |
| $R_{t-1}$            |                      | 0.0221**     | .              | .                      |
| $\Delta c_{t-1}$    |                      | -0.1933**    | .              | .                      |
| $V_{t-1} (R_t)$      |                      | .            | 0.8835**       | .                      |
| $V_{t-1} (\Delta c_t)$ |                      | .            | .              | 0.3445**               |
| $Cov_{t-1} (R_t, c_t)$ |                      | .            | .              | .                      |
| $(\varepsilon_{t-1}^R)^2$ |                      | .            | .              | 0.0636**               |
| $(\varepsilon_{t-1}^{\Delta c})^2$ |                      | .            | .              | 0.2312**               |
| $\varepsilon_{t-1}^R \varepsilon_{t-1}^{\Delta c}$ |                      | .            | .              | 0.1213**               |
| Log likelihood       | 2813.880            | .            | .              | .                      |

* (**) denotes that the coefficient is significant at the 5% (1%) confidence level.
not significant. In addition, the estimation gives us the following parameters

\[ q^- = 112.29 \text{ and } q^+ = -102.88 \]

Note however that these two coefficients are not statistically significantly different from each other, which makes it difficult to infer some value for the utility parameter \( \sigma \) (cf. Appendix C). With these coefficients it is possible to estimate the evolution of the difference between the reference level and actual consumption and the probability to fall under the reference level in the next period (cf. Figure 9). We can see that this probability varies from one period to another, but stays very close to its mean of 50.11%.

4.8 Loss aversion model B (reference level based on past consumption)

The estimation of the loss aversion model with a reference level based on past consumption is presented in Table 6. As for previous models, the coefficients of the consumption growth and of the conditional second moments are similar to Model 1. In the uncon-
Table 6: Estimated loss aversion model B

<table>
<thead>
<tr>
<th></th>
<th>Unconstrained model 3</th>
<th>Constrained model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( R_t ) ( \Delta c_t ) ( V_t (R_{t+1}) ) ( V_t (\Delta c_{t+1}) ) ( \text{Cov}<em>t (R</em>{t+1}, \Delta c_{t+1}) )</td>
<td></td>
</tr>
<tr>
<td>( c )</td>
<td>0.0149 0.0029**</td>
<td>8.34e-5 3.68e-5 5.25e-5</td>
</tr>
<tr>
<td>( R_{t-1} )</td>
<td>-0.0223 0.0240**</td>
<td></td>
</tr>
<tr>
<td>( \Delta c_{t-1} )</td>
<td>0.5159 -0.1943**</td>
<td></td>
</tr>
<tr>
<td>( V_{t-1} (R_t) )</td>
<td>3.4726 0.8566**</td>
<td></td>
</tr>
<tr>
<td>( V_{t-1} (\Delta c_t) )</td>
<td>223.8821 0.1874**</td>
<td></td>
</tr>
<tr>
<td>( \text{Cov}_{t-1} (R_t, \Delta c_t) )</td>
<td>-323.9910 0.4006**</td>
<td></td>
</tr>
<tr>
<td>( (\varepsilon_{t-1})^2 )</td>
<td>-0.8172 0.1111**</td>
<td></td>
</tr>
<tr>
<td>( (\varepsilon_{\Delta_t})^2 )</td>
<td>0.3169 0.2609**</td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_{-1}^{R_t} \varepsilon_{t-1}^{\Delta c_t} )</td>
<td>-43.0290 0.1702**</td>
<td></td>
</tr>
<tr>
<td>( \eta_{t} )</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td>( \eta_{t} \text{Cov}_{t-1} (R_t, \Delta c_t) )</td>
<td>-11.6878</td>
<td></td>
</tr>
<tr>
<td>( \eta_{t}^{R_t} \varepsilon_{t-1}^{\Delta c_t} )</td>
<td>-2.6755</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>2830.753</td>
<td></td>
</tr>
</tbody>
</table>

|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |

\* (**) denotes that the coefficient is significant at the 5% (1%) confidence level.

Strained model, all coefficients are non significant. The results of the constrained model are more disturbing because the coefficients are non significant, which is very different from the previous models. In all other models, we observe a stability of the consumption growth dynamic coefficients between the constrained and the unconstrained model. Here, the coefficients change significantly.

Figure 10 presents the expected excess returns given by the unconstrained model and by the restricted model. We observe a rank correlation of -0.07, which is statistically not significant. In addition, the estimation of the constrained model gives us the
following parameters

\[ q^- = 91.07 \text{ and } q^+ = -38.40 \]

These two coefficients are not statistically significantly different from each other, which makes it difficult to infer some value for the utility parameter \( \sigma \) (cf. Appendix C). The unconstrained model give us \( \eta = 0.9593 \) whereas the constrained model gives \( \eta \simeq 1 \) which is equivalent to a model where the reference level is equal to the present consumption.

### 4.9 Tests on the excess returns dynamic

As section 2.3 explain, it is possible to use a formal test to check if the dynamic of the expected returns does follow the one predicted by one particular model. For that, I test the set of restrictions \( g'A = 0 \) with a Wald test and a Likelihood ratio test. The results are displayed in Table 7.

The results clearly show that null hypothesis are strongly rejected for the CRRA model and for the external habit formation model. Therefore, the dynamic implied by these models does not formally correspond to the observed dynamic. The internal
**Figure 11: Gap with reference level and probability to fall under it (loss aversion B)**

*Left panel: gap between the consumption and its reference level. Right panel: probability for the consumption to be under its reference level in the next period.*

<table>
<thead>
<tr>
<th></th>
<th>Wald test</th>
<th>Likelihood Ratio test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test stat.</td>
<td>p-value</td>
</tr>
<tr>
<td>Constant relative risk aversion</td>
<td>24.1077</td>
<td>0.0022</td>
</tr>
<tr>
<td>Internal habit formation</td>
<td>17.1301</td>
<td>0.0166</td>
</tr>
<tr>
<td>External habit formation</td>
<td>29.1585</td>
<td>0.0021</td>
</tr>
<tr>
<td>Loss aversion A</td>
<td>17.8322</td>
<td>0.1209</td>
</tr>
<tr>
<td>Loss aversion B</td>
<td>23.3300</td>
<td>0.0251</td>
</tr>
</tbody>
</table>

* (* *) denotes that the coefficient is significant at the 5% (1%) confidence level.
habit formation model and the loss aversion model with a reference level based on past consumption are accepted at a 1% confidence level by both tests. The loss aversion model with a reference level defined by the expected consumption is not rejected at a 10% confidence level by Wald test and at a 1% confidence level by the Likelihood ratio test. On the basis of these tests, the loss aversion model with expected consumption seems to be the one which is the most in adequacy with the observed dynamic of excess return. It is however still rejected by the likelihood test at a 5% confidence level. The internal habit formation model and the loss aversion model based on past consumption are the next plausible candidates, but they are both rejected at a 5% confidence level. Finally, the constant relative risk aversion model and the external habit formation model are strongly rejected.

5 Conclusion

The results of the estimated linear SDF models show that the constant relative risk aversion model and the external habit model are strongly rejected by the data. The theoretical excess returns generated by these models do not correspond to the observed dynamic. I also find evidence that the conditional covariance between excess returns and consumption growth seems to be negatively related with excess returns, which is in contradiction with those two models. This finding is in-line with Duffee (2005) results. Finally, the constraints implied by constant relative risk aversion and external habit formation are formally rejected by both Wald and Likelihood ratio test. The validity of the loss aversion model using past consumption as reference level is also doubtful. Although its restrictions cannot be rejected at a 1% confidence level, the values of the estimated parameters cast some doubt on it.

These conclusions let us with two remaining candidates to explain the excess return dynamic: the internal habit formation model and the loss aversion model with expected consumption as reference level. Both models cannot be rejected at a 1% confidence
level and have plausible parameters. My preference goes to the loss aversion model because the Wald test does not reject its constraints at a 10% level, whereas the internal habit formation model is rejected at a 5% level by both tests. This suggests that loss aversion models do play a role in the expected excess returns movements and that they might be a suitable answer to the equity premium puzzle. Consequently, the loss aversion component should be included in consumption-based model in this context. Note that the results presented here also confirm those of Chen and Ludvigson (2004), who find that data are better described by internal habit formation than by external habit formation.

One interesting follow-up to this paper would be to use the VAR-GARCH-M framework to test the original specification of the loss aversion model, where the loss function is defined on wealth rather than on consumption. If the change in wealth better explains the data, then one should remember Barberis et al. (2001) justification for using wealth and study "feelings unrelated to consumption" (Barberis et al. 2001 p. 6) to understand the factors driving the expected excess returns. Without looking to far away from the original consumption-based model, the results presented here suggest that consumption growth variance, i.e. the uncertainty about consumption, might play a role in the excess return dynamic. This factor does not appear in any of the proposed models and thus might constitute a starting point for developing a more adequate model.

A Approximation of the logarithms

A.1 Logarithm of a sum

Campbell, Lo and McKinlay (1997) show that we can approximate the logarithm of a sum by the sum of logarithm. First consider
\[\ln (A + B) = \ln A \left(1 + \frac{B}{A}\right) = a + \ln \left(1 + e^{b-a}\right) \tag{53}\]

where \(a = \ln A\) and \(b = \ln B\). The second part of this equation can be approximated by a standard Taylor approximation around its mean. Define \(x = b - a\) and \(f(x) = \ln (1 + e^x)\) The Taylor approximation yields \(f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x})\) with \(f'(\bar{x}) = e^\bar{x}/(1 + e^\bar{x})\). Plugging this into equation (53) gives the final result

\[\ln (A + B) \simeq \kappa + \rho a + (1 - \rho) b\tag{54}\]

with \(\rho = 1/(1 + e^\bar{x})\) and \(\kappa = -\ln \rho - (1 - \rho) \ln (1/\rho - 1)\). We have that \(0 < \rho < 1\)

### A.2 Logarithm of a subtraction

Similarly for a subtraction, we have that

\[\ln (A - B) = \ln A \left(1 - \frac{B}{A}\right) = a + \ln \left(1 - e^{b-a}\right) \tag{55}\]

Define \(g(x) = \ln (1 - e^x)\) which implies \(g'(\bar{x}) = -\frac{e^\bar{x}}{1-e^\bar{x}}\) and thus

\[\ln (A - B) \simeq \varsigma + \psi a - (\psi - 1) b\tag{56}\]

with \(\psi = 1/(1 - e^\bar{x})\) and \(\varsigma = -\ln \psi - (1 - \psi) \ln (1 - 1/\psi)\). Note that this approximation is possible only if \(A - B > 0\), which implies that \(b - a < 0\) and thus \(\psi > 1\).
B Linearization of the internal habit marginal utility

From the definition of the habit level in equation (5), we can compute the derivative of the habit with respect to consumption:

\[
\frac{\partial H_{t+i}}{\partial C_t} = \begin{cases} 
\theta_i H_{t+i} C_t^{-1} & \text{if } i \leq k \\
0 & \text{if } i > k 
\end{cases} \tag{57}
\]

Plugging this in equation (16)

\[
U'(C_t) = (C_t - H_t)^{-\gamma} - \sum_{i=0}^{k} \delta^i (C_{t+i} - H_{t+i})^{-\gamma} \theta_i H_{t+i} C_t^{-1} \tag{58}
\]

The first order condition of the C-CAPM states that the marginal utility of consumption is equal to the Lagrangian of the maximization problem. The Lagrangian is equal to the marginal utility of relaxing the budget constraints by one unit, which correspond to marginal utility of one additional unit of initial wealth. By definition, this marginal utility is always positive, which implies that the marginal utility of consumption is also positive and the logarithm of it exists. Taking the log of this difference yields (see Appendix A)

\[
\ln U'(C_t) = \varsigma_1 - \gamma \psi_1 \ln (C_t - H_t) + (\psi_1 - 1) c_t - \\
- (\psi_1 - 1) \ln \sum_{i=0}^{k} \delta^i (C_{t+i} - H_{t+i})^{-\gamma} \theta_i H_{t+i} \tag{59}
\]
The log of the sum in the last term can also be linearized with a Taylor first order approximation (see Appendix A). We have that

\[
\ln \sum_{i=0}^{\infty} \delta^i (C_{t+i} - H_{t+i})^{-\gamma} \theta_i H_{t+i} = \kappa_1 + (1 - \rho_1) \left( \ln \theta_0 - \gamma \ln (C_t - H_t) + h_t \right) \tag{60}
\]

\[
+ \rho_1 \ln \sum_{i=1}^{\infty} \delta^i (C_{t+i} - H_{t+i})^{-\gamma} \theta_i H_{t+i}
\]

Solving this equation forward yields

\[
\ln \sum_{i=0}^{\infty} \delta^i (C_{t+i} - H_{t+i})^{-\gamma} \theta_i H_{t+i} = \sum_{i=0}^{k} \left( \prod_{j=0}^{i-1} \rho_j \right) \left( \kappa_i + (1 - \rho_i) \left( i \ln \delta + \ln \theta_i - \gamma \ln (C_{t+i} - H_{t+i}) + h_{t+i} \right) \right) \tag{61}
\]

Collecting all constant terms yields in \( \kappa'' \)

\[
\ln \sum_{i=0}^{\infty} \delta^i (C_{t+i} - H_{t+i})^{-\gamma} \theta_i H_{t+i} = \kappa'' + \sum_{i=0}^{k} \left( \prod_{j=0}^{i-1} \rho_j \right) (1 - \rho_i) (h_{t+i} - \gamma \ln (C_{t+i} - H_{t+i})) \tag{62}
\]

Since, by definition, the consumption is always greater than the habit, we can also approximate the logarithm of the difference between them

\[
\ln (C_t - H_t) = \varsigma_2 + \psi_2 c_t - (\psi_2 - 1) h_t \tag{63}
\]

Taking the logarithm of equation (17) gives

\[
h_t = \theta_0 c_t + \theta_1 c_{t-1} + ... + \theta_k c_{t-k} = L(\theta, c_t) \tag{64}
\]

and thus

\[
\ln (C_t - H_t) = \varsigma_2 + \psi_2 c_t - (\psi_2 - 1) L(\theta, c_t) \tag{65}
\]
Plugging this results into equation (62) gives

\[ \ln \sum_{i=0}^{\infty} \delta^i \left( C_{t+i} - H_{t+i} \right)^{-\gamma} \theta_i H_{t+i} = \]

\[ = \kappa' - \sum_{i=0}^{k} \left( \prod_{j=0}^{i-1} \rho_j \right) \left( (1 - \rho_i) \left( \gamma \psi_2 c_{t+i} - \left( 1 + \gamma (\psi_2 - 1) \right) \right) \right) \]

This equation tells us that \( \ln \sum_{i=0}^{\infty} \delta^i \left( C_{t+i} - H_{t+i} \right)^{-\gamma} \theta_i H_{t+i} \) is a linear function of \( c_{t+i} \) with for all \( i \in [-k,k] \). This can be simply rewritten as

\[ \ln \sum_{i=0}^{k} \delta^i \left( C_{t+i} - H_{t+i} \right)^{-\gamma} \theta_i H_{t+i} = \kappa' + \sum_{i=-k}^{k} b_i c_{t+i} \] \hspace{1cm} (67)

Finally combining this equation with equation (59) also gives a linear function:

\[ \ln U' (C_t) = \kappa + \sum_{i=-k}^{k} a_i c_{t+i} \] \hspace{1cm} (68)

which implies that

\[ m_{t+1} = \sum_{i=-k+1}^{k+1} a_i \Delta c_{t+i} \] \hspace{1cm} (69)

C Linearization of the loss aversion marginal utility

Rewriting equation (35) gives

\[ w' (Z) = \begin{cases} 
-\theta \ln \left( V (C_t) - V (X_t) \right) & \text{if } C_t \geq X_t \\
\ln \lambda - \theta \ln \left( V (X_t) - V (C_t) \right) & \text{if } C_t < X_t 
\end{cases} \] \hspace{1cm} (70)
Since both $V(C_t) - V(X_t)$ in the upper part of this equation and $V(X_t) - V(C_t)$ in the under part of it are by definition positive, we can approximate them with a Taylor approximation (cf. Appendix B). We get that

$$\ln (V(C_t) - V(X_t)) = \zeta_3 + \psi_3 \ln V(C_t) - (\psi_3 - 1) \ln V(X_t) \text{ if } C_t \geq X_t \quad (71)$$

$$\ln (V(X_t) - V(C_t)) = \zeta_4 + \psi_4 \ln V(X_t) - (\psi_4 - 1) \ln V(C_t) \text{ if } C_t < X_t \quad (72)$$

with $\psi_3 = 1/\left(1 - e^{\ln V(C_t) - \ln V(X_t)}\right)$ and $\psi_4 = 1/\left(1 - e^{\ln V(X_t) - \ln V(C_t)}\right)$ where $\ln V(C_t) - \ln V(X_t)$ is the average over all $C_t \geq X_t$ and $\ln V(X_t) - \ln V(C_t)$ is the average over all $C_t < X_t$.

Using the fact that $V()$ is a power utility function, we get that

$$\ln V(C_t) = (1 - \sigma) c_t - \ln (1 - \sigma) \quad (73)$$

$$\ln V(X_t) = (1 - \sigma) x_t - \ln (1 - \sigma) \quad (74)$$

and

$$\ln (V(C_t) - V(X_t)) = \zeta'_3 + (1 - \sigma) (\psi_3 c_t - (\psi_3 - 1) x_t) \quad (75)$$

$$\ln (V(X_t) - V(C_t)) = \zeta'_4 + (1 - \sigma) (\psi_4 x_t - (\psi_4 - 1) c_t) \quad (76)$$

where $\psi_3 = 1/\left(1 - e^{(1-\sigma)(c-x)}\right)$ and $\psi_4 = 1/\left(1 - e^{(1-\sigma)(x-c)}\right)$. If the expected gain is greater than the expected loss, then $\psi_3 > \psi_4$.

Plugging these last equations into equation (70) gives

$$u'(C_t) = \begin{cases} 
  k^+ + q^+ c_t + p^+ x_t & \text{if } c_t \geq x_t \\
  k^- + q^- c_t + p^- x_t & \text{if } c_t < x_t 
\end{cases} \quad (77)$$
where $k^+ = k - \theta q_3^4$, $q^+ = -(\sigma + \theta q_1 (1 - \sigma) \psi_3)$, $p^+ = \theta (1 - \sigma) q_3 (\psi_3 - 1)$, $k^- = k - \theta q_4^4 + q \ln \lambda$, $q^- = -(\sigma + \theta q_1 (1 - \sigma) (\psi_4 - 1))$ and $p^- = -\theta q_1 (1 - \sigma) \psi_4$. Note that if the expected gain is greater than the expected loss (i.e. $\psi_3 > \psi_4$), then we have that $q^+ < q^-$ if $\sigma < 1$ and $q^+ > q^-$ if $\sigma > 1$.

## D Covariance with loss aversion function

Let us define $X$ and $Y$ as two joint normally distributed random variables and $Z$ as a binary variable such as $Z = 1$ if $Y < \bar{Y}$ and 0 otherwise. The variable $B$ is a binary variable that takes value $B^-$ if $Z = 1$ and $B^+$ otherwise. Let us define $\Phi$ as the probability that $Z = 1$ and $\phi$ as the probability that $Y = \bar{Y}$. Let us start by the identity

$$
\text{Cov} (X, BY) = E (\text{Cov} (X, BY | Z)) + \text{Cov} (E (X | Z), E (BY | Z)) \quad (78)
$$

The first part of the right had side of equation (78) is equal to

$$
E (\text{Cov} (X, BY | Z)) = (\Phi B^- + (1 - \Phi) B^+) \text{Cov} (X, Y) \quad (79)
$$

To compute the second part, first note that

$$
E (BY | Z) = E (B | Z) E (Y | Z) + \text{Cov} (B, Y | Z) \quad (80)
$$

Since $B$ is a constant when $Z$ is known, then $\text{Cov} (B, Y | Z) = 0$ and

$$
\text{Cov} (E (X | Z), E (BY | Z)) = \text{Cov} (E (X | Z), E (B | Z) E (Y | Z)) \quad (81)
$$
Then define the two residuals $e^x = X - E(X)$ and $e^y = Y - E(Y)$, which implies that

$$
E(X | Z) = E(X) + E(e^x | Z) \quad (82)
$$
$$
E(Y | Z) = E(Y) + E(e^y | Z) \quad (83)
$$

Plugging these definition in the previous equations yields

$$
Cov(E(X | Z), E(BY | Z)) = Cov(E(X) + E(e^x | Z), E(B | Z)(E(Y) + E(e^y | Z))) \quad (84)
$$

Since $E(X)$ and $E(Y)$ are constant, this equation becomes

$$
Cov(E(X | Z), E(BY | Z)) = E(Y) Cov(E(e^x | Z), E(B | Z)) +
$$
$$
+ Cov(E(e^x | Z), E(B | Z)E(e^y | Z)) \quad (85)
$$

Furthermore, using the fact that $Cov(a,b) = E(ab) - E(a)E(b)$ and that, by the law of iterated expectation, $E(E(e^x | Z)) = E(e^x) = 0$, we can simplify the previous equation to

$$
Cov(E(X | Z), E(BY | Z)) = E(Y) E(E(e^x | Z)E(B | Z)) +
$$
$$
+ E(E(e^x | Z)E(B | Z)E(e^y | Z)) \quad (86)
$$

Developing the expectation term, we get

$$
Cov(E(X | Z), E(BY | Z)) = \Phi B^* E(e^x | 1)(E(Y) + E(e^y | 1)) +
$$
$$
+ (1 - \Phi) B^* E(e^x | 0)(E(Y) + E(e^y | 0)) \quad (87)
$$

For the next step, let us assume that $X$ and $Y$ jointly normally distributed. This
implies that
\[ E (e^x | e^y) = \frac{\text{Cov}(e^x, e^y)}{V(e^y)} e^y \]  
(89)

Taking the conditional expectation on \( Z \) yields
\[ E (e^x | Z) = \frac{\text{Cov}(e^x, e^y)}{V(e^y)} E (e^y | Z) \]  
(90)

The last term of this equation is the expectation of a truncated normal distribution and is equal to (see Greene 2000, p. 899)
\[ E (e^y | Z) = \begin{cases}  
-V^{1/2}(e^y) \frac{\phi}{\Phi} & \text{if } Z = 1 \\
V^{1/2}(e^y) \frac{\phi}{1-\Phi} & \text{if } Z = 0 
\end{cases} \]  
(91)

which implies by plugging equations (90), (91) and (62) in equation (87)
\[ \text{Cov}(E(X|Z), E(BY|Z)) = \phi \left( \frac{B^+ - B^-}{V^{1/2}(e^y)} E(Y) + \frac{B^+}{1 - \Phi} + \frac{B^-}{\Phi} \right) \text{Cov}(e^x, e^y) \]  
(92)

Finally, combining equations (79) and (92), we get the final expression for the covariance in equation (78)
\[ \text{Cov}(X, BY) = \left( \phi \frac{B^+ - B^-}{V^{1/2}(e^y)} E(Y) + \frac{\Phi^2 + \phi^2}{\Phi} B^- + \frac{(1 - \Phi)^2 + \phi^2}{1 - \Phi} B^+ \right) \text{Cov}(e^x, e^y) \]  
(93)

References


