Copenhagen Business School

M.Sc. Finance & Strategic Management

August 2011

The equity premium puzzle and myopic loss aversion

An empirical study of a non-standard preference structure in Sweden

By Signe Læssøe and Martin Diernisse

Advisor: Lisbeth Funding la Cour

Department of Economics

120 standard pages / 272,964 characters

"vi er mennesker –
hvorfor bebrejde hinanden,
at vi ikke er andet"

*"we are human -
why blame each other
that this is all we are"*

- Gustav Munch-Petersen (Danish poet, 1912-1938)

Executive summary

In 1985, Mehra & Prescott introduced the so-called equity premium puzzle, which captured that the historic equity premium in the US for the period 1889-1978, could by no means be explained with a traditional asset pricing model, based on expected utility theory. The authors concluded that the discrepancy had to be explained, by investors being much more risk averse than had previously been assumed, or by concluding that the traditional model did not capture actual investor behavior.

This thesis investigates an alternative to the standard investor preferences assumed in the model tested by Mehra & Prescott (1985), namely myopic loss aversion as proposed by Benartzi & Thaler (1995). The model is based on prospect theory of Tversky & Kahneman (1979; 1992), as well as established concepts from the field of behavioral finance. It is descriptive in nature, as it is based on experimental studies of human decision making under risk, rather than relying on assumptions of a purely rational individual.

First we expand the study of the equity premium puzzle to the small open economy of Sweden, using the original method of Mehra & Prescott (1985), and the statistical reformulation of the puzzle by Kocherlakota (1996). We start by investigating the period 1919-2010, and reach the surprising conclusion that the equity premium puzzle is not present in Sweden for that period. Based on the previous findings of the puzzle, we re-evaluate our period of analysis, and find that the years before 1925 are highly atypical. We thus eliminate the years from our dataset, and for the revised period from 1925-2010 we find an equity premium puzzle similar to the one found in the US. The result is found to be robust against variations in data input, as well as further changes of the sample period.

Next, we investigate whether myopic loss aversion is able to serve as a potential solution to the equity premium puzzle in Sweden. We find that the puzzle can be explained by assuming that investors are loss averse, meaning that a loss hurts them twice as much as a gain pleases them, and that they are myopic in their investment strategy, meaning that they on average evaluate their portfolios too often.

Our conclusion is supported by observations of optimal asset allocation, and we find that the implied equity premium would fall, if investors were to evaluate the return on their equity investments, in a more aggregated manner. Finally, we test the sensitivity of our method, against different changes in input variables, and over different sample periods, and again we find that our conclusion is robust.

Table of Contents

Part I: The equity premium puzzle

Part II: Myopic loss aversion

Part III: Empirical analysis and results

1. Introduction

Conventional wisdom dictates that economic agents require compensation for carrying risk. Hence, it is of little surprise that investors historically have demanded higher compensation for holding stocks compared to relatively risk-free fixed-income assets. The nature of this return premium is however somewhat more elusive.

Traditional financial theory rests on the basic assumption that investors are rational and value optimizing, and that they will always seek to maximize their expected utility. The value they place on an asset must therefore be tied to the utility it provides them with. If investors demand a premium for investing in equity compared to a risk-free asset, it must be because stocks provide them with less utility.

In 1985 Mehra & Prescott set out to test the empirical validity of this utility-based explanation of the equity premium. They employed a widely recognized asset pricing model based on standard expected utility theory, and evaluated its explanatory power by pinning its results against the actual observed premium within the US market.

The authors found that the traditional model failed miserably. Despite an observed equity premium of 6.18%, their model could only justify a risk premium of 1.4%, even when investors were assumed to be highly risk averse. The rational and utility optimizing investors would thus have to be *extremely* risk averse, for the model to justify the observed equity premium as compensation to the investors for carrying risk.

Mehra & Prescott (1985) dubbed this conundrum the equity premium puzzle, and the authors concluded that there could only be two explanations for their finding: either investors were much more risk averse than had previously been assumed in financial theory, or the traditional asset pricing model was flawed in its assumptions.

Since the first formulation of the puzzle, there have been countless attempts to solve it. Some have attempted to vindicate the notion of the rational investor, by pointing to inconsistent assumptions about market perfection in the model applied by Mehra & Prescott, or to biases in their data. Others have sought to explain the observed investor behavior by suggesting that the rational and utility optimizing investor actually has an alternative preference structure.

One attempt to replace the traditional preference structure was provided in 1995 by Benartzi & Thaler, who introduced a possible solution to the puzzle called myopic loss aversion. Their proposition replaced the normative investor preferences with a descriptive and empirically based preference structure, observed by Tversky & Kahneman (1979; 1992) in experiments with decision making under risk. Benartzi & Thaler (1995) applied these observed preferences in an attempt to model the behavior of US investors, and were able to account for the observed equity premium over the period 1926-1990.

Even though many attempts have been made to solve the equity premium puzzle, the vast majority of them have tested their hypotheses on US data, including the study of Benartzi & Thaler (1995). In this paper we seek to rectify this shortcoming of the literature within the field, by investigating the equity premium puzzle in a non-US country, namely Sweden for the period 1919 to 2010.

The puzzle has been confirmed in countries outside the US, however studies in this field are quite scarce. Consequently, we first seek to determine whether the equity premium puzzle is present in Sweden during the period in question. Secondly, we test the myopic loss aversion hypothesis of Benartzi & Thaler (1995) on our Swedish dataset, to investigate whether the alternative preference structure can explain the observed behavior of Swedish investors.

1.1 Problem statement

It is a well established empirical fact that investors all over the world require compensation for investing in equity relative to a risk-free alternative. The ability of standard expected utility theory to explain the equity premium outside the US, has however received far less attention in academia. Given the unique world-dominating status of the US economy, it would be interesting to expand the empirical investigation of the equity premium puzzle and myopic loss aversion to a country with different characteristics.

We have chosen to focus our efforts on the small open economy of Sweden, which also differentiates itself from the US by having a much more egalitarian societal structure. Furthermore, data of a remarkable quality is available for Sweden for a period of more than 90 years.

Our first research question aims at contributing to the field, by expanding the analysis of the equity premium puzzle to a different economic setting than the US:

i. *Can the utility-based model of Mehra & Prescott (1985) explain the observed equity premium in Sweden, i.e. is there an equity premium puzzle?*

Based on the previous merits of this model, we do find it somewhat likely that the utility-based model will fail in this endeavor. Our second research question is therefore concerned with testing the potential solution of Benartzi & Thaler (1995) on a non-US dataset:

ii. Can an alternative investor preference structure, in the form of myopic loss aversion, provide an explanation to a Swedish equity premium puzzle?

The aim of this paper is to contribute to academia by expanding the empirical investigations of the puzzle, as well as testing a potential solution in a country other than the US. Besides the obvious academic interest in testing an existing hypothesis in a new context, the nature of the equity premium puzzle and myopic loss aversion makes this effort particularly interesting. As they are both concerned with modeling the behavior and preferences of economic agents, researchers in a multitude of academic fields could potentially benefit from further insights into the modeling of human behavior.

1.2 Methodology

1.2.1 Structure

To answer our research questions in a manner most logical to the reader, we have structured our analysis in three main sections:

Part I will start by providing the formal deduction of the equity premium puzzle and present the findings from the original article by Mehra & Prescott (1985). The contribution of Mehra & Prescott will then be supplemented by a reformulation of their method based on a statistical approach, and this is followed by a review of the limited findings of the equity premium puzzle internationally. Following the establishment of the puzzle, we will review the key literature in the field of potential solutions to the equity premium puzzle. Lastly we will present myopic loss aversion as a possible solution.

Part II will give a thorough introduction to myopic loss aversion. We start by examining the axioms behind the standard preference structure assumed under standard expected utility theory, and subsequently we provide an insight into the axioms' empirical credentials. We will then move on to describe the alternative prospect theory preferences assumed under myopic loss aversion, and provide a technical formulation of the preference structure to be used in the myopic loss aversion model. Following the establishment of the alternative investor preferences, we will move on to the actual intuition behind myopic loss aversion, which is drawn from the field of behavioral finance. Finally, we

combine the preference structure of prospect theory, and the concepts from behavioral finance, to define the actual formulation of myopic loss aversion.

Part III contains our analyses of the equity premium puzzle in Sweden and of myopic loss aversion as a possible solution to. It starts with a detailed description of the data series used for our study, including the calculations and modifications that have been necessary. Following this is a technical specification of the methodology used in the two parts of the analysis, and a presentation of our results. The chapter ends with tests of the robustness of our findings and finally provides a conclusion to the paper.

1.2.2 Theoretical foundation

The theoretical foundation of this paper's analysis of the equity premium puzzle in Sweden is primarily the original formulation of the puzzle by Mehra & Prescott (1985; 2003). This is supplemented by a statistical reformulation of the puzzle by Kocherlakota (1996), as well as the original formulation of the so-called risk-free rate puzzle by Weil (1989).

For the test of myopic loss aversion we have relied on the original framework of Benartzi & Thaler (1995), which again is based on the preference structure of cumulative prospect theory developed by Tversky & Kahneman (1979; 1992).

1.2.3 Delimitations

The purpose of this paper is to provide empirical testing of two existing academic topics, namely the presence of the equity premium puzzle in a non-US country, and the ability of myopic loss aversion to provide a possible solution to the puzzle. Hence there are a number of issues that we will leave for other research papers to investigate. Given that our contribution is primarily an empirical investigation, we will not provide a complete review of all the extensive literature written on the issue of the equity premium puzzle, nor on the equity premium itself. We will of course provide an overview of the main literature in the field and discuss the validity of the various attempts that have been made to solve the puzzle, but for a detailed review of this line of literature we refer to other publications on the matter (see chapter 4).

For our empirical examination of the equity premium puzzle, we will confine ourselves to establish whether it is present in Sweden for the sample period, and hence we will not attempt to explain variations in the equity premium puzzle or observed investor risk aversion over time. We do provide a brief overview of our results for eight different periods, but this is only to show that the finding of the equity premium puzzle in itself is robust over different sample periods.

Regarding our test of the explanatory power of myopic loss aversion, we make the assumption that the preference structure of Swedish investors can be described by prospect theory, despite it being based on experiments with US investors. We will not attempt to give a detailed analysis of the possible cognitive differences between the American investors and their Nordic counterparts. We do however believe that the success (or failure) of myopic loss aversion on Swedish data will shed new light on the general applicability of prospect theory in describing human decision processes.

Part I: The Equity Premium Puzzle

In 1985, Rajnish Mehra and Edward C. Prescott published their famous article "The equity premium – a puzzle" on the apparent failure of traditional finance theory. They showed that a generally accepted asset pricing model based on utility theory was unable to replicate the observed equity premium in the US. The model did rightly predict the presence of an equity premium, but the implied level of risk aversion needed to explain the observed premium was considerably larger than empirical estimates.

Investors simply seemed to demand an unreasonably high return for carrying the extra risk from stock investments relative to risk-free assets – at least with the definition of risk provided by utility theory. The authors concluded that the observed equity premium was a puzzle: investors were either much more risk averse than previously estimated, or the traditional utility theory did not describe investor behavior as well as it had been assumed until then. The controversial claims of Mehra & Prescott (1985) sparked a heated academic debate, which after more than two and a half decade is still far from being settled.

Part one of this paper will review the original model and findings of Mehra & Prescott (1985), and provide an overview of the relevant literature regarding the equity premium puzzle. The review is structured in three parts: Chapter 2 derives the model of Mehra & Prescott and presents their formulation of the puzzle based on their updated article from 2003. Chapter 3 provides the statistical reformulation of the puzzle by Kocherlakota (1996), introduces the risk-free rate puzzle of Weil (1989), and presents some international evidence of its presence. Chapter 4 reviews the most prominent attempts to solve the puzzle by relaxing the three basic assumptions behind the Mehra-Prescott model, and by arguing for the inherent presence of an *ex post* bias in all studies of equity premia. Following part one, we present and review myopic loss aversion as possible solution to the puzzle in part two.

2. The model of Mehra & Prescott

The equity premium puzzle is essentially concerned with the premium investors get for holding equities relative to holding a risk-free asset. As equities are a more risky investment, a rational and risk-averse investor would be expected to demand a premium for holding a stock relative to a risk-free asset such as a T-bill. Mehra & Prescott (1985) investigated the expected size of this premium using a traditional consumption-based capital asset pricing model (C-CAPM) adapted from Lucas (1978). They found that

the model was incapable of explaining the observed equity premium without assuming extreme levels of risk aversion among investors.

The Lucas model is based on classic utility theory, which models investor behavior by assuming that investors are rational and that they always seek to maximize their expected utility (Lucas, 1978). The implication for the pricing of financial assets is that the equilibrium asset price is where the expected marginal utility from an investment at time *t* equals the marginal utility forfeited at time *0*. In other words, the investor must be compensated with expected future utility to give up current utility.

It is assumed that marginal utility varies over time, i.e. it is state dependent. When total consumption is low, marginal utility from consumption is high, and vice versa. Put simply, an investor will get higher utility from receiving $100 when his bank account is empty relative to when it is full. The implication is that an asset, which pays off when the investor is already at a high level of consumption, will provide him with *less* marginal utility than an asset that pays off when his consumption is low. The investor will therefore be willing to forfeit less utility to obtain the former asset, and it will thus trade at a lower price.

As attractive assets will trade at a higher price and thus yield a lower return, it means that counter-cyclical assets offer a lower expected return than their pro-cyclical equivalents. Equities typically have the highest return when consumption is already high (boom periods), while the relatively risk-free fixed income instruments pay out regardless of the overall economic conditions[1]. Hence, marginal utility from equities will be lower than marginal utility from the risk-free asset, so they will trade at a lower price per unit of payoff. A low price yields a high return, and this is the source of the equity premium under the Lucas model.

The original model by Lucas (1978) is a so-called pure-exchange consumption-based capital asset pricing model (C-CAPM), and is based on three fundamental assumptions:

  i.   Trading is instant and costless
 ii.   Markets are perfect and efficient so asset prices in equilibrium reflect all available information
iii.   Investors are rational and identical at the aggregate level.

Mehra & Prescott (1985) made only one alteration to the model in order to account for the increase in consumption over time. Where Lucas assumed that the economy's overall consumption *level* follow a stochastic process, Mehra & Prescott instead assumed that the *growth rate* of consumption follows a

---

[1] Assuming the issuing country does not default

stochastic process. This allows the overall consumption in the model to increase over time. For the remainder of the paper, we will refer to this model as the Mehra-Prescott model.

The basic implication of the model of Mehra & Prescott is that the risk of an asset - and thereby its price - can be determined by its covariance with the investor's consumption, the key assumption being that the investor treats current and future consumption as two different goods. Again, since markets are understood to be perfect and complete, this assumption means that in equilibrium, the asset is priced so that the loss in marginal *current* utility from buying the asset, equals the present value of the expected marginal *future* utility from holding the asset (Lucas R. E., 1978). The resulting asset pricing relationship states that the price of an asset equals the risk free rate, plus an expression for its covariance with consumption growth[2].

2.1 Deduction of the model

As stated above, the Mehra-Prescott model assumes a frictionless economy with one representative investor. This investor seeks to optimize the sum of expected utility from all future consumption. Formally stated his investment decision is determined by maximizing the following expression:

$$(2.1) \qquad E_0\left\{\sum_{t=0}^{\infty}[\beta^t U(c_t)]\right\}, \qquad 0 < \beta < 1$$

where $U(c_t)$ is the utility derived from consumption c at time t. β is a subjective discount factor between 0 and 1, which captures the impatience of the investor, meaning that he discounts future utility just because he has to wait for it. $\beta = 0$ means that future consumption has no value at all and $\beta = 1$ means that he is indifferent between one unit of real current consumption and one unit of real future consumption. E denotes that the investor seeks to optimize *expected* utility at time *t+1*, when the investment decision is taken at time *0*. It is assumed that the investor has infinite life, i.e. that he will always seek to optimize future utility regardless of his life expectancy. This is in fact not an unreasonable assumption, if his affection towards his children means that he gets the same expected utility from their consumption as he would have gotten from his own (Mehra & Prescott, 2003).

The utility function of the investor is restricted to display constant relative risk aversion, meaning that the investor has a single attitude towards risk regardless of any external factors. This is also known as an

---

[2] Equivalent to the traditional CAPM model, which defines asset price as the risk-free rate plus an expression for its covariance with market return

isoelastic utility function. The relationship is derived in appendix A, and the utility function is formally stated as:

$$(2.2) \qquad\qquad U(c, \alpha) = \frac{c^{1-\alpha} - 1}{1 - \alpha}, \qquad 0 < \alpha < \infty$$

Here α is the *coefficient of relative risk aversion* (CRRA) of the investor, as defined by the Arrow-Pratt definition of relative risk aversion. The use of a constant to capture the risk attitude of the investor, implies a reciprocal relationship between the coefficient of relative risk aversion and the *elasticity of intertemporal substitution* (EIS = α^{-1} as seen in appendix A). The EIS captures how an investor chooses to allocate his consumption between different time periods, where a low value indicates a strong preference for distributing consumption equally over time. A highly risk-averse investor (high alpha) will seek to smooth his consumption over different states of the economy to minimize variations in utility. Because of the reciprocal relationship with EIS however, he will consequently also seek to smooth consumption over different time periods[3].

When the investor wishes to distribute consumption equally over time, it implies that he dislikes consumption growth. This seems counterintuitive and Mehra & Prescott (2003) also point out that there is no a priori reason why this should be the case. This feature of the Mehra-Prescott model has later been altered by Epstein & Zin (1989; 1991), who separated CRRA and EIS in their generalized expected utility theory. Their contribution to solving the equity premium puzzle, as well as other attempts to separate the CRRA and the EIS are reviewed in section 4.2.3.

The behavior of the investor is guided by the goal of maximizing his total expected utility, and his specific investment decision is made by weighing the forfeited current utility against expected future utility. To invest in an asset at time t the investor has to pay a price p. To do so the investor has to give up p_t units of consumption and thereby give up utility equal to $p_t U'(c_t)$, with U' denoting marginal utility.

When the investor sells the asset at time *t+1*, he receives $p_{t+1} + y_{t+1}$ units of consumption, where p_{t+1} is the price, and y_{t+1} is the dividend of the asset at time *t+1*. The expected future utility from the investment is therefore equal to $\beta E[(p_{t+1} + y_{t+1})U'(c_{t+1})]$. This is interpreted as total expected pay-off on the asset times the expected marginal utility from consumption at time *t+1*, discounted by the impatience factor β.

---

[3] A high CRRA implies a low EIS, which means a strong preference for a smooth consumption pattern over time.

In equilibrium the utility forfeited at time t must equal the expected future utility at time $t+1$ in order for the investment to be made. This leads to the fundamental pricing relation of the Mehra-Prescott model:

$$(2.3) \qquad\qquad p_t U'(c_t) = \beta E[(p_{t+1} + y_{t+1})U'(c_{t+1})]$$

The expected pay-off of the asset at time t can be rewritten in terms of a return: $R_{e,t+1} = \frac{p_{t+1} + y_{t+1}}{p_t}$. Appendix B shows that the two expressions can be rearranged to define the expected return on an asset under the Mehra-Prescott model:

$$(2.4) \qquad\qquad R_{e,t+1} = R_{f,t+1} + Cov_t \left[\frac{-U'(c_{t+1}), R_{e,t+1}}{E_t[U'(c_{t+1})]} \right]$$

where $R_{f,t}$ is the risk-free rate at time t. This states that the expected return on an asset equals the risk-free rate plus a risk premium, which depends on the covariance between the returns of the asset and the marginal utility of consumption at the time. As mentioned earlier this is the case because it is assumed that marginal utility varies with the level of consumption. When an investor is prospering, the marginal utility from additional consumption will be lower than if he was in a financial rough spot, so assets that pay off when times are good provide less expected utility, and thus command a higher expected return. This is equivalent to the high expected return on high-beta stocks under the standard CAPM, which pay off relatively more when the market return is high.

We now introduce consumption growth into the model. In the Mehra-Prescott model this is defined as $x_{t+1} = \frac{c_{t+1}}{c_t}$. We show in appendix C that introducing this term to the fundamental pricing relationship of equation (2.3), yields the following expression for the expected return on equity:

$$(2.5) \qquad\qquad E(R_{e,t+1}) = \frac{E_t(x_{t+1})}{\beta E_t(x_{t+1}^{1-\alpha})}$$

As well as the equivalent expression for the risk-free rate:

$$(2.6) \qquad\qquad R_{f,t+1} = \frac{1}{\beta E_t(x_{t+1}^{-\alpha})}$$

As mentioned earlier, the Mehra-Prescott model assumes that the growth rate of consumption is stochastic, as opposed to Lucas (1978) who assumed that the *level* followed a stochastic process. The Mehra-Prescott model assumes that growth rates are described by a lognormal distribution, and

appendix C shows how this assumption leads to the following expression for the expected return on equity at time *t+1*:

$$(2.7) \qquad E_t(R_{e,t+1}) = \frac{e^{\mu_x + 1/2\sigma_x^2}}{\beta e^{(1-\alpha)\mu_x + 1/2(1-\alpha)^2\sigma_x^2}}$$

Which can be rearranged to:

$$(2.8) \qquad \ln E_t(R_{e,t+1}) = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2 + \alpha\sigma_x^2$$

Where $\mu_x = E(\ln x)$ represents the mean consumption growth, and $\sigma_x^2 = Var(\ln x)$ is the variance of consumption growth. In the same manner it is shown that the return of the risk-free asset is defined as:

$$(2.9) \qquad R_{f,t+1} = \frac{1}{\beta e^{-\alpha\mu_x + 1/2\alpha^2\sigma_x^2}}$$

Which can be rearranged to:

$$(2.10) \qquad \ln R_{f,t+1} = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2$$

From this it follows that the equity premium must equal:

$$(2.11) \qquad \ln E(R_{e,t+1}) - \ln(R_{f,t+1}) = \alpha\sigma_x^2$$

Hence the equity risk premium in the Mehra-Prescott model is defined as the representative investor's risk aversion times the variance of consumption growth. This is the essential prediction of the utility based C-CAPM, and this is the relationship that Mehra & Prescott (1985) tested empirically. We will now turn to a more detailed description of their findings.

2.2 Empirical findings of Mehra & Prescott

In their original article, Mehra & Prescott (1985) tested their model on US data for the period 1889 to 1978. Their data series were the real annual return on the S&P 500 (total return index), growth in real per capita consumption, and the real return on 90-day government T-bills (representing the risk-free asset).

Below are the sample statistics for Mehra & Prescott's original data set:

Table 2.1: Sample statistics

Mean real risk-free rate	R_f	0.80%
Mean real equity return	$E(R_e)$	6.98%
Mean real growth rate of consumption	$E(x)$	1.80%
St.dev. real growth rate of consumption	$\sigma(x)$	3.60%
Mean equity premium	$E(R_e) - R_f$	6.18%

In order to apply their model, Mehra & Prescott had to set values of alpha and beta. These are subjectively determined by the investor, so it is not possible to observe their exact values. General estimations of alpha at the time ranged between one and two (Mehra & Prescott, 1985) and a later consensus estimate has been close to three (Mehra & Prescott, 2003). Their hypothesis was that the model would require a much higher level of risk aversion to explain the observed equity premium, so to stack the deck in favor of the model they chose to apply $\alpha = 10$.

Though such a high level of risk aversion would help explain the equity premium, it would also inadvertently drive up expected returns. To bring expected returns to as realistic a level as possible, Mehra & Prescott (1985) chose to set the impatience factor, beta at 0.99. This means that the investor is almost indifferent between current and future consumption, so the required risk-free rate is kept at a minimum. We will return to this issue in the review of the risk-free rate puzzle in section 3.2.

The assumed levels of risk aversion and impatience (α and β) were inserted in equation (2.10) along with the mean and variance of consumption growth, to find the implied risk-free rate:

$$\ln R_f = -\ln(0.99) + 10 * 0.018 - \frac{1}{2} * 10^2 * 0.036^2 = 0.120$$

$$\Rightarrow \boldsymbol{R_f = 12.7\%}$$

Their utility-based model thus predicted a risk-free rate of 12.7% compared to the observed level of 0.8%. Unreasonable as it may seem, it was not unexpected given the high level of risk aversion that was assumed needed in order to explain the equity premium. The expected return on equity was found by rearranging equation (2.11).

$$(\boldsymbol{2.11}) \qquad \ln E(R_e) - \ln(R_f) = \alpha\sigma_x^2 \Longleftrightarrow$$

$$\ln E(R_e) = \ln(R_f) + \alpha\sigma_x^2$$

The implied risk-free rate was then used to find the expected return on equity:

$$\ln E(R_e) = 0.120 + 10 * 0.036^2 = 0.132$$

$$\Rightarrow \boldsymbol{E(R_{re}) = 14.1\%}$$

The model of Mehra & Prescott was thus able to explain an equity premium of 1.4%[4], which was far lower than the observed level of 6.18%. Despite assuming an implausibly high level of investor risk aversion, their model did not succeed in explaining the observed equity premium. This is what Mehra & Prescott (1985) dubbed the equity premium puzzle. Their findings are summarized below:

Table 2.2: Results of Mehra & Prescott (1985)

	Risk-free rate	Equity return	Equity premium	**Puzzle**
Observed	0.80%	6.98%	6.18%	**4.78%**
Mehra-Prescott model	12.7%	14.1%	1.4%	

Source: Mehra & Prescott (1985)

2.3 Chapter summary

The consumption-based capital asset pricing model (C-CAPM) of Lucas (1978) rests on a basic assumption of rational and value optimizing investors. Mehra & Prescott (1985) decided to test a slightly modified version of the model (referred to in this paper as the Mehra-Prescott model) on US data for the period 1889-1978, to see whether the traditional model was able to explain the observed equity premium as compensation to the investors for carrying risk. They noted that empirical estimates of the coefficient of relative risk aversion (CRRA) for investors were in the range of 1-3, but decided to set it equal to ten to give as much credit to the model as possible.

They found that even with this implausibly high level of risk aversion, the return on equity had been much too high to be justified as a rational compensation for carrying risk. Their model could only account for a premium of 1.4%, even though the observed equity premium had been 6.18% during the period.

Mehra & Prescott (1985) concluded that the difference of 4.78% must either be explained by investors being much more risk averse than had previously been assumed, or by accepting that a model based on rational agents acting to optimize their expected utility, is not appropriate for describing the behavior of real-life investors.

---

[4] 14.1% - 12.5% = 1.4%

3. Further empirical validation of the puzzle

The existence of the puzzle has been confirmed in different ways since Mehra & Prescott published their original article. The following section will review two alternative methods for determining the puzzle, namely the statistical approach of Kocherlakota (1996) and the formulation of the risk-free rate puzzle by Weil (1989). Afterwards we will seek to further certify its presence, by reviewing findings of the puzzle outside of the US.

3.1 A statistical approach

The model of Mehra & Prescott (1985) has the somewhat problematic feature that it relies on subjective estimations of the investor's risk aversion (α) and impatience (β). As mentioned in the previous chapter, this was solved in the original study by assuming extreme levels of both factors to give as much credit to the model as possible. However in 1996, Kocherlakota introduced a new method for investigating the explanatory power of the Mehra-Prescott model, without having to rely on these subjective estimations.

Instead of assuming certain levels of α and β, he asserted that the Mehra-Prescott model should provide the same expected return on equity as on risk-free assets, once they are both adjusted for the factors that are assumed to have an effect on them, namely: consumption growth, impatience, and risk aversion. The relationship is deduced in appendix D, and is stated below:

$$(\textbf{3.1}) \qquad \beta E_t \left[\left(\frac{c_{t+1}}{c_t} \right)^{-\alpha} R_{e,t+1} \right] = \beta E_t \left[\left(\frac{c_{t+1}}{c_t} \right)^{-\alpha} \right] R_{f,t+1}$$

Equation (3.1) states that the expected equity return and the risk-free rate should be the same when both are corrected for the effects of risk aversion (α), impatience (β), and consumption growth $\left(\frac{c_{t+1}}{c_t} \right)$.

Consequently, after these adjustments, the expected return on equity minus the risk-free rate (i.e. the equity premium) should be zero. Appendix D shows that equation (3.1) therefore can be reduced to:

$$(\textbf{3.2}) \qquad E_t \left[\left(\frac{c_{t+1}}{c_t} \right)^{-\alpha} (R_{e,t+1} - R_{f,t+1}) \right] = 0$$

This relationship states that the equity premium $(R_{e,t+1} - R_{f,t+1})$ should be zero, once consumption growth and risk aversion (α) are taken into account. This is exactly the same intuition as we saw in equation (2.11) in the deduction of the Mehra-Prescott model: The model predicts that the equity premium can be explained by risk aversion and consumption growth.

As consumption growth $\left(\frac{c_{t+1}}{c_t}\right)$, equity return ($R_{e,t+1}$), and the risk-free rate ($R_{f,t+1}$) can all be observed, the only variable in equation (3.2) is risk aversion (α). This was the groundbreaking feature of the Kocherlakota method: If Mehra & Prescott's utility-based model was able to explain the observed equity premium, then equation (3.2) should hold for reasonable levels of risk aversion (α). If equation (3.2) did not hold for alpha values between one and three, then the puzzle was confirmed without making a rigid assumption about a certain value of α and β as done by Mehra & Prescott.

Kocherlakota (1996) restated this expression into a testable H_0 hypothesis, to validate the existence of the equity premium puzzle: If the Mehra-Prescott model is empirically valid, then the left-hand side of equation (3.2) should not be significantly different from zero for reasonable levels of risk aversion. This is expressed in equation (3.3):

$$(\mathbf{3.3}) \qquad \left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}(R_{e,t+1} - R_{f,t+1})\right] = e_{t+1}$$

Kocherlakota calculated e_{t+1} for every year in the Mehra-Prescott dataset, to see whether their utility-based model could explain the equity premium in that year. He did the calculations for values of α ranging from 0.0 to 10.0, and found the mean value $\overline{e_{t+1}}$ over all the years for each value of α. The value of $\overline{e_{t+1}}$ should not be significantly different from zero for $1 < \alpha < 3$, if the Mehra-Prescott model was to succeed and the equity premium puzzle to be dismissed. Based on the results of Mehra & Prescott (1985), Kocherlakota (1996) expected the utility-based model to understate the observed equity premium, (i.e. $\overline{e_{t+1}} > 0$), so he applied a one-tailed t-test to investigate its validity.

A sample of his findings is reported in Table 3.1, which shows that $\overline{e_{t+1}}$ for values of α smaller than 8.5, were in fact significantly different from zero[5]. This meant that based on the observed equity returns, risk-free rate, and consumption growth, the Mehra-Prescott model required a CRRA of 8.5 to explain the equity premium in the original dataset. This clearly supported the existence of the equity premium puzzle: The observed equity premium could only be explained by the traditionally founded model, if investors had been much more risk averse than generally assumed.

---

[5] 90 degrees of freedom, 95% confidence level (one-tailed)

Table 3.1: Results of Kocherlakota (1996)

α	$\overline{e_{t+1}}$	t-stat
0.0	0.0594	3.345
2.5	0.0528	2.890
5.0	0.0433	2.370
7.5	0.0357	1.824
8.0	0.0341	1.715
8.5	**0.0326**	**1.607**
9.0	0.0310	1.501
10.0	0.0279	1.291

Source: Kocherlakota (1996) p. 50.

3.2 The risk-free rate puzzle

Further evidence in support of the puzzle was provided by Weil (1989). He is credited for introducing the risk-free rate puzzle, which can be seen as a theoretical consequence of the equity premium puzzle. The findings of Mehra & Prescott (and Kocherlakota) could not decisively dismiss the validity of the utility based asset pricing model, as the true level of risk aversion is unobservable. They could only conclude that the required level of risk aversion seemed highly implausible.

Weil's insight was that if investors were in fact as risk averse as the Mehra-Prescott model suggested, then the historic risk-free rate seemed puzzling low. Using $\alpha = 10$ (which was still insufficient to explain the observed equity premium) the Mehra-Prescott model predicted a risk-free rate of 12.7% well above the observed level of 0.8%. Hence if the equity premium puzzle was solved by simply increasing assumed investor risk aversion, it would just lead to a new question: If investors were in fact as risk averse as traditional finance theory seems to assume, why did they not demand a higher risk-free return?

The nature of the risk-free rate puzzle can be shown by decomposing equation (2.10):

$$(\mathbf{2.10}) \qquad \ln R_f = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2$$

Recall that β captures the *impatience discount factor* of the investor, and that $0 < \beta < 1$, as the investor is assumed to prefer current consumption over future consumption (see section 2.1). In a world without consumption growth ($\mu_x = 0$), and with perfect information about the future (i.e. no risk, $\alpha = 0$), equation (2.10) becomes: $\ln R_f = -\ln\beta$. So the risk-free rate must be equal to β^{-1}. Investors would in other words only need to be compensated for the inconvenience of postponing their consumption.

The second term, $\alpha\mu_x$ arises from *consumption growth*. If consumption is expected to increase in the future, an investor with isoelastic risk preferences would prefer to shift some of this consumption to the present by borrowing against his future consumption[6]. This would fulfill his desire to smooth his consumption pattern over time and economic states. The higher his risk aversion is and the higher his expectations to the consumption growth rate are, the more he will wish to borrow. If consumption is expected to grow fast (high μ) and the investor has a strong preference for keeping a steady level of consumption (high α), he will have a strong incentive to borrow against his future consumption to shift future consumption to the present. The demand for loans will lead to a higher risk-free interest rate, hence the positive correlation between this second term and the risk-free rate.

The third term, $-\frac{1}{2}\alpha^2\sigma_x^2$, arises from the desire to save, as a precaution against future decreases in consumption. It depends on the investor's risk aversion, along with the risk of future consumption, which is captured by its standard deviation σ. The effect is the opposite than above: If there is a high risk that consumption will decrease in the future (high σ), and the investor is highly averse to such decreases (high α), he will save up to safeguard himself against a possible future disruption in his consumption pattern. A risk averse investor in an uncertain world, would thus have a high precautionary savings rate. The term increases the demand for savings along with the level of investor risk-aversion, and the uncertainty about future consumption. This drives down the risk-free rate, and thus we see a negative relationship between this third term and the risk-free rate.

Equation (2.10) implies that when consumption is growing at the observed level of 1.8%, and the growth rate has a standard deviation of 3.6%, it takes a high risk-free rate to deter a risk averse investor from borrowing against his future consumption. In other words, consumption has been growing sufficiently fast that the precautionary savings effect $(-\frac{1}{2}\alpha^2\sigma_x^2)$ is completely dominated by the desire to shift future consumption to the present $(\alpha\mu_x)$. An investor would have to be *extremely* risk averse to use saving instead of borrowing as measure to smooth consumption[7].

With a moderate level of risk aversion ($\alpha = 3$) the risk-free rate under the Mehra-Prescott model as captured by equation (2.10) yields an interest rate of 6%[8]. This is far above the observed rate of 0.8%.

---

[6] Recall from section 2.1 that a high CRRA implies a low EIS, which means that the investor will desire the same level of consumption in all time periods

[7] Mehra-Prescott (2003) showed that it takes $\alpha = 48$ and $\beta = 0.55$ to replicate the historic interest rate

[8] $\ln R_f = -\ln(0.99) + 3 * 0.018 - \frac{1}{2} * 3^2 * 0.036^2 = 0.058 \Rightarrow R_f = 6.0\%$

The isoelastic preference structure of the Mehra-Prescott model implies that the investor seeks to smooth his consumption over time, and not just over different states of the economy, i.e. he dislikes consumption growth. Consumption has however grown historically, so the investor should want to borrow to transfer some of his future consumption to the present. This would lead to a high demand for loans, thereby driving up the risk-free rate, which is essentially the risk-free rate puzzle of Weil (1989).

3.3 The Equity premium puzzle internationally

The existence of an equity premium has been thoroughly established in the US, as well as in the rest of the world. The most credited work in the field has undoubtedly been done by Dimson, Marsh, & Staunton (2008; 2011), who have meticulously constructed a database of international equity premia.

It should be noted that this does not necessarily document the presence of a puzzle. The Mehra-Prescott model also predicts the presence of an equity premium, and recall that the equity premium puzzle is a strictly quantitative anomaly. Given the size of the international equity premia, and the merits of the model in the US, it does however seem somewhat implausible that international data should redeem the validity of the utility based model. We return to a test of the Mehra-Prescott model on Swedish data, in the third part of this paper.

A brief overview of some estimates of international equity premia is presented below[9]:

---

[9] See also Jorion & Goetzmann (1999) and Hassan & van Biljon (2010)

Table 3.2: International estimates of equity premia

	Dimson, Marsh & Staunton (2011) 1900-2010	Barro (2006) 1954-2004	Kyriacou, Madsen & Mase (2006) 1900-2002
Australia	6.7		6.9
Belgium	2.9		
Canada	4.2	5.0	6.2
Denmark	2.8		
Finland	5.9		
France	6.0	7.2	9.9
Germany	5.9	8.0	5.2
Ireland	3.0		7.6
Italy	5.8	5.1	8.7
Japan	5.9	8.3	
Netherlands	4.2		6.8
New Zealand	4.1		
Norway	3.0		
South Africa	6.2		
Spain	3.2		4.1
Sweden	4.3		
Switzerland	3.4		
UK	4.3	7.9	5.2
US	5.3	7.6	6.9
Mean	**4.6%**	**7.0%**	**6.8%**

Source: Dimson, Marsh & Staunton (2011); Barro (2006); Kyriacou, Madsen & Mase (2006).

The existence of an equity premium outside the US is hard to dispute, though the size of the premia differs substantially across the studies. The primary explanation for this is the choice of time period, and the treatment of anomalous periods such as years of wartime, hyperinflation, and the Great Depression also factors in. We will return to this issue in regards to our own dataset in chapter 9.

Even though the existence of an equity premium is well documented in the literature, few have actually tested the explanatory power of consumption-based asset pricing models in countries other than the US. One of these studies was presented by Barro (2006), who examined the puzzle in seven developed economies[10] for the period 1954-2004. He found that a utility-based model similar to the one used by

---

[10] Canada, France, Germany, Italy, Japan, UK, and the US (G7)

Mehra & Prescott (1985) could only justify a premium of 0.16%, despite an average observed equity premium of 7.0%, when applying $\alpha = 4$.

Similarly, he also found that the model predicted an average risk-free rate of 12.7% for the seven countries, as opposed to the observed 2%. Barro's results have the limitation of only being reported on an aggregate level for all seven countries, but fortunately he provided sufficient data for us to perform the calculations ourselves. We have applied the same values for α and β as Mehra & Prescott (1985), and the results are reported in Table 3.3:

Table 3.3: The equity premium puzzle in the G7 countries, 1954-2004

Country	Observed equity premium	Equity premium Mehra-Prescott model	**Puzzle**
Canada	5.0	0.7	**4.3%**
France	7.2	0.4	**6.8%**
Germany	8.0	0.7	**7.3%**
Italy	5.1	0.6	**4.5%**
Japan	8.3	1.7	**6.6%**
UK	7.9	0.4	**7.5%**
US	7.6	0.6	**7.0%**

Source: Barro (2006) and own calculations.

The study of Barro (2006) confirms that the equity premium puzzle appears to be an international phenomenon, and further substantiates the original finding of the puzzle.

Erbas & Mirakhor (2007) also found a similar puzzle in a total of 53 emerging and mature markets, though they do not report the exact size. Their work was also conducted at the aggregate level, but unfortunately their data does not allow for a breakdown, or recalculation of their results.

Based on the work of Kocherlakota (1996), Weil (1989), and Barro (2006) we conclude that Mehra & Prescott's finding of an equity premium puzzle is well established for US data. All in all, the international findings of equity premia suggest that the puzzle also exists in countries outside the US, though few studies seem to actually have investigated the explanatory power of utility-based models, including that of Mehra & Prescott.

3.4 Chapter summary

The chapter showed that the findings of Mehra & Prescott (1985) have been confirmed in various ways since their formulation of the equity premium puzzle. One alternative restatement of the puzzle was provided by Kocherlakota (1996), who introduced a reformulation of the original framework, which statistically tested how low the risk aversion could be for the model to succeed. He found that the model of Mehra & Prescott (1985) required value for alpha of at least 8.5, which was considerably higher than the maximum empirically founded estimate of 3. Hence, he was able to confirm the finding of the equity premium in the original dataset without making assumptions about risk aversion and impatience.

Another confirmation of the existence of the puzzle was provided by Weil (1989), with the introduction of the risk-free rate puzzle. Weil observed that the equity premium puzzle could not be solved by simply increasing estimates of investor risk aversion, as it would result in implied risk-free rates that were high above observed levels. Highly risk averse investors with utility theory preferences, would not have been content with the low observed risk-free rates. This observation effectively dismissed that a simple increase of assumed investor risk aversion could be the solution to the equity premuim puzzle.

The existence of the equiy premium puzzle has also found some support intertationally. The equity premium itself has been documented in most economies through highly regarded studies like Dimson, Marsh, & Staunton (2008; 2011), and the sheer size of these premia seem ominous to the success of the Mehra-Prescott model. A model similar to theirs have in fact been tested by Barro (2006) in the G7 countries for the period 1954-2004, and it failed miserably in explaining the observed equity premia.

We concluded that the existence of the equity premium puzzle has been firmly established, though its international presence has received somewhat limited coverage in the literature.

Chapter 4 will provide an overview of the possible solutions to the puzzle that have been provided so far, before we move on to part II of this paper, which introduces myopic loss aversion and the preference structure that it relies on.

4. Potential solutions to the puzzle

During the last quarter of a century there have been countless attempts to explain the equity premium puzzle, but none of them have found sufficiently wide-spread support to be characterized as the decisive solution. In this section we will present the most prominent studies in this line of research. Given the vast amount of literature in the field we will not attempt to provide a full review of all articles written on the matter, but we will cover the most influential and interesting studies. For a more complete review we recommend Kocherlakota (1996), Mehra & Prescott (2003), and Mehra (2008).

The literature that deals with potential solutions to the puzzle, generally points to one of two basic explanations: either the data used by Mehra & Prescott (1985) was biased, or else their model in itself was flawed. We will start by reviewing the data used to study the equity premium, and then we move on to review the studies that have attempted to falsify the model by discrediting its underlying assumptions.

4.1 Biases in data

One possible explanation to why Mehra & Prescott (1985) found a puzzle, could be that their dataset was biased. This does not undermine their findings, but it has obvious consequences for the conclusion, namely whether the puzzle was a testament to the general failure of utility theory, or just the result of special circumstances in the US during the period they investigated.

The equity premium is by no means a local phenomenon as shown in Table 3.2, and there is also some empirical evidence of the puzzle internationally as shown in Table 3.3. The geographically confined nature of the data used Mehra & Prescott, does therefore not seem to hold the explanation to the puzzle, but our study will also provide further insight to this possible shortcoming of the original study[11].

As another critique, several scholars have pointed to the *ex post bias* as a potential explanation to the high observed equity premium. All studies of the equity premium are per definition done *ex post*, while the investment decisions of agents are done *ex ante*. If the realized return on stocks (bonds) has consistently been higher (lower) than what was expected at the time of investment, then a high observed equity premium may have been due to an extra *unexpected* gain (loss) to the investors, and is therefore not the result of an irrational risk profile. If this is in fact the case, then the equity premium

---

[11] Jorion & Goetzmann (1999) claimed that the success and stability of the US stock market weakens the general inferences from studies based on this market alone, but other studies have in fact found amble documentation of the premium globally

could simply be the result of windfall gains and should not be interpreted as the premium required by investors for holding stocks. The high return on equity would thus have been purely coincidental, and is not the result overly risk averse investors.

Fama & French (2002) proposed that the equity premium in the period after WWII (1950-2000) could be partly explained by unexpectedly high capital gains on stocks. They used a dividend growth model and an earnings growth model to estimate the *expected return* on equity during the period. They found that the realized return was more than double of what could be forecasted by using the two models.

The former point was also proposed by Madsen & Dzhumashev (2009), who found that the increased inflation following the abandonment of the gold standard was largely unanticipated by the investors. This drove down the real return to bonds and thereby increased the *ex post* equity premium.

Dimson, Marsh, & Staunton (2008) proposed a possible source of unexpected capital gains on stocks from the closer integration of goods markets and capital markets during the 20th century. If barriers disappeared faster than investors anticipated, causing diversification to become less costly, then this could have driven up stock prices unexpectedly. Abel (2002) proposed a model with doubtful and pessimistic investors, and was thereby able to explain the high *ex post* premium. This model was however later tested empirically by Giordani & Söderlind (2006), who found very little support for the explanation.

While the studies above can possibly provide a partial solution to the equity premium puzzle, and no study to our knowledge has dismissed the equity premium on grounds of the *ex post bias.* Even if the *ex ante* equity premium was in fact only half of the *ex post* premium (as Fama & French find), we still have an equity premium of over 3% in the Mehra-Prescott dataset. This reduces the failure of the consumption based asset pricing model, but its prediction is still off by more than 100% (1.4% explained vs. 3.1% observed). Furthermore, a study like that of Fama & French would only provide a valid estimate of the expected premium if investors did in fact apply the two models in determining their required return, which has yet to be confirmed.

Biases in the Mehra-Prescott dataset do therefore not seem to offer a satisfactory explanation to the equity premium puzzle, though they might be able to account for part of the puzzle[12].

---

[12] Other biases such as *survivorship bias* and a bias in the form of *omitted markets* have also been proposed, but they have only been found to overstate the equity premium with around 0.1%-point (Dimson, Marsh, & Staunton, 2008).

4.2 Questioning the assumptions behind model

The findings of Mehra & Prescott (1985) appear to be quite robust. The puzzle has been found outside the US and over varying time periods, and general biases in the datasets do not appear able to fully explain their findings. It therefore seems likely that the conclusive solution to the puzzle (if one such exists) must be found by reviewing the assumption behind the model. Numerous attempts have been made to solve the puzzle in this manner, but none of them have yet been credited for solving it once and for all.

As described in section 2.1 the model of Mehra & Prescott (1984) rests on three basic assumptions: That markets are complete, that trading is costless, and that investor preferences are represented by the standard utility function shown in equation (2.1). In order for the model to fail one of these assumptions must be flawed. The following section reviews a number of the most influential studies that claim to provide a possible explanation to the puzzle by relaxing one or more of the assumptions.

4.2.1 Assumption 1: Markets are complete

This assumption is implicit in all models that assume a representative investor. In order for this investor to exist, it must be possible for all the heterogeneous agents to equalize their marginal rate of substitution over different states by trading in a perfect capital market. After trading in the market the heterogeneous individuals become *marginally* homogeneous and the representative investor emerges (Mehra & Prescott, 2003). The presence of such a market is therefore essential for using per capita consumption in the model as the authors do.

If an investor is not able to trade efficiently, his consumption pattern becomes much more volatile than per capita consumption, and the model no longer holds at the aggregate level. A violation of this assumption leads to a higher explanatory power of the Mehra-Prescott model and hence, diminishes the size of the equity premium puzzle. Recall from equation (2.11) in section 2.1 that in the utility-based model of Mehra & Prescott, the equity premium is driven by investor risk-aversion (CRRA) and the volatility of consumption. If markets are not complete, individual consumption volatility will be higher than the per capita measure used by Mehra & Prescott, and this would lead to a higher predicted equity premium without increasing the assumed risk aversion. A higher volatility thus means a higher expected equity premium and a possible elimination of the puzzle.

As pointed out by previous studies, the main caveat of the assumption is that not all income shocks are diversifiable through market trading. If the individual investor is unable to "share" an income shock with

28

all the other investors, then his consumption is more volatile than per capita consumption and the explanatory power of the Mehra-Prescott model increases. An example of such an uninsurable income shock could be the risk of getting fired. Even if the shock is dampened by unemployment benefits, it will still decrease the investor's income subtantially.

Weil (1992) and Mankiw (1986) both explored the effect of such shocks in a two-period version of the utility-based model employed by Mehra & Prescott (1985). They suggested that uninsurable uncertainty about period two will increase the precautionary savings rate in period one, and thereby drive down the risk-free rate. This would offer an explanation of the low risk-free rate, given that investors are in fact as risk averse as the utility-based model suggests, which would effectively solve the risk-free rate puzzle.

Lucas (1994) however found, that the equity premium puzzle is persistent under different assumptions of market incompleteness, including undiversifiable income risks[13], and argued that a two-stage model is too simple. Kocherlakota (1996) supported her argument with the introduction of *dynamic self-insurance*. Here, individuals are able to compensate for incomplete capital markets by lending and borrowing amongst themselves. In the two-period model, the entire income shock must be absorbed in period two, while individuals under dynamic self-insurance can smooth changes in consumption over multiple periods. Arguably this resembles the situation faced by real-life investors more closely.

It should be noted that self-insurance does require that income shocks are non-permanent, in order for the investor to be able to absorb a shock without changing his consumption. Heaton & Lucas (1996) found though, that such shocks to labor income are generally not permanent, and this conclusion was further substantiated in two of their subsequent studies (Heaton & Lucas, 1997; 2000).

Constantinides, Donaldson & Mehra (2002) relaxed the assumption about the one representative investor, by dividing individuals into three generations and analyzing a representative investor from each. The young generation has very limited funds to invest in stocks and is also faced with borrowing constraints[14], and this generation is therefore largely unable to access the stock market. As they grow older, more and more of their wealth is derived from investment returns, until it approaches 100% at retirement. At this point the correlation between return on investments and consumption therefore

---

[13] Others include constraints on short sales and borrowing. Telmer (1993) investigated the representative agent theory under a combination of incomplete markets and heterogeneous agents and arrived at the same result as Lucas.
[14] In a perfect market they would be able to borrow against their future income (and thereby increase the interest rate), but in reality this is not possible, as human capital generally does not serve as collateral in a modern economy for reasons of moral hazard and adverse selection (Davis & Willen, 2001).

approaches one. The implications are twofold: first, the correlation between equity return and consumption is much higher for market participants than for the population as a whole (explaining the high return on stocks), and second, investors are willing to accept a low return on bonds to insure themselves against fluctuations in equities. However appealing this argument may be, it has unfortunately not yet undergone sufficient empirical testing due to its complicated data requirements.

Based on the seemingly non-permanent nature of undiversifiable income shocks and the ability of investors to self-insure, violations of the complete market assumption do not seem to have provided a definite solution to the puzzle. The framework of Constantinides, Donaldson, & Mehra (2002) may prove to be such a solution, but its empirical foundation is still too vague to draw any decisive conclusion in this regard. So for the time being we conclude that relaxing the Mehra-Prescott model's assumption of complete markets does not seem to present a satisfactory explanation to the equity premium puzzle.

4.2.2 Assumption 2: No trading barriers

With the assumption of a perfect market, the Mehra-Prescott model also assumes a market without trading barriers in the form of transaction costs or constraints (e.g. on borrowing). The question is yet again, whether the assumption is strong enough to discredit the model and provide an explanation for the apparent failure of the traditional utility-based paradigm. If trading is sufficiently costly it could disrupt the market from clearing and dismiss the appropriateness of analyzing a representative investor. Further, substantial borrowing constraints would drive down the demand for loans and thereby explain the low observed risk-free rate even if investors are as risk averse as utility theory seems to suggest.

Aiyagari & Gertler (1991) studied transaction costs as a possible explanation to the high historical equity premium in the US. They provided a solution to the equity premium puzzle, where the high equity premium arises from a combination of a demand for liquidity, and transactions costs on stocks. Uninsured investors in the economy (which are assumed to exist) will self-insure in the market, but as only stocks are assumed to carry a transaction cost, the primary instrument for this purpose would be the risk-free asset (T-bills). They found support for this argument in the higher turnover of T-bills relative to stocks, and further suggested that this could give rise to a liquidity premium on the risk-free asset. Both effects would drive down the expected risk-free rate, without making changes to the investors' preference structure, and hence solve the risk-free rate puzzle. Their model had some success, and they found that it was able to explain about half of the equity premium (Aiyagari & Gertler, 1991).

Jang et al. (2007) refined the previous studies by introducing two scenarios with different transaction costs and liquidity constraints, namely a bull market and a bear market. They found that this improved the explanatory power of their model, but they were still not able to explain the observed equity premium in full.

Common for most of the research that suggest possible solutions the puzzle by introducing market frictions is however, that the models require very high levels of frictions to succeed (up to 10% of trading value). In a thorough review of this line of literature, Heaton & Lucas (2007) found that current studies of market imperfections and transaction costs cannot explain the equity premium puzzle. Aiyagari & Gertler (1991) and Kocherlakota (1996) reached a similar conclusion. Kocherlakota further argued that stocks and bonds must be subject to highly differentiated trading costs for the expected equity premium to increase, and he did not find empirical support of this being the case.

Thus, the existence of market imperfections including trade barriers has not yet proved able to fully explain the presence of the equity premium puzzle either.

4.2.3 Assumption 3: Standard preference structure

So far we have looked at biases in data, and two violations of assumptions regarding perfect market efficiency, as possible solutions to the puzzle. We now turn to the assumption that has been investigated by the largest number of studies, namely that of a standard preference structure (see section 2.1). We start by reviewing some of the most recognized alterations to the standard preference structure, and then we turn our attention to the more fundamental changes that were introduced under myopic loss aversion. Myopic loss aversion replaces the normative preference structures with the inherently descriptive prospect theory preferences, and thereby provides a possible solution to the equity premium puzzle. A detailed description of myopic loss aversion is provided in part two, and the ability of myopic loss aversion to explain the equity premium puzzle in Sweden, is examined in chapter 11 of this paper.

Generalized Expected Utility (GEU)

As we noted in section 2.1, the model of Mehra & Prescott assumes that the coefficient of relative risk-aversion (CRRA $= \alpha$) is the reciprocal of the elasticity of intertemporal substitution (EIS $= \alpha^{-1}$). The interpretation is that the investor is averse towards variations in his consumption between different economic states, but it also implies that he is averse to changes to his consumption over time, i.e. that he dislikes consumption growth. This seems counterintuitive, and Mehra & Prescott also pointed out that there is no apparent reason why this should be the case (2003). When investors in the Mehra-

Prescott model dislike variations in consumption, but observe that consumption has grown over time, they have little incentive to save. In fact they are more likely to borrow against future consumption to smooth their consumption pattern, as we saw in the formulation of the risk-free rate puzzle. The absent demand for savings (or high demand for loans), drives the expected risk-free rate up, high above historic levels (e.g. 12.7% with $\alpha = 10$ and $\beta = 0.99$). The disentanglement of CRRA from EIS could thus provide a solution to the risk-free rate puzzle, by reducing the demand for savings, and in turn provide further insights into the nature of the equity premium puzzle.

The generalized expected utility model (GEU) of Epstein & Zin (1989; 1991) separated the coefficient of relative risk aversion (CRRA $= \alpha$) from the elasticity of intertemporal substitution (EIS $= \rho^{-1}$). As α and ρ can vary independently, the model allowed both CRRA and EIS to be high (the isoelastic utility function of equation (2.2) assumes that $\alpha = \rho$). When the two factors are separated, the investor is no longer averse to consumption growth, and will borrow only to smooth consumption over different economic states and not over time in general. This will increase the savings rate (or decrease the demand for loans), and drive down the expected risk-free rate.

Mehra & Prescott (2003) and Kocherlakota (1996) pointed out that the findings of Epstein & Zin (1991) unfortunately had a shortcoming, as their solution did require both CRRA (α) and EIS (ρ^{-1}) to be high. This solved the risk-free rate puzzle, as risk aversion could be high without driving down the risk-free rate, but it did not explain the equity premium puzzle. Investors still needed to be highly risk averse for the expected utility model to work. Furthermore, Campbell (2001) found that the EIS is actually quite small, and hence re-introduced the risk-free rate puzzle. GEU does therefore not seem to hold an immediate solution to the two puzzles.

Habit formation

Habit formation represented another attempt to solve the two puzzles by separating the CRRA from the EIS. First formulated by Constantinides (1990), it stipulated that the utility of consumption in period t is dependent on the consumption in period *t-1* (the Mehra-Prescott model assumed no dependence). Formally, it defined *momentary utility* in period t as a *decreasing* function of consumption in period *t-1* and *marginal utility* as an *increasing* function of consumption in *t-1*.

In other words, a high level of past consumption decreases utility from current consumption (decreases momentary utility), and it increases the desire for maintaining the high level of consumption (increases marginal utility). This seems intuitively appealing, as it captures that an affluent lifestyle in one period

will leave the investor craving a similar standard of living in the succeeding period, while appreciating his current living conditions less. Constantinides found that the model could explain the equity premium puzzle in the Mehra-Prescott dataset with CRRA = 2.81 and a quite small EIS (0.09).

The problem with the model of Constantinides (1990), as pointed out by Kocherlakota (1996) and Mehra & Prescott (2003), is that it requires investors to be highly averse to even small decreases in consumption. For habit formation to explain the puzzles, investors must be willing to pay a substantial premium to avoid even minor decreases in consumption. Hence, the model can provide a reasonable aversion to wealth risk, but instead it introduces a high aversion to consumption risk. Athanasoulis & Sussmann (2007) showed that habit formation is indistiguishable from simply increasing the CRRA gradually over time. They found that habit formation had little explanatory power, if the CRRA was held constant. Habit formation does therefore not seem able to explain the equity premium puzzle.

The model fares somewhat better in terms of explaining the risk-free rate puzzle. In order to be certain of maintaining current consumption, the investor will require a quite high savings rate. As with GEU this can explain the low historic interest rates, but again it requires that investors are implausibly risk-averse, hence leaving the equity premium puzzle intact. Further, Heaton (1993) pointed out that the model introduces a highly volatile interest rate, and this counters with the empirically low volatility of the short-term interest rates.

Recent research have proposed a more sophisticated approach, which incorporates habit formation along with incomplete markets and heterogeneous agents. Pijoan-Mass (2007) proposed one such model, and had considerable success in explaing the observed equity premium. Habit formation should therefore not be rejected as a possible solution to the equity premium puzzle just yet.

"Keeping up with the Joneses"

The concept of "keeping up with the Joneses" (KUJ) captures that the investor does not solely obtain utility based on his own consumption, but also from his consumption *relative* to the average investor in the economy (the "Joneses"). It was coined by Duesenberry (1949) and first applied by Abel (1990) in relation to the equity premium puzzle. Marginal utility is here an increasing function of average consumption, as the desire of the investor to consume increases with the consumption obtained by his peers. This is similar to the assumed effect under habit formation where marginal utility increases with past consumption. Like habit formation and GEU, KUJ incorporates time-separable utility functions, i.e. allows for a separation of CRRA and EIS.

The model of Abel (1990) compared individual consumption in period *t* to the average consumption in period *t-1*, as he argued that investors are primarily concerned about falling behind the average investor[15]. Gali (1994) modified Abel's model to compare consumption in the same time period, and Kocherlakota (1996) combined the two models to incorporate both current and lagged relative consumption.

The implication of KUJ for the equity premium puzzle is that the investor does not need to be highly averse towards his own consumption risk, if he is sufficiently averse towards variations in average consumption. This is equivalent to investors under habit formation being highly averse towards any decrease in their consumption over time. By investing in stocks, the investor in KUJ faces the possibility of incurring a loss that decreases his consumption relative to the average investor, and this will make stocks unattractive to him (Abel, 1990).

The prediction of KUJ does however not offer much insight into the nature of the equity premium puzzle, as it simply replaces a high risk aversion towards individual consumption risk with a high risk aversion towards societal consumption risk (Kocherlakota, 1996). This is essentially the same problem that the habit formation model of Constantinides suffered under, i.e. replacing one high risk aversion with another. Mehra & Prescott (2003) further noted that because the model is concerned with the *ratio* of individual consumption relative to average consumption (and not the absolute difference), the definition of the equity premium in KUJ can actually be reduced to equation (2.11), and is thus indistinguishable from the original Mehra-Prescott model.

DeMarzo, Kaniel, & Kremer (2004) suggested another interesting take on the KUJ preference structure. They proposed an economy of *N* communities that each have a scarce resource, which is only valued by the investors of that community. It results in the price of the resource being highly correlated with the affluence of the community, and *this* is why investors would care about their relative consumption. The alternative rationale behind the KUJ framework may seem trivial at first, but it implies that an investor who is averse towards societal consumption *still has a standard preference structure*. He does not get disutility from jealousy or loss of status caused by falling behind his peers, but because his affluent peers drive up the price of the good he desires, and this simply reduces his consumption. Gómez, Priestley, & Zapatero (2009) tested the two rationales (Abel vs. DeMarzo et al.) on international data, but were

---

[15] Abel therefore called it "*catching* up with the Joneses"

unable to dismiss either of them. They did however find that both models had some explanatory power with regards to past asset returns.

It is important to note that the model of DeMarzo et al. still violates the assumptions behind the Mehra-Prescott model. Their framework assumes incomplete markets and investors faced with borrowing constraints, so their adaptation simply shifts the KUJ theory to an investigation of the first two assumptions instead. KUJ under these terms may prove to fare better than seems to have been the case with the theory's attempt to justify a non-standard preference structure.

Myopic Loss Aversion

Myopic loss aversion also provided a potential solution to the equity premium puzzle by changing the assumed preference structure of the investor, but it did so in a profoundly different way than the explanations described above. Instead of disentangling CRRA and EIS, and introducing various sources of high investor risk, myopic loss aversion assumed a preference structure based on experimental evidence and behavioral finance.

Introduced by Benartzi & Thaler (1995), the model explained investors' unwillingness to bear risk as a combination of loss aversion and mental accounting. The authors argued that the combination of the two effects could explain the high observed equity premium, by modeling the fact that financial investors are averse to losses (and not risk) and that they evaluate their portfolios too often. Both phenomena are well-known characteristics of decision making under risk in experimental settings, but it must be stressed that the combination of the two provided a descriptive model, and not a normative contribution like the other potential explanations of the equity premium puzzle provided in this chapter. However, this does not make the explanation any less intriguing nor does it diminish its potential as a possible solution to the puzzles.

The model leans on two principles from behavioral finance, namely *loss aversion* and *mental accounting*. Loss aversion is a tendency that was discovered through behavioral experiments, and it implies that investors are more sensitive to decreases in their wealth, relative to increases. In other words, if an individual is faced with a loss and a gain of equal size, the loss will yield a higher negative value than the corresponding positive value associated with the gain (Kahneman & Tversky, 1979).

Mental accounting basically relates to the unconscious steps that people employ to organize, evaluate, and keep track of their investments, including how often these steps are performed. More formally it refers to the cognitive procedure of individuals, specifically investors evaluating and structuring

prospects and investments (Thaler R. H., 1999). Rather than evaluating decisions within their broader context, individuals tend to evaluate them one at a time and then assign them to separate mental accounts. For financial investors, this implies that securities are often assessed singularly rather than as a component of the entire portfolio. Furthermore, investors are inclined to make frequent evaluations of their transactions, leading to short-term trading decisions (Thaler, Tversky, Kahneman, & Schwartz, 1997).

The implication for the equity premium puzzle is as follows: Given that stock returns are more volatile than the risk-free return, the investor will observe losses on his stock investments relatively more often, if he evaluates them too frequently. If the investor has prospect theory preferences, he will be hurt twice as hard from these observed losses as he will benefit from the observed gains. This would lead him to demand a higher return on stocks, and thus provide an explanation to the equity premium puzzle. In the end, Benartzi & Thaler (1995) were able to model the observed equity premium in the US for the period 1926-1990, whilst employing a low risk-free rate.

The remainder of the paper is dedicated to investigating the explanatory power of myopic loss aversion on a non-US economy, namely Sweden. As the model assumes a profoundly different preference structure than expected utility theory, we will start by taking a closer look at the assumptions behind standard preferences and review their empirical validity.

4.3 Chapter summary

Vast amounts of literature have tried to provide a solution to the equity premium puzzle, but none have yet received general recognition as a definite solution. The possible solutions suggest two basic ways of explaining the puzzle: Either the data is biased, or the model is flawed.

The potential bias that has received the most attention is the *ex post* bias, which captures that all studies of the equity premium are done *ex post,* while the investment decision of investors is taken *ex ante*. If investors have somehow received a higher return on their equity investments (or a lower return on their risk-free holdings) than they had anticipated, then the high observed equity premium cannot be interpreted as a demand by the investors to carry risk. Studies of the ex post bias have been able to explain about half of the equity premium puzzle, but none have dismissed the puzzle entirely on this basis.

The majority of studies have been concerned with investigating the three main assumptions behind the Mehra-Prescott model, namely that markets are complete, that trading is costless and without barriers, and that investors have standard preferences according to expected utility theory. Studies of the first two assumptions about market perfection, and lack of friction, have had some success in explaining the puzzle, but they still are far from providing a credible solution with empirical support. Research into the assumed preference structure of the Mehra-Prescott model seems to have fared somewhat better.

Most attempts to modify the standard preference structure have been concerned with separating the coefficient of relative risk aversion, and the elasticity of intertemporal substitution, to remove the inherent aversion towards consumption growth in the Mehra-Prescott model. Unfortunately most of the potential solutions end up substituting a high aversion towards consumption growth, with a high aversion towards some other factor, which may not have more intuitive or empirical appeal.

Myopic loss aversion also proposes a non-standard preference structure as a solution to the puzzle. It has no normative foundation, but is based on experimental evidence of human decision making under uncertainty.

Part II will provide a deeper understanding of the foundation and the intuition behind prospect theory and myopic loss aversion, before the theory is tested empirically in Sweden in part III.

Part II: Myopic Loss Aversion

5. The failure of the standard preference structure

Myopic loss aversion rests on the foundation of prospect theory, which was formulated as an alternative to the standard preference structure assumed by utility theory. Where utility theory was conceived as a normative model evolving around a rational and idealized decision maker, prospect theory was introduced as a descriptive model focused on the actual behavior of humans.

In order to outline the premise of prospect theory, we will start by examining the empirical shortcomings of standard preferences under expected utility theory, which was the theoretical foundation of the utility-based model of Mehra & Prescott (1985).

5.1 The axioms of expected utility theory

Theoretically, competitive environments ensure that rational and value optimizing individuals and organizations will outperform and outlive irrational players, and hence markets are assumed to be dominated by rational decisions. Being a normative model, expected utility theory therefore relies on the assumption that people make rational choices that maximize their utility. While this assumption has intuitive appeal, empirical research indicates that violations of logic choices are systematic, widespread, and rather fundamental and thus profoundly discredits the axioms of utility theory. We will examine these violations in more detail in the following sections, to provide a further insight into the issued with the preference structure used in the original Mehra-Prescott framework.

At the core of expected utility theory are five essential decision principles, which are assumed to govern the preferences people exhibit, when choosing amongst alternatives (Tversky & Kahneman, 1986);

i. *Cancellation*. If prospect A is preferred to B, then a chance to receive A if it snows tomorrow is preferred to a chance to receive B if it snows.

ii. *Transitivity*. Agents' preferences are consistent so that if prospect A is preferred to B, which in turn is preferred to C, then transitivity governs that A will be preferred to C.

iii. *Dominance*. If prospect A yields outcomes at least as high as B in all states, and higher in at least one state, then A dominates B and is preferred.

iv. *Invariance*. The preference order of a series of prospects is independent of the manner in which the prospects are presented or framed. As a consequence, two alternative versions of the same random choice should be treated alike and result in the same decision.

v. *Risk aversion*. Agents are risk averse, i.e. they will always prefer a riskless prospect to any risky prospect yielding the same expected value.

These tenets have varying intuitive appeal, but are all crucial for the premise of a rational decision maker and consequently essential for the workings of expected utility theory, and the model of Mehra & Prescott (1985). The following section will address the empirical validity of each assumption.

5.2 Violations of the axioms

Through their experiments of investor sentiment, Tversky & Kahneman (1979; 1981; 1986) were able to employ cognitive psychology to describe how investors actually form opinions and beliefs about assets, and how they evaluate risks. Consequently, they were able to solidify their critique of all five axioms on grounds of them being normatively essential but empirical invalid. To illustrate their findings regarding the shortcomings of utility theory, we turn to a series of their experimental evidence. It should be noted that the experiments presented are merely chosen to illustrate the conclusions of the authors and should not be interpreted as the sole empirical foundation of prospect theory.

5.2.1 Certainty effect and probability sizes

The main argument behind *cancellation*[16] is that only one state will eventually be realized, so prospects may be evaluated separately for each state. The available options should therefore only be evaluated on the basis of the states in which they yield different outcomes, and the characteristics that they share may be "cancelled" out without it changing the preference order. Despite it being an axiom of utility theory, cancellation has been rejected by several authors (Allais, 1979; Ellsberg, 1961), and the following examples will illustrate why.

An important violation of cancellation comes from the so-called *certainty effect*. While standard utility theory assumes that different outcomes are weighted by their individual probabilities, studies show that people are prone to give excessive weight to outcomes that they consider certain relative to outcomes that are merely probable.

---

[16] If A is preferred to B the investor will prefer to receive A over B in the event of X

Consider the following experiment of Kahneman & Tversky (1979). Here they provided *N* number of respondents with two problems that each offered two different options:

Problem 1:

Option	Outcome	Probability	
	2,500	0.33	
A	2,400	0.66	[18]
	0	0.01	
B	2,400	1.00	**[82]**
	0	0.00	
N = 72			

Problem 2:

Option	Outcome	Probability	
	2,500	0.33	
C	0	0.67	**[83]**
D	2,400	0.34	[17]
	0	0.66	
N = 72			

Source: Kahneman & Tversky (1979) p. 265-66.

The figures in the brackets indicate the percentage of people who chose the option in question. Respondents clearly preferred option B and C respectively, and an analysis of the individual preference patterns revealed that 61% of respondents choose this combination of options. In order to make the relevant inference, let us first consider Problem 1. Given the condition of utility theory that $u(0) = 0$, we note that respondent preferences imply the following relationship:

$$B > A \Leftrightarrow u(2,400) > 0.33u(2,500) + 0.66u(2,400)$$

Removing, or "cancelling" the 66% chance of receiving 2,400, leaves us with the following:

$$\Leftrightarrow 0.34u(2,400) > 0.33u(2,500)$$

In Problem 2 we see the reverse inequality:

$$C > D \Leftrightarrow 0.33u(2,500) > 0.34u(2,400)$$

We may therefore conclude that people behave irrationally in regards to the assumptions of utility theory. The difference is in the character of the two prospects, where the first represents a sure gain and the latter a probable gain and thus demonstrates the certainty effect.

The example above is constructed to illustrate the violation of *cancellation*. According to the cancellation axiom, the choice preference ought not to change when the individual probabilities are reduced by a common factor. As we see above, individuals do however react differently when reducing the probability of a certain outcome versus an already uncertain outcome by the same factor. Reducing the probability of 100% in B by any given amount, would hurt the decision maker more than reducing any other

probability by the same amount, because it eliminates the option of having a sure gain. Therefore, the certainty effect fundamentally violates the cancellation effect.

Tversky & Kahneman (1979) also showed that even when the certainty effect is removed, and the proportions of both probabilities and outcomes are identical, the preference order will not necessarily be the same. Once the probabilities became very small, the focus changed to monetary payoff, and people had a tendency to attribute too much weight to the small probability. As respondents clearly did not evaluate prospects purely based on the *absolute* probabilities, the cancellation axiom of utility theory was contradicted.

5.2.2 Intransitive preferences

Recall that the *transitivity* axiom dictates that if option A > C and C > E, then A > E. Empirical observations however, indicate that there is an inherent variability in the evaluative process of individuals. Transitivity is likely to hold when the prospects in question are evaluated separately, but it fails when the value of one choice is influenced by the available alternatives. To understand the intuition behind this statement, we turn to a study by Tversky (1969).

An experiment was designed to examine preferences between simple gambles, and the respondents were asked to choose between two of the gambles below, in varying combinations:

Problem 3:

Option	Probability	Outcome	Expected value[17]
A	0.29	5.00	1.46
B	0.33	4.75	1.58
C	0.38	4.50	1.69
D	0.42	4.25	1.77
E	0.46	4.00	1.83
N = 22			

Source: Tversky (1969) p. 33.

A considerable number of respondents displayed intransitive preferences. In particular, people used a so-called *lexicographic semi-order heuristic* when choosing between prospects. Investors initially compared the probabilities to win in the respective gambles, and if this difference was large enough, the respondents chose the prospect that offered the higher probability to win. On the other hand, if the difference was small, the gamble with the highest outcome was chosen. Consequently, people who

---

[17] Respondents were only provided with the probability and the outcome, the expected value is merely included here for the comfort of the reader

preferred A to C and C to E, subsequently chose E over A due to the larger difference between their respective probabilities. These findings were later verified by other scholars (Budescu & Weiss, 1987; Lindman & Lyons, 1978), who ultimately disproved descriptive validity of the transitivity axiom.

5.2.3 Combining risky prospects

To illustrate the empirical violations of the last three tenets (*dominance*, *invariance*, and *risk aversion*), Tversky & Kahneman (1981) orchestrated an experiment concerning decisions between combinations of risky prospects. Initially, respondents were asked to make a choice between the following two options:

Problem 4:

Option	Outcome	Probability	
A	240	0.25	[0]
	-760	0.75	
B	250	0.25	**[100]**
	-750	0.75	
N = 86			

Source: Tversky & Kahneman (1981) p. 454.

Not surprisingly, all respondents made the rational decision of selecting option B, which clearly dominated option A.

Subsequently, Tversky & Kahneman presented the respondents with a problem that involved two consecutive choices, one between favorable prospects and one between unfavorable prospects. It was assumed that the two gambles would be played out independently of one another.

Problem 5 - Decision i:

Option	Outcome	Probability	
C	240	1.00	**[84]**
	0	0.00	
D	1,000	0.25	[16]
	0	0.75	
N = 150			

Problem 5 - Decision ii:

Option	Outcome	Probability	
E	-750	1.00	[13]
	0	0.00	
F	-1,000	0.75	**[87]**
	0	0.25	
N = 150			

Source: Tversky & Kahneman (1981) p. 454.

In this problem, the majority of respondents revealed a preference for option C and F, with 73% of the participants choosing that particular combination. Interestingly, that combination is the exact equivalent

of the unattractive option A in Problem 4[18]. The more desirable option B is the combination of D and E[19], but this was only chosen by 3% of respondents.

Now let us see how these experimental results affect the three remaining assumptions of expected utility theory:

The first axiom to consider is *invariance*, which dictates that the appeal of a prospect should not rely on the way in which it is presented or framed. We saw that respondents exhibited different preferences towards the same outcome, due to the way in which the problem was framed. The two representations in Problem 4 and 5 should have elicited the exact same choices; however the vast majority of subjects altered their decision significantly, causing the invariance axiom to fail unequivocally.

Next, we may consider the assumption of *dominance*. Recall that the dominance tenet dictates that if one option is better than another in one state, and at least as good in all other states, then the dominant option should be preferred. The fact that respondents evaluated decisions i) and ii) in Problem 5 separately, caused them to select a dominated option over the dominant one, and consequently displayed behavior that was inconsistent with the assumptions of utility theory. The experiment illustrates how failures of *invariance* are likely to produce violations of *dominance*.

Interestingly, the experiment also showed that people faced with several prospects fail to make their decision based on *final wealth*. Instead of considering the final states that were illustrated in Problem 4, the respondents evaluated each prospect in isolation, thereby treating them as individual gains and losses relative to their initial reference point. We return to the implications of this finding in chapter 6.

Finally, Problem 4 and 5 illustrate how individuals do not obey the laws of *risk aversion* in the form applied by utility theory. In terms of decision i), 84% of the participants displayed *risk aversion* by selecting option C that offered a sure gain. For the second decision however, 87% of respondents demonstrated *risk seeking* behavior by choosing option F instead of taking the sure loss. Tversky & Kahneman (1981) noted that risk aversion should be substituted with *loss aversion*.

As the intuition behind this finding is crucial for the remainder of this paper, we consider a final experimental choice set concerning the behavior of risk aversion surrounding gains and losses respectively.

---

[18] Option C + Option F = $(240, 1.00) + (-1000, 0.75) = (240, 0.25; -760, 0.75) =$ Option A
[19] Option D + Option E = $(1,000, 0.25) + (-750, 1.00) = (250, 0.25; -750, 0.75) =$ Option B

5.2.4 Gains versus losses

In their 1979 paper, Tversky & Kahneman included the following experiment. Respondents were presented with two problems that were identical apart from the signs being reversed:

Problem 6: Positive prospect

Option	Outcome	Probability	
A	4,000	0.80	[20]
	0	0.20	
B	3,000	1.00	**[80]**
	0	0.00	
N = 95			

Problem 6': Negative prospect

Option	Outcome	Probability	
C	-4,000	0.80	**[92]**
	0	0.20	
D	-3,000	1.00	[8]
	0	0.00	
N = 95			

Source: Kahneman & Tversky (1979) p. 268.

From Problem 6, we see another representation of the certainty effect. As displayed in Problem 1, individuals have a strong preference for certain gains, with 80% of respondents choosing option B. When faced with a choice between negative prospects however, the participants changed their preferences and became risk seeking, which is a direct violation of utility theory that assumes that individuals are always risk averse.

In order to explore the consequences of removing the certainty effect, Tversky & Kahneman presented the same participants with Problem 7 as seen below:

Problem 7: Positive prospect

Option	Outcome	Probability	
E	4,000	0.20	**[65]**
	0	0.80	
F	3,000	0.25	[35]
	0	0.75	
N = 95			

Problem 7': Negative prospect

Option	Outcome	Probability	
G	-4,000	0.20	[42]
	0	0.80	
H	-3,000	0.25	**[58]**
	0	0.75	
N = 95			

Source: Kahneman & Tversky (1979) p. 268.

While the preferences between the negative prospects were still the mirror image of the preferences between positive prospects, the results exhibited in Problem 7 indicate that when probabilities decrease, people shift their focus from variance to expected value. Within the domain of gain, the difference in variance between (4,000, 0.20) and (3,000, 0.25) was not enough to outweigh the difference in expected value, and participants became less risk averse than observed in Problem 6. Correspondingly, respondents became less risk seeking in the domain of losses, by avoiding the option that represented the largest potential loss despite offering a lower probability. The experiment ultimately showed that certainty increases the desirability of gains as well as the aversiveness of losses.

All in all, the experiments performed by Tversky & Kahneman (1969; 1979; 1981) clearly illustrated the failures of the axioms of standard utility theory, and thereby profoundly discredited the axioms behind the standard preference structure in utility theory. The authors consequently set out to formulate a descriptive theory on decision making under risk that could model the observed preferences of real economic agents. Myopic loss aversion adapts the resulting prospect theory preferences to replace the standard preferences in the Mehra-Prescott model.

5.3 Chapter summary

As seen in chapter 4, some authors have argued that the solution to the equity premium puzzle must be found through the introduction of an alternative preference structure, compared to the one applied under traditional expected utility theory. This chapter has presented five central axioms that ultimately represent the foundation for the preference structure used under the model of Mehra & Prescott (1985), namely: cancellation, transitivity, dominance, and invariance.

We illustrated how Tversky & Kahneman (1969; 1979; 1981) have conducted a wide range of experimental research that clearly demonstrated systematic violations of all the five axioms, thereby irrevocably rejecting standard expected utility theory as an empirically valid model.

All in all, the experiments showed that individuals do not evaluate prospects based on objective probabilities of the potential outcomes, as they have a tendency to put too much emphasis on small probabilities. Also, the way in which prospects are framed may influence people's preference order. Finally, we saw that individuals make choices separately and not based on measures final wealth, and that they appear to be loss averse rather than risk averse, in that they treat gains and losses differently.

Tversky & Kahneman (1979; 1992) subsequently formulated an alternative theory to account for the observed behavior amongst decision makers. This theory, namely prospect theory, represents a crucial part of the potential solution to the equity premium puzzle, which we test empirically in this paper, and it will be the focal point of the next chapter.

6. Prospect Theory

Prospect theory was introduced by Kahneman and Tversky (1979), and it represented a deductively founded alternative to the standard preferences, assumed in the Mehra-Prescott model (1985). The theory pointed to loss aversion rather than risk aversion, and it employed investors' subjective decision weights, whilst dismissing the notion of objective probabilities for different states or outcomes.

As illustrated in the previous chapter, the authors used extensive experimental evidence to formulate their descriptive theory on how individuals actually evaluate risk. Investigations of decision making under uncertainty revealed that investors do not evaluate their positions at absolute levels of wealth, but rather gains and losses relative to a reference point. People were in fact found to be much more sensitive to losses – even small losses – than to gains of the same magnitude.

Kahneman & Tversky (1981) inferred that choice processes with any level of risk occur in two phases; the *framing phase* where the agent attempts to organize and classify available prospects, and the subsequent *evaluation phase* where a decision is made regarding the best alternative. The process is applied by Benartzi & Thaler (1995) in their model of myopic loss aversion, so the following sections will address the cognitive steps that decision makers go through, as it is a key part of our methodology.

6.1 The Framing Phase

The framing phase refers to the investor's initial analysis of any available prospect. It captures how the investor weighs potential consequences of the prospect's different outcomes, estimates their probabilities, and finally forms a subjectively adjusted representation of the prospect.

As we observed in the experiments from chapter 5, individuals for example have a tendency to ignore very small probabilities. This tendency comes from the fact that investors are generally inclined to simplify their decisions, and during the framing phase they may modify their perceptions of the available choices.

Recall that one of the axioms behind rational choice theory dictates *invariance*, meaning that decisions are made independently of the description of the problem. Just as Tversky & Kahneman (1981) observed violations of this axiom, Barberis & Thaler (2002) also found that the behavior of the decision maker is indeed influenced by the *formulation* of the issue. In addition to the effects discussed earlier, their findings also pointed to the importance of the *individual framing* of each investor. People make decisions

based on personal "references" that depend on individual norms and habits, so the same prospect may yield dissimilar utilities for different investors. Consequently, the preference order of a given number of prospects is likely to be highly influenced by the representation of said prospects as well as the personal characteristics of the investor. Apart from being difficult to model, this predisposition amongst investors effectively makes market-wide rationality implausible.

In chapter 7, we will return to a more detailed description of the behavioral patterns that are observed during the framing phase, namely mental accounting and choice bracketing. For now it is sufficient to note that the framing phase is where a given prospect is coded by the decision maker, and where it is transformed into a more simple and accommodating representation. It is this representation that is used for the evaluation phase, which we will turn to next.

6.2 The Evaluation Phase

At the end of the framing phase, the investor has interpreted available information and formed a new and subjectively adapted version of the prospect. This version is then transferred to the evaluation phase where all of the adapted prospects are evaluated based on their subjective value, and the prospect providing the highest value is chosen.

The investor is assumed to maximize his prospective value, and as in utility theory he employs a value function to estimate the utility from the available prospect. There are however two distinct differences between the two approaches.

Firstly, where utility theory utilizes objective probabilities of an outcome, prospect theory uses subjective decision weights. Secondly, investors with prospect theory preferences base their decisions on the value of relative outcomes – meaning gains or losses – rather than estimating the utility of final states of wealth (Tversky & Kahneman, 1981).

Consequently, the modeling of the investor's value function and his subjective decisions weights are crucial elements in the formulation of prospect theory. The following sections will address these aspects.

6.3 The Value Function

Where standard utility theory employs the objective probability of a particular event, prospect theory multiplies each possible outcome with a *decision weight.* The decision weight captures the *individually perceived* probabilities, and thereby quantifies the perceived utility impact of a potential event. Consequently, the overall value function under prospect theory is defined as follows:

$$(6.1) \qquad\qquad V(x,p;y,q) = \pi(p)v(x) + \pi(q)v(y)$$

Where x and y denote the possible relative outcomes, and p and q represent their objective probabilities. V denotes the value of prospect, while v represents value of each of the possible outcomes. The value of an outcome that has no probability of occurring is zero; $v(0) = 0$. Also, for the extreme probabilities of 0% and 100%, the decision weights will be 0 and 1 respectively; $\pi(0) = 0$ and $\pi(1) = 1$.

The combination of the value function v and the subjective decision weights π make up the overall *prospect value*. In other words, the value of the prospect is defined as the sum of the objectively expected value multiplied by a subjective decision weight for each possible outcome within that prospect.

Following their experimental contributions, Kahneman & Tversky (1979) suggested that the shape of the investor's value function must possess three characteristics:

i. Utility is derived based on the investor's initial reference point
ii. Investors are generally loss averse
iii. Investors display diminishing sensitivity

In order to understand the implications for the prospective value function, let us first consider the characteristics of the standard utility function as it is depicted in Figure 6.1:

**Figure 6.1: Standard utility function**
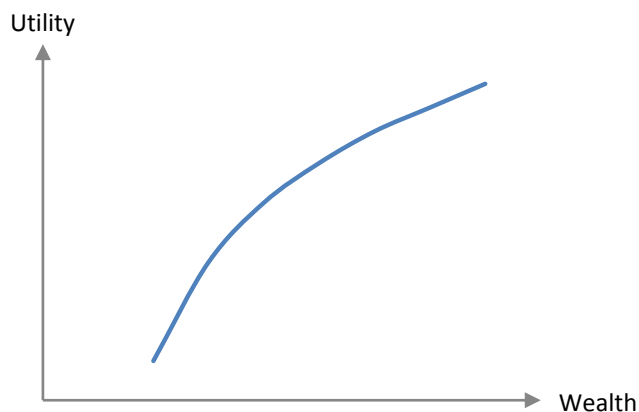
Under traditional utility theory, the investor will evaluate a prospect based on final states. We therefore see value (or utility) as a function of overall wealth. In other words, the choice amongst alternative

48

investment opportunities will rely solely on their effect on the investor's final aggregate level of wealth. As we mentioned earlier, standard utility theory assumes that agents are always risk averse. Consequently, we observe a concave value function where the coefficient of relative risk aversion (CRRA) represents the curvature of the function.

Next consider Figure 6.2, which offers an illustration of a hypothetical prospective value function that encompasses the effects of reference dependence, diminishing sensitivity, and loss aversion.

Figure 6.2: Prospective value function
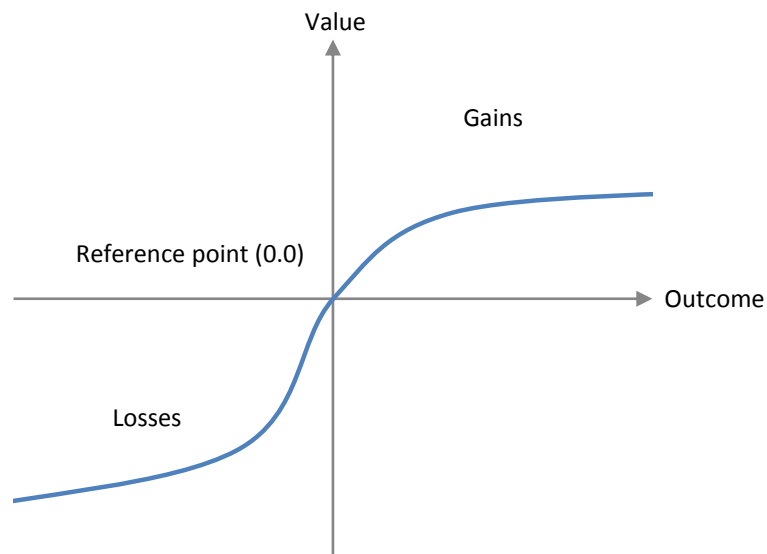
Reference dependence entails that value is a function of a reference point (0.0), meaning that individuals either obtain a gain or a loss from a given prospect. Within the reference dependent model, value is treated as a divergence from the investor's initial asset position, or the *status quo* position (Kahneman & Tversky, Prospect Theory: An Analysis of Decision under Risk, 1979).

Consequently, Figure 6.2 shows value as a function of the outcome instead of final wealth as we saw in Figure 6.1. It is important to note the differing x-axes, because it implies that utility is derived from the magnitude of the positive or negative change from the reference point, and not from the absolute level of wealth obtained. We will return to some of the cognitive manifestations of reference dependence in chapter 7.

Loss aversion is the second characteristic of the value function, and as illustrated earlier, it implies that provided a loss and a gain of equal size, the loss will yield a higher disutility than the positive utility derived from a similar gain.

Tversky & Kahneman (1979) found that most individuals are disinclined towards symmetric gambles $(x, 0.5; -x, 0.5)$ and that the aversion towards such gambles increases even further as the value (x) increases. So, assuming that $x > y \geq 0$, then $(y, 0.5; -y, 0.5)$ is preferred over $(x, 0.5; -x, 0.5)$. Referring back to the original value function (6.1), were therefore see that:

$$V(x, 0.5; -x, 0.5) < V(y, 0.5; -y, 0.5)$$

$$\Leftrightarrow \pi(0.5)v(x) + \pi(0.5)v(-x) < \pi(0.5)v(y) + \pi(0.5)v(-y)$$

As $\pi(0.5)$ is a common factor for all four terms, we remove it from the equation altogether, which leaves us with:

$$v(x) + v(-x) < v(y) + v(-y)$$

$$\Leftrightarrow v(x) - v(y) < v(-y) - v(-x)$$

Setting $y = 0$, we see that $v(x) < -v(-x)$, meaning that the value of a positive outcome $v(x)$ is smaller than the value of a negative outcome $v(-x)$ for any given level of x. This supports the theory that the value function is steeper in the area of losses than in the domain of gains.

As observed in Figure 6.2 the prospective value function has a kink at the origin, which represents this difference between a loss and a gain. Essentially, the ratio of the two slopes at the origin provides a measure of loss aversion, an important input for our subsequent analyses. Empirical estimates of loss aversion fall around 2, meaning that the disutility of a loss is twice as great as the utility of an equivalent gain (Tversky & Kahneman, 1991; Kahneman, Knetsch, & Thaler, 1991).

Diminishing sensitivity is the final characteristic of the value function, and it basically refers to the fact that the difference in value between a gain of 100 and 200 was found to be greater than the difference between a gain of 1,100 and a gain of 1,200 (Tversky & Kahneman, 2000). The inference is that the value associated with a gain increases at a decreasing rate, i.e. the marginal utility from gains decreases with the absolute size of the gain. This causes the value function to display concavity in the domain of gains. Likewise, individuals become less aggravated by their marginal loss as it increases, i.e. marginal disutility from losses decreases with the absolute size of the loss, thereby creating convexity in the domain of losses.

In order to understand the intuition behind diminishing sensitivity and its effect on the graphical interpretation of the prospective value function, consider the following set of choice problems devised by Tversky & Kahneman (1979):

Problem 8:			
Option	Outcome	Probability	
A	6,000	0.25	[18]
	0	0.75	
B	4,000	0.25	[82]
	2,000	0.25	
	0	0.50	
N = 68			

Problem 9:			
Option	Outcome	Probability	
C	-6,000	0.25	[70]
	0	0.75	
D	-4,000	0.25	[30]
	-2,000	0.25	
	0	0.50	
N = 64			

Source: Kahneman & Tversky (1979) p. 278.

By employing equation (6.1) Tversky & Kahneman were able to make inferences about the pattern of prospective value. The respondents' preferences in Problem 8 implied the following relationship:

$$A < B \iff V(6{,}000, 0.25) < V(4{,}000, 0.25; 2{,}000, 0.25)$$

$$\iff \pi(0.25)v(6{,}000) < \pi(0.25) * [v(4{,}000) + v(2{,}000)]$$

$$\iff v(6{,}000) < v(4{,}000) + v(2{,}000)$$

This illustrates why the value function demonstrates *concavity* in the domain of gains, as the potential gain of 6,000 provides less utility for the investor than the combination of two separate but smaller gains of 4,000 and 2,000 respectively.

Meanwhile, the reaction to Problem 9 implied the reverse relationship:

$$C > D \iff V(-6{,}000, 0.25) > V(-4{,}000, 0.25; -2{,}000, 0.25)$$

$$\iff \pi(0.25)v(-6{,}000) > \pi(0.25) * [v(-4{,}000) + v(-2{,}000)]$$

$$\iff v(-6{,}000) > v(-4{,}000) + v(-2{,}000)$$

The inequality sign has reversed, indicating that respondents become risk seeking when faced with negative prospects, and the value function must therefore display *convexity* in the domain of losses.

To summarize, the value function is: i) defined on the grounds of deviations from the reference point, ii) steeper for losses than for gains, and iii) commonly concave in domain of gains and convex in domain of losses.

Collectively, these three features lead to a value function with an asymmetric S-shape as the one illustrated in Figure 6.2. Due to the use of decision weights, the actual scaling is more difficult to estimate than in classic utility theory, however a large number of empirical studies have found that most of the observed utility functions were indeed convex for losses and concave for gains, while a very small number of individuals exhibited risk aversion for both losses and gains. In addition, the studies found that the functions were indeed considerably steeper for losses than for gains (Barnes & Reinmuth, 1976; Green, 1963; Swalm, 1966).

Having discussed the characteristics of the prospective value function, let us now turn to the nature of the subjective decision weights, and the inherent weighting function.

6.4 The Weighting Function, π

Equally important to the prospective value is the actual weighting function, which transforms objective probabilities into subjective perceptions, or more formally the original probabilities (p) into decision weights $\pi(p)$. Recall from equation (6.1) that these weights are then assigned by the investor to the given outcomes $v(x)$, which are subsequently summed to obtain the perceived prospect value.

The basic difference between decision weights and probabilities is that the latter measures the objective likelihood of a given number of outcomes, while the decision weights measure the *impact* of these events on the desirability of the prospects at hand, i.e. how much a given outcome will affect the individual decision maker. Consequently, the decision weights are not themselves measures of probability, but rather *emphasis parameters* derived from the original probabilities. This is an important distinction.

As decision weights are not a probability measure, they are not to be interpreted as measures of degree or belief. In some cases, i.e. basic intuitive experiments, decision weights can be expressed as functions of the stated probabilities, but more generally speaking the weights will be influenced by additional factors. One example is the certainty effect that we observed in the last chapter. Because agents are strongly preferential towards certain outcomes when they are positive, the decision weights are irrational and therefore appear non-linear (Barberis & Thaler, A survey of behavioral finance, 2002).

To understand the patterns of the weighting function, we must return to the findings from the previous section. As with the value function, three essential characteristics also govern the weighting function:

i.   In the domain of gains, investors become risk averse.

ii. In the realm of losses, investors are risk seeking.

iii. Investors put too much emphasis on small probabilities.

In an empirical study, Tversky & Kahneman (2000) observed that people's preference order shifts as the multitude, as well as the nature, of prospects change. For positive prospects with low probabilities, 78% of respondents were risk seeking, while the equivalent losses urged 80% to become risk averse. The mirrored effect was seen for higher probabilities, where 88% of the participants displayed risk aversion when faced with positive prospects, and 87% were risk seeking for negative prospects.

The implication of the results is that the weighting function must encompass what Tversky & Kahneman (2000) referred to as the *fourfold pattern of risk attitudes*; for small probabilities investors are risk seeking in the domain of gains and risk averse in the domain of losses. For higher probabilities the pattern is reversed. This leads to four distinct patterns in the description of an investor's risk attitude. The pattern is a key empirical generalization of decision making under risk, and it has been validated through several other experimental contributions (Cohen, Jaffray, & Said, 1987; Wehrung, 1989).

We will illustrate the graphical implications of the fourfold pattern of risk attitudes in the next section, which addresses the technical formulation of prospect theory, namely the explicit functional forms of the weighting- and the value function.

6.5 The technical functions of prospect theory

In traditional utility theory, the assumption of preference homogeneity implies that the investor has a constant relative risk aversion (CRRA), which leads to a one-part power function (see Figure 6.1). Based on their empirical evidence however, Tversky & Kahneman (1992) found that value is in fact derived through a two-part power function, which differentiates between positive and negative outcomes. So the value function that assigns a value v to an outcome x was defined as follows:

$$(\mathbf{6.2}) \qquad v(x) = \begin{cases} x^a & if \ x \geq 0 \\ -\lambda(-x)^\beta & if \ x < 0 \end{cases}$$

Note that x represents the relative outcome based on the initial reference point, or status quo, meaning that it is a measures of the difference in wealth with respect to the last time wealth was measured. The

exponents α and β[20] capture the curvature of the value function for positive and negative outcomes respectively. As investors are observed to display diminishing sensitivity, both these factors are smaller than 1, because the function is marginally decreasing. Finally, λ denotes the aggravation parameter, or the coefficient of loss aversion, which is essentially the difference in the slopes of the positive and the negative part of the value function.

Through nonlinear regressions of their experimental observations, Tversky & Kahneman (1992) estimated λ to be 2.25, while the median exponents, α and β, were projected at a value of 0.88.

Having defined the value function for the individual outcomes, we now turn to the technical formulation of the weighting function, which is slightly more complex.

The weighting function used in original prospect theory (Kahneman & Tversky, 1979) transformed each probability separately, meaning that π was defined purely as a function of p. In 1992, the authors extended their original contribution and introduced cumulative prospect theory. The extension was a response to the works of several scholars (Quiggin, 1982; Schmeidler, 1989; Yaari, 1987; Weymark, 1981) who had advanced the notion of employing cumulative rather than individual probabilities. Consequently, the modified version of prospect theory incorporated a new rank-dependent representation.

In the updated version, the weighting function relied on the entire cumulative distribution. In other words, the investor's perception of a given outcome will depend on its rank within the entire range of possible outcomes.

In order to define the cumulative function of a prospect, all the outcomes must be ranked in increasing order to illustrate the distribution of the entire gamble. The position - or the rank - of the individual outcome is denoted i and lies between $-m$ and n.

The prospect itself is denoted (x_i, p_i), with x representing the outcomes and p their respective probabilities. The subscript does not only signify the rank, but also the nature of the prospect, namely whether it is positive or negative. Outcomes that lie within the domain $0 \leq i \leq n$ are considered strictly positive, while strictly negative outcomes are in the domain $-m \leq i \leq 0$.

---

[20] Note that within the framework of Benartzi & Thaler, α and β refer to the curvature of the prospective value function for positive and negative outcomes respectively. They should be confused with the α and β applied by Mehra & Prescott.

As decision makers have different preferences for gains than for losses, the transformation from probability to decision weight must follow different routines for gains and losses respectively. Therefore, we introduce the capacities w^+ for positive outcomes and w^- for negative outcomes. These capacities are defined by the entire probability distribution of the prospect, meaning the sum of all the individual probabilities, and they are used to facilitate the transformation from probability p to decision weight π.

The decision weight associated with each outcome may be interpreted as the marginal contribution of that particular outcome defined in terms of the capacities. For positive outcomes, the decision weight π_i^+, is therefore the difference between the capacities of the events where the outcome is *at least* as good as x_i, and events the outcome is *strictly* better than x_i. Correspondingly, for negative outcomes, π_i^- is equal to the range of outcomes that are at least as bad as x_i minus the range of outcomes that are strictly worse than x_i (Tversky & Kahneman, 2000).

For positive prospects $(0 \leq i \leq n - 1)$ we therefore see the following function:

$$(\mathbf{6.3}) \qquad\qquad \pi_i^+ = w^+(p_i + \ldots + p_n) - w^+(p_{i+1} + \ldots + p_n)$$

Correspondingly, when prospects are negative $(1 - m \leq i \leq 0)$ the weighting function is:

$$(\mathbf{6.4}) \qquad\qquad \pi_i^- = w^-(p_{-m} + \ldots + p_i) - w^-(p_{-m} + \ldots + p_{i-1})$$

Where w^+ and w^- are conditioned by: $w^+(0) = 0$, $w^-(0) = 0$ and $w^+(1) = 1$, $w^-(1) = 1$ meaning that they are both strictly increasing functions.

To sum up, the decision weight represents the marginal contribution of the ith outcome, x_i, to the overall prospect. At the outer margins – represented by the most extreme values of the ranked prospects – the decision weight and the capacity will coincide, as no outcomes are better than x_n or worse than x_{-m} respectively.

As the decision weight is an increasing function of the probabilities, the value of π will increase along with the probability of a particular outcome, and it holds that at the margins $\pi(0) = 0$ and $\pi(1) = 1$, meaning that an outcome with zero probability will have zero weight, and that a sure outcome will also be weighted as such.

The next step is then to find the capacities to apply for the two domains (gains and losses). Recall that capacities were defined as cumulative functions of the pure probabilities tied to each outcome. Tversky

& Kahneman (2000) consequently used the following one-parameter approximations to define the capacities for the respective domains of gains and losses:

$$(6.5) \qquad w^+(p_i) = \frac{p_i^{\gamma}}{\left(p_i^{\gamma} + (1 - p_i)^{\gamma}\right)^{\frac{1}{\gamma}}} \qquad\qquad w^-(p_i) = \frac{p_i^{\delta}}{(p_i^{\delta} + (1 - p_i)^{\delta})^{1/\delta}}$$

Where γ and δ represent the two single factors that were used to model the behavior of the weighting functions. The functions were tested against a series of aggregate and individual data, for probabilities within the range of 5% - 95%, and the median values of γ and δ obtained through regression were 0.61 and 0.69 respectively.

In 1994, Camerer & Ho tested equation (6.5) through a number of studies on risky choices, and they estimated γ to be 0.56, which was relatively close to Tversky & Kahneman's value of 0.61 (Camerer & Ho, 1994).

Figure 6.3 is an adaption from Tversky & Kahneman (2000, p. 60) and it shows a plot of the weighting function for gains (w^+) and for losses (w^-) based on the median estimates of γ and δ in equation (6.5):

Figure 6.3: Weighting functions
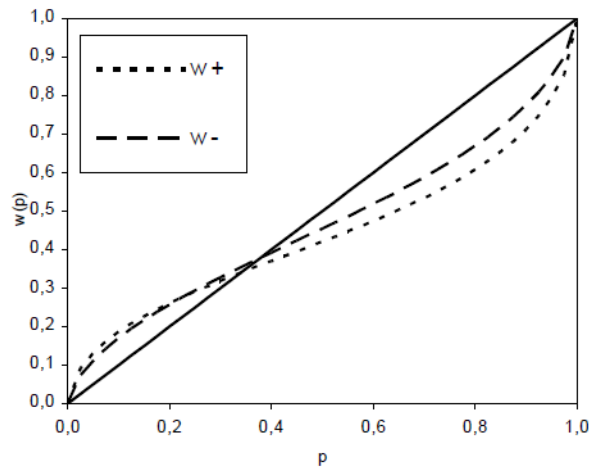
Source: Tversky & Kahneman (1992) p. 313.

As we can see, the functions follow the fourfold pattern of risk attitude, which dictates that people put too much weight on low probabilities, whilst underestimating high probabilities. In other words, the decision weights exceed the corresponding probabilities for the low range of p, thereby creating a concave pattern. For probabilities in the mid to high range, the decision weights understate the actual likelihood of an event, which creates a convex pattern.

While the weighting functions for gains (w^+) and for losses (w^-) are relatively similar, we observe that positive prospects yielded a more curved function compared to the negative prospects (i.e. $\gamma < \delta$). This shows us that risk aversion for gains was more pronounced than risk seeking for losses, for probabilities within the mid to high range.

Having defined the two-part value function (5.4) as well as the cumulative weighting function (5.5), we can now term prospective value V as the sum of prospects $v(x_i)$ timed their decision weights π_i:

$$(\mathbf{6.6}) \qquad V = \sum_{i=-m}^{n} \pi_i v(x_i)$$

To illustrate the practical implications of the functions introduced in this section, let us consider a simple numerical example:

An investor is faced with a prospect that is made up of two different potential outcomes. He has a 10% chance of gaining \$100 (outcome A) and a 90% chance of gaining \$10 (outcome B). The first step is to establish how much value the investor would get from each of the two outcomes, and for this purpose he employs function (6.2) for positive outcomes: $v(x) = x^a$. The value of outcome A is therefore $100^{0.88} = 57.54$ and the value of outcome B is $10^{0.88} = 7.59$. Due to the diminishing sensitivity characteristic of the investor's value function, the perceived value of the outcomes is lower than their actual monetary value.

The next step is to figure out how much each of the outcomes contributes to the investor's overall perception of the prospect, or in other words how much weight he attributes to each of them respectively. Both outcomes are positive, meaning that they each represent a gain compared to his initial asset position. Recall that the decision weight of a positive outcome x, is the difference between the capacities of the outcomes that *at least* as good as x, and the outcomes that are *strictly* better than x.

As the gain of \$100 is the best possible outcome, no other outcomes are strictly better than it. Therefore, the inherent decision weight for outcome A is $\pi^+ = w^+(p)$. To find the capacity and thereby the decision weight, he applies equation (6.5):

$$\pi_A = w^+(0.10) = \frac{0.10^{0.61}}{(0.10^{0.61} + (1 - 0.10)^{0.61})^{\frac{1}{0.61}}} = 0.186 \approx 18.6\%$$

As outcome B is the worst possible outcome, the probability for seeing an outcome that is at least as good is 100%. We know that $w^+(1) = 1$, and because there only are two possible outcomes, then the

capacity for outcome A represents all the outcomes that are strictly better than outcome B. So the decision weight is found via equation (6.3):

$$\pi_A = w^+(1) - w^+(0.10) = 1 - 0.186 = 0.814 \approx 81.4\%$$

As dictated by the fourfold pattern of risk, the investor overstates the small probability of outcome A, whilst understating the large probability of outcome B. Finally equation (6.6) helps us to find the prospect's overall value by combining the values and the decision weights of the respective outcomes:

$$V = (18.6\% * 57.54) + (81.4\% * 7.59) = 16.89$$

All in all the investor's prospectively utility of a prospect that offers outcome A and B is $16.89.

These are the technical formulations of the prospect theory preferences that are employed in Benartzi & Thaler's (1995) model for myopic loss aversion. In the following chapter, we will the cover the behavioral elements of the myopic loss aversion model, and then move on to our own empirical examination of the equity premium puzzle in Sweden, as well as the ability of myopic loss aversion to provide a solution to such a puzzle.

6.6 Chapter summary

We have looked at the alternative preference structure, which was introduced as part of prospect theory by Tversky & Kahneman (1979; 1992). The authors argued that investors make decisions about prospects in two phases; a framing phase, and an evaluation phase. Within these phases, prospects are initially modified into more simple representations, and subsequently evaluated based on their overall characteristics, e.g. the return on an investment is either classified as a gain or a loss, and it is then evaluated accordingly.

The descriptive theory was based on a variety of experimental studies, and it dictated that the overall prospective utility of any given prospect, is given by two functions, namely a value function and a weighting function. We saw how the prospective value function is governed by three main characteristics; reference dependence, loss aversion, and diminishing sensitivity. These features cause the value function to be flat and concave in the domain of gains, and steep and convex in the domain of losses, which gives it an asymmetric S-shape.

Furthermore, the weighting function is similarly governed by what is referred to as the fourfold-pattern of risk. It refers to the fact that individuals are observed to be risk seeking in the domain of gains and risk averse in the domain of losses for small probabilities, and for higher probabilities the pattern is reversed.

Consequently, people will tend to modify probabilities according to their own subjective perception, via so-called emphasis parameters that are known as decision weights.

In the following chapter, we look further into the behavioral aspects that drive our main analysis of solving the equity premium puzzle, namely myopic loss aversion.

7. Myopic Loss Aversion

In its essence, myopic loss aversion is defined within the realm of behavioral finance, which revolves around concepts from the psychology of decision making. Assumptions regarding rational behavior are relaxed and economic theory and observations about human psychology are integrated. One of the concepts that were formulated based on such observations is *loss aversion*, which we have illustrated through experimental evidence in earlier chapters. Another is *mental accounting*, which captures the way people organize their thoughts, particularly in relation to decision making. The following sections will address the significance of these two concepts, as well as the effect achieved by combining them in pursuit of a solution to the equity premium puzzle.

7.1 Loss aversion

As we have seen in previous chapters, people exhibit loss aversion under experimental settings. To understand the behavioral foundation of loss aversion, we briefly turn to two manifestations of this asymmetry of value perception, namely the *endowment effect* and the *status quo bias*.

One effect of loss aversion is that the negative utility that comes from losing a valued good is higher than the utility gain associated with receiving it. Thaler (1980) dubbed this phenomenon the *endowment effect*, because the value of a good changes once an individual owns it, or once it is incorporated into that person's endowment. Furthermore, empirical studies have found that individuals appear strongly inclined towards retaining the *status quo* when faced with decision problems where they had already been endowed with a good.

Kahneman, Knetsch & Thaler (1991) performed a series of classroom experiments to document the endowment effect, as well as the status quo bias. Let us consider one of those experiments:

A collection of students was divided into three groups; sellers, buyers, and choosers. The first group – the *sellers* - was presented with a decorated mug with a retail value of $5, with the instruction that they could either (x) sell the mug at each of a series of prices ranging from $0.25 to $9.25 or (y) keep it. The second group – the *buyers* – was offered to buy a mug at the same set of prices. Finally, the *choosers* were given the option of either receiving the mug or the cash for each price within the same range. The experiment created median reservations prices of $7.12 for the sellers, $3.12 for the choosers, and $2.87 for the buyers.

Despite objectively being in the same situation as the sellers, the choosers interestingly acted more like buyers. The reason is not the difference in value, but the fact that their *reference states* differ. Recall from our discussion on the prospective value function, that utility is established based on an initial reference point. This means that people will define value based on the potential divergence their initial asset position – or their status quo position.

Consider Figure 7.1, which illustrates the initial state and the potential reference points for the choice between (x) to receive cash and (y) the mug:

Figure 7.1: Multiple reference points for choice between x and y
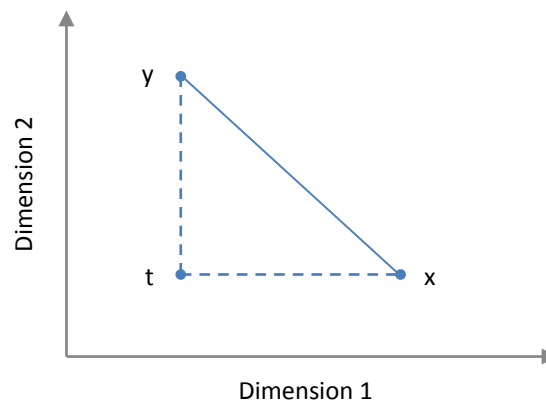
Source: Tversky & Kahneman (1991) p. 1042.

The choosers' reference state is (t) as they have the choice between option (x) and (y), that both dominate (t) on dimension 1 and 2 respectively. The sellers are already in possession of the mug, so their reference point is (y). From there they may choose to sell the good, and thereby travel to (x). Because this action entails a loss in terms of dimension 2 – the mug is given up – loss aversion kicks in.

In this experiment, the endowment effect was instantaneously created by giving an individual property rights over a good. The effect caused the sellers' reservation price to be slightly more than twice that of the choosers' who in turn acted much more like the buyers (Tversky & Kahneman, 1991). Furthermore, this experiment also supports a loss aversion parameter of 2.25, as used in the cumulative version of prospect theory (Tversky & Kahneman, 1992).

In a second experiment, two groups of students were asked to fill out a questionnaire. As compensation, the first group was presented with a mug, while the second group received a chocolate bar. At the end of the experiment, the students were shown the alternative gift, and given the choice of switching.

Referring back to Figure 7.1, the dimensions represent the two alternative gifts, and the students' reference states are represented by (x) and (y). All the participants faced a loss within one of the dimensions, so loss aversion caused a bias amongst the students towards maintaining the status quo, and as it turned out 90% of the participants retained their gift.

To sum up, we have illustrated two manifestations of the asymmetry between gains and losses that is essentially brought on by loss aversion. We now turn to another concept within behavioral finance, namely mental accounting.

7.2 Mental accounting

Mental accounting is defined as the: "*set of cognitive operations used by individuals and households to organize, evaluate, and keep track of financial activities*" (Thaler, 1999, p. 183). The implicit coding that occurs when individuals evaluate these transactions is highly subjective, and it therefore drives attitudes towards risk and consequently investors' decision making.

In essence, mental accounting is applied during the framing phase, and Tversky & Kahneman define a mental account as a *frame for evaluation* (1981). Recall from section 6.1, that the framing phase is where prospects are coded by the decision maker, and transformed into an accommodating representation. Through this coding, the individual will assign the prospect to a mental account, and subsequently evaluate it within the confines of this particular account.

Tying back to the experimental evidence from earlier, the difference with which gains and losses are treated, stem from the fact that they are effectively posted to different mental accounts – which means that they are being evaluated as separate events. In other words, the available prospects are grouped into sets, and observations within the individual sets are considered in conjunction with one another. This grouping is known as *choice bracketin*g, and it represents a cornerstone within mental accounting.

7.2.1 Choice bracketing

When an individual is faced with a large number of choices, he may choose to evaluate the consequences of all of them taken together. In doing so, the individual is applying a *broad bracketing*, because he assesses the combined effect of all the decisions he needs to make. This behavior will generally serve to maximize the total potential utility, because prospects are being chosen based on their effect on overall wealth. This is the decision process assumed in the utility-based model of Mehra & Prescott (1985). If on the other hand, an individual applies *narrow bracketing* he only evaluates a smaller

selection of his options at a time. His choices may hence be suboptimal because he is missing the big picture effect of his decisions (Read, Loewenstein, & Rabin, 1999).

Consider the decision on whether or not to have dessert after a meal. If the choice is made in isolation – meaning one meal at a time – then the pleasure of enjoying dessert could easily outweigh the minor effect on a person's physical health and appearance. If on the other hand the choices of 1,095 desserts (3 meals a day for a year) are considered in conjunction, then the health issues become much more severe, and may just outweigh the aggregated pleasure of enjoying something sweet after every meal. If broad bracketing was applied, then the combined effects of the choices would be taken into consideration, and the individual may have rejected the option all together.

Choice bracketing also refers to the *frequency* with which the portfolio of choices is evaluated. Frequent evaluations that occur due to the shortsightedness – or myopia - of an individual also fall under the term narrow bracketing (Read, Loewenstein, & Rabin, 1999). Thaler however refers to the phenomenon as *narrow framing* (1999), because from his perspective it relates to the framing phase of prospect theory. As our study is based on prospect theory, we will use the latter term throughout the remainder of this paper.

Under traditional utility theory, individuals can only apply broad framing, because transactions must be treated in conjunction with one another to maximize overall utility. In other words, the theory assumes that investors will evaluate a new prospect based on how it contributes to the overall risk levels of the existing portfolio. Empirical observations found in experimental settings however, have indicated that people will sometimes evaluate new gambles in isolation, separately from their other risks, meaning that they effectively apply narrow framing.

An illustration of narrow framing can be seen in one of the choice problems from Tversky & Kahneman (1981) that we presented in section 5.2.3. Recall that in Problem 5, individuals were asked to make a pair of coexisting decisions – one was a favorable prospect and the other was an unfavorable prospect. Because people failed to evaluate the combined effect of the two decisions, they ultimately chose the combination that offered to lowest overall utility.

The fact that agents act as if their utility is derived directly from the outcome of a single gamble, despite it being part of a portfolio of gambles that collectively determine the overall wealth level, is therefore a direct contradiction of the premise of utility theory. Consequently, this cognitive pattern is included in the set of alternative preferences structures presented by the model of myopic loss aversion.

Tying back to the idea of reference dependence, we may note that whenever the investor assesses his portfolio, he effectively evaluates the state of his investment compared to his previous reference point – meaning the value of the portfolio from the last time he evaluated it. Once this new evaluation has occurred, the investor will essentially recalibrate his status quo to the new value of his investment. Consequently, Benartzi & Thaler (1995) argued that the risk attitude of loss averse investors depend on the frequency with which they reset their references point, meaning how often they "*count their money*" (Thaler, 1999, p. 200).

To sum up, the implications for the remainder of this paper is that broad framing by investors will ultimately lead to higher utility, as more investments are evaluated collectively over a longer period of time. Investors are however typically observed to be myopic, and their narrow framing leads them to reject strategies that are superior over a longer horizon, due to frequent evaluations. This offers an explanation to the seemingly irrational behavior that leads to the equity premium puzzle.

7.3 Combining loss aversion and myopia

In 1995, Benartzi & Thaler proposed that a combination of loss aversion and mental accounting – more specifically myopia – could potentially solve the equity premium puzzle. They introduced a descriptive model that evolved around the tendency amongst investors to be more sensitive to losses and to apply narrow evaluations frames.

The foundation for myopic loss aversion was laid by Samuelson in 1963. In a now famous encounter the author asked a colleague whether he would accept a bet over a coin toss that offered a 50% chance to win $200 and a 50% chance to lose $100. The colleague turned down the bet by arguing that: "*I won't bet because I would feel the $100 loss more than the $200 gain. But I'll take you on if you promise to let me make 100 such bets*" (Samuelson, 1963, p. 108). Samuelson commented that his colleague's behavior was irrational, by reasoning that the willingness to accept multiple wagers whilst refusing a single bet was inconsistent with utility maximization under traditional expected utility theory.

The colleague's rationale for refusing the bet is an illustration of the intuition behind the concept of loss aversion. To capture this notion, consider this simple utility function (Benartzi & Thaler, 1995):

$$U(x) = \begin{cases} 2.5x, & x < 0 \\ x, & x \geq 0 \end{cases}$$

where x represents the change in wealth relative to the status quo. A negative outcome hurts him 2.5 times more than the benefit he gets from a positive outcome. Assuming that Samuelson's colleague had this utility function, he would decline a single bet as it would provide him with negative utility:

$$[0.5 * \$200] + [0.5 * 2.5 * -\$100] = -25$$

Given that each bet would be evaluated separately, they would all yield negative utility and should be rejected. If on the other hand, the colleague would not have to watch the bets being played out, he could accept two bets as these would now supply a positive utility:

$$[0.25 * \$400] + [0.50 * \$100] + [0.25 * 2.5 * -\$200] = 25$$

This example illustrates how loss averse individuals are willing to take on more risk, if they evaluate their performance less frequently (i.e. display non-myopic behavior), and this is exactly the argument put forth by Benartzi & Thaler.

In essence, the authors sought to illustrate the drivers of demand in equity markets by employing prospect theory preferences in terms of decision making under uncertainty. They argued that an investor, who can choose between a risky asset that is expected to provide a yearly return of 7% with a standard deviation of 20%, and a risk-free asset offering 1%, will base his decision on the inherent evaluation horizon. The risky asset will appear more attractive the longer the investor intends to hold on to it, as long as he does not evaluate the investment too frequently. The high volatility will cause the risky asset to have a negative return from time to time, but this will *only* be observed by the investor if he evaluates his portfolio sufficiently often. If he refrains from doing so he will merely observe the higher average expected return of the risky asset, and it will be more attractive than its risk-free alternative. This is equivalent to the willingness of Samuelson's colleague to accept the bet if he did not have to observe the individual outcomes of the coin toss.

When an investor narrows the framing of his investment decisions, the frequency with which he assesses his investments increases. If an investor is highly myopic, he may evaluate his portfolio on a daily basis, and there will be many days where the return on his equity investment is lower than the return he gets on his savings account or his bond investment. In other words, the likelihood of observing a negative return will increase along with the frequency of evaluations, leading to suboptimal investment decisions (Thaler, Tversky, Kahneman, & Schwartz, 1997).

So, if we accept the premise that losses hurt more than gains, myopic investors will obtain less utility from owning stocks compared to non-myopic investors, despite having the same level of risk aversion. This is because the aggravation parameter λ from the prospective value function (6.2) is applied much more often under narrow framing. If investors were to apply broader brackets, stocks would appear much more appealing, thereby driving down the spread between returns on stocks relative to bonds (Read, Loewenstein, & Rabin, 1999). Consequently, Benartzi & Thaler (1995) argued that the use of prospect theory must be accompanied by a specification of the frequency with which returns are evaluated, namely an *evaluation horizon*.

Note that while the evaluation horizon is not the same as the planning horizon, or the investment horizon (retirement in 30 years), investors act is if these are the same. Consequently, estimating the investors' evaluation period becomes the same as finding their implied investment horizons.

In summary, Benartzi & Thaler (1995) argued that the equity premium puzzle and its magnitude may be explained by investors being myopic and loss averse – leading to shorter investment horizons than those stipulated by classic expected utility theory.

7.4 Chapter summary

Having introduced the prospective preference structure in the previous chapter, we moved on to the actual concept of myopic loss aversion, as well as its theoretical foundation within behavioral finance, consisting of loss aversion and mental accounting. Two effects of loss aversion were discussed, in the form of the endowment effect and the status quo bias, which both capture the reference dependence of individuals. Furthermore, we showed that experimental studies indicate that people are hurt twice as much by as loss, as they are pleased by a gain.

The second component of myopic loss aversion, mental accounting, mainly revolves around the concept of choice bracketing, as myopic – or shortsighted - individuals are thought to apply narrow bracketing in their investment decisions. This leads to shorter evaluation horizons, which in turn makes the investor more risk averse, as the chance of him observing a loss on his investment increases. This is particularly the case for volatile stock investments, and myopic investors are thereby assumed to require a higher premium for investing in equity compared to a stable risk-free alternative.

The following chapter will address the empirical results of Benartzi & Thaler (1995), as they attempted to test the explanatory power of myopic loss aversion as a solution to the equity premium puzzle in the US.

8. Myopic loss aversion empirically

8.1 Myopic loss aversion in the US

The study of Benartzi & Thaler (1995) was essentially based on three questions:

i.   How often would an investor with prospect theory preferences have to evaluate his portfolio in order to be indifferent between the historical distribution of returns on stocks and risk-free assets?

ii. Taking this evaluation period as given, what combination of stocks and risk-free assets would maximize his prospective utility?

iii. What would be the effect on the equity premium, if the evaluation period increased?

The questions represented an attempt to model the effects of myopic loss aversion amongst investors. The utility from holding a financial asset comes from the changes in that asset's value. The loss averse investor will therefore require a higher compensation for holding volatile assets, because they are more likely to cause him a negative return. Furthermore, if the investor evaluates his portfolio often, the short-run fluctuations in asset returns will effectively lower his prospective utility of holding that asset.

Benartzi & Thaler (1995) hypothesized that the high equity premium observed in the US could be explained by investors being loss averse and shortsighted. The authors argued that the premium required for holding stocks was directly related to how often the investor evaluated his portfolio. Consequently, the first step of Benartzi & Thaler's model was to derive an implied evaluation period based on historical monthly returns on stocks, bonds, and bills in the US for the period 1926-1990.

8.1.1 Implied evaluation horizon

In effect, Benartzi & Thaler (1995) computed the prospective utility of holding a pure-stock portfolio and compared it to holding a portfolio of risk-free assets for evaluation periods starting at one month and then increasing one month at a time. The implied evaluation period was found by locating the equilibrium horizon where an investor with prospect theory preferences would have been indifferent between holding stocks or risk-free assets based on historical data.

The implied evaluation horizon was calculated for four different combinations of returns series. The CRSP[21] stock index was compared both with T-bill returns, and with five-year bond returns, and both of these comparisons were done in nominal as well as in real terms.

While the authors reported on the results of all four combinations for the sake of illustrating robustness, they argued that the combination of nominal return series with five-year bond returns representing the risk-free asset should be given the most weight.

As for T-bills versus bonds, Benartzi & Thaler (1995) argued that for long-term private investors seeking to invest in a risk-free asset, bonds represented the closest substitute. We on the other hand, have chosen to employ short-term government notes as our proxy for the risk-free asset in our main analysis. We do so to maintain the link between our analysis of the equity premium puzzle using the Mehra-Prescott model (1985), the Kocherlakota method (1996), and myopic loss aversion in Sweden. To ensure that our methods are robust however, we do run our model with 10-year government bonds as the risk-free asset. We will return to this discussion in our results section.

Benartzi & Thaler's argument for preferring nominal returns to real ones was twofold: Firstly, they referred to the fact that returns are predominately reported in nominal terms, e.g. in annual reports. As the model was purely descriptive, investor behavior was more likely to be driven by nominal accounts of their returns.

Secondly, Benartzi & Thaler (1995) found that historically, the prospective utility of holding T-bills would be negative if the average investor had based his evaluations on real returns. As any rational individual would refrain from holding an investment that he perceived to yield negative utility, the authors argued that evaluations must have been based on nominal returns.

Table 8.1 presents the results of Benartzi & Thaler (1995) for all four combinations:

Table 8.1: Evaluation horizons, months

	Nominal	Real
Stocks versus 5-year bonds	13	10 - 11
Stocks versus T-bills	12	9 - 10

Source: Benartzi & Thaler (1995) p. 83.

[21] The Center for Research in Security Prices

For nominal returns, the evaluation horizon was found to be approximately a year, with 13 months for bonds and 12 months for bills. Benartzi & Thaler argued that with tax filings and pension reports taking place once a year, this evaluation period appeared to be realistic. Thaler (1999) later noted that their results implied that if the most prominent evaluation period for investors is once a year, then the equity premium puzzle is effectively solved.

8.1.2 Optimal asset allocation

Having found an implied evaluation horizon of about a year, Benartzi & Thaler (1995) went on to derive the optimal asset allocation for this evaluation period, in order to facilitate an empirical validation of their method. Their argument was that if value-maximizing investors had in fact employed an evaluation horizon of roughly a year, then the asset allocation that would yield the highest utility for that horizon should correspond to the observed asset allocation amongst US investors.

Setting the evaluation horizon equal to twelve months, the authors computed the prospective utility of each portfolio mix between 100% in bonds and 100% in equity with increments of 10%. The results are illustrated in Figure 8.1, which shows that portfolios that had between 30% and 55% in stock all yielded roughly the same prospective value.

Figure 8.1: Prospective utility as a function of asset allocation
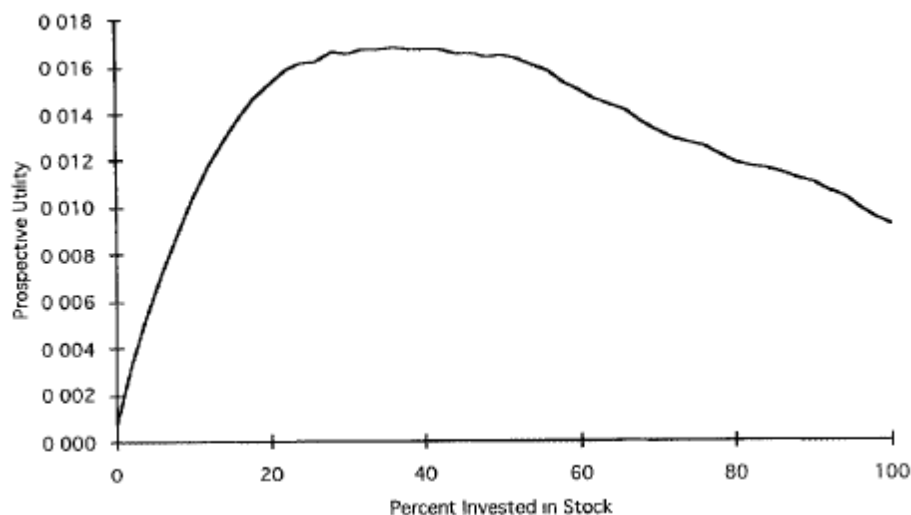
Source: Benartzi & Thaler (1995) p. 85.

Benartzi & Thaler (1995) referred to reports on institutional investor behavior, which indicated that funds (primarily pension funds) on average invested 53% of their assets in stocks. This served as support

69

of their own findings, as the value was within the range where investors would optimize their expected prospective utility given a yearly evaluation horizon.

8.1.3 Implied equity premia

The final step of the study was to calculate the implied equity premia for varying evaluation horizons to illustrate the decreasing pattern predicted by theory. As monthly equity returns are positive on average, the effect of aggregating returns into longer intervals would be a decreased probability of seeing a drop in equity value. As the evaluation horizon increases, the investor is therefore less likely to observe a loss on his stock portfolio, and his loss aversion will kick in less often. As he will effectively experience less discomfort from holding equity, stocks become more attractive and his required premium falls.

Using prospect theory preferences, Benartzi & Thaler (1995) found that the implied equity premium given a one-year evaluation horizon was 6.5%. Recall that Mehra & Prescott found a premium of 6.18% over the period 1889-1978, and with Benartzi & Thaler using data from 1926-1990, the figures are fairly similar. Our own observations regarding the choice of sample period, showed that equity premium calculations are particularly sensitive to these kinds of changes, but we will return to this point later on.

Table 8.2 presents a summary of Benartzi & Thaler's results regarding implied equity premia:

Table 8.2: Implied equity premia

Evaluation horizon, years	1	2	5	10	20
Implied equity premium	6.5%	4.7%	3.0%	2.0%	1.4%

Source: Benartzi & Thaler (1995) p. 83.

As seen in the table, the premium would fall from 6.5% a year to 4.7% if the evaluation horizon was doubled to 2 years. Correspondingly, a twenty-year horizon would imply an equity premium of 1.4%. Effectively, the cost of evaluating yearly for someone with a twenty-year investment horizon is 5.1% per year, as long as the market collectively behaves irrational. Because the investor refrains from investing in stocks, which offer a much higher return than the risk-free asset, he is effectively letting his loss aversion get the better of him. The difference therefore represents the price an investor pays for exercising excessive vigilance.

To sum up, Benartzi & Thaler (1995) concluded that the equity premium observed in the US could be explained by investors evaluating their portfolios once a year, and by people being twice as sensitive to their losses compared to their gains. While these results were intriguing, the model was merely descriptive, and the solution remained circumstantial. In the next section, we will discuss some of the

existing support and critique of myopic loss aversion as a potential solution to the equity premium puzzle.

8.2 Further empirical evidence

8.2.1 Imposing frames

Following the publication of Benartzi & Thaler's (1995) study, several attempts were made to establish the prevalence of myopic loss aversion through experimental studies. One of these contributions was a study by Gneezy & Potters (1997). The authors investigated how manipulations of the frequency with which feedback is provided will affect decision making under uncertainty. Just as Benartzi & Thaler (1995), they argued that the dynamic and individual aggregation rules that people apply as part of their mental accounting, have substantial influence over their inherent risk attitudes.

Gneezy & Potters (1997) sought to establish whether the combination of loss aversion and a person's evaluation horizon – or the width of their framing – represents a consistent choice anomaly in experimental settings. Hence, the authors set out to test whether a manipulation of the evaluation frames would cause respondents to display behavior consistent with myopic loss aversion:

In fourteen sessions, two groups of subjects were presented with a sequence of twelve identical and independent rounds of a lottery. At the beginning of the first nine rounds, the participants were given roughly one dollar to bet with, while for the three last rounds they were left with the money they had earned throughout the first part of the lottery. Each round offered a probability of 1/3 of winning two and a half times the betted amount, and a 2/3 probability of losing the amount.

The first group of subjects was defined as high-frequency, to represent narrow framing. These participants were supplied with feedback information after each round, meaning that they had the chance to evaluate their choice of how much to gamble after seeing the outcome of each individual round. Meanwhile, the other participants in the low-frequency group were only given information about the aggregate outcome after three rounds, and were thus not able to adjust their choice more than three times throughout the lottery. If myopic loss aversion did indeed represent a consistent behavioral pattern, then the low-frequency group ought to bet more of their endowment compared to the first group.

The results were that the high-frequency groups consistently chose to bet less than the groups that were imposed with broader evaluation frames, assumingly because they were deterred by the occurrence of

losses. Furthermore, Gneezy & Potters (1997) found that the difference in the average bets across the two groups was highly statistically significant.

In summary, Gneezy & Potters' study showed that if a long-term horizon is imposed externally, then people become inclined to take on more risk, effectively validating the existence of myopic loss aversion as a consistent choice anomaly.

This tendency was also observed in an experimental study by Thaler, Tversky, Kahneman & Schwartz (1997). In their experiment, participants were divided into three groups and asked to make investment decisions between stocks and bonds. The information made available to each group meant that they were effectively making their decisions at different frequencies, namely eight times a year, once a year, and once every five years. The two groups that were investing under long-term conditions (one and five years respectively) invested approximately two-thirds of their funds in stocks. Interestingly, the participants that were imposed with the frequent evaluation condition only chose to invest about 40% of their assets in equities.

The results essentially supported the notion that more frequent evaluations by the investor, effectively makes equity less attractive to him, which will drive up the equity premium as observed by Benartzi & Thaler (1995).

8.2.2 Introducing market interaction

So far the supporting evidence of myopic loss aversion has predominantly been experimental. This tendency has three disadvantages: First, individuals are only faced with a particular decision problem once, so they are not given the opportunity to move along a learning curve based on their own experience or by learning from others. Secondly, while individual violations of the axioms of expected utility are visible in experimental settings, the effect may be washed out at the aggregate level. Finally, people may act differently in real market situations compared to being confronted with basic decision problems in a laboratory setup.

As an answer to some of these concerns, Gneezy, Kapteyn & Potters (2003) published a study that aimed at capturing the effects of introducing market interaction in the form of a competitive environment, rather than just relying on the patterns of individual decision making. The authors organized experimental markets, where subjects were observed whilst buying and selling a risky asset. As in Gneezy & Potters' experiment from 1997, the participants were divided into high- and low-frequency groups. Again the first group was provided with feedback on their performance after each trading

period, and the second group was asked to commit their investment for three periods. They were only being informed of the cumulative return after those three periods, so that gains and losses could not be attributed to a particular round.

As in their previous study, Gneezy, Kapteyn & Potters (2003) found that people who evaluated the consequences of their choices on more aggregated terms (i.e. over longer periods), were much more willing to invest in the risky asset. The authors ultimately found that investors, who are supplied with more information about their portfolio as well as more flexibility in terms of changing that portfolio, are less likely to take on risk.

8.2.3 Myopic loss aversion outside the US

A key issue with myopic loss aversion is that few contributions within academia have attempted to test its validity outside the US. One exception comes in the form of a recent empirical study by Kliger & Levit (2009). The pair chose to utilize a special feature of the Tel Aviv stock exchange (TASE) that is used to enhance liquidity. When some securities become too illiquid, the exchange may choose to shift them from daily to weekly trading. As this feature effectively imposes a new trading frequency, Kliger & Levit argued that the inverse relationship between risk aversion and the evaluation horizon dictated by myopic loss aversion, should be visible under these conditions.

Recall that the investor will require a lower return if he employs broad framing for his investment decisions. Kliger & Levit (2009) therefore hypothesized that for Benartzi & Thaler's method to be validated, the return on stocks evaluated on a weekly a basis should be lower than the return on daily-evaluated stocks. In addition, they noted that as risk aversion is predicted to fall under broader framing, then the sensitivity of the investor towards relevant economic events should also decrease. Hence, they argued that price fluctuations should be lower for weekly-traded securities than for their daily-traded counterparts.

Kliger & Levit (2009) tested both of these hypotheses, and found that returns were indeed lower when securities were switched from daily to weekly trading and vice versa, and they also found that price fluctuations stemming from economic events were much less pronounced for the stocks that were evaluated less frequently. Ultimately, the authors concluded that their results represented a strong support for myopic loss aversion in real financial markets.

While these results are quite encouraging, there remains plenty of room for more empirical validation of myopic loss aversion as a solution to the equity premium puzzle, in particular for countries outside the

US. Before presenting our own analysis for Sweden however, we turn briefly to some of the more broadly founded critique of myopic loss aversion and the method of Benartzi & Thaler in particular.

8.3 A critical view

8.3.1 Institutional investors

One caveat to consider with the evidence presented in the previous section, is the role of intuitional investors such as pension funds, life insurance companies, mutual funds, as well as private hedge funds. Most financial literature refers to the preferences of private investors, and individual decision making, despite the fact that the largest players in the stock market are in fact investment managers who act in the place of private passive investors. It could be argued that professional fund managers have less disutility from losses, as they do not have to suffer the losses personally. Specifically, Locke & Mann (2005) argued that any research that chooses to ignore the role of professional traders should be considered severely flawed, as individual private investors are too far removed from the price discovery process to have any substantial impact on markets prices.

Benartzi & Thaler (1995) themselves argued that investment managers are often governed by short-term and rather complex incentive systems, which give rise to short horizons as long-term results are overshadowed by short-term losses. Furthermore, the authors also commented that agency problems may cause investment managers to display myopic loss aversion. In support of these arguments, Haigh & List (2005) found that traders recruited from the Chicago Board of Trade in fact exhibited behavior consistent with myopic loss aversion, and that they did so to an even greater extent than students under experimental settings.

In addition, a recent paper by Eriksen & Kvaløy (2010) dealt with the issue by employing prospect theory along with the methods of Benartzi & Thaler (1995) to investigate how professionals deal with risk when handling other people's money. The study found the prevalence of what was referred to as "myopic investment management", using experimental evidence to show that investment managers exhibit myopic loss aversion even in the absence of agency problems. The study concluded that investment managers' objectives are in fact aligned with their clients' objectives, even when there are no monetary incentives to secure such an alignment.

8.3.2 Disappointment aversion

The main assumption behind Benartzi & Thaler's (1995) results is that investors have prospect theory preferences. This assumption has been challenged by Ang, Bekaert & Liu (2005), who pointed to a different phenomenon from behavioral finance for solving the equity premium puzzle, namely disappointment aversion.

The two models have a number of features in common. Firstly, both of them attempt to provide behavioral finance explanations to the equity premium puzzle. Secondly, they both rely on the argument that investor preferences are defined against a reference point, as we have discussed earlier. Thirdly, they assume that people apply narrow framing in terms of their investment decisions, in particular the decision between investing in risky versus risk-free assets. Finally, both concepts maintain that investors focus on the prospects for returns rather than considering the covariance with consumption, as framed in the model of Mehra & Prescott (1985).

Just as Benartzi & Thaler (1995), Ang et al. (2005) attempted to provide a more formal treatment of portfolio choice, but rather than relying on prospect theory for modeling the preference structure of investors, the authors used the disappointment aversion framework as presented by Gul (1991). The framework evolved around a new alternative preference structure, which was essentially a one-parameter extension of the standard expected utility framework. Rather than defining gains and losses relative to the initial asset position, disappointment aversion used the *certainty equivalent*[22] as the reference point of the investor.

In other words, the certainty equivalence of an investment becomes the benchmark, and outcomes that are better than this benchmark will please the investor, whilst outcomes that are lower the certainty equivalent will represent a disappointment. Because the reference point relates to the investor's expectations rather than his initial asset position, it is assumed to evolve endogenously in the model.

Consider an individual with an evaluation horizon of twenty years. If he was to evaluate his twenty-year return based on whether he had lost or gained money compared to his initial asset position from the beginning of that period, then stocks would always provide a gain. If on the other hand, the investor was to constantly adjust his expectation levels to the amount of guaranteed money that he would trade for the investment for the given horizon, then his reference point would also evolve over time.

---

[22] The certainty equivalent is the guaranteed amount of money that the investor would find just as desirable as a risky asset

Ang et al. (2005) essentially argued that under myopic loss aversion, the initial reference point becomes more and more irrelevant, the longer the evaluation horizon becomes. Consequently, they objected that the heavy focus on the evaluation frequency meant that the model would fail over longer horizons, ultimately rendering it invalid.

As a response to these allegations, Fielding & Stracca (2007) chose to perform a sensitivity analysis of the two behavioral finance explanations of the equity premium puzzle. Based on the critical stance of Ang et al. (2005) against the heavy focus on evaluation horizons, Fielding & Stracca decided to concentrate on the role of the horizon itself in determining the size of the equity premium. More specifically, they wanted to test whether myopic loss aversion would be plausible as an explanation for investors with longer time horizons. The authors tested two things:

i.   What would happen to the explanatory power of *myopic loss aversion* if the aggravation parameter λ from the prospective value function was larger than assumed by Benartzi & Thaler?

ii. At what time horizons could reasonable parameters for the degree of *disappointment aversion* explain the historic equity premium?

Fielding & Stracca (2007) followed the methods prescribed by the respective studies, and found that even with an increased aggravation parameter, myopic loss aversion was superior in terms of explaining the premium on equity for shorter horizons. Meanwhile, disappointment aversion dominated for investment periods of ten years or longer. The authors hypothesized that the reference point may evolve as a function of the time horizon, and that myopic loss aversion is the driver of investment behavior until the horizon becomes so long that the expected return on the risky asset is very high.

Fielding & Stracca (2007) concluded that the two behavioral approaches are not necessarily mutually exclusive, and that the high historic equity premium may be explained by both myopic loss aversion and disappointment aversion. Furthermore, as of yet there remains to be presented any experimental evidence to support the risk preferences presented by disappointment theory. We therefore argue that whilst Ang et al. (2005) presented some intriguing points, their work should be seen as complementary to myopic loss aversion, and further testing of the model of Benartzi & Thaler (1995) continues to be of academic interest.

8.3.3 The role of consumption

Finally, we will address the critique presented by Barberis, Huang & Santos (2001). The authors claimed that the main weakness of Benartzi & Thaler's framework was that the investor only gets direct utility from the changes in the value of his financial investments, and not from consumption or total wealth.

Barberis, Huang & Santos (2001) argued that a framework of investor behavior should return to the original premise of standard expected utility theory, namely that the investor is affected by his consumption level. The authors proposed that a realistic framework should describe the joint properties of stock returns and consumption growth. They subsequently presented a model where the investor gets utility from consumption, as well as getting utility from fluctuations in the financial component of his wealth.

The basic premise of myopic loss aversion was maintained, as the authors assumed that investors apply narrow framing when assessing the risk in financial markets, and the framework was equally based on the preference structure presented in prospect theory, with the investor being loss averse. The main difference between the two models was that the framework of Barberis, Huang & Santos (2001) employed a dynamic measure of narrow framing, which was essentially based on aggregate consumption.

The authors found that their model could explain the high historic premium, and that the narrow framing premise was valid even when consumption growth was smooth and only weakly correlated with stock returns.

In a recent paper however, Barberis & Huang (2008) themselves noted that their framework remains imperfect. As they had not been able to present an actual value function, it is difficult to validate the soundness of the preference parameters of the model. Furthermore, the dynamic feature of narrow framing had been scaled according to a neutral factor based on aggregate consumption, and this factor was admittedly based on ad-hoc assumptions rather than experimental evidence.

While the concerns of Barberis, Huang & Santos (2001) may be valid, we have chosen to stay true to the model presented by Benartzi & Thaler (1995). The goal of our contribution is to test the original framework in a new country, as it continues to represent the main contribution within the field of behavioral finance in terms of explaining the equity premium puzzle. More specifically, we feel that introducing the original framework into an international arena will help to understand the role and the

impact of myopic loss aversion in financial markets and provide new insights to investor behavior outside the US.

8.4 Chapter summary

This chapter has shown that by assuming that investors have prospect theory preferences, and that they evaluate their portfolios roughly once a year, Benartzi & Thaler (1995) found that myopic loss aversion was able to explain the equity premium puzzle in the US.

Following their findings, several experimental studies have been performed to test the validity of myopic loss aversion as a descriptive theory, and they all found support of the behavioral patterns that Benartzi & Thaler had used in their model.

Some critical considerations were taken into account, namely the role of institutional investors, the use of loss aversion versus disappointment aversion, as well as the influence of consumption on the investor's utility. While particularly the two latter points offered some interesting perspectives, we saw that neither of them have yet disproved the validity of myopic loss aversion as a potential solution to the equity premium puzzle.

We now turn to part three of this paper, which presents our analyses of the equity premium puzzle in Sweden, as well as the use of myopic loss aversion as a solution to the puzzle. First we comment on the data used in our study.

Part III: Empirical analysis and results

9. Data

The original study of Mehra & Prescott (1985) employed five data series from the US, for the period 1889-1978:

i. Annual average S&P Composite Stock Price Index

ii. Real annual dividends for the S&P index

iii. Real per capita consumption on non-durables and services

iv. Consumption deflator series based on the consumer price index

v. Annual yield on the risk-free securities (government debt with a maturity of 60-90 days)

In their study on myopic loss aversion in the US, Benartzi and Thaler (1995) used monthly return series for stocks, T-bills, and 5-year bonds for the period 1926-1990, which they obtained from CRSP[23]. Furthermore, the authors used monthly inflation figures, to obtain real return series for all three asset classes.

Our analyses of the equity premium puzzle and myopic loss aversion in Sweden are based on equivalent data series, but for the period 1919-2010. As the yearly return series used on the first part of our analysis are aggregated based on the monthly returns used in the second part, the following discussions on data apply to both parts of our study.

Most studies of the equity premium puzzle outside the US suffer under the limited availability of high quality data before the 1970s, but thanks to the remarkable work of Frennberg & Hansson (1992)[24] and Krantz & Nilsson (1975), Swedish data stands out in this regard. The Riksbank[25] (2008) has continued the work of Frennberg & Hansson to cover the period 1990-2006, and we have extended the series to 2010.

Due to the rather long horizon of our study it has been necessary to construct the majority of the data series from multiple sources. In the following section we will go though the steps we took in constructing each series, and discuss any potential issues.

---

[23] The Center for Research in Security Prices
[24] The meticulous studies of international equity premia by Dimson et al. (2008; 2011) also rely on their dataset
[25] The Swedish Central Bank

9.1 Equity returns

The main challenge with obtaining equity return for such a long data series, is to determine a total return index that includes reinvested dividends (series i and ii of the Mehra-Prescott dataset). This has however been done on Swedish data by Frennberg & Hansson (1992), who constructed a total monthly return index for the period January 1919 to December 1989. Their work was continued by the Riksbank (2008), and the two periods are composed in a combined data series provided by the Riksbank for the period 1919-2006. Total return data for 2007-2010 was obtained via Datastream. In annualizing the figures, we have calculated the annual return index as the arithmetic average of the monthly return index for each year.

9.2 Per capita consumption

The data on annual private consumption is drawn from two separate sources. From 1918 to 1950 we rely on the extensive work of Krantz & Nilsson (1975) and from 1950 to 2010 we have used data from the International Monetary Fund (IMF).

There is a discrepancy between the two sources that should be noted. Krantz & Nilsson report real consumption whereas IMF provides nominal figures, and this is somewhat troubling. Krantz & Nilsson deflate their consumption series with multiple inflation indices for the different components of consumption (food, firewood, etc.) based on their weight in the given year. It has not been possible to obtain similar indices and weights for the period after 1970, so the process could not be replicated for the IMF data. Unfortunately it was not possible to reverse the calculations for the period before 1950 either with acceptable accuracy, so the two series are converted into real figures with inconsistent deflators.

The two sources overlap in the period from 1950 to 1970, so we compared the growth rate in the twenty- year period to see if the inconsistency was unacceptable. We found that the IMF data displayed lower annual real consumption growth rates, with an average figure of 2.26% compared to the 2.46% reported by Krantz & Nilsson. This is a difference worth noting, but we do not believe that it undermines the use of the data series, as it is by far the best sources available. For the overlapping period of 1950-1970 we use the IMF data due to its higher transparency.

9.3 Consumer deflator series

The consumer deflator series (inflation) is based on the impressive Swedish Consumer Price Index series stretching back to 1290, which determines the annual CPI for the entire 920 year period. The data series was obtained through the Riksbank (2008).

For our analysis of myopic loss aversion, an estimate of inflation on a monthly basis is needed. As records of monthly inflation are not available before 1957 (IMF via Bloomberg), our sensitivity analyses concerning real versus nominal returns are based solely on the period from 1957-2010. These tests also serve to validate our findings for a shorter sample period.

9.4 Risk-free return

The choice of the risk-free asset is often a somewhat controversial one. A common proxy in finance literature is the return on a 30-day T-bill or a similar short-term government debt instrument, and this was also used by Mehra & Prescott in their original study (1985). It could be argued though, that for most investors this is not the risk-free alternative to holding stocks, as household investors in the US actually hold virtually no short-term government debt (Mehra, 2008).

As we have chosen to follow the original framework of Mehra & Prescott (1985), our primary analysis of the equity premium puzzle in Sweden will apply the yield on short-term government debt as the risk-free rate. To verify its robustness we have also run the model using the return on government bonds with an average maturity of ten years.

To ensure transparency between the two sections, we also use T-bills as the risk-free asset in our main analysis of myopic loss aversion, despite the fact that is represents a divergence from the original study of Benartzi & Thaler (1995). We do however subsequently perform sensitivity tests, where we apply a return series on long-run government bonds as the risk-free asset.

From 1919 to 1982 our risk-free index is based on the short-run discount rate of the Riksbank. For the remainder of the period, the calculation of the return index is based on the yield of a 30-day government note. Both series are obtained through the Riksbank for the period 1919-2006, and obtained via

Datastream for the period 2007-2010. The index calculations are done by Frennberg & Hansson, the Riksbank, and us[26].

9.5 Long-run government bonds

We have chosen to expand the dataset of Mehra & Prescott (1985) with the return series of a long-run risk-free asset, as this has been suggested as an alternative to T-bills as the risk-free rate. We use the monthly series in our analysis of myopic loss aversion, where it is used to test the robustness of our results.

Again we rely on the work of Frennberg & Hansson (1992) and the Riksbank (2008). From 1919 to 1950 the return is based on the monthly yield on consols (government perpetuities), and from 1950 to 2010 it comes from the yield on government bonds with average maturity of ten years. From 1919 to 1990 the index calculations were provided by Frennberg & Hansson (1992) and from 1991 to 2006 they were done by the Riksbank (2008). For the remainder of the period we have performed the index calculations ourselves.

Unfortunately, our predecessors have been quite modest with providing detailed information about their calculations, so the continuation of the series proved to be somewhat of a less trivial task than anticipated. The basic task is to calculate a total monthly return based on 1/12 of the annualized yield, plus a capital gain/loss. The capital gain/loss can be determined precisely by calculating the value of the bond before and after any change in the yield, but this method is computationally heavy and requires knowledge of the coupon rates for all the bonds in question. The other possibility is to approximate the price change based on the duration of the bond[27].

Frennberg & Hansson (1992) appear to have calculated the return index directly, based on estimations of the coupon rates, but as they relied on decade-wide coupon rates, their calculations must still have been approximations. We tried to obtain the precise calculations from the authors, but a brief email correspondence with Professor Hansson did not provide further insights[28].

---

[26] Monthly return is simply calculated as 1/12 of the annual yield. Due to the very short maturity of the notes, we have not included price changes in the return calculation. We have replicated the index calculations of our predecessors with this method, and they appear to have used exactly the same approach.

[27] Convexity considerations could also be included, but as government bonds generally have very low convexity, we will not consider this any further

[28] The professor was very helpful but as he noted, "it is around 20 years since we made the decisions on how to construct the bond returns!"

The Riksbank (2008) did not include comments on their choice of methodology for the period 1990-2006 either. Again we contacted the author, but yet another brief email correspondence did not provide any decisive answer. It did however appear as if the Riksbank had approximated the price changes, rather than basing his calculations on specific coupons rates.

Given the scarcity of a methodic description, we chose to approximate the capital gains/losses based on the duration of the bonds. We therefore defined the monthly return based on the following relationship:

$$(9.1) \qquad\qquad R_t = \frac{r_{t-1}}{12} - D_m * (r_t - r_{t-1})$$

where D_m is the modified duration, r is the annualized yield, and R is the total return.

To verify that our approach did not cause a structural shift in the return calculations for the two sub periods 1919-2006 and 2007-2010, we replicated the return calculations for the period 1990-2006 and compared them to the work of the Riksbank (2008). We did not replicate the index calculations for the Frennberg-Hansson period, as they seemed to have applied a different method than the Riksbank and ourselves.

Below we have plotted the return on the Riksbank index, against the return on our replicated index for the same period. If the calculations were identical the scatter plot should have shown a straight line with a slope of one and intersect in (0.0). As Figure 9.1 shows, this was not the case.

Figure 9.1: Replicated index plotted against the Riksbank index
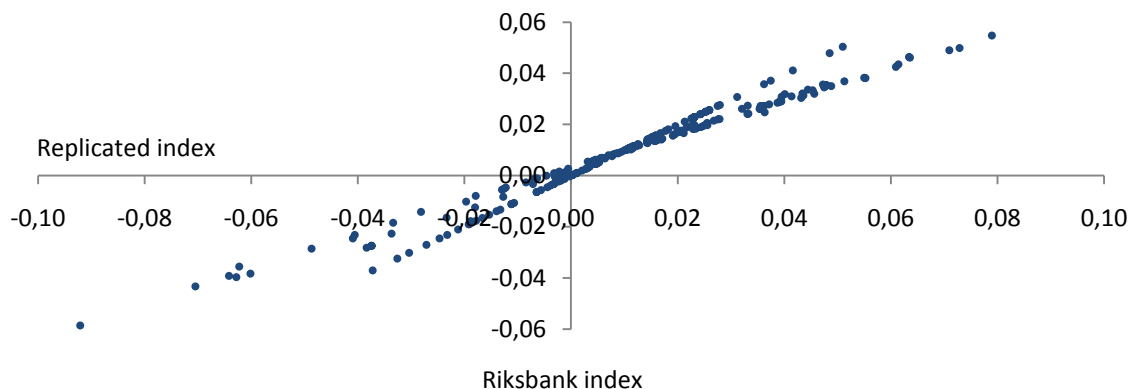
The plot shows a clear pattern, but unfortunately it is not the one we hoped for. Some of our return calculations seem to consistently understate both positive and negative returns.

To establish whether the discrepancy is constant over the period, we divided our replicated index by the Riksbank index for each month. Months where the two methods yielded the same return, would thus give the result 1, and variations around this value would show where our calculations under- or overstated the Riksbank index. If our calculations were identical for the entire period, the result would be a horizontal line, intersecting the y-axis in 1. The result is shown below:

Figure 9.2: Replicated index divided by the Riksbank index

Year:Month
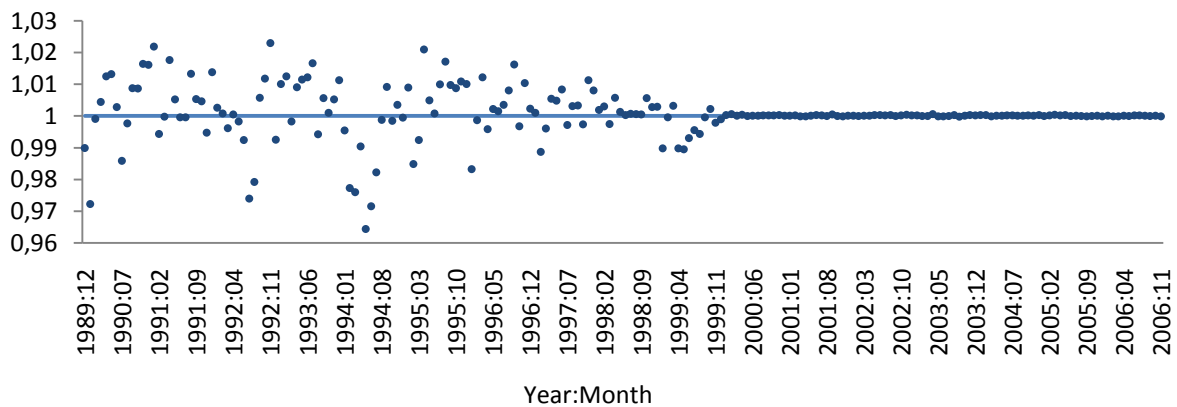
The chart shows that our replications are somewhat off for the period before 2000, but interestingly they are extremely similar for the period 2000-2006. A regression of the Riksbank returns on our replicated returns for the period after 2000 yielded a slope of 0.991 with intersect at 0.009, which is quite close to the expected values of 1 and 0. This indicates that the Riksbank changed its methodology at the turn of the millennium to one practically identical to ours. For the period 2000-2006 our calculations overstate his average annualized return calculations by merely 0.01%-point.

We therefore conclude that our method for extending the total return series is in line with the methodology of our most recent predecessor. Regarding apparent changes in the applied return calculations after the Frennberg-Hansson period and after the year 2000, we trust the good judgment of the Riksbank.

10. The Equity premium puzzle in Sweden

10.1 Methodology

To ensure transparency between previous studies of the equity premium puzzle and the conclusion of our analysis, we have applied the original method of Mehra & Prescott (1985). As the methodology suffers the limitation that it requires values for α (the coefficient of relative risk aversion) and β (the impatience of the investor), which are notoriously difficult to observe, we have also applied the statistical approach of Kocherlakota (1996). Recall from section 3.1 that his framework does not rely on the same subjective estimations, but instead tests what α must be for the Mehra-Prescott model to explain observed equity premia.

We start by applying the approach by Mehra and Prescott (1985). The first step is to determine the risk-free rate predicted by the Mehra-Prescott model by inserting the observed mean and standard deviation of consumption growth in equation (2.10). To ease comparison between our findings and the original statement of the puzzle, we use the same estimations of the CRRA ($\alpha = 10$) and the subjective impatience factor ($\beta = 0.99$).

$$(\mathbf{2.10}) \qquad \qquad \ln R_f = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2$$

Based on this implied real risk-free rate, we can find the expected real return on equity from the rearrangement of equation (2.11):

$$(\mathbf{2.11}) \qquad \qquad \ln E(R_e) = \ln(R_f) + \alpha\sigma_x^2$$

The difference between the risk-free rate and the expected return on equity represents the equity premium. This can then compared to the observed equity premium in Sweden during the period, and the presence of the equity premium puzzle can either be confirmed or rejected.

To further validate our findings from the Mehra-Prescott method, we apply the statistical approach of Kocherlakota (1996), which does not rely on fixed estimations of the two factors alpha and beta. Instead of assuming a certain CRRA as Mehra & Prescott do, recall that the method of Kocherlakota argues that the equity premium should not be significantly different from zero for reasonable levels of risk aversion (1≤α≤3). Equation (3.3) states the testable hypothesis:

$$(\mathbf{3.3}) \qquad \left[\left(\frac{c_{t+1}}{c_t} \right)^{-\alpha} (R_{e,t+1} R_{f,t+1}) \right] = e_{t+1}$$

where $\left(\frac{c_{t+1}}{c_t} \right)$ is consumption growth in a given year, and R_e and R_f are the real annual returns on equity and a risk-free investment in that year. The result e, is the expected equity premium for that year, given the value of alpha, and the calculation is done for every year for increasing values of alpha. For each value of alpha the mean value of e for all years is then calculated, and subsequently tested using a t-statistic to determine whether it is significantly different from zero. If the utility-based model of Mehra & Prescott (1985) is appropriate for estimating the equity premium, the sample mean should not be significantly different from zero for reasonable levels of risk aversion.

Having covered the methodology for establishing whether or not there exists and equity premium puzzle in Sweden, we now turn to the results from our initial analysis.

10.2 Results, 1919-2010

As outlined, we have analyzed the equity premium puzzle in Sweden using the methods of Mehra & Prescott (1985) and Kocherlakota (1996). The sample statistics for the entire period are provided below:

Table 10.1: Sample statistics, 1919-2010

Mean real risk-free rate	R_f	2.23 %
Mean real equity return	$E(R_e)$	6.40%
Mean real growth rate of consumption	$E(x)$	2.28%
St.dev. real growth rate of consumption	$\sigma(x)$	3.99%
Mean equity premium	$E(R_e) - R_f$	4.17%

Data: Sweden, 1919-2010, real returns.

To follow the approach of Mehra & Prescott, we apply $\alpha = 10$ and $\beta = 0.99$. Inserting these figures into equation (2.10), along with the mean and variation of consumption growth, yields the following risk-free rate:

$$(\mathbf{2.10}) \qquad \ln R_f = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2$$

$$\ln R_f = -\ln(0.99) + 10 * 0.0228 - \frac{1}{2} * 10^2 * 0.0417^2 = 0.1582 \quad \Longrightarrow \quad \boldsymbol{R_f = 17.14\%}$$

Applying equation (2.11) in the same manner, yields an expected equity return of:

$$(2.11) \qquad\qquad \ln E(R_e) - \ln(R_f) = \alpha\sigma_x^2$$

$$\text{Ln}E(R_e) = 0.1582 + 10 * 0.0417^2 = 0,1741 \quad \Longrightarrow \quad \boldsymbol{E(R_e) = 19.01\%}$$

Allowing for rounding errors this gives an expected equity premium of 1.88%, which is well below the observed equity premium of 4.17% shown in Table 10.1. Recall from section 3.2 that the high return levels produced by the model is a result of the risk-free rate puzzle. Using the method of Mehra & Prescott (1985), we can thus document an equity premium puzzle of 2.30% in Sweden during the period from 1919 to 2010.

As Mehra & Prescott (1985) documented a puzzle in the order of 4.78%, this first test seems to indicate that the equity premium puzzle is much smaller in Sweden, than what was found by the original study in the US. Before jumping to conclusions, we turn to the test of Kocherlakota.

Table 10.2 reports our results for the period using the statistical approach of Kocherlakota (1996). In each year and for values of alpha ranging from 0.0 to 50.0, we have applied equation (3.3) to test whether the Mehra-Prescott model is able to explain the equity premium. If this was the case, then the average value of e for all years (the values displayed in italics in Table 10.2), should not be significantly different from zero, for $\alpha \leq 3$.

Table 10.2: Kocherlakota's alpha, 1919-2010

Alpha	0	0.5	1.0	1.5	**2.0**	2.5	3.0
1919	*-0.178*	*-0.170*	*-0.161*	*-0.154*	***-0.146***	*-0.139*	*-0.132*
1920	*-0.114*	*-0.108*	*-0.102*	*-0.097*	***-0.091***	*-0.086*	*-0.081*
…	*…*	*…*	*…*	*…*	***…***	*…*	*…*
2010	*0.281*	*0.277*	*0.274*	*0.270*	***0.266***	*0.263*	*0.259*
Mean	0.042	0.040	0.038	0.037	**0.035**	0.033	0.031
St.dev.	0.184	0.185	0.186	0.188	**0.190**	0.192	0.195
T-stat	2.165	2.066	1.963	1.858	**1.750**	1.640	1.528

Data: Sweden, 1919-2010, real returns

The Kocherlakota method determines that the Mehra-Prescott model is able to explain the equity premium in Sweden for $\alpha \geq 2$ using a one-tailed t-test[29].

It could be argued that the one-tailed test applied by Kocherlakota is inappropriate, as e appears negative in some years (e.g. 1919 and 1920). This would indicate that the model of Mehra & Prescott

---

[29] 95% confidence and 90 degrees of freedom

actually *over*states the observed equity premium in some periods. Consequently, we also ran the model using a two-sided t-test, which yielded $\alpha \geq 0.5$.

As alpha values between 1 and 3 are considered reasonable, the result suggests that the Mehra-Prescott model is able to explain the equity premium in Sweden during the period – or even that the model overstated the observed equity premium, as it resulted in an alpha value less than 1 when subjected to a two-sided t-test. Using the method of Kocherlakota we therefore arrive at the surprising conclusion that there is no equity premium puzzle in Sweden between 1919 and 2010.

Given our finding of a puzzle using the Mehra-Prescott method, along with the dismal performance of the model on US data, we decided to investigate the matter a bit further. We therefore revisited our data to check for any anomalies.

10.3 Re-examination of data

Based on the surprising result of the Kocherlakota test, we decided to investigate the data to see if the period 1919-2010 contained any obvious anomalies in the composition of the equity premium, i.e. whether changes were driven by an increase in equity return or a decrease in risk-free return. Figure 10.1 shows the annual equity premium, equity return, and risk-free return for the entire period from 1919 to 2010.

Figure 10.1: Annual equity return, risk-free return and equity premium
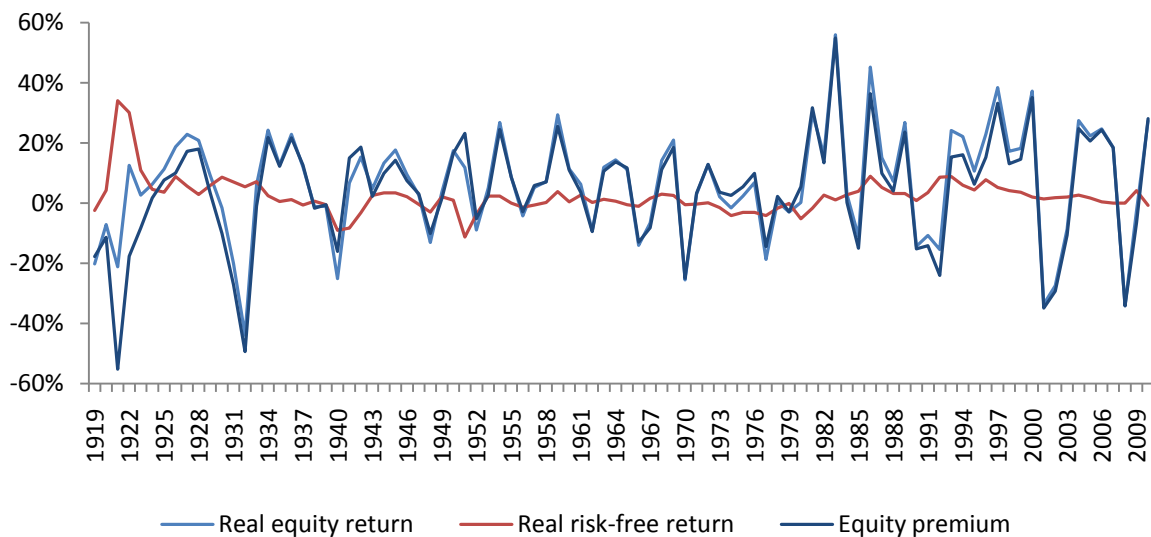
The figure shows that variations in the equity premium have largely been driven by changes in equity return, as the real risk-free return has been quite stable. One period stands out however, namely 1921-1922 where the large drop in the equity premium was caused by a sharp increase in the real risk-free return. Closer analysis revealed that the real risk-free return in the two years was 34.0% and 30.1% respectively, driven by high levels of deflation. Figure 10.2 shows the distribution of annualized real risk-free rates in a histogram.

Figure 10.2: Annualized real risk-free return
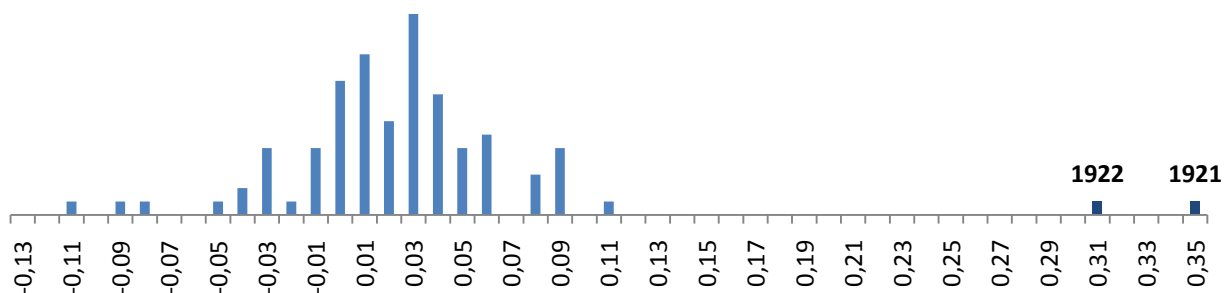
1921 and 1922 clearly stand out from the rest of the period in terms of real risk-free return, and consequently in their impact on the equity premium. Though the years seem highly unusual compared to the rest of the period, they should of course still be included if they can be considered representative. In the following section we will provide a brief look into the economic history of Sweden to investigate the nature of the seemingly odd returns of 1921 and 1922.

10.3.1 The Swedish deflation crisis of the early 1920's

Sweden had remained neutral during the course of WWI and had enjoyed a substantial surplus on their trade balance from exports to the war-engaged nations[30]. When the war came to an end and the industries of the warring nations returned to peacetime production, the flow of goods was reversed, and consequently the value of the krona (SEK) dropped relative to the dollar and gold.

A group of economists was put together to solve the problem, and they recommended that Sweden should seek to restore exchange rates to their pre-war levels, by increasing interest rates and implementing measures to decrease wages. The fiscal policies were also adjusted accordingly, and the large defense budget that had been sustained to ensure neutrality was cut drastically. These deflationary policies were passed in the late spring of 1920, but before they could be put into effect, the increased supply of goods from abroad had halted the inflation that had built during the war, and from July of 1920

---

[30] The historic review in this section is based on Larsson (1991) and Schön (2007)

Sweden experienced deflation. The domestic industries had built a considerable stock of supplies to ensure production during the war, and when the deflation hit, these were quickly dumped in the market to limit losses, which increased deflation further.

The general belief in the merits of the free market meant that the Swedish authorities were only concerned with monetary policies, and were not expected to intervene through fiscal policies. Hence they forged ahead with the measures to strengthen the krona, and implemented policies to increase deflation even further. The effects came promptly with an annual deflation of almost 20% in 1921, while unemployment rates increased from 4-5% to 25%, and *real* per capita consumption decreased by 12.7% in 1921 alone.

It seems fair to state that there has been a fundamental shift in Swedish economic policies since the deflation crisis of the early 1920's. At the time, monetary policies were the only concern of the government, but after the deflation crisis, fiscal considerations began to play an ever larger role. The shift already became apparent during the crisis in the early 1930's, where the Swedish government made considerable investments in housing construction to stimulate the economy (Larsson, 1991). We believe that the extreme real risk-free returns of the early 1920's in combination with the paradigm shift that took place shortly after the deflation crisis, calls for an elimination of the period from our dataset.

10.4 Results, 1925-2010

To maintain continuity in our dataset we have eliminated the period from 1919-1924, to ensure that the effects of the deflation crisis are eliminated. The following section repeats our analysis based on the modified dataset from 1925-2010.

Table 10.3: Sample statistics, 19125-2010

Mean real risk-free rate	R_f	1.43 %
Mean real equity return	$E(R_e)$	7.16%
Mean real growth rate of consumption	$E(x)$	2.17%
St.dev. real growth rate of consumption	$\sigma(x)$	3.37%
Mean equity premium	$E(R_e) - R_f$	5.73%

Data: Sweden, 1925-2010, real returns.

Inserting the figures in equation (2.10)[31] and (2.11)[32] as before, yields a risk-free return of 18.59% and an expected equity return of 19.94%. The Mehra-Prescott model can thereby explain an equity premium of 1.36%, which is far lower than the observed equity premium of 5.73%. Using data from 1925 instead of 1919 thus increases the puzzle from 2.30% to 4.37%, which seems more in line with Mehra & Prescott's finding of 4.78% in the US (1985).

As shown in Table 3.2, Dimsom et al. (2011) find an equity premium in Sweden of 4.3% for the period 1900-2010, but the deviation from our result of 5.73% is likely to stem from the unstabile period before 1925. For the subperiod 1961-2010 they report an equity premium of 6.1%, and a rerun of our model for this subperiod shows a similar premium of 6.0%.

The statistical results for the revised period are reported below:

Table 10.4: Kocherlakota's alpha, 1925-2010

Alpha	0	2.5	5	7.5	**8**	8.5	9
1925	*0.076*	*0.068*	*0.060*	*0.054*	***0.053***	*0.051*	*0.050*
1926	*0.100*	*0.085*	*0.073*	*0.062*	***0.060***	*0.059*	*0.057*
…	*…*	*…*	*…*	*…*	***…***	*…*	*…*
2010	*0.281*	*0.263*	*0.246*	*0.230*	***0.227***	*0.224*	*0.221*
Mean	0.057	0.050	0.043	0.036	**0.034**	0.033	0.032
St.dev.	0.174	0.174	0.178	0.185	**0.187**	0.189	0.192
T-stat	3.039	2.627	2.203	1.780	**1.697**	1.614	1.533

Data: Sweden, 1925-2010, real returns.

For the period 1925 to 2010 we find that only values of $\alpha \geq 8$ can produce an equity premium that is not significantly different from zero using a one-tailed t-test[33]. Recalling that Kocherlakota (1996) found that the Mehra-Prescott dataset required $\alpha \geq 8.5$, this result seems more in line with the results found on US data, than $\alpha \geq 2$ (or 0.5), which we found for the entire period.

Again we looked at the validity of using a one-tailed t-test, i.e. only testing whether the Mehra-Prescott model understates the observed equity premium. We found just twelve incidents where the model did *not* understate the observed premium, out of almost 2000 calculations. Hence, we believe that it is appropriate to follow the original approach of Kocherlakota and apply a one-tailed test for the revised

---

[31] $\ln R_f = -\ln(0.99) + 10 * 0.0217 - \frac{1}{2} * 10^2 * 0.0337^2 = 0.1705 \Rightarrow R_f = 18.59\%$

[32] $\ln E(R_e) = 0.1705 + 10 * 0.0337^2 = 0.132 \Rightarrow E(R_{re}) = 19.94\%$

[33] 84 degrees of freedom, 95% confidence level

period. Should we have applied a two-tailed test, the required value of alpha would have been 6, which is still twice as high as the acceptable level.

After correcting our dataset for the highly atypical period before 1925, we conclude that both the Mehra-Prescott model and the test of Kocherlakota show the existence of an equity premium puzzle in Sweden during the period 1925-2010. The results for the period 1919-2010 and 1925-2010 are summarized in the table below, along with the original findings of Mehra & Prescott:

Table 10.5: Results, overview

		Risk-free rate	Equity return	Equity premium	**Puzzle**	**Alpha**
US 1889-1978	Observed	0.80%	6.98%	6.18%	**4.78%**	**8.5**
	Mehra-Prescott model	12.7%	14.1%	1.4%		
Sweden 1919-2010	Observed	2.23 %	6.40%	4.17%	**2.30%**	**2.0**
	Mehra-Prescott model	17.1%	19.0%	1.9%		
Sweden 1925-2010	Observed	1.43%	7.16%	5.73%	**4.37%**	**8.0**
	Mehra-Prescott model	18.6%	19.9%	1.4%		

10.5 Robustness of results

To test the robustness of an equity premium puzzle in the Swedish dataset, we have redone the calculations three times: With nominal figures instead of real, with 10-year bonds instead of T-bills, and for varying time periods to see whether the result for the revised period is merely a coincidence.

Nominal vs. real

It can be argued that investors are focused on nominal returns instead of real returns, as done by Benartzi & Thaler (1995). We therefore repeated our analysis of the equity premium puzzle using nominal figures. The results are reported below:

Table 10.6: Results, nominal returns

		Risk-free rate	Equity return	Equity premium	**Puzzle**	**Alpha**
Sweden 1925-2010	Observed	5.14%	11.08%	5.94%	**2.84%**	**6**
	Mehra-Prescott model	70.26%	73.36%	3.09%		

Data: Sweden, 1925-2010, nominal returns

The puzzle is considerably smaller than the one found using real data, and it may an indication of the behavioral explanation to the puzzle provided in this paper. Nonetheless, the puzzle is still clearly documented using both the method of Mehra-Prescott (1985) as well as that of Kocherlakota (1996).

Bond vs. T-bills

Following the discussion in section 9.4 regarding the appropriate choice of risk free rate, we ran the model once more using government bonds with an average maturity of ten years instead of the short term T-bills. As expected, it decreased the size of the premium, as the longer maturity on bonds should lead investors to demand a higher return for holding them, and thereby decreasing the spread to the equity return.

The observed equity premium decreased from 5.73% to 4.39%, and the size of the puzzle decreased from to 4.73% to 3.03%. The minimum acceptable alpha under Kocherlakota's method decreased from 8 to 4.5, which is also an expected effect from increasing the risk of the "risk-free" asset as we do by replacing bills with bonds. Still, the finding of the equity premium puzzle persists when we apply 10-year bonds as the risk-free alternative to equities.

Table 10.7: Results, 10-year bonds as risk-free asset

		Risk-free rate	Equity return	Equity premium	**Puzzle**	**Alpha**
Sweden 1925-2010	Observed	2.78%	7.16%	4.39%	**3.03%**	**4.5**
	Mehra-Prescott model	18.59%	19.94%	1.36%		

Data: Sweden, 1925-2010, 10-year bond returns.

Varying starting years

The revision of our dataset showed that the choice of starting year can make a big difference for the findings. To verify that our results for the period 1925-2010 are not just coincidental, we reran the model for seven additional time periods[34], increasing the beginning year gradually from 1925 to 1960. The results from using the original approach, with T-bills as the proxy for the risk-free asset, and calculations being done in real terms, are shown in the figures below:

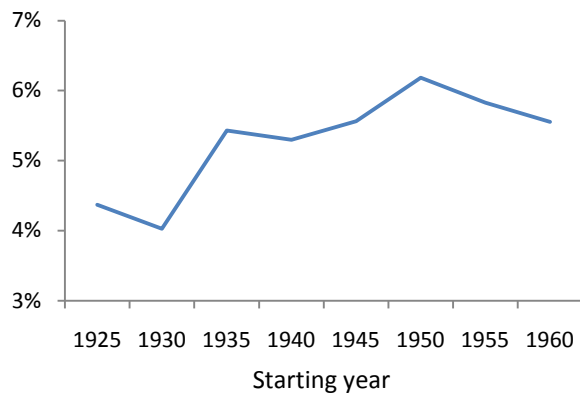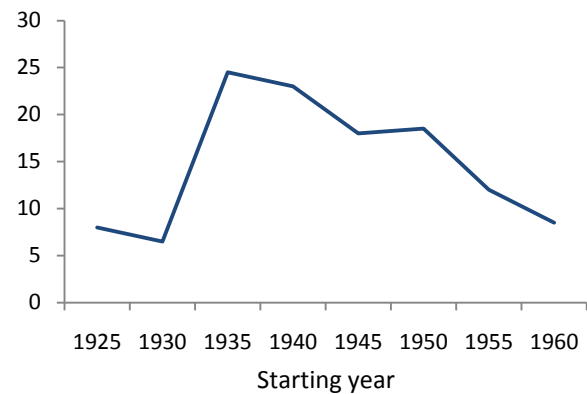[34] 1925-, 1930-, 1935, 1940-, 1945-, 1950-, 1955-, 1960-2010

 As Figure 10.3 shows, the finding of an equity premium puzzle appears to be consistent, regardless of the choice of starting year. The alpha values in Figure 10.3 indicate that the utility-based model of Mehra & Prescott is unable to explain the equity premium for any of the time periods, with a reasonable level of risk aversion. The lowest value of alpha (6.5) is found for the period 1930-2010, and it peaks at 24.5 for the period 1935-2010. A likely explanation for this sudden change in apparent risk aversion, is the great stock market crash of the late 1920's. The active fiscal policy implemented after the deflation crisis in the early 1920's succeeded in dampening the effects of the Great Depression in Sweden, and the full effects were not felt until 1932 (Larsson, 1991). The consequences of the Great Depression can also be seen in the substantial drop in equity return in Figure 10.1. Such an event would most likely cause investors to become more risk averse, hence increasing alpha.

The risk aversion needed to explain the equity premium appears to decrease for the periods starting after 1935, but the inference from this is not necessarily that investors have increased their risk tolerance. As the method of Kocherlakota is a statistical approach that relies on a t-statistic, it is quite sensitive to the decreasing number of observations from postponing the starting year. A rerun of the model using quarterly data available from 1970, showed an average required alpha of 10 for the periods 1965-, 1970-, 1975-, and 1980-2010, which supports the persistent presence of the equity premium puzzle in Sweden over time.

We conclude that our finding of an equity premium puzzle in Sweden is robust over different sample periods, as the minimum value for alpha is well above the maximum acceptable level of 3 for all eight periods plus the four periods investigated using quarterly data. In regards to the brief discussion above on the apparent variation in risk aversion over time, we caution the reader to draw any decisive conclusions based on Figure 10.3 and 10.4. The purpose is merely to show that the observed equity

premium cannot be explained by the utility-based model of Mehra & Prescott for reasonable levels of risk aversion in any of the periods.

10.6 Chapter summary

We started by testing the ability of the Mehra-Prescott model to explain the equity premium in Sweden for the period 1919-2010, using both the original methodology of Mehra & Prescott (1985), as well as the reformulated model of Kocherlakota (1996). To our surprise we found that the latter method dismissed the existence of an equity premium puzzle in the dataset, and that the Mehra-Prescott method found a quite small puzzle.

Given the general empirical merits of the model, we decided to investigate our dataset for dubious data points before jumping to conclusions. We found that for the period 1919-2010, changes in the equity premium had mainly been driven by changes in real equity return, as the real risk-free return had been quite stable during the period. Two years stood out however, namely 1921 and 1922, where the real risk-free rate soared around 30%, driven by high levels of deflation.

We therefore excluded the period 1919-1924 from our dataset, and ran our calculations again. We now found that the Mehra-Prescott method yielded an equity premium puzzle of 4.37%, and that the test of Kocherlakota required an alpha of 8.0 to explain the size of the observed equity premium. This clearly demonstrated the existence of the equity premium puzzle in Sweden for the revised period 1925-2010, and the finding proved to be robust to various changes in the setup. This included changes in the choice of sample period, which provides further support to the exclusion of the anomalous years before 1925.

We now turn to our analysis of myopic loss aversion, as a possible solution to the observed equity premium puzzle in Sweden for the period 1925-2010.

11. Myopic Loss Aversion in Sweden

Before reporting on our results for using myopic loss aversion as a potential solution to the equity premium puzzle in Sweden, we turn to a brief description of the mechanics behind our methodology.

11.1 Methodology

Recall from chapter 8, that the study of Benartzi & Thaler (1995) involved three main steps, namely 1) finding the implied evaluation horizon, 2) establishing the optimal asset allocation, and 3) finding the implied equity premium for varying horizons.

As an important note, we have chosen to make a small divergence from the method applied by Benartzi & Thaler (1995). In their original study, the authors used bootstrapping to simulate their return series. We have however chosen to base our calculations on the actual empirically observed return series, partly based on the fact that we observed autocorrelation in our return series. For at discussion on our motivation behind this alteration, as well as our test of autocorrelation, please refer to appendix E.

11.1.1 Evaluation horizon

To find the implied evaluation horizon of Swedish investors, we establish the prospective utility of holding stocks versus bonds based on historical data.

Calculating prospective utility involves several computationally heavy steps, so we employ VBA programming in Excel (see appendix F) to calculate the individual utility values.

In the following section, we will explain the mechanics for each of these steps:

First the monthly returns for stocks and bonds are set up in respective columns, and from these the monthly portfolio returns may be calculated, based on the relevant allocation between the two assets.

As we wish to calculate prospect utility for different evaluation horizons H, we use the monthly portfolio returns to compute compounded returns over the relevant horizon. If we denote the monthly returns r_i, we can calculate each of our compounded H-month returns x_i as follows:

$$(\mathbf{11.1}) \qquad x_i = \prod_{i=1}^{i=H}(1 + r_i) - 1$$

As noted earlier, prospect theory dictates that positive and negative outcomes of the same magnitude do not yield the same value.

To find the utility obtained from each of our H-month returns x_i, we therefore employ the value function from chapter 6:

$$(6.2) \qquad v(x) = \begin{cases} x^a & if \ x \geq 0 \\ -\lambda(-x)^\beta & if \ x < 0 \end{cases}$$

As we reported earlier, Tversky & Kahneman (1992) provided parameters that were used to describe the representative decision maker with prospective preferences. We consequently use the value of 2.25 for λ and 0.88 for both α and β[35], which enables us to find the perceived value of each H-month return, namely $v(x_i)$.

Once the relevant series of portfolio returns is found, the returns are sorted in ascending order. Each outcome $v(x_i)$ is denoted with a rank, i starting from 0 for the worst return and ending with $n - 1$ for the best return.

As $v(x_i)$ represents a possible outcome, the next step is to find the probability p of seeing each these outcomes.

We let p_i denote the probability of seeing a return that is at least as good as $v(x_i)$ and let p_i^* be the probability of obtaining a return that is strictly better than $v(x_i)$.

For the domain of gains we therefore note that:

$$(11.2) \qquad p_i^+ = \frac{(n-i)}{n} \qquad\qquad p_i^{*+} = \frac{(n-i-1)}{n}$$

And for the domain of losses we see that:

$$(11.3) \qquad p_i^- = \frac{(i+1)}{n} \qquad\qquad p_i^{*-} = \frac{i}{n}$$

Once we have defined the individual probabilities, we are able to find the individual decision weights π_i that must be attached to $v(x_i)$. We do so, by applying the following formulas:

$$(11.4) \qquad \pi_i^+ = w^+(p_i) - w^+(p_i^*) \qquad\qquad \pi_i^- = w^-(p_i) - w^-(p_i^*)$$

---

[35] Recall that within the framework of Benartzi & Thaler, α and β refer to the curvature of the prospective value function for positive and negative outcomes respectively. They should be confused with the α and β applied by Mehra & Prescott.

Recall that w is the nonlinear transform of the cumulative distribution of entire returns series. In chapter 6 we saw that these transforms were calculated separately for gains and losses, with the following functions:

$$(6.5) \qquad w^+(p_i) = \frac{p_i^\gamma}{\left(p_i^\gamma + (1-p_i)^\gamma\right)^{\frac{1}{\gamma}}} \qquad\qquad w^-(p_i) = \frac{p_i^\delta}{(p_i^\delta + (1-p_i)^\delta)^{1/\delta}}$$

From Tversky & Kahneman (1992), we use the values 0.61 and 0.69 for γ and δ respectively.

Having calculated $w(p_i)$ and $w(p_i^*)$ for each outcome, we can indentify each of the individual decision weights π_i and attach them to the outcomes $v(x_i)$. Now we have the impact of each outcome, and by summing these we can calculate the overall prospective utility V with the value function from chapter 6:

$$(6.6) \qquad\qquad V = \sum_{i=-m}^{n} \pi_i v(x_i)$$

To find the implied evaluation horizon, we must repeat this calculation for portfolios consisting of 100% stocks and 100% risk-free assets for varying time horizons. We start by setting the horizon $H = 0$, and then increasing it incrementally one month at a time. By plotting the prospective utility of the two portfolios against the different horizons, we can locate where the functions intersect, which indicates the horizon where the portfolios are equally attractive to the investor. In other words, we identify the implied evaluation horizon.

11.1.2 Optimal asset allocation

Just as Benartzi & Thaler (1995), we seek to validate our implied evaluation horizon by determining the optimal asset allocation for the horizon in question. As in the previous section, we calculate the prospective utility of holding an asset portfolio for a specific evaluation horizon. We take the implied evaluation horizon as a given, which allows us to calculate the utility of holding portfolios with mixed asset allocations rather than just holding a pure stock or a pure bond portfolio.

We calculate the utility of holding portfolios with stocks representing between 0 - 100% of the investment in increments of 5%. Plotting the prospective utility as a function of asset allocation, as done by Benartzi & Thaler (1995) in Figure 8.1 will allow us to determine the amount that should optimally be invested in stocks compared to bonds, by a value maximizing investor.

Once we have determined the optimal asset allocation, we evaluate our findings based on actual observed asset allocations within Sweden.

11.1.3 Implied equity premium

The final step is to find the implied equity premium, and compare it with our results from chapter 10. The idea is to find the point where prospective utility of investing in bonds is equal to the utility of investing in stocks for a given evaluation horizon. To do so, a small premium[36] must be added to each of the monthly bond returns, until an investor with prospect theory preferences would be indifferent between holding stocks or bonds.

Effectively, we use VBA programming (appendix E) to calculate the prospective utility of holding either 100% stocks or 100% risk-free assets for varying evaluation horizons. We use the Solver function in Excel to find the monthly premium that needs to be added to the bond returns, by minimizing the difference between prospective utility from bonds and stocks, using the premium as the variable input.

Once the monthly premium has been established for a particular evaluation horizon, it is added to the risk-free series. Then we compound the monthly returns for bonds (plus the premium) and stocks, and subsequently calculate the mean annual returns for both returns series. The difference between these mean values represents the implied equity premium.

We repeat this process for evaluation periods starting at 6 months and increasing with one-month increments until 24 months. Subsequently, the implied premia are plotted against their respective evaluation horizons, allowing us to infer on the pattern of their relationship. Finally, we run the process for 5, 10 and 20 years to compare with the results of Benartzi & Thaler (1995) (see Table 8.2).

The following section will present the results of our analysis.

11.2 Myopic loss aversion as an explanation to the Swedish equity premium puzzle

The main analysis employs monthly nominal return series for Swedish stocks and 30-day T-bills from January 1925 to December 2010. Following the main results section, we will present a range of sensitivity tests with varying data input to report on the robustness of our method.
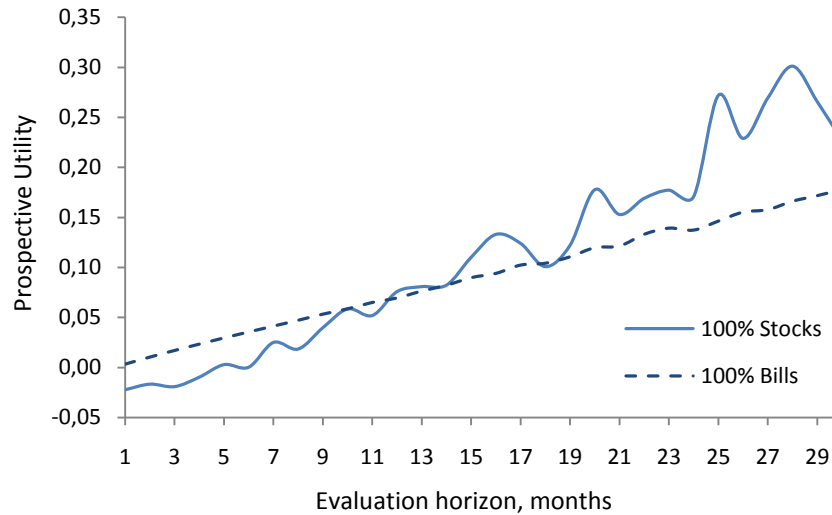
---

[36] Note that for horizons where bonds are most attractive than stocks, the premium is negative.

11.2.1 Evaluation horizon

In Figure 11.1, the prospective utility of holding a pure stock portfolio and a pure bills portfolio is plotted against evaluation horizons spanning from 1 to 30 months:

Figure 11.1: Prospective utility versus evaluation horizon

Data: Stocks vs. 30-day T-bills, nominal terms, 1925-2010.

We see that while both value functions are positively correlated with longer evaluation horizons, the prospective utility of holding a pure stock portfolio increases significantly over time. The utility of the stock portfolio is quite volatile, which results in several intersections between the two functions, namely at 12, 14 and 18 months.

As neither of the two latter intersections appear to be part of the function's upward trend, we argue that the 12-month evaluation horizon represents the period where investors would find the two portfolios equally attractive. Consequently, if we accept that investors have prospect theory preferences, the observed equity premium in Sweden over the last century can be explained by individuals evaluating their portfolios on a yearly basis.

Referring back to Table 8.1, we see that this 12-month horizon is equal to the one that Benartzi & Thaler (1995) found for nominal returns on stocks and T-bills in the US. The question is whether the period is plausible for the Swedish market.

Naturally, it does not make sense to assume that one single evaluation period would apply to all investors. Benartzi & Thaler themselves noted that even a single investor may: "*employ a combination of*

evaluation periods, with casual evaluation every quarter, a more serious evaluation annually, and evaluations associated with long-term planning every few years" (1995, p. 83).

Intuitively however, it seems likely that investors on average would evaluate their portfolios on a yearly basis. As argued by Benartzi & Thaler (1995), private investors are exposed to a range of reporting that is done once a year, including broker reports, tax filings, and pension accounts, along with other financial statements. As for institutional investors, their focus is mainly on annual reports, which may be cause for a restructuring of the underlying investments.

While our preliminary results are – at least intuitively - consistent with observed behavior, they do not offer a solid solution to the equity premium puzzle on their own. In the following section, we will therefore perform a sanity check of our initial results, by relying on the assumption that investors are value maximizers.

11.2.2 Optimal asset allocation

To validate our findings from the previous section, we take the evaluation horizon of 12 months as given, and estimate the optimal asset allocation for a yearly investment. Figure 11.2 shows a plot of prospective utility against the percentage of the portfolio that is allocated to stocks versus risk-free assets:

**Figure 11.2: Asset allocation for a 12-month horizon**
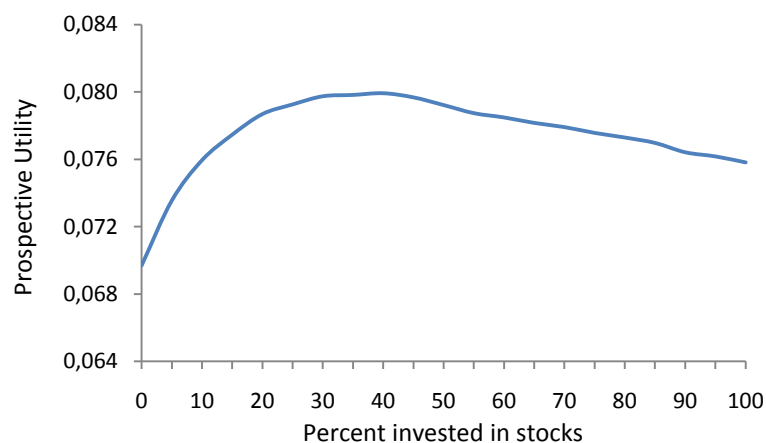
Data: Stocks vs. 30-day T-bills, nominal terms, 1925-2010.

Whereas Benartzi & Thaler (1995) found that a 30-55% stock allocation would yield approximately the same prospective utility for US investors, Figure 11.2 implies that Swedish investors could maximize their yearly investments by allocating somewhere between 25-45% to equity.

The result is encouraging for the ability of myopic loss aversion to explain the equity premium puzzle in Sweden. First of all, the shape of the function is very similar to that of Benartzi & Thaler (see Figure 8.1). Secondly, we find that our results are fairly consistent with figures for observed asset allocation in Sweden as illustrated in Figure 11.3:

Figure 11.3: Observed asset allocation in Sweden, 1990-1998
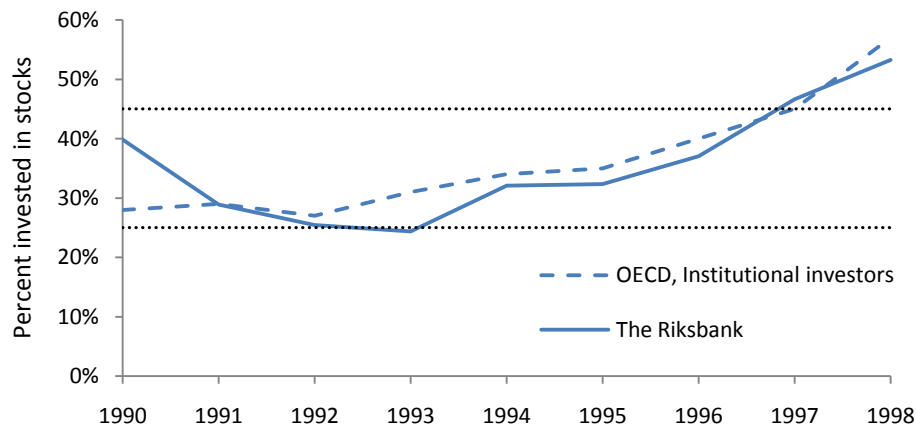
The graph illustrates observed asset allocation as reported by two sources. The dashed line is based on figures from an OECD report on the portfolio composition of institutional investors (OECD, 2001). The solid line is based on records from statistical yearbooks from the Riksbank (1990-1998), and it reflects the proportion of the year-end market value of the Swedish stock market to amount outstanding on bonds and bills combined.

While the allocation appears fairly volatile, the average percent invested in equity was 36% over the period for both series. Furthermore, investor surveys from 2006 to 2011 showed that institutional investors in the Sweden, had allocated between 26% and 33% of their assets to shares (Kirstein A/S, 2006-2010).

In general, records on asset allocation amongst Swedish investors are not readily available, and we have not been able to find figures for the entire sample period. However, if we assume that the past twenty years are roughly representative, then our results appear to be supported.

As the final part of our initial analysis, we use our model to calculate the implied equity premium for a range of evaluation horizons.

11.2.3 Implied equity premium

Figure 11.4 shows the implied equity premium for evaluation horizons spanning from 6 to 24 months:

Figure 11.4: Implied equity premium versus evaluation horizon
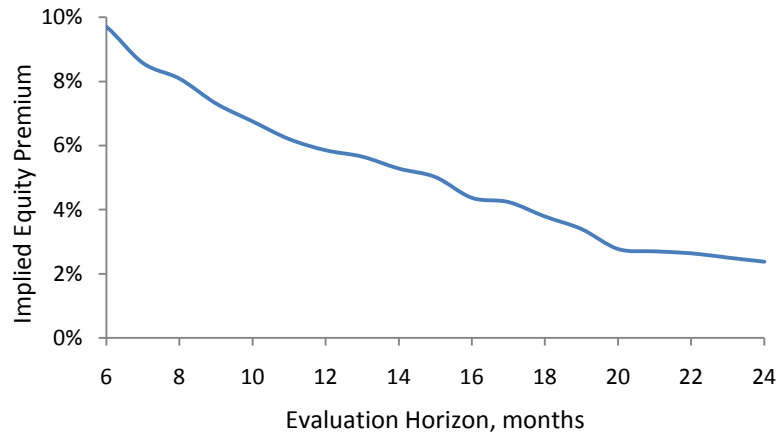
Data: Stocks vs. 30-day T-bills, nominal terms, 1925-2010.

As expected, the graph clearly illustrates the inverse relationship between equity premium and the evaluation horizon. We observe that the function is very steep for the short horizons, indicating that the investor could lower the premium considerably by employing broader bracketing in regards to his investment decisions.

Table 11.1 shows how the implied premium behaves over longer time periods:

Table 11.1: Implied equity premia

Years	1	2	5	10	20
Implied equity premium	5.9%	2.4%	1.1%	0.4%	0.1%

Data: Stocks vs. 30-day T-bills, nominal terms, 1925-2010

Comparing these results with those of Benartzi & Thaler (Table 8.2) we observe a slightly lower implied premium for a one-year horizon, with 6.5% and 5.9% comparatively. The difference becomes more prominent when we look at the long-term horizons. Where the US study found an equity premium of 4.7% for a two year-horizon, we observe an implied premium of just 2.4%. The difference may be caused by more fluctuations in the return data employed in the original study, compared to the volatility on the returns on Swedish asset.

Recall that the investor requires a high compensation for holding stocks, because the volatility will cause fluctuations in his wealth and therefore trigger his loss aversion more often. If he evaluates the

performance of his investment often, these fluctuations will cause him considerable discomfort because he is more sensitive to losses than to gains. If the investor was to evaluate once every decade instead of once a year, the likelihood of experiencing a negative return on stocks would be much lower, and he would therefore require a lower premium on his equity holdings.

Our results indicate a change in evaluation horizon from once a year to once every ten years could reduce the equity premium from 5.9% to 0.4%. For someone with a ten-year investment horizon, the price of evaluating yearly is therefore 5.5% per year, given that the rest of the market were to retain their trading behavior. Effectively, this difference represents the implied cost of myopia.

Finally as a sanity check of our method, recall from section 10.4 that we found an observed equity premium of 5.73% in Sweden for the period 1925-2010. By employing the method of Benartzi & Thaler, we found that individuals with prospect theory preferences would require a premium of 5.9% if they evaluated their investments on a yearly basis. The similarity of these two figures offers a strong support for the ability of myopic loss aversion to explain the equity premium puzzle, with an evaluation horizon of one year. Not only have we confirmed that the equity premium is likely to have been slightly less than 6% over the period in question, but more importantly our method also serves to solidify myopic loss aversion as a solution to the equity premium puzzle outside the US.

11.3 Sensitivity analyses

So far our analyses have proved to validate the explanatory power of myopic loss aversion, as a solution to the equity premium puzzle outside the US. To further test the robustness of the method, we now turn to a number of sensitivity analyses.

11.3.1 Choice of risk free asset

As mentioned earlier, the debate regarding the proper proxy for a risk-free asset is ongoing. Benartzi & Thaler used five-year government bonds, arguing that private investors are more likely to view bonds as a risk free alternative, than the short-term government notes.

To ensure that our method is robust, we therefore ran the model with the monthly returns on 10-year government bonds, which are available for the entire period 1925-2010. Figure 11.5 shows prospective utility of holding a pure bond portfolio against the utility of only holding equity.

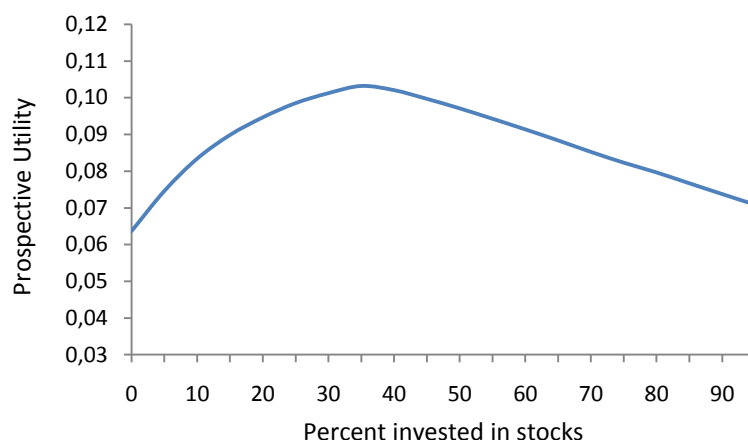## Figure 11.5: Stocks versus 10-year bonds, 1925-2010
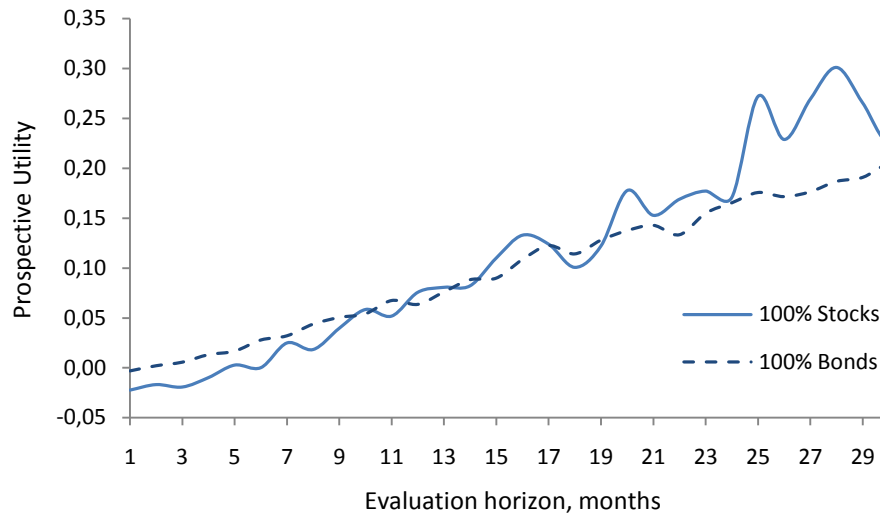
Data: Stocks vs. 10-year bonds, nominal terms, 1925-2010.

While the function for the pure stock portfolio is of course identical to the one we saw in our main analysis, the risk free function has become a bit more volatile. As the return on 10-year bonds per definition is more volatile than the return on short-term T-bills, we would expect that the function would be less well behaved, and that is exactly what we observe.

We also see that the utility function for bonds is steeper than that for bills. Again, this is what we expected, as the average nominal return on bonds was 7% per year over the period, compared to 5% per year on the 30-day T-bills. Recall that an investor with prospect theory preferences will increase his utility from investing in volatile assets, if he applies broader bracketing. If he has a longer evaluation horizon, it is less likely that he will observe a loss during one of his evaluations. Consequently, a longer horizon would allow him to extract more utility from the higher average return on the bonds, without suffering as much from the higher volatility compared to bills.

As in our main analysis, the two functions do not have one clear intersection point. In fact, the increased volatility of the bond returns has caused the functions to intersect roughly five times over the 30 different horizons. To compare with the results from our main analysis, we note that a yearly horizon appears to be fairly probable. Consequently, we try to take this horizon as a given and turn to the implied asset allocation for a one-year evaluation horizon for bonds instead of bills:

Figure 11.6: Asset allocation for a 12-month horizon

Data: Stocks vs. 10-year bonds, nominal terms, 1925-2010.

Recall that the utility function for a portfolio consisting of stocks and the short-term risk free asset had a maximum utility when 25-45% of the portfolio was allocated to equity (Figure 11.2).

For the bond series, we see a clearer pattern (Figure 11.6). The maximum utility for an investor with prospect theory preferences comes from allocating 30% of the assets to stocks and 70% to 10-year government bonds. The volatility should make it less attractive than 30-day bills, however as noted earlier, it has had a substantially higher return as well. The two effects may have evened each other out, and the results are consistent with the observed asset allocation of Swedish investors (see section 11.2.2).

We also ran a test of the implied premium under varying evaluation horizons, and it displayed the same pattern as in our main analysis, meaning that is decreased as evaluation horizons were increased.

All in all, our method appears to be robust against changes in the risk free proxy. The next section covers two sensitivity tests; nominal versus real returns, and a shorter sample period.

11.3.2 Nominal versus real returns, 1957-2010

We have been able to obtain monthly inflation data for Sweden for the period 1957-2010. Consequently, we perform a sensitivity analysis to see if our results change dramatically for a shorter sample period, and whether there is a difference between using nominal and real return series.

First, we run the model on stocks versus the 30-day T-bills in nominal as well as real terms:
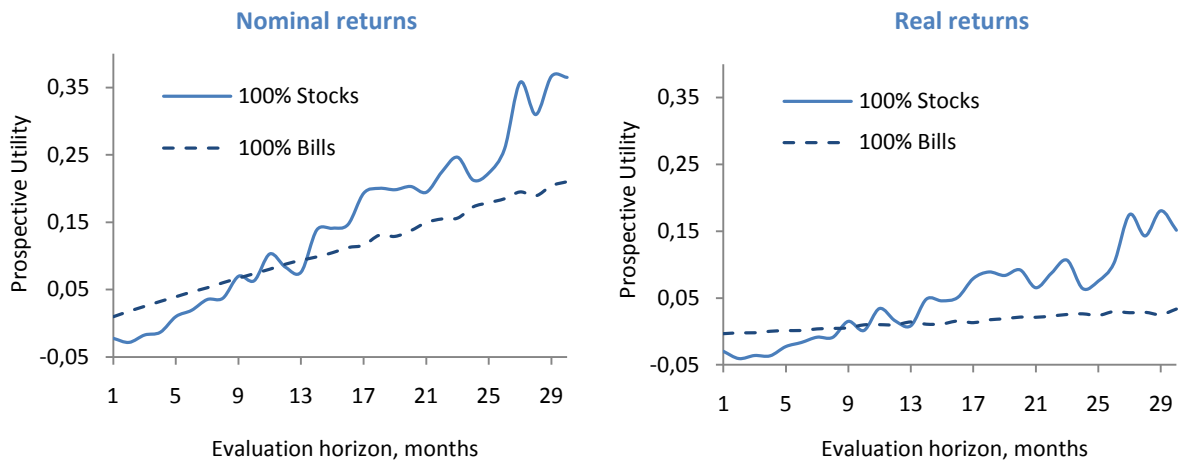
Data: Stocks vs. 30-day T-bills, 1957-2010.

Nominal returns, T-bills

For the *nominal* returns, the functions appear very similar to the ones we observed over the entire period in our main analysis. The utility from bills increase steadily along with the evaluation horizon, and the function for equity is fairly volatile with a steeply increasing pattern.

The two functions intersect three times at 9, 11, and 13 months respectively, which all correspond reasonably well with the yearly evaluation horizon observed initially. If we assume that a horizon of 11 months represents the period where investors would find the two portfolios equally attractive, we find that the investor would optimize his investment by allocating 25-35% of his portfolio to stocks. Again, these results support our initial findings.

Real returns, T-bills

As we would expect, the functions that illustrate the *real* return series appear flatter than their nominal counterparts. When the monthly returns on short-term notes are corrected for inflation, many of them become negative. Because the calculation of prospect utility weighs losses more than twice as much as gains, we therefore see a much more subdued function for the T-bills.

Again, we see intersections a 9, 11, and 13 months, however the function's overall upward trend seems to indicate that an investor looking at real returns would be indifferent between investing in stocks and bills at a horizon of roughly 10 months. If we take this period as a given, we find that the optimal asset allocation assigns 30-35% to stocks, which again seems reasonable.

Next we turn to the results obtained from running the model for the short period with stocks and 10-year bonds in nominal as well as real terms:

Figure 11.8: Stocks versus bonds, nominal versus real terms, 1957-2010
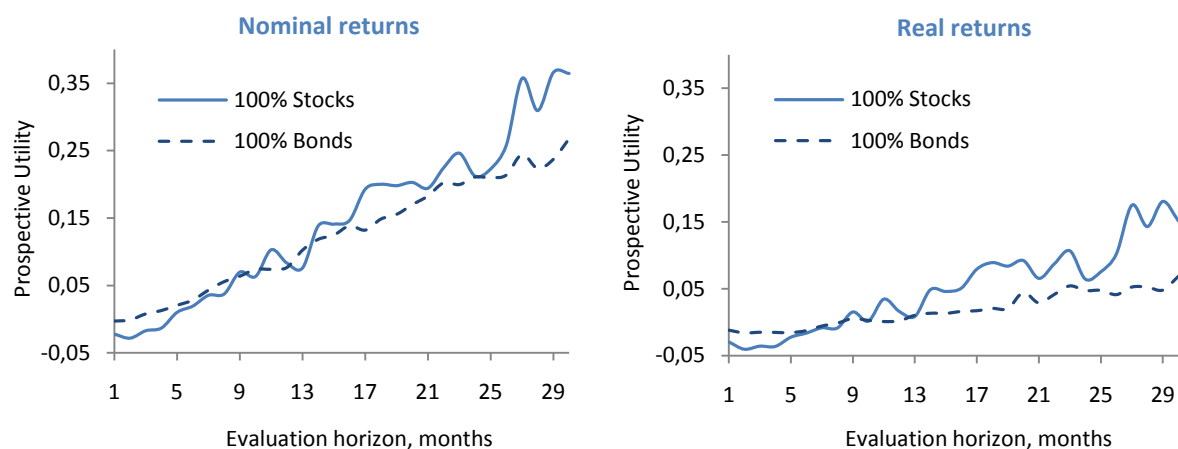
Data: Stocks vs. 10-year bonds, 1957-2010.

Nominal returns, 10-year bonds

Again, the utility function for *nominal* returns on 10-year bonds has more fluctuations than that of the short-term bill, which makes it more difficult to make exact inferences about the implied evaluation horizon. While we do see the functions touching at 16 and 24 months, the actual intersections again occur at 9, 11, and 13 months.

If we assume an evaluation horizon of 11 months, the investor would maximize his utility by allocating 35-40% of his portfolio to stocks, which again corresponds with real world observations.

Real returns, 10-year bonds

Finally, Figure 11.8 also illustrates the prospective utility of holding stocks and bonds in *real* terms. As expected, the function for bonds is considerably flatter than it was in nominal terms, and it also appears slightly steeper than the function for real bill returns (Figure 11.7).

The two real return functions intersect at 9 and 11, and they touch at the 13 month evaluation horizon. Just as we inferred for stocks versus bills in real terms, we argue that a 10 month evaluation horizon appears probable, and taking this as a given, we find an optimal allocation to stocks of 30%.

Table 11.2 summarizes our results of the main analysis, as well as the sensitivity tests performed so far:

Table 11.2: Summary of results

	1925-2010		1957-2010, nominal		1957-2010, real	
Stocks versus:	T-bills	Bonds	T-bills	Bonds	T-bills	Bonds
Evaluation horizon, months	12	12	11	11	10	10
Stock allocation	25-45%	30%	25-35%	35-40%	30-35%	30%

While our results in the last section were slightly less conclusive than our previous analysis, we do not find that any of them invalidate neither our initial findings nor our main hypothesis. Having tested how sensitive our model is to changes in data input, we now turn to a discussion of one of the main assumptions behind the model.

11.3.3 Sensitivity of the loss aversion parameter

As a final assessment of our model's robustness, we turn our focus to the aggravation parameter. In our main analysis, we employed a fixed value of $\lambda = 2.25$ as done by Benartzi & Thaler (1995). Recall that the value was established via experimental evidence, and presented in the updated cumulative version of prospect theory (Tversky & Kahneman, 1992).

There has not however been a true consensus on the actual value of the loss aversion parameter. Consequently, we have tried to insert varying coefficients of loss aversion (λ) into our model, and recalculate the implied evaluation horizon and stock allocation for stocks versus bills in nominal terms as in our main analysis. The results are presented in Table 11.3:

Table 11.3: Varying coefficients of loss aversion

λ	1.5	2.0	2.25	2.5	3.0
Evaluation horizon, months	5	11	12	17	24
Stock allocation	100%	35-45%	25-45%	15-20%	10%

Data: Stocks vs. 30-day T-bills, nominal terms, 1925-2010

The results for the middle value (2.25) represent our findings from the main analysis (section 11.2). From there we have run the model with small variations in λ (2.0 and 2.5), and with more extreme values (1.5 and 3.0). A loss aversion parameter of 2 has been suggested by many authors, and the results are very encouraging. We see an evaluation horizon of 11 months, which supports the idea that investors evaluate their portfolios roughly once a year. The figures for the optimal asset allocation are also very

accommodating, with an equity weight between 35-45%, which fits well with our main results and with the observed behavior amongst Swedish investors (see section 11.2.2). Consequently, our results appear to confirm that the loss aversion parameter has been roughly 2 for the period in question, and that Tversky & Kahneman's findings seem to hold for analyzing investor behavior outside the US.

For the other degrees of loss aversion, the model behaves as we would expect it to. We see that the implied evaluation horizon falls as investors become less loss averse and stocks become more attractive because the losses experienced due to volatility hurt less. For a coefficient of 1.5, the investor would in theory invest exclusively in stocks, which is very improbable in real life.

Following this intuition, we also see that the evaluation horizon increases substantially for loss aversion parameters of 2.5 and 3.0, to every seventeenth month for the first and to every second year for the latter. The occasional negative stock returns will hurt substantially more than the corresponding positive returns, and equity becomes highly unattractive. An investor with a loss aversion parameter of 3 should only invest 10% of his portfolio in stocks, which again fits poorly with empirical observations.

In summary, the model behaves exactly as we expect it to, and we find that the results for optimal asset allocation only fit the actual observed behavior for loss aversion factors of 2.0 and 2.25. This gives further support to the existing belief that investors are hurt roughly two times more by losses than they are pleased by gains.

11.4 Chapter summary

Our main analysis showed that myopic loss aversion is able to explain the observed the equity premium puzzle in Sweden, by assuming that investors have prospect theory preferences, and that they evaluate their portfolios on a yearly basis.

We saw that the optimal asset allocation predicted by the model for this evaluation horizon was 25-45% to stocks, which fitted well with observed behavior of Swedish investors over the last twenty years. As expected, we also saw that the implied equity premium was inversely related to the length of the average investor's evaluation horizon, and our results indicated that a change in horizon from once a year, to once every ten years, could reduce the equity premium from 5.9% to 0.4%.

To test the robustness of our method, we ran the model using a different proxy for the risk-free asset, for nominal as well as real returns, and for a short sample period. All in all, we found that the model was able to endure all these input variations, without changing the overall finding.

Finally, we ran a sensitivity test with varying coefficients of loss aversion, which indicated that highly loss averse investors of ($\lambda = 3$) should more or less refrain from investing in stocks, while individuals with a loss aversion parameter of just 1.5, could maximize their investment by allocating their entire portfolio to equity. Our results supported the notion that the general Swedish investor is indeed hurt roughly two times more by losses than he is pleased by gains.

All in all, our findings served as a support of myopic loss aversion as a solution to the equity premium puzzle in Sweden. In the next and final chapter, we will conclude on the major findings of the paper.

12. Conclusion

The main goal of this paper has been to investigate the existence of an equity premium puzzle in Sweden, and to see whether myopic loss aversion could serve as a potential solution to such as puzzle.

Initially, we conducted a study of the equity premium puzzle using two different methods, and we performed a thorough review of the data employed in our analysis. Using our entire dataset from 1919-2010, we arrived at the surprising conclusion that a puzzle did not seem to be present in Sweden. Based on the compelling documentation of the equity premium puzzle in the US, and findings of the puzzle in all G7 countries, we decided to investigate the matter further.

We found that two of the early years in our dataset were highly atypical for the period, in that the negative equity premium in those years was driven by extremely high real risk-free rates. Following an investigations of Swedish economic policies at the time, we decided to exclude the period 1919-1924 from our dataset. For the revised period, we found compelling evidence of an equity premium puzzle in Sweden, using both the original method of Mehra & Prescott (1985), and the statistical reformulation of the puzzle by Kocherlakota (1996).

We tested the robustness of our method, by replacing T-bills with 10-year government bonds as the proxy for a risk-free alternative to equity investments. As expected it diminished the size of the puzzle, but both methods still clearly validated its presence.

As the result of our analysis had proven to be sensitive to the choice of sample period, we ran our calculations for seven different starting years, from 1930 to 1960, and found the puzzle to be present for all periods. For the period after 1970, we ran our model using the Kocherlakota (1996) method on quarterly data instead of annual, and still found that the puzzle persisted.

Having thoroughly confirmed the existence of an equity premium puzzle in Sweden, we went on to investigate the explanatory power of myopic loss aversion as a solution. We used the descriptive model of Benartzi & Thaler (1995), which is based on a prospective preference structure observed in experimental settings by Tversky & Kahneman (1979; 1992).

We found that the model was able to explain the puzzle, by assuming that Swedish investors are hurt twice as much by losses as they are pleased by gains, and that they evaluate their portfolios frequently, namely once a year.

Supporting this conclusion, we saw that the optimal asset allocation predicted by the model (25-45% to stocks), corresponded with observed levels of asset allocation amongst Swedish investors over the last twenty years. Furthermore, we noted that the implied premium predicted by the model was inversely related to the length of the investor's evaluation horizon, and our results indicated that a change in horizon from once a year, to once every ten years, could reduce the equity premium from 5.9% to 0.4%.

While the contribution of this thesis is mainly of academic interest, it may also have a practical relevance. One implication concerns the potential for manipulating the evaluation period, amongst fund managers. If a client was provided with less frequent information concerning the performance of his investments, or prevented from continually adjusting his portfolio, he may be induced to evaluate the return of his investments on a more aggregated basis. This would essentially serve to the investor's advantage, as he would be less inclined to cut his losses based on occasional setbacks, and thereby increase his returns in the long run.

Our contribution has focused on applying the frameworks presented by Mehra & Prescott (1985), Kocherlakota (1996), and Benartzi & Thaler (1995), regarding the equity premium puzzle and myopic loss aversion in a non-US country, namely Sweden. While our analyses provided conclusions quite similar to those found in the US, there still remains a vital task of testing the explanatory power of myopic loss aversion across international boundaries. As shown in Table 3.3, the findings of Barro (2006) suggest that the equity premium puzzle seems to be present in all seven G7 countries, so future contributions could potentially investigate how myopic loss aversion fares in terms of explaining the puzzle across these dispersed financial markets. Also, as Barro (2006) only applied the model of Mehra & Prescott (1985), further studies of the puzzle in these countries, could potentially apply the statistical method of Kocherlakota (1996) to thoroughly validate the gravity of the individual puzzles.

Finally, while our contribution has been focused on myopic loss aversion as the sole solution to the puzzle, our literature review revealed that other studies have indeed been, at least partly, successful in explaining the equity premium, though primarily in the US. An important next step could therefore be an attempt to bridge the gap between the normative and the descriptive solutions to the equity premium puzzle. Whereas for instance habit formation was introduced as a normative theory, it does have features that relate to the field of behavioral finance, as it could be interpreted as a form of reference dependence. Perhaps it could be combined with a descriptive theory such as myopic loss aversion, to provide a more nuanced solution to the equity premium puzzle.

13. Bibliography

Abel, A. B. (2002). An exploration of the effects of pessimism and doubt on asset returns. *Journal of Economic Dynamics & Control , 26*, 1075-1092.

Abel, A. B. (1990). Asset prices under habit formation and Catching up with the Joneses. *A.E.R. Papers and Proceedings , 80* (2).

Aiyagari, S. R., & Gertler, M. (1991). Asset returns with transaction costs and uninsured individual risk. *Journal of Monetary Economics , 27*, 311-331.

Allais, M. (1979). The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American School. In M. Allais, & O. Hagen, *Expected Utility Hypotheses and the Allais Paradox.* Dordrecht: Reidel.

Ang, A., Bekaert, G., & Liu, J. (2005). Why stocks may disappoint. *Journal of Financial Economics , 76*, 471-508.

Athanasoulis, S. G., & Sussmann, O. (2007). Habit formation and the Equity-Premium Puzzle: A sceptical view. *Annals of Finance* (3), 193-212.

Barberis, N., & Huang, M. (2008). The Loss Aversion/Narrow Framing Approach to the Equity Premium Puzzle. In R. Mehra, *Handbook of the Equity Risk Premium* (pp. 199-236). Elsevier BV.

Barberis, N., & Thaler, R. (2002). *A survey of behavioral finance.* Cambridge, MA: National Bureau of Economic Research.

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics* , 1-53.

Barnes, J., & Reinmuth, J. (1976). Comparing Imputed and Actual Utility Functions in a Competitive Bidding Setting. *Decision Sciences , 7*, 801-812.

Barro, R. J. (2006, August). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics* , 823-866.

Benartzi, S., & Thaler, R. (1995). Myopic Loss Aversion and the Equity Premium Puzzle. *The Quarterly Journal of Economics , 110*, 73-92.

Budescu, D. V., & Weiss, W. (1987). Reflection of transitive and intransitive preferences: a test of prospect theory. *Organizational Behavior and Human Decision Processes , 39*, 184-202.

Camerer, C., & Ho, T. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty , 8* (2), 167-196.

Campbell, J. Y. (2001). Asset pricing at the millenium. *Journal of Finance , 55*, 1515-1567.

Cohen, M., Jaffray, J. Y., & Said, T. (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational Behavior and Human Decision Processes , 39*, 1-22.

Constantinides, G. M. (1990). Habit Formation: A resolution to the Equity Premium Puzzle. *The Journal of Political Economy , 98* (3), 519-543.

Constantinides, G. M., Donaldson, J. B., & Mehra, R. (2002). Junior can't borrow: A new perspective on the equity premium puzzle. *The Quarterly Journal of Economics , 117* (1), 269-296.

Davis, S. J., & Willen, P. (2001). Income shocks, asset returns and portfolio choice. In Z. Bodie, B. Hammond, & O. Mitchell, *Innovations in Retirement Financing.* Philadelphia: University of Philadelphia Press.

DeMarzo, P. M., Kaniel, R., & Kremer, I. (2004). Diversification as a public good: Community effects in portfolio choice. *The Journal Of Finance , 59* (4), 1677-1715.

Dimson, E., Marsh, P., & Staunton, M. (2008). The worldwide equity premium: A smaller puzzle. In R. Mehra, *Handbook of the Equity Risk Premium* (pp. 469-501). Elsevier B.V.

Dimson, E., Marsh, P., Staunton, M., Holland, D., & Matthews, B. (2011). *Credit Suisse Global Investment Returns Yearbook.* Credit Suisse.

Duesenberry, J. S. (1949). *Income, saving, and the theory of consumer behavior.* Cambridge, MA: Harvard University Press.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics , 75*, 643-669.

Epstein, L. G., & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica , 57*, 937-969.

Epstein, L. G., & Zin, S. E. (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy , 99*, 263-286.

Erbas, N., & Mirakhor, A. (2007). The equity premium puzzle ambiguity aversion, and institutional quality. *IMF Working Paper* (230).

Eriksen, K. W., & Kvaløy, O. (2010). Myopic Investment Management. *Review of Finance* (14), 521-542.

Fama, E. F., & French, K. R. (2002). The equity premium. *The Journal of Finance , 57* (2), 637-659.

Fielding, D., & Stracca, L. (2007). Myopic loss aversion, disappointment aversion, and the equity premium puzzle. *Journal of Economic Behavior & Organization , 64*, 250-268.

Frennberg, P., & Hansson, B. (1992). Swedish stocks, bonds, bills and inflation (1919-1990). *Applied Financial Economics , 2*, 79-86.

Gali, J. (1994). Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices. *Journal of Money, Credit, and Banking , 26* (1), 1-8.

Giordani, P., & Söderlind, P. (2006). Is there evidence of pessimism and doubt in subjective distributions? Implications for the Equity Premium Puzzle. *Journal of Economic Dynamics & Control , 30*, 1027-1043.

Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics , 102*, 631-645.

Gneezy, U., Kapteyn, A., & Potters, J. (2003). Evaluation Periods and Asset Prices in a Market Experiment. *The Journal of Finance* , 821-837.

Gómez, J.-P., Priestley, R., & Zapatero, F. (2009). Implications of keeping-up-with-the-Joneses for the equilibrium cross section of stock returns: International evidence. *The Journal of Finance , 64* (6), 2703-2737.

Green, P. (1963). Risk Attitudes and Chemical Investment Decisions. *Chemical Engineering Progress , 59*, 35-40.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica , 59*, 667-686.

Haigh, M. S., & List, J. A. (2005). Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *The Journal of Finance* , 523-534.

Hassan, S., & van Biljon, A. (2010). The equity premium and risk free rate puzzles in a turbulent economy: evidence form 105 years of data in South Africa. *South African Journal of Economics , 78* (1), 23-39.

Heaton, J. (1993). An empirical investigation of asset pricing with time nonseparable preferences and time aggregation. *manuscript, Massachusetts Institute of Technology* .

Heaton, J. C., & Lucas, D. J. (1997). Market Frictions, Savings Behavior and Portfolio Choice. *Journal of Macroeconomic Dynamics , 1*, 76-101.

Heaton, J. C., & Lucas, D. J. (2000). Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk. *Journal of Finance , 5*.

Heaton, J., & Lucas, D. (2007). Can heterogeneity, undiversified risk, and trading frictions solve the equity premium puzzle? In Mehra, & Rajnish, *Handbook of the Equity Risk Premium* (p. Ch 13). Elsevier.

Heaton, J., & Lucas, D. J. (1996). Evaluating the effect of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy , 103* (3), 443-487.

Jang, B.-G., Koo, H. K., Liu, H., & Loewenstein, M. (2007). Liquidity premia and transaction costs. *The Journal of Finance , 67* (5), 2329-2366.

Jorion, P., & Goetzmann, W. (1999). Global stock markets in the twentieth century. *The Journal of Finance , 54* (3), 953-980.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica , 47* (2), 263-292.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *The Journal of Economic Perspectives , 5*, 193-206.

Kirstein A/S. (2006-2010). *Nordic Investor Survey.* Copenhagen.

Kliger, D., & Levit, B. (2009). Evaluation periods and asset prices: Myopic loss aversion at the financial marketplace. *Journal of Economic Behavior & Organization , 71*, 361-371.

Kocherlakota, N. R. (1996). The equity premium: It's still a puzzle. *Journal of Economic Literature , 34*.

Krantz, O., & Nilsson, C.-A. (1975). *Swedish national product 1861-1970 - new aspects on methods and measurements.* Kristianstad: Kristianstads Boktryckeri.

Kyriacou, K., Madsen, J. B., & Mase, B. (2006). Does inflation exaggerate the equity premium? *Journal of Economic Studies , 33* (5), 344-356.

Larsson, M. (1991). *En svensk ekonomisk historia 1850-1985.* Stockholm: SNS Forlag.

Lindman, H., & Lyons, J. (1978). Stimulus Complexity and Choice Inconsistency among Gambles. *Organizational Behavior & Human Performance , 21* (2), 146-159.

Locke, P. R., & Mann, S. (2005). Professional trader discipline and trade disposition. *Journal of Financial Economics , 76* (2), 401-444.

Lucas, D. J. (1994). Asset pricing with undiversifiable income risk and short sale contraints: Deepening the equity premium puzzle. *Journal of Monetary Economics , 34*, 325-341.

Lucas, R. E. (1978). Asset Prices in an Exchange Economy. *Econometrica , 46* (6), 1429-1445.

Madsen, J. B., & Dzhumashev, R. (2009). The equity premium puzzle and the ex post bias. *Applied Financial Economics , 19*, 157-174.

Mankiw, N. G. (1986). The equity premium and the concentration of aggregate stocks. *Journal of Financial Economics , 17*, 211-219.

Mehra, R. (2008). *Handbook of the Equity Risk Premium.* Amsterdam: Elsevier.

Mehra, R., & Prescott, E. C. (1985). The Equity Premium - A Puzzle. *Journal of Monetary Economics , 15*, 145-161.

Mehra, R., & Prescott, E. C. (2003). The Equity Premium in Retrospect. *Working paper, National Bureau of Economic Research , NBER Working Paper No. 9525*.

OECD. (2001). *Recent Trends: Institutional Investors Statistics.*

Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization , 3*, 323-343.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice Bracketing. *Journal of Risk and Uncertainty , 19*, 171-197.

Samuelson, P. A. (1963). Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia , XCVIII*, 108-13.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica , 57*, 571-587.

Schön, L. (2007). *En modern svensk ekonomisk historia.* Stockholm: SNS Forlag.

Swalm, R. O. (1966). Utility Theory - Insights into Risk Taking. *Harvard Business Review , 44*, 123-136.

Telmer, C. I. (1993). Asset-pricing puzzles and incomplete markets. *The Journal of Finance , 48*, 1803-1832.

Thaler, R. H. (1999). Mental Accounting Matters. *Journal of Behavioral Decision Making , 12* (3), 183-206.

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Ecocnomics , 112* (2), 647-661.

Thaler, R. (1980). Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization* , 39-60.

The Riksbank. (1990-1998). *Statistical Yearbook.* Stockholm: Norstedts Tryckeri.

The Riksbank. (2008). Stock prices, returns and bonds 1856-2006. (D. Waldenström, Ed.) Stockholm.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review , 76* (1), 31-48.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty , 5*, 297-323.

Tversky, A., & Kahneman, D. (2000). *Choices, Values and Frames* (8th ed.). Cambridge, New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1991). Loss Aversion and Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics , 106* (4), 1039-1061.

Tversky, A., & Kahneman, D. (1986). Rational Choice and the Framing of Decisions. *Journal of Business , 59* (4), 251-278.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science , 211* (4481), 453-458.

Wehrung, D. A. (1989). Risk taking over gains and losses: A study of oil executives. *Annals of Operations Research , 19*, 115-139.

Weil, P. (1992). Equilibrium asset prices with undiversifiable labor income risk. *Journal of Economic Dynamics and Control , 16*, 769-790.

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics , 24*, 401-421.

Weymark, J. A. (1981). Generalized Gini Inequality Indices. *Mathematical Social Sciences* , 409-430.

Yaari, M. E. (1987). The Dual Theory of Choice Under Risk. *Econometrica , 55*, 95-115.

14. Appendices

14.1 Appendix A – The isoelastic utility function

The Arrow-Pratt coefficient of relative risk aversion (CRRA) is defined as $R(c)$

$$(\mathbf{A.1}) \qquad R(c) = \frac{-c_t U''(c_t)}{U'(c_t)}$$

Mehra & Prescott assume investor preferences that behave in accordance with the isoelastic utility function stated in equation (2.2)

$$(\mathbf{2.2}) \qquad U(c, \alpha) = \frac{c^{1-\alpha} - 1}{1 - \alpha}, \qquad 0 < \alpha < \infty$$

Its first and second derivatives are

$$(\mathbf{A.3}) \qquad U'(c_t) = (1 - \alpha)c^{-\alpha}\frac{1}{1 - \alpha} = c_t^{-\alpha}$$

$$(\mathbf{A.4}) \qquad U''(c_t) = -\alpha c_t^{-\alpha-1}$$

It can then be shown that the coefficient of relative risk aversion is a constant, as

$$(\mathbf{A.5}) \qquad R(c) = \frac{-c_t(-\alpha c_t^{-\alpha-1})}{c_t^{-\alpha}} = \frac{\alpha c_t c_t^{-\alpha-1}}{c_t^{-\alpha}} = \frac{\alpha c_t^{-\alpha}}{c_t^{-\alpha}} = \alpha$$

The elasticity of intertemporal substitution (EIS) is defined as

$$(\mathbf{A.6}) \qquad EIS = \frac{U'(c_t)}{-c_t U''(c_t)} \qquad or \qquad \frac{1}{\alpha}$$

i.e. the reciprocal of the CRRA.

14.2 Appendix B – Expected return on a unit of equity

The basic pricing relationship is stated in equation (2.3)

$$(\mathbf{2.3}) \qquad \mathbf{p_t U'(c_t) = \beta E[(p_{t+1} + y_{t+1})U'(c_{t+1})]}$$

$$(\mathbf{B.2}) \qquad 1 = \beta E_t \left[\frac{(p_{t+1} + y_{t+1})U'(c_{t+1})}{p_t U'(c_t)} \right]$$

$$(\mathbf{B.3}) \qquad 1 = \beta E_t \left[\frac{p_{t+1} + y_{t+1}}{p_t} \frac{U'(c_{t+1})}{U'(c_t)} \right]$$

Defining return on equity as $R_{e,t+1} = \frac{p_{t+1} y_{t+1}}{p_t}$ yields

$$(\mathbf{B.4}) \qquad 1 = \beta E_t \left[R_{e,t+t} \frac{U'(c_{t+1})}{U'(c_t)} \right]$$

As the return on a risk-free investment at time, t+1 is known at time, t, the equivalent relationship for the risk-free asset is

$$(\mathbf{B.5}) \qquad 1 = \beta E_t \left[\frac{U'(c_{t+1})}{U'(c_t)} \right] R_{f,t+t}$$

Inserting the expectations for each of the parts in equation (B.4) yields

$$(\mathbf{B.6}) \qquad 1 = \beta \left[E_t \left(\frac{U'(c_{t+1})}{U'(c_t)} \right) E_t(R_{e,t+1}) + Cov\left(\frac{U'_{t+1}}{U'_t}, R_{e,t+1} \right) \right]$$

$$(\mathbf{B.7}) \qquad 1 = \beta E_t \left(\frac{U'(c_{t+1})}{U'(c_t)} \right) E_t(R_{e,t+1}) + \beta Cov\left(\frac{U'_{t+1}}{U'_t}, R_{e,t+1} \right)$$

Rearranging the relationship for the risk-free asset in equation (B.5)

$$(\mathbf{B.8}) \qquad 1 = \beta E_t \left(\frac{U'(c_{t+1})}{U'(c_t)} \right) R_{f,t+t} \quad <=> \quad \frac{1}{R_{f,t+t}} = \beta E_t \left(\frac{U'(c_{t+1})}{U'(c_t)} \right)$$

Inserting this in equation (B.7)

$$(\mathbf{B.9}) \qquad 1 = \frac{1}{R_{f,t+t}} E_t(R_{e,t+1}) + \beta Cov\left(\frac{U'_{t+1}}{U'_t}, R_{e,t+1} \right)$$

$$(\mathbf{B.10}) \qquad R_{f,t+t} = E_t(R_{e,t+1}) + R_{f,t+t} \beta Cov\left(\frac{U'_{t+1}}{U'_t}, R_{e,t+1} \right)$$

Rearranging equation (B.5) once again

$$\textbf{(B. 11)} \quad 1 = \beta E_t \left(\frac{U'(c_{t+1})}{U'(c_t)} \right) R_{f,t+t} \quad <=> \quad R_{f,t+t} = \frac{1}{\beta E_t \left(\frac{U'(c_{t+1})}{U'(c_t)} \right)}$$

$$<=> \quad R_{f,t+t} = \frac{1}{\beta} E_t \left(\frac{U'(c_t)}{U'(c_{t+1})} \right)$$

Inserting this is equation (B.10)

$$\textbf{(B. 12)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + \frac{1}{\beta} E_t \left(\frac{U'(c_t)}{U'(c_{t+1})} \right) \beta Cov \left(\frac{U'(c_{t+1})}{U'(c_t)}, R_{e,t+1} \right)$$

$$\textbf{(B. 13)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + E_t \left(\frac{U'(c_t)}{U'(c_{t+1})} \right) Cov \left(\frac{U'(c_{t+1})}{U'(c_t)}, R_{e,t+1} \right)$$

Marginal utility, $U'(c_t)$ is known at time t:

$$\textbf{(B. 14)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + \frac{U'(c_t)}{E[U'(c_{t+1})]} Cov \left(\frac{U'(c_{t+1})}{U'(c_t)}, R_{e,t+1} \right)$$

Asserting that $Cov(aY, bX) = abCov(Y, X)$

$$\textbf{(B. 15)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + \frac{U'(c_t)}{E[U'(c_{t+1})]} \frac{1}{U'(c_t)} Cov(U'(c_{t+1}), R_{e,t+1})$$

$$\textbf{(B. 16)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + \frac{1}{E[U'(c_{t+1})]} U'(c_t) \frac{1}{U'(c_t)} Cov(U'(c_{t+1}), R_{e,t+1})$$

$$\textbf{(B. 17)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + \frac{1}{E[U'(c_{t+1})]} U'(c_t) \frac{1}{U'(c_t)} Cov(U'(c_{t+1}), R_{e,t+1})$$

$$\textbf{(B. 18)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + Cov \left(\frac{1}{E[U'(c_{t+1})]} U'(c_{t+1}), \frac{1}{E[U'(c_{t+1})]} R_{e,t+1} \right)$$

$$\textbf{(B. 19)} \quad R_{f,t+t} = E_t(R_{e,t+1}) + Cov \left(\frac{U'(c_{t+1}), R_{e,t+1}}{E[U'(c_{t+1})]} \right)$$

$$\textbf{(B. 20)} \quad E_t(R_{e,t+1}) = R_{f,t+t} - Cov \left(\frac{U'(c_{t+1}), R_{e,t+1}}{E[U'(c_{t+1})]} \right)$$

Which yields **equation (2.4):**

$$\textbf{(2. 4)} \quad \boldsymbol{E_t(R_{e,t+1}) = R_{f,t+t} + Cov \left(\frac{-U'(c_{t+1}), R_{e,t+1}}{E[U'(c_{t+1})]} \right)}$$

This is a somewhat more intuitive representation than (B.20). $Cov[c_t, R_{e,t}]$ is positive, as high equity returns give high consumption. When consumption c_t, is high then marginal utility of consumption $U'(c_t)$ is low, so $Cov[U'(c_t), R_{e,t}]$ is negative. The expected return on equity from equation (2.4) therefore has the intuitive interpretation as the risk-free rate plus a risk-premium depending on the asset return's covariance with consumption.

14.3 Appendix C – Introducing growth in the asset pricing equation

The basic pricing relationship is stated in equation (2.3):

$$(\mathbf{2.3}) \qquad \boldsymbol{p_t U'(c_t) = \beta E[(p_{t+1} + y_{t+1})U'(c_{t+1})]}$$

$$(\mathbf{C.2}) \qquad p_t = \frac{1}{U'(c_t)}\beta E[(p_{t+1} + y_{t+1})U'(c_{t+1})]$$

$$(\mathbf{C.3}) \qquad p_t = \beta E\left[(p_{t+1} + y_{t+1})U'(c_{t+1})\frac{1}{U'(c_t)}\right]$$

$$(\mathbf{C.4}) \qquad p_t = \beta E\left[(p_{t+1} + y_{t+1})\frac{U'(c_{t+1})}{U'(c_t)}\right]$$

From equation (A.3) we know that $U'(c_t) = c_t^{-\alpha}$

$$(\mathbf{C.5}) \qquad p_t = \beta E\left[(p_{t+1} + y_{t+1})\frac{c_{t+1}^{-\alpha}}{c_t^{-\alpha}}\right]$$

We introduce consumption growth x_{t+1} defined as $\frac{c_{t+1}}{c_t}$

$$(\mathbf{C.6}) \qquad p_t = \beta E[(p_{t+1} + y_{t+1})x_{t+1}^{-\alpha}]$$

Mehra & Prescott assume that price p_t is a linear function of dividends y_t

$$(\mathbf{C.7}) \qquad p_t = wy_t$$

Recall that return on equity $R_{e,t+1} = \frac{p_{t+1} + y_{t+1}}{p_t}$. Combining this with equation (C.7) yields

$$(\mathbf{C.8}) \qquad R_{e,t+1} = \frac{wy_{t+1} + y_{t+1}}{wy_t} = \frac{y_{t+1}(w+1)}{wy_t} = \frac{y_{t+1}}{y_t}\frac{w+1}{w}$$

In equilibrium consumption c_t must equal asset payoff y_t, so $\frac{y_{t+1}}{y_t} = \frac{c_{t+1}}{c_t}$ = consumption growth, x_{t+1}

$$(\mathbf{C.9}) \qquad R_{e,t+1} = x_{t+1}\frac{w+1}{w}$$

The expected equity return is therefore

$$(\mathbf{C.9}) \qquad E(R_{e,t+1}) = E(x_{t+1})\frac{w+1}{w}$$

Combining equation (C.6) with (C.7) to determine $\frac{w+1}{1}$

$$(\mathbf{C.10}) \qquad wy_t = \beta E[(wy_{t+1} + y_{t+1})x_{t+1}^{-\alpha}]$$

$$(\mathbf{C.11}) \qquad wy_t = \beta E[(w+1)y_{t+1}x_{t+1}^{-\alpha}]$$

$(\textbf{C. 12})\quad w = \beta E\left[(w + 1)\dfrac{y_{t+1}}{y_t}x_{t+1}^{-\alpha}\right]$

Substituting with the expression for consumption growth x_{t+1} (as $\dfrac{y_{t+1}}{y_t} = \dfrac{c_{t+1}}{c_t}$)

$(\textbf{C. 13})\quad w = \beta E[(w + 1)x_{t+1}x_{t+1}^{-\alpha}]$

$(\textbf{C. 14})\quad w = \beta E(w + 1)x_{t+1}^{1-\alpha}$

$(\textbf{C. 15})\quad \dfrac{w}{w + 1} = \beta E x_{t+1}^{1-\alpha}$

$(\textbf{C. 16})\quad \dfrac{w + 1}{w} = \dfrac{1}{\beta E x_{t+1}^{1-\alpha}}$

Inserting this is equation (C.9)

$(\textbf{C. 17})\quad E(R_{e,t+1}) = E(x_{t+1})\dfrac{1}{\beta E x_{t+1}^{1-\alpha}}$

This yields **equation (2.5)**

$(\textbf{2. 5})\quad E(R_{e,t+1}) = \dfrac{E_t(x_{t+1})}{\beta E_t(x_{t+1}^{1-\alpha})}$

The equivalent expression for the risk-free asset can be found using the pricing relationship in equation (C.6)

$(\textbf{C. 6})\quad p_t = \beta E[(p_{t+1} + y_{t+1})x_{t+1}^{-\alpha}]$

The price of the risk-free asset is defined as q

$(\textbf{C. 19})\quad q_t = \beta E[(p_{t+1} + y_{t+1})x_{t+1}^{-\alpha}]$

The price of the asset at time *t+1* is known, $q_{t+1} = p_{t+1} + y_{t+1} = 1$

$(\textbf{C. 20})\quad q_t = \beta E[1x_{t+1}^{-\alpha}]$

Equivalent to return on equity, the risk-free return is defined as $R_{f,t+1} = \dfrac{p_{t+1} + y_{t+1}}{p_t}$ but as $p_{t+1}+y_{t+1} = 1$, it reduces to $\dfrac{1}{q_1}$. Combining this with equation (C.20) yields **equation (2.6)**

$(\textbf{2. 6})\quad R_{f,t+1} = \dfrac{1}{\beta E[1x_{t+1}^{-\alpha}]}$

Mehra & Prescott assume that growth rate of consumption x_t and dividends y_t are lognormally distributed. As the mean of a lognormal distribution, $\bar{x} = e^{\mu_x + 1/2\sigma^2}$, the expected growth rate must be

$(\textbf{C. 22})\quad E(x_{t+1}) = e^{\mu_x + 1/2\sigma_x^2}$

and

(C. 23) $\quad E(x_{t+1}^{1-\alpha}) = (e^{\mu_x+1/2\sigma_x^2})^{1-\alpha} = e^{(1-\alpha)\mu_x+(1-\alpha)^2 1/2\sigma_x^2} \quad as \quad (x^a)^b = x^{ab}$

where $\mu_x = E[\ln(x_t)]$, $\sigma_x^2 = Var[\ln(x_t)]$, and $\ln(x_t)$ is the continuously compounded growth rate of consumption. Inserting these in the expression for expected return on equity, equation (2.5)

(2. 5) $\quad E(R_{e,t+1}) = \dfrac{E_t(x_{t+1})}{\beta E_t(x_{t+1}^{1-\alpha})}$

gives **equation (2.7)**

(2. 7) $\quad E(R_{e,t+1}) = \dfrac{e^{\mu_x+1/2\sigma_x^2}}{\beta e^{(1-\alpha)\mu_x+(1-\alpha)^2 1/2\sigma_x^2}}$

Asserting that $\ln\left(\dfrac{x}{y}\right) = \ln(x) - \ln(y)$ allows (2.7) to be rearranged to

(C. 25) $\quad \ln[E(R_{e,t+1})] = \ln(e^{\mu_x+1/2\sigma_x^2}) - \ln(\beta e^{(1-\alpha)\mu_x+(1-\alpha)^2 1/2\sigma_x^2})$

Asserting that $\ln(xy) = \ln(x) + \ln(y)$

(C. 26) $\quad \ln[E(R_{e,t+1})] = \mu_x + 1/2\sigma_x^2 - [\ln(\beta) + (1-\alpha)\mu_x + (1-\alpha)^2 1/2\sigma_x^2]$

(C. 27) $\quad \ln[E(R_{e,t+1})] = \mu_x + \dfrac{1}{2}\sigma_x^2 - \ln(\beta) - \mu_x(1-\alpha) - \dfrac{1}{2}(1-\alpha)^2\sigma_x^2$

(C. 28) $\quad \ln[E(R_{e,t+1})] = \mu_x + \dfrac{1}{2}\sigma_x^2 - \ln(\beta) - \mu_x + \alpha\mu_x - \dfrac{1}{2}(1-\alpha)^2\sigma_x^2$

(C. 29) $\quad \ln[E(R_{e,t+1})] = \mu_x + \dfrac{1}{2}\sigma_x^2 - \ln(\beta) - \mu_x + \alpha\mu_x - \dfrac{1}{2}\sigma_x^2(\alpha^2 - 2\alpha + 1)$

(C. 30) $\quad \ln[E(R_{e,t+1})] = \mu_x + \dfrac{1}{2}\sigma_x^2 - \ln(\beta) - \mu_x + \alpha\mu_x - \dfrac{1}{2}\sigma_x^2\alpha^2 + \alpha\sigma_x^2 - \dfrac{1}{2}\sigma_x^2$

which reduces to **equation (2.8)**

(2. 8) $\quad \ln[E(R_{e,t+1})] = -\ln\beta + \alpha\mu_x - \dfrac{1}{2}\alpha^2\sigma_x^2 + \alpha\sigma_x^2$

The equivalent expression for the risk-free rate is found using the same property of the lognormal distribution as in (C.22):

(C. 32) $\quad E(x_{t+1}^{-\alpha}) = e^{-\alpha\mu_x+1/2\alpha^2\sigma_x^2}$

Inserting this in the expression for the risk-free return, equation (2.6)

(2. 6) $\quad R_{f,t+1} = \dfrac{1}{\beta E[1 x_{t+1}^{-\alpha}]}$

yields **equation (2.9)**

$$(2.9) \quad R_{f,t+1} = \frac{1}{\beta e^{-\alpha\mu_x + 1/2\alpha^2\sigma_x^2}}$$

which can be rewritten as

$$(C.34) \quad \ln(R_{f,t+1}) = \ln(1) - [\ln(\beta) - \alpha\mu_x - 1/2\alpha^2\sigma_x^2]$$

to give **equation (2.10)**

$$(2.10) \quad \ln(R_{f,t+1}) = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2$$

The equity premium is therefore found by subtracting the risk-free rate of equation (2.10) from the expected equity return of equation (2.8)

$$(C.36) \quad \ln[E(R_{e,t+1})] - \ln(R_{e,t+1}) = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2 + \alpha\sigma_x^2 - (-\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2)$$

$$(C.37) \quad \ln[E(R_{e,t+1})] - \ln(R_{e,t+1}) = -\ln\beta + \alpha\mu_x - \frac{1}{2}\alpha^2\sigma_x^2 + \alpha\sigma_x^2 + \ln\beta - \alpha\mu_x + \frac{1}{2}\alpha^2\sigma_x^2)$$

which reduces to **equation (2.11)** as the expression for the equity premium

$$(2.11) \quad \ln[E(R_{e,t+1})] - \ln(R_{e,t+1}) = \alpha\sigma_x^2$$

14.4 Appendix D – The Kocherlakota method

Kocherlakota (1996) combines the expression for the expected return on equity of equation (2.5) and the expression for the risk-free rate of equation (2.6) to arrive at equation (3.2)

$$(2.5) \quad E_t(R_{e,t+1}) = \frac{E_t(x_{t+1})}{\beta E_t(x_{t+1}^{1-\alpha})}$$

$$(2.6) \quad R_{f,t+1} = \frac{1}{\beta E_t[1 x_{t+1}^{-\alpha}]}$$

For equity we re-arrange equation (2.5)

$$(D.1) \quad E_t(R_{e,t+1}) = E_t(x_{t+1}) \frac{1}{\beta E_t(x_{t+1}^{1-\alpha})}$$

$$(D.2) \quad E_t(R_{e,t+1}) \, \beta E_t(x_{t+1}^{1-\alpha}) \, E_t(x_{t+1}^{-1}) = 1$$

Asserting that $x^m * x^n = x^{m+n}$

$$(D.3) \quad E_t(R_{e,t+1}) \, \beta E_t(x_{t+1}^{-\alpha}) = 1$$

$$(D.4) \quad \beta E_t(x_{t+1}^{-\alpha} R_{e,t+1}) = 1$$

Recall that growth rate of consumption $x_{t+1} = \frac{c_{t+1}}{c_t}$

$$(D.5) \quad \beta E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha} R_{e,t+1}\right] = 1$$

For the risk-free asset we re-arrange equation (2.6)

$$(D.6) \quad \beta E_t(x_{t+1}^{-\alpha}) \, R_{f,t+1} = 1$$

Once again we insert the expression for consumption growth

$$(D.7) \quad \beta E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right] R_{f,t+1} = 1$$

Equation (D.5) and (D.7) are then combined to yield **equation (3.1)**

$$(3.1) \quad \beta E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha} R_{e,t+1}\right] = \beta E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right] R_{f,t+1}$$

Which can be rearranged to H_0 in the Kocherlakota (1996) method

$$(D.8) \quad \beta E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha} R_{e,t+1}\right] - \beta E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right] R_{f,t+1} = 0$$

$$(D.9) \quad E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha} R_{e,t+1}\right] - E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right] R_{f,t+1} = 0$$

$(\textbf{D. 10})\quad E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right]E_t(R_{e,t+1}) - E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right]R_{f,t+1} = 0$

$(\textbf{D. 11})\quad E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\right]\left[E_t\left(R_{e,t+1}\right) - R_{f,t+1}\right] = 0$

This yields **equation (3.2)**

$(\textbf{3. 2})\quad \boldsymbol{E_t\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\alpha}\left(R_{e,t+1} - R_{f,t+1}\right)\right] = 0}$

14.5 Appendix E - Autocorrelation and a discussion of bootstrapping

The idea behind bootstrapping is that a theoretically infinite number of observations can be drawn from a limited dataset, without having to rely on parametric assumptions, e.g. a normal distribution of the data. This is done by drawing an observation from the dataset T times, and replacing the observation after each draw. A simulated dataset with the same statistical characteristics as the original data can then be constructed from a large number of draws, by putting 1/T weight on each observation.

The method requires however, that the observations in the original dataset are independent and identically distributed. This requirement proved a problem in the model of Benartzi & Thaler (1995), as return series often display serial correlation. The authors handled the issue by drawing n-month blocks of data and calculating the return on these successive months before the block was replaced in the dataset. This means that the return calculation for a portfolio, over a six-month period, was done by drawing a block of six consecutive months somewhere in their dataset, instead of finding the return over six randomly drawn months. In this way, Benartzi & Thaler (1995) were able to contain the structure of their dataset, while still being able to increase the number of data points significantly.

To test if our dataset necessitated the same simulation approach as applied by Benartzi & Thaler (1995), we tested our monthly nominal equity return series for autocorrelation using a Breusch-Godfrey's test. The result was a LM test statistic of 26.22 for first-order autocorrelation, which translates into a p-value of <0.001 for accepting H_0 (no autocorrelation). The first-order autocorrelation coefficient was found to be 0.156, as seen in Figure 14.1 below

Figure 14.1: Coefficient of autocorrelation
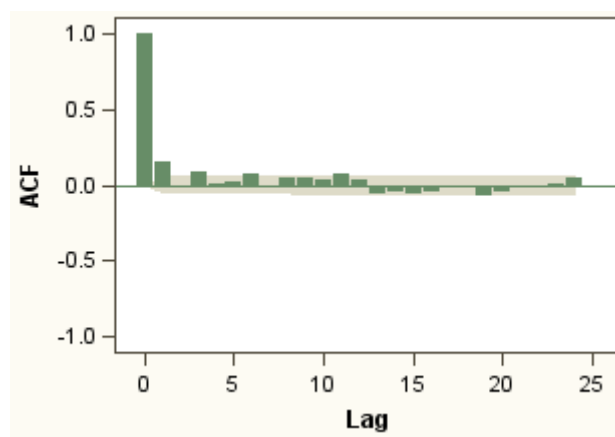
In line with the previous discussions of the importance of the choice of time period, we tested the return series starting in 1935 and 1957 for autocorrelation. Both series showed p-values < 0.001 for accepting H_0, with first-order autocorrelation coefficients of 0.147 and 0.158 respectively. Hence, we find that our data contains autocorrelation, and that the finding seems to be consistent for varying time periods. This needs to be considered in our analysis.

As stated earlier we have taken a different approach to preserving the structure of our dataset compared to Benartzi & Thaler's original study. As the practice of bootstrapping is not an uncontroversial statistical method, we discussed the case with Professor Søren Feodor Nielsen[37]. We agreed that due to the length of our sample period, the amount of data points (21 years more than in Benartzi & Thaler's study) would be sufficient to show the trend that we are investigating, without having to employ data simulations. Because we wish to make inferences about the actual historical investor experience, and ensure that the structure of the data is fully preserved, we have thus chosen to use the original empirically observed return series without any statistical alterations.

---

[37] Center for Statistics, Copenhagen Business School

14.6 Appendix F – Visual Basics programming

Main function:

```
Function ProspectUtility(Dates As Range, StockReturns As Range, Stockweight As Double, BondReturns As Range,
Bondweight As Double, alpha As Double, beta As Double, gamma As Double, lambda As Double, delta As Double,
InvestmentHorizon As Integer) As Variant

Dim i, N, m As Integer
N = Dates.Rows.Count

m = Application.WorksheetFunction.RoundDown(N / InvestmentHorizon, 0)

Dim CompDates()
ReDim CompDates(m - 1)

    For i = 1 To m
        CompDates(i - 1) = Dates(InvestmentHorizon * i)
    Next i


Dim StockReturnsComphorizon()
ReDim StockReturnsComphorizon(m - 1)

    StockReturnsComphorizon = CompoundReturns(StockReturns, InvestmentHorizon, N)


Dim BondReturnsComphorizon()
ReDim BondReturnsComphorizon(m - 1)

    BondReturnsComphorizon = CompoundReturns(BondReturns, InvestmentHorizon, N)


Dim PortfolioReturns()
ReDim PortfolioReturns(m - 1, 1)

    PortfolioReturns = PortfolioReturn(CompDates, StockReturnsComphorizon, BondReturnsComphorizon, Stockweight,
    Bondweight, m)


Dim v()
ReDim v(m - 1, 1)

    For i = 0 To m - 1
        v(i, 0) = PortfolioReturns(i, 0)

    If PortfolioReturns(i, 1) >= 0 Then
        v(i, 1) = PortfolioReturns(i, 1) ^ alpha

    Else:
        v(i, 1) = -lambda * (-PortfolioReturns(i, 1)) ^ beta

    End If

    Next i
```

```vbnet
Dim Sortetv()
ReDim Sortetv(m - 1, 1)

    Sortetv = DualSorter(v, 1)


Dim P, Px, Wp, Wpx, WPdiff, Vg As Double

    i = 0
    ProspectUtility = 0

    For i = 0 To m - 1

    If Sortetv(i, 1) >= 0 Then
        P = (m - i) / m
        Px = (m - i - 1) / m

    Else
        P = (i + 1) / m
        Px = i / m

    End If


    If Sortetv(i, 1) >= 0 Then
        Wp = (P ^ gamma) / ((P ^ gamma + (1 - P) ^ gamma)) ^ (1 / gamma)
        Wpx = (Px ^ gamma) / ((Px ^ gamma + (1 - Px) ^ gamma)) ^ (1 / gamma)

    Else
        Wp = (P ^ delta) / ((P ^ delta + (1 - P) ^ delta)) ^ (1 / delta)
        Wpx = (Px ^ delta) / ((Px ^ delta + (1 - Px) ^ delta)) ^ (1 / delta)

    End If


WPdiff = Wp - Wpx

Vg = WPdiff * Sortetv(i, 1)


ProspectUtility = ProspectUtility + Vg

Next i

End Function
```

Sub functions:

```vbnet
Function CompoundReturns(Data, H As Integer, ObsCount)

Dim i, j
Dim N, m

    N = ObsCount

    m = Application.WorksheetFunction.RoundDown(N / H, 0)

Dim rtn()
ReDim rtn(m - 1)

    If N = m Then
```

```vba
        For j = 0 To m - 1
            rtn(j) = Data(j + 1)

    Next j

Else

    Dim C()
    ReDim C(H - 1)

        For j = 0 To (m - 1) * H Step H
        For i = 0 To H - 1
            C(i) = 1 + Data(j + i + 1)

        Next i

        rtn(j / H) = Application.WorksheetFunction.Product(C) - 1

    Next j

End If

CompoundReturns = rtn

End Function
```

```vba
Function  PortfolioReturn(Dates, Asset1_Returns, Asset2_Returns, Asset1_Weight As Double, Asset2_Weight As
Double, N As Integer)

Dim i As Integer

Dim Pfrtn()
ReDim Pfrtn(N - 1, 1)

    For i = 0 To N - 1
        Pfrtn(i, 0) = Dates(i)
        Pfrtn(i, 1) = Asset1_Weight * Asset1_Returns(i) + Asset2_Weight * Asset2_Returns(i)

    Next i

PortfolioReturn = Pfrtn

End Function
```

```vba
Function DualSorter(ByRef arrArray, DimensionToSort)

    Dim row, j, StartingKeyValue, StartingOtherValue, _
        NewStartingKey, NewStartingOther, _
        swap_pos, OtherDimension
    Const column = 1

    If DimensionToSort = 1 Then
        OtherDimension = 0
    ElseIf DimensionToSort = 0 Then
        OtherDimension = 1
    Else
        Response.Write "Invalid dimension for DimensionToSort: " & _
                "must be value of 1 or 0."

        Response.End
```

```vba
        End If

    For row = 0 To UBound(arrArray, column) - 1

        StartingKeyValue = arrArray(row, DimensionToSort)
        StartingOtherValue = arrArray(row, OtherDimension)

        NewStartingKey = arrArray(row, DimensionToSort)
        NewStartingOther = arrArray(row, OtherDimension)

        swap_pos = row

        For j = row + 1 To UBound(arrArray, column)
            If arrArray(j, DimensionToSort) < NewStartingKey Then
                swap_pos = j
                NewStartingKey = arrArray(j, DimensionToSort)
                NewStartingOther = arrArray(j, OtherDimension)
            End If

        Next

        If swap_pos <> row Then
            arrArray(swap_pos, DimensionToSort) = StartingKeyValue
            arrArray(swap_pos, OtherDimension) = StartingOtherValue

            arrArray(row, DimensionToSort) = NewStartingKey
            arrArray(row, OtherDimension) = NewStartingOther

        End If

    Next

DualSorter = arrArray

End Function
```