Higher-moment Stochastic Discount Factor Specifications and the Cross-Section of Asset Returns

Dengli Wang

Thesis presented to Dublin City University

in fulfilment of the requirements for the degree of

Doctor of Philosophy

Dublin City University

Business School

Supervisor: Dr. Valerio Potì

September, 2013

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have excercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed :_______________________(Dengli Wang)

ID No.:___ 56118473________ Date: ____12/09/2013_______

Acknowledgements

First and foremost, I would like to thank Dr. Valerio Potì for being my supervisor at Dublin City University. Without his constructive critiques and recommendations, this thesis would not exist. In fact, his generous patience, unfailing support and precious guidance encourage me to achieve this work. I also want to thank Dr. Cal Muckley, for his effort to bring me to Ireland. His unique approach to research and the academic advice have proven invaluable. Furthermore, I am grateful to Dr. Hiroyuki Kawakatsu, Lecturer of Econometrics, for his willingness to challenge the statistical and econometrical techniques used in this thesis.

Second, I would like to thank all sponsors and supporters of my PhD research. I thank Dublin City University Business School. Because of the scholarships, I got the opportunity to work and live an unforgettable four year in the great city of Dublin. I would also like to acknowledge DCU for being a heaven of intellectual freedom.

Sincere thanks also to my Chinese colleagues, more or less in order of appearance on the scene of my life in Dublin, Decai Jiao, Yurong Liu, Jie zhu, Yan Zhu, and Qiang Zen. Especially, Yurong Liu, your and your wife's kind and considerate help release lots of my pressure and homesick, when I was so far away from the hometown in China.

Like many other nice things that have happened since I have met my wife, the existence of this thesis owes a lot to Mrs. Tingyu Jiang. Her contagious encouragement supports me firmly during the difficult time of writing this thesis.

Finally, I would express my grateful thanks to my parents. Because of you, I have the chance to enjoy living in this world.

Abstract

The stochastic discount factor model provides a general framework for pricing assets. A suitably specified discount factor encompasses most of the theories currently in use, including the CAPM, consumption CAPM, higher-moment CAPM and their conditional versions. In this thesis, we focus on the empirical admissibility of alternative SDFs under restrictions that ensure that investors' risk-preferences are well behaved. More innovatively, we explore whether the SDF implied by the 3 and 4-moment CAPM is plausible under restrictions that are weaker than those considered by Dittmar (2002) yet sufficient to rule out implausible curvature of the representative investor's utility functions. We find that, even under these weaker restrictions, the 3 and 4-moment CAPM cannot solve well known puzzles which plague the empirical performance of extant rational asset pricing models, even though the higher order terms do generate considerable additional explanatory power. Faced with this difficulty, we then explore whether the failure to fully account for cross-sectional differences in average returns can be explained by the presence of either transaction costs or a behavioural component of the SDF, reflecting investors' systematic mistakes in processing information. We find evidence of both problems, though our analysis is not conclusive in this respect. Finally, in a more applied exercise, we apply the SDF-framework to test whether Chinese fund managers generate superior investment performance, and find that Chinese fund managers have not achieved better performance than the individual investors under either the unconditional or the conditional measure.

Table of Contents

1. Introduction and Overview

1.1. Introduction

Starting with the establishment of portfolio theory (Markowitz, 1952), financial researchers have devoted considerable efforts to the investigation of the trade-off between return and risk for traded assets. A legitimate pricing model is equivalent to a pricing function that projects uncertain future payoffs into present prices that are conditionally determined in the observable information set. Cochrane (2001) partitions asset pricing into two folds: absolute asset pricing and relative asset pricing. Absolute pricing involves pricing each asset with respect to its exposure to fundamental sources of macroeconomic risk. For example, the fundamental sources can be GDP, oil price, industry production, and consumption. Relative asset pricing is less ambitious. It aims to price an asset with reference to the prices of other assets. The CAPM (i.e. the "beta" model) and the Black-Scholes option pricing theory, respectively, are two classical examples. No matter which approaches we want to use, however, a generalized framework, namely the pricing model representation based on the stochastic discount factor (henceforth SDF) or pricing kernel, can be applied.

For example, consider the paradigm proposed by Arrow-Debreu for the study of equilibrium in financial markets. In this setup, as a fundamental mechanism to study investor's wealth allocation between consumption and investment across time and states of nature, asset pricing models focus on the relation between future payoffs and current prices. In a complete market, for each state of nature at each time date, there exists a state price. And for each traded asset, the sum of all its possible future payoffs weighted by the relevant state prices equals the current price of this asset. This approach can be renormalized in the framework of the SDF pricing model. As pointed out by, among others, Cochrane (2001) and Smith and Wickens (2002), most asset pricing models can be treated as particular versions of the SDF model. These models include the original capital asset pricing model (CAPM) introduced by Sharpe (1964), Lintner (1965) and Black (1972), the European option pricing model developed by Black and Scholes (1973), and the general equilibrium consumption-based inter-temporal capital asset pricing model of Breeden and Litzenberger (1978) and Lucas (1978). Precisely, if the SDF is linearly related to the market portfolio, then stock returns can be described by the static single-factor CAPM; If one constructs a SDF which prices an asset continuously, then it also can be applied to price the related option (Cochrane, 2001); If the economy has a representative agent with a well-defined utility function, then the SDF is related to the growth of marginal utility of aggregate consumption and stock returns can be explained under the consumption-based CAPM. Furthermore, most models proposed by the recently emerged behavioural asset pricing literature, notably the behavioural asset pricing theorem (Shefrin, 2010), also can be formalized in terms of this framework.

As pointed by Campbell (2000), the challenge for researchers is to understand the economic forces that determine the SDF, or the reward for investors' bearing particular risks. Although we have made progress in this field in the past 30 years, our understanding is still far from perfect. A number of anomalies persist in asset pricing (Lewellen & Nagel, 2006). For example we do not yet fully understand why, for many years, small stocks have outperformed large stocks (the "size effect", see, (Fama & French, 1992)), why firms with high book-to-market (B/M) ratios outperform those with low B/M ratios (the "value premium", see, (Fama & French, 1992), and (Loughran & Ritter, 1997)), why stocks with high returns in the previous years outperform those with low prior returns (‘‘momentum phenomenon’’, see, (Jegadeesh & Titman, 1993), and (Jegadeesh & Titman, 1999)), why the link between consumption growth and the equity premium or interest rates is so weak ("equity premium puzzle", see, (Mehra & Prescott, 1985), and (Campbell & Cochrane, 2000); and "risk-free rate puzzle", see, (Weil, 1989)).

Any measurable random variable can be described by its moments. For example, the mean of a random variable is its first moment, variance is its second moment centred around the mean, skewness is a centred third moment and kurtosis is a centred fourth moment. In this thesis, I make large use of the notion of moments in discussing and characterizing the multivariate distribution of equity returns. I study mainly first, second, third and fourth moments. I pay special attention to their interaction, as emphasized by modern asset pricing theory, and I discuss the portfolio, investment and risk management implications of alternative models equilibrium first moment, i.e. equilibrium mean returns.

In the next section, I define and contrast unconditional and conditional moments from the viewpoint of the task represented by the modelling of the multivariate and multi-period distribution of asset returns. In 1.3, I outline the old and new paradigms of asset returns. In Section 1.4, I specify the main research questions. In Section 1.5, I explain the motivations of this study. In Section 1.6, I outline the structure of this thesis. In Section 1.7, I provide an overview of the main theoretical results and empirical findings of this thesis and I highlight their contribution to the extant financial and econometric literature. Section 1.8 concludes.

1.2. Conditional vs. Unconditional Moments

Time series of asset returns can be seen as realizations y_1, y_2, ... y_t of a multivariate random variable y drawn from a joint probability distribution $p(y_1, y_2, ... y_t)$. Similarly, for practical modelling purposes, future returns can be seen as realizations y_{t+1} of a random variable drawn from the conditional probability distribution $p(y_{t+1}|y_1, y_2, ... y_t)$. Loosely speaking, stationary series have time invariant moments. Strictly stationary series are realizations of random variables drawn from a time invariant probability distribution and, therefore, all their moments are time invariant. Covariance stationary, also known as wide sense or weakly stationary, series have finite and time invariant first, second and cross-second moments (e.g., respectively, the mean, variance and autocovariances/autocorrelations). Thus, strictly stationary series with *finite* first and second moments are also covariance stationary but not vice versa. Since independence of two random variables refers to the possibility of writing their joint density function as the product of their marginal densities, serial independence requires that all the cross-moments between any polynomial of current and past realizations be zero. It therefore requires independence between all the moments. Formally, for any random process, and hence also for any return y_t, serial independence (i.e. independence between y_t and y_{t-i}) means that $E[g(y_t)h(y_{t-i})] = E[g(y_t)]E[h(y_{t-i})]$ for any integer i, implying that $Cov\left[g(y_t), h(y_{t-i})\right] = 0$ for any measurable function g and h and, therefore, for any cross-moment of y_t and y_{t-i}. Autocorrelation is one possible source of serial dependence in returns. It implies linear dependence of the mean of the process on past realizations, and it therefore corresponds to dependence in the first moment. More general forms of serial dependence introduce linear relations between different moments. These might appear as non-linearities in the dependence in first moments.

One way to summarize serial dependence and co-dependence between moments is to let the corresponding polynomials of returns $g(y_t)$ be determined by data generating processes similar to those commonly used for returns, e.g. auto-regressive moving averages (ARMA). Serial dependence between moments can then be modelled as dependence between polynomials of current returns and past return realizations. For example, with $g(y_t) = y_t$, the expectation of y_t conditional on its past history $\mu_{yt} = E_{t-1}(y_t)$, i.e. the first conditional centred moment of y_t, can be defined as a function of y_{t-i} ($i > 0$). Using a simple autoregressive specification, we might let $g(y_t) = a + bg(y_{t-1}) + \varepsilon_t$ or $y_t = a + bg(y_{t-1}) + \varepsilon_t$, where a and b are constants and $\varepsilon_t = y_t - \mu_t$ is a conditionally zero-mean return innovation. In this specification, the first moment is a function of the past realization of the process, i.e. $\mu_{yt} = E_{t-1}(y_t) = a + by_{t-1}$. Similarly, with $g(y_t) = \varepsilon_t^2$, the conditional expectation $\sigma_{yt}^2 = E_{t-1}(\varepsilon_t^2) = (y_t - \mu_t)^2 = (y_t - a - by_{t-1})^2$ is the conditional variance of the return process and it depends upon the past history of the latter.

Specifications like these introduce the distinction between conditional and unconditional moments and allow the former to be time-varying. In Finance, this distinction is important unless we assume that assets are held for a long period of time. In this case the relevant conditioning information set is far in the past and its influence on conditional expectations is negligible, e.g. for (stationary) series for which $E[g(y_t)]$

exists, $\lim_{k\to+\infty} E_{t-k}[g(y_t)] = E[g(y_t)]$. The distinction between conditional and unconditional variance was emphasized by Engle (1982).

If a process is covariance stationary, its unconditional moments exist and they are the mean of the conditional moments over all the possible realizations of the process itself, i.e.

$$
\begin{aligned}
E\left\{E_{t-1}\left[g(y_t)\right]\middle| t-1=1, \quad 2, \quad \ldots \quad +\infty\right\} &= E\left\{E_{t-k}\left[g(y_t)\right]\middle| k\in\left(-\infty,+\infty\right)\right\} \\
&= E\left\{E_{t-k}\left[g(y_t)\right]\middle| k=1, \quad 2, \quad \ldots \quad +\infty\right\}
\end{aligned}
. \qquad (1\text{-}1)
$$

Thus, if a process is co-variance stationary, its unconditional variance exists and is the expectation of the time-t conditional variance conditional upon all the possible realizations, i.e. letting $g(y_t) = \varepsilon_t^2$, the unconditional variance of y_t is

$$
\sigma_y^2 = E\left\{E_{t-k}\varepsilon_t^2 \middle| k=1, \quad 2, \quad \ldots. \quad +\infty\right\}. \qquad (1\text{-}2)
$$

As pointed out by Loretan and Phillips (1994), the existence of unconditional moments critically depends on the shape of the density in the tails, i.e. if the density function does not decline rapidly enough as we move away from the centre of the distribution some of the moments might not exist. For example, returns with finite conditional variance might display infinite unconditional variance. The density in the tails of a distribution, i.e. its “thickness” and the related height of the peak towards the central part of the distribution (leptokurtosis), is captured by the fourth moment, the kurtosis.

1.3. Old and New Paradigm

The last few decades have witnessed a radical transformation in the way financial theory and financial econometrics researchers model asset returns. In the old paradigm, returns were thought to be independently drawn from an underlying joint distribution with time-invariant moments and all the moments were supposed to exist. In other words, returns were assumed to be independently and identically distributed (henceforth, *i.i.d.*). This view was particularly common in the fifties and sixties and it is summarized by the random walk representation (see Malkiel (1999) for a discussion) of the asset price process with constant drift and white noise error[1]. In the random walk model of asset prices, returns have finite moments of any order and conditional and unconditional moments are the same. Normality of the error term, moreover, implies that the entire multivariate distribution of asset returns can be described by its first and second moments. This, in turn, implies that rational investors should only be concerned about the mean and variance of their portfolios, leaving no room for any role of higher moments in the portfolio optimization problem, as in Markowitz (1952) mean-variance-portfolio theory. In such a setting, broadly corresponding to the static Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965), only the first two moments of the multivariate return distribution have asset pricing implications.

This paradigm came under intense scrutiny, especially in the 1980s and 1990s. Four main issues about the distribution of asset returns drew the attention of the empirical

---

[1] Formally: $\frac{dP}{P} = \mu dt + \varepsilon\sigma\sqrt{dt}$ with $\mu = E\left(\frac{dP}{P}\right)$, $\sigma^2 = E\left(\frac{dP}{P} - \mu\right)^2$, P is the price, ε *i.i.d* and $E(\varepsilon) = 0$. Here $E()$ represents both the unconditional and conditional expectation operator.

financial literature, namely whether financial series are independently distributed, whether they are identically distributed over time, whether all the moments of the asset returns distribution exist and whether returns are normally distributed. A large body of evidence, as summarized in Pagan (1996), has since then made clear that, while high frequency returns are virtually serially uncorrelated and lower frequency returns are generally little auto-correlated, there is considerable serial dependence in higher moments. For example, there is overwhelming evidence of conditional heteroskedasticity and time variation in second moments. Furthermore, evidence on return predictability suggests that first moments are time-varying. For example, while monthly returns are generally found to be largely unpredictable[2], there is evidence that annual returns are somewhat predictable and returns at five-year horizons are very predictable (Fama and French (1989) and Cochrane (1999b)) using forecasting variables such as the dividend yield, the price earning ratio and other functions of stock prices normalized by an appropriate divisor to make them stationary. This suggests that the mean of the return process is time varying and driven by a slow moving state variable.

Subsequent studies in the empirical finance literature have reported evidence of two types of asymmetries in the distribution of stock returns. The first is skewness, i.e. $E_t(\varepsilon_{i,t+1}^3)$, or asymmetry in the distribution of individual stock returns, which has been reported and studied by numerous authors over the last three decades. See, among others, Simkowitz and Beedles (1978) and Singleton and Wingender (1986). The second type of asymmetry is in the joint distribution of stock returns. One possible

[2] Monthly and higher frequency stock returns typically have slight, statistically significant predictability with coefficient of determination R^2 of about 1 percent.

source of such an asymmetry is coskewness, i.e. $E_t(\varepsilon_{i,t+1}\varepsilon_{i,t+1}^2)$, where $\varepsilon_{i,t+1}$ and $\varepsilon_{j,t+1}$ are two zero-mean return innovations. Evidence that stock returns exhibit some form of asymmetric co-dependence has been reported by several authors in recent years, see for example Erb *et al.* (1994), Longin and Solnik (2001), Ang and Bekaert (1999, 2002), Ang and Chen (2002), Campbell *et al.* (2002) , and Bae *et al.* (2003). The presence of either of these asymmetries violates the assumption of normally distributed portfolio returns, which underlies traditional mean-variance analysis (see Ingersoll (1987)).

Pagan (1996), Campbell *et al.* (1997) and Cochrane (1999b), among many others, provide a summary of the main stylized empirical features of the multivariate distribution of asset returns, such as serial dependence, time variation in first, second and higher moments and non-normality. As these features have become common characteristics of models of asset returns, the old paradigm has been gradually abandoned in favour of a richer one. In this new paradigm, the multivariate distribution of asset returns cannot be described simply by its first and second moments and conditional and unconditional moments are not in general the same.

1.4. The Fundamental Research Questions

The research questions that I address in this thesis are both theoretical and empirical in nature. They all concern the asset pricing problem, i.e. the description and explanation of observed mean returns, including the fundamental questions concerning the relationship between the representative investor's risk preference and the distribution of asset returns. The key question, however, is whether investors are rewarded not only for holding portfolios that perform poorly when aggregate returns are low, as in Sharpe

(1964) and Lintner (1965) CAPM, but also for holding portfolios that perform poorly when volatility is high or even for holding portfolios that perform poorly when skewness is low. One way of reformulating this question is to ask whether asset coskewness and cokurtosis, in addition to the covariance with the market portfolio, explains the cross-section of average asset returns. In the empirical investigation of this issue, I focus on the explanatory power of coskewness and cokurtosis for the cross-section of average returns on a particular set of benchmark assets, i.e. the portfolios formed sorting by industry the NYSE, AMEX and NASDAQ stocks included in the database of the Center for Research on Security Prices (CRSP) of the University of Chicago. A further research question related to the asset pricing problem is whether investors' sentiment and market frictions play important roles.

1.5. Motivations

The investigation of the asset pricing problem is motivated by its profound implications for capital budgeting, portfolio selection and portfolio management. There is ongoing debate on the ability of the CAPM to explain the cross-section of average asset returns. In particular, there is puzzling evidence on the limited ability of theoretically motivated risk factors to drive out the explanatory power of firm characteristics such as size and book-to-market ratio, see Fama and French (1992, 1993, 1995), momentum, see Jagadeesh and Titman (1993), coskewness (i.e. systematic skewness), see Harvey and Siddique (2000), cokurtosis, see Dittmar (2002), and industry, see Moskowitz and Grinblatt (1999) and Dittmar (2002). The evidence on the asymmetry of the multivariate distribution of asset returns suggests that, if investors' preferences are not restricted to be defined only over the first two moments, expected returns might depend on higher order odd moments. This possibility motivates the study of the explanatory power of asset coskewness and cokurtosis in the cross-section of excess returns, as in Harvey and Siddique (2000), Dittmar (2002), Post *et al.* (2008) and Potì and Wang (2010). I focus on the cross-section of portfolios[3] formed sorting stocks according to the industry in which the issuing firm operates because this characteristic has been used less frequently in the extant empirical literature as a sorting criterion and it is known, see Dittmar (2002), for producing a very dispersed (and therefore challenging) cross-section of average returns.

As stressed by Grinblatt and Titman (1989), empirical studies in finance must rely on non-experimental data. As a consequence, there is a concern that established statistical and econometric methods may suffer from biases. Since empirical research is often motivated by the successes or failures of past investigations, there is a danger of data-snooping, e.g. Ferson and Harvey (1999). Following a similar line of argument, Lewellen *et al.* (2010) suggest that asset pricing models should not be judged by their success in explaining average returns on portfolios for which a few popular factors are known to explain most of the time-series and cross-sectional variation. For example, nearly 92 percent time variation of portfolio constructed on the basis of size and Book-to-Market ratio can be explained by the three-factor model that includes market return, SMB and HML as factors (Fama & French, 1992). Then tests of whether pricing models capture cross-sectional variation of asset returns on the size-B/M portfolios is more or less equivalent to searching for factors which are highly correlated with the SMB and HML factors. This critique is similar in spirit to the point made by Farnsworth

[3] I thank K. French for making this data publicly available for download.

et al. (2002b), and was also raised in the seminal paper by Roll (1977). Since the strong factor structure involved in size-B/M portfolios is less of a theory than an empirical observation, one way of making pricing model tests more convincing is to expand the set of test portfolios beyond the size-B/M sorting (Ferson & Siegel, 2009). Hence, in order to minimize data snooping bias, I use portfolios of CRSP stocks sorted by industry, alongside portfolios sorted by size and book-to-market, and experiment with a dataset of Chinese mutual funds.

1.6. Structure of the Thesis

The first part of this thesis is devoted to the discussion of the extant literature on asset pricing. The second part presents original, largely empirical results on asset pricing and on fund evaluation. In particular, my novel contributions appear in Chapters 3 to 6 and in Chapter 7 I discuss their implications for asset pricing.

I first discuss, in Chapter 2, the modern view of the multivariate distribution of asset returns within the conceptual framework of modern asset pricing theory. Chapter 3 focuses on the cross-sectional dimension of the asset pricing problem and, in particular, on whether coskewness helps explain the cross-section of average returns and raised the so called ‘coskewness puzzle’. This Chapter is based on an article, i.e. Potì and Wang (2010), recently published in the Journal of Banking and Finance. In chapter 4, we show that augmenting the (C)CAPM with sentiment, and thus allowing for systematic investor error in forming beliefs about the distribution of returns, permits to largely reconcile investors’ optimizing behaviour with the cross-section of average returns. In Chapter 5, we comparas e several competing pricing kernels using a modified version of

Hansen-Jagannathan distance (Hansen & Jagannathan, 1997), which not only accounts for the conditional information but also recognizes the existence of transaction costs. Chapter 6 is devoted to the study of the open-end fund performance in the Chinese market. Chapter 7 summarizes the main findings, provides a discussion of their implications, outlines directions for future research and draws together the conclusions.

1.7. Main Findings and Contributions

My thesis contributes to depict a representation of the multivariate distribution of stock returns where the relations between moments and their conditional dynamics are important in explaining their cross-sectional differences. More innovatively, we explore whether the SDF implied by the 3 and 4-moment CAPM is plausible under restrictions that are weaker than those considered by Dittmar (2002) yet sufficient to rule out implausible curvature of the representative investor's utility functions. We find that, even under these weaker restrictions, the 3 and 4-moment CAPM cannot solve well known puzzles which plague the empirical performance of extant rational asset pricing models, even though the higher order terms do generate considerable additional explanatory power. In chapter 3, our findings confirm that the quadratic and cubic market factors help explain observed stock returns. They play an important role in the pricing of certain payoffs, including strategies characterized by relatively high SRs, such as those spanned by a fine industry-level diversification, most notably until the late 90s, or by dynamic portfolios managed on the basis of available conditioning information, as well as momentum portfolios. They do so, however, by generating high levels of SDF volatility. To rationalize this evidence within a higher moment CAPM framework, we would need to postulate implausibly high levels of investors' risk aversion. We conclude, therefore, that the 3M and 4M-CAPM provide at best a partial explanation of the differences in average returns on stocks and stock strategies. This gives rise to a coskewness (and cokurtosis) puzzle. The solution of the latter requires an explanation, different from the 3M and 4M-CAPM, for why the quadratic and cubic market factors are priced in the cross-section of stock returns.

Faced with this difficulty, we then explore whether the failure to fully account for cross-sectional differences in average returns can be explained by the presence of either transaction costs or a behavioural component of the SDF, reflecting investors' systematic mistakes in processing information. In chapter 4, we show that augmenting the (C)CAPM with sentiment, and thus allowing for systematic investor error in forming beliefs about the distribution of returns, permits to largely reconcile investors' optimizing behaviour with the cross-section of average returns. In fact, The Sentiment-(C)CAPM and Sentiment-3M(C)CAPM are empirically more successful, and most of the increase in the explanatory power is due to the inclusion of sentiment. This implies that investors must either commit systematic errors, at least ex-post, in assessing the joint distribution of stock returns and aggregate consumption or they must behave in a way that, at the aggregate level, is inconsistent with expected utility maximization and with standard risk aversion assumptions.

In chapter 5, we compare several competing pricing kernels using a modified version of Hansen-Jagannathan distance (Hansen & Jagannathan, 1997), which not only accounts for the conditional information but also recognizes the existence of transaction costs. We follow the approach done by He and Modest (1995) and Luttmer (1996), shows how the Hansen-Jagannathan volatility bounds can be derived for economies with the kinked budget constraints that arise from proportional transaction costs. Since these bounds do not depend on a particular model for the stochastic discount factor, but only on the form of the budget constraint, this analysis provides a robust way to quantify the extent to which market frictions affect inferences about important features of asset pricing models. In fact, we find that, if the market is frictionless (i.e. if we assume no transaction costs), the volatility of the admissible nonparametric pricing kernels is so high that even the models with nonlinearity terms cannot match it whereas a number of the estimated pricing kernels can do so when transaction costs are assumed to be 0.8 percent per each one-way trade.

In the final part, we follow apply the SDF approach, as in Chen and Knez (1996), to analyse the open-end fund performance in the Chinese market. In a similar spirit as Dahlquist and Söderlind (1999), we test whether Chinese open-end fund managers process information more professionally than individual investors do, so as to generate significant abnormal returns, and also try to observe whether their performance can be replicated by employing mechanic and easily replicable strategies that make use of publicly available information. The results of our analysis show the performance of such fund managers in an unfavourable light. The fund managers of the selected open-end funds in fact generate no superior performance. Even the naive buy-and-hold trading strategy is able to replicate their performance easily. Then the question that really needs to be answered, but that we leave for future research, is how the phenomenon of the poor fund performance can coexist with the rapid expansion of the open-fund industry in the Chinese market.

1.8. Conclusions

In this Chapter, I first introduced some preliminary material on the distinction between conditional and unconditional moments and how it arises in the transition from the old to the new paradigm of asset returns. I then specified the fundamental research questions of this thesis and discussed the motivations that drive their investigation. In particular, I discussed the connection between the research questions and unresolved issues in asset pricing and second moment modelling. Finally, I outlined the structure of the thesis and I summarized the main findings and their contribution to the extant literature.

2. Literature Review: Asset Pricing

2.1. Introduction

In this chapter, I review the literature on the distribution of asset returns and on the closely related topic of asset pricing. Rather than attempting to list all the countless contributions, my aim is to show how the discovery of moment dynamics and their role in asset pricing unfolded over the transition from the old view of asset returns, based on the random walk model and on the identity between conditional and unconditional moments, to the new paradigm that allows for time-varying conditional moments and returns predictability. I first show, however, how all asset pricing models can be derived as specializations of a common analytical framework, the general SDF model with possibly time-varying risk premia and returns predictability. I then discuss at more length select specifications. In particular, I focus on specifications that allow for systematic skewness and kurtosis to play a role in asset pricing.

The Chapter is organized in eight main sections. The next two sections introduce the SDF approach to asset pricing and the risk-neutral density. In Section 2.4, I provide a brief account of no-arbitrage moment restrictions on the SDF. In Section 2.5, I discuss the relationship between the SDF and popular linear pricing models. Section 2.6-2.8 review the developments in the theory of efficient markets and rational asset pricing, and the literature focusing on coskewness and higher-moments as priced risk exposures. Section 2.9 discusses the alternative behavioural approach to asset pricing and Section 2.10 illustrates the dichotomy between absolute and relative pricing. Section 2.11

discusses the importance of transaction costs. The last section summarizes the chapter and draws together the conclusions.

2.2. Pricing the Contingent Claims through SDF

Under the assumption of no-arbitrage (henceforth NA), an asset with payoffs that always dominate the payoffs of another asset in every state of world must have a greater price. The absence of arbitrage opportunities implies the Law of One Price (here forth LOP), i.e. two assets that have exactly identical payoffs across every state of nature must be priced identically at every point in time. This assumption is more restrictive compared with the LOP. The violation of this relation creates a riskless opportunity for making unlimited profit.

Pliska (1997) provides a systematic analysis of NA in a one-period complete market setting. The initial state of the system is at time t, and the terminal time is $t+1$. Trading and consumption can be done only at these times. $\Omega = \{\omega_1, \omega_2, ..., \omega_d\}$ denotes the possible states of the world[4]. A probability measure is defined on Ω with $P(\omega) > 0$ for all $\omega \in \Omega$. Since P is a probability measure then we have $\sum_{\omega\in\Omega} P(\omega) = 1$. This complete probability space can be denoted as (Ω, F_t, P), where F_t is the sigma-filtration, with $F_{t_1} \subseteq F_{t_2}$ when $t_1 < t_2$. Therefore, the filtration F_t represents the information available up to and including time t. The risk-less asset is modelled as a bank account process, denoted by $B = \{B_t, B_{t+1}\}$, where $B_t = 1$, and

[4] Here, d could be infinite in a Hillbert Space.

$B_{t+1} = B_t(1 + r_{t+1})$ for every state of the world ω at time t, earning the risk-free interest rate r_{t+1}. In the market there are N risky securities, $S^1, S^2, ..., S^N$. For each security, the price in a particular state of world ω at time $t+1$ will be denoted $S^i_{t+1}(\omega)$. An investor's portfolio will be made of the bond B_t and several risky securities S^i_t. A linear pricing measure is a non-negative vector $\pi = (\pi(\omega_1), \pi(\omega_2), ..., \pi(\omega_d))$ [5] such that for every trading strategy $H_t = (H^1_t, H^2_t, ..., H^N_t)$, where the elements of H_t denote the value initially invested in each particular asset, generates a value process $V_{t+1}(\omega) = H^0_t B_{t+1} + \sum_{i=1}^{N} H^i_t S^i_{t+1}(\omega)$ with the current value $V_t = \sum_{\omega\in\Omega} \pi(\omega) V_{t+1}(\omega) \Big/ B_{t+1}$. If we let $V^*_{t+1}(\omega) = V_{t+1}(\omega) / B_{t+1}$, then we have,

$$V_t = \sum_{\omega\in\Omega} \pi(\omega) \ \ V^*_{t+1}(\omega) = \sum_{\omega\in\Omega} \pi(\omega) \frac{V_{t+1}(\omega)}{B_{t+1}}. \tag{2-1}$$

Notice that, if in Equation (2-1) we consider the strategy that invests only in the bank account and nothing in the risky assets, we get

$$V_t = \sum_{\omega\in\Omega} \pi(\omega) \ \ V^*_{t+1}(\omega) = \sum_{\omega\in\Omega} \pi(\omega) \frac{H^0_t B_{t+1}(\omega)}{B_{t+1}} = H^0_t \sum_{\omega\in\Omega} \pi(\omega), \tag{2-2}$$

[5] This is an implication of Riesz's Representation Therom. Denote by H a Hilbert space, and let H^* be its dual space, then for any linear function $\varphi : H \to \mathbb{R}, \forall \varphi \in H^*$, we have $\varphi(x) = < y, x >$, where $x, y \in H$ and $<,>$ denotes the inner product on the linear space H. The assumption of the existence of the linear function $\varphi : H \to \mathbb{R}$ in fact requires the holding of the law of one price, i.e. $\varphi(x_1) = \varphi(x_2), if \ \ x_1 = x_2$. But $\varphi(x_1 - x_2) = 0 \to x_1 = x_2$ is not necessary.

but $V_t = H_t^0$, which implies $\sum_{\omega\in\Omega} \pi(\omega) = 1$, and therefore π can be treated as a probability measure. Moreover, if in (2-1) we consider the strategy that invests in just one particular asset S^n (thus $H_j = 0$ for all $j \neq n$), then

$$S_t^n = \sum_{\omega\in\Omega} \pi(\omega) S_{t+1}^{n,*}(\omega) = \sum_{\omega\in\Omega} \frac{\pi(\omega)}{B_{t+1}} S_{t+1}^n(\omega), \text{ for all } n = 1,..., N. \qquad (2\text{-}3)$$

Equation (2-3) says that the initial price of the asset equals the expected value of its discounted future price, under the linear pricing measure π. It can be proved that, if the no-arbitrage opportunity condition holds, the linear pricing measure π for each state of world is positive. Moreover, the linear pricing measure π is unique if the market is complete, i.e. if any contingent claim is attainable. Equation (2-3) can be rewritten as follows,

$$\begin{aligned} S_t^n &= \sum_{\omega\in\Omega} \frac{\pi(\omega)}{B_{t+1}} S_{t+1}^{n,*}(\omega) = E[\frac{\pi(\omega)}{P_{\omega\in\Omega}(\omega) * B_{t+1}} S_{t+1}^n(\omega) \mid F_t], \\ &= \int_{\omega\in\Omega} \frac{\pi(\omega)}{B_{t+1} f_{\omega|F_t}(\omega)} S_{t+1}^n(\omega) f_{\omega|F_t}(\omega) d\omega \ \text{ for all } n = 1,..., N. \end{aligned} \qquad (2\text{-}4)$$

where $f_{\omega|F_t}(\omega)$ is the conditional probability density function, under the information set F_t, for each state of world ω. In fact, both $E[\frac{\pi(\omega)}{P_{\omega\in\Omega}(\omega) * B_{t+1}} S_{t+1}^n(\omega) \mid F_t]$ and $\int_{\omega\in\Omega} \frac{\pi(\omega)}{B_{t+1} f_{\omega|F_t}(\omega)} S_{t+1}^n(\omega) f_{\omega|F_t}(\omega) d\omega$ make the same statement, the former in discrete time and the latter under the continuous time measure. If, following Cochrane (2001), we scale the contingent claims prices by the inverse probability of each state, the resulting expression provides the SDF specification so widely used in finance. That is, $\frac{\pi(\omega)}{P_{\omega\in\Omega}(\omega \mid F_t) * B_{t+1}}$ or, equivalently, $\frac{\pi(\omega)}{B_{t+1} f_{\omega|F_t}(\omega)}$ can be seen as the value taken by the SDF or pricing kernel, m_{t+1}, in state of the world ω and equation (2-4) is simply written as

$$S_t^n = E[m_{t+1} S_{t+1}^n(\omega) \mid F_t], \quad \text{for all } n = 1, ..., N. \tag{2-5}$$

Equivalently but more conveniently,

$$E[m_{t+1} R_{t+1} \mid F_t] = 1, \tag{2-6}$$

where R_{t+1} is the gross return for any asset S^n in market at time $t+1$, and F_t denotes the information set available to investors at t. Of course, condition (2-6) also can be represented using vector notation,

$$E[m_{t+1} R_{t+1} \mid F_t] = \tau^{N\times 1}. \tag{2-7}$$

Here, R_{t+1} is the vector of asset returns, and $\tau^{N\times 1}$ is the $N \times 1$ vector of units. Asset prices then equal the expectation of the product between the SDF and the payoffs. As explained by Campbell (2000), linear factor pricing models are associated with SDF specifications linear in a set of linear factors. For example, the well-known classical consumption-based pricing model is associated with a SDF structure specified as the inter-temporal marginal rate of substitution of the representative-agent investor, which is linear in the consumption growth of such agent.

2.3. Risk-neutral pricing density

$\pi_{t+1}(\omega_i)/B_{t+1}$ is the state-contingent price $p_t(\omega_i)$, which represent the current price of a security that secures one dollar once a certain state (ω_i) occurs at a future date $t+1$, and nothing otherwise. As discussed in section 2.2 of this chapter, the relation among state-price contingent prices $p_t(\omega_i)$, the risk-neutral pricing density $\pi(\omega_i)$, and the pricing kernel $m_{t+1}(\omega_i)$ is,

$$p_t(\omega_i) = \frac{\pi(\omega_i)}{B_{t+1}} = m_{t+1}(\omega_i) f_{t+1}(\omega_i), \tag{2-8}$$

here, $f_{t+1}(\omega_i)$ is the physical probability density at time $t+1$. If the non-constant information set F_t at time t is available, then $f_{t+1}(\omega_i)$ is the conditional density, otherwise the unconditional density is used instead in empirical test. The sum of such state-contingent prices across all states at $t+1$ has to equal the current price of a risk-free bond that pays off one unit at maturity $t+1$ for sure, i.e.

$$\int_{w_i \in \Omega} p_t(\omega_i) d\omega_i = \int_{w_i \in \Omega} [\frac{\pi(\omega_i)}{B_{t+1}} \times 1] d\omega_i = E[m_{t+1} \mid F_t] = 1/B_{t+1}\ . \tag{2-9}$$

Since B_{t+1} is observable at t and can be replaced by risk-free rate $e^{r_{f,t}}$, the availability of either $p_t(\omega_i)$ or m_{t+1} permits the estimation of the other. Breeden and Litzenberger (1978) provide an explicit formula for the estimation of $p_t(\omega_i)$:

$$p_t(\omega_i) = e^{-r_{f,t}} \pi(\omega_i) = \frac{\partial^2 C_t}{\partial K^2} |_{S_{t+1}(\omega_i)=K}, \tag{2-10}$$

i.e. the second derivative of a European call price C_t taken with respect to its strike price K is the state-contingent price $p_t(\omega_i)$ of the future asset price ending up at exactly the strike price of the option. The measure $\pi(\omega)$ satisfies the probability axioms, since it is positive and $\sum_{\omega\in\Omega} \pi(\omega) = 1$, and therefore can be referred to as the risk-neutral pricing density,

$$\pi(\omega_i) = e^{r_{f,t}} \frac{\partial^2 C_t}{\partial K^2} |_{S_{t+1}(\omega_i)=K}, \tag{2-11}$$

It satisfies, under no-arbitrage, the following pricing condition,

$$S_t = e^{-r_{f,t}} \int_{\omega\in\Omega} \pi(\omega_i) S_{t+1} d\omega_i = e^{-r_{f,t}} E^{\pi}[S_{t+1}]. \tag{2-12}$$

Here, $E^{\pi}[\cdot]$ is the expectation operator under the risk-neutral probability $\pi(\omega)$. According to equation (2-12), in a world in which investors are risk-neutral, all risky assets - including options - must yield an expected return equal to the risk-free rate of interest. Estimating the risk-neutral pricing density from option prices is equivalent to estimating the SDFs using the prices and payoffs of assets priced by each SDF. Generally, the approaches for estimating the risk-neutral pricing density can be categorized into two different classes: parametric estimation and non-parametric estimation (see Cont (1997), and Jackwerth (1999)).

Concerning parametric estimation, expansion methods, generalized distribution methods and mixture methods are the three main subgroups. More specifically, starting with a simple distributional assumption, expansion methods typically add correction terms into the normal or lognormal density, then estimate a flexible-shape version of the risk-neutral distribution in order to fit the observed option prices. Fore example, Abadir and Rockinger (1997) use confluent hypergeometric functions as a basis for the risk-neutral density and derive the option price function across strike prices. Abken *et al.* (1996a, 1996b) use instead four-parameter Hermite polynomials. Potters *et al.* (1998) use cumulant expansions to add a single correction term to a normal distribution, which adjusts for the kurtosis of the risk-neutral distribution. However, this approach may generate negative values for the risk-neutral density, which violate the positivity constraint. Without adding correction terms, generalized distribution methods concentrate more on inherent flexibility, abandoning familiar two-parameter normal and lognormal distributions by adding one or two additional parameters beyond mean and variance (see, Aparicio and Hodges (1998), Posner and Milevsky (1998)). The mixture methods can even achieve greater flexibility than the previous two. For example, Melick and Thomas (1997) apply the mixture methods, which combine three lognormal distributions, to evaluate American options on crude oil futures. Ritchey (1990) modifies log returns by a mixture of normal distributions to generate option prices. Although the number of parameters expands quickly under mixture methods, this approach is more capable of generating a wider variety of shapes for the probability distributions than generalized distributions.

Non-parametric estimation tries to achieve greater flexibility in fitting the risk-neutral distribution to option prices. Rather than requiring a parametric form of the distribution, it allows the use of more general functions. The most commonly used non-parametric method is kernel regression, which is conceptually related to regressions in that they both try to fit a function to observed data. The main difference is that kernel regression does not specify the parametric form of the function. As a data-intensive method, kernel regression assumes each data point is the center of the region where the true underlying functions lie. The further away a point in the support of the density is from the observed data point, the less likely it is that the true function goes through that distant point. Take the observe implied volatilities $\sigma(K_i)$ across strike prices K_i as an example, the kernel regression in one dimension is

$$\sigma(K_i)=\frac{\sum_{i=1}^{n}k(\frac{K-K_i}{h})\sigma(K)}{\sum_{i=1}^{n}k(\frac{K-K_i}{h})} \tag{2-13}$$

where h is the bandwidth, which governs the smoothness of the kernel regression, and the kernel $k(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, which as often the case takes the the form of a normal distribution probability density function with zero mean and unit standard deviation, measures the corresponding drop in likelihood when we move away from the data point. Aït - Sahalia and Lo (1998) use kernel regression to estimate the risk-neutral density or state-price density (SPD) based on five-dimensional data points consisting of stock price, strike price, time to expiration, interest rate and dividend yield. They also experiment with a lower-dimension problem, in which, for simplicity, the three dimensions are forward price, strike price, and time to expiration. Pritsker (1998) employs their methodology in order to investigate the risk-neutral distribution of interest rates..

2.4. Moment Restrictions on SDF

According to Equation (2-6), the expected return on the asset i is given by

$$E_t(R_{i,t+1}) = \frac{1}{E_t(m_{t+1})} - \frac{cov_t(m_{t+1},\, R_{i,t+1})}{E_t(m_{t+1})}\,. \tag{2-14}$$

where $cov_t(\cdot)$ denotes the conditional covariance operator. The first moment of the SDF is determined by the risk-free rate. The above relation holds also when R_{t+1} is the risk-free return $R_{f,t+1}$, which equals $e^{r_{f,t}}$ in continuous time. Since the conditional covariance between SDF and risk-free rate is, by definition, zero, we have that

$$E_t(m_{t+1}) = \frac{1}{R_{f,t+1}}\,. \tag{2-15}$$

In other words, the expected SDF is the price of a short-term risk-free zero-coupon bond with a unit payoff at time $t+1$. This characterizes the SDF as a random variable, i.e.

$m_{t+1} = \frac{1}{R_{f,t+1}} + \varepsilon_{t+1}$, where the conditional mean of the error term is zero, i.e. $E_t(\varepsilon_{t+1}) = 0$.

The restriction in (2-14) also applies to excess returns, i.e.

$$\begin{aligned} E_t(R_{i,t+1} - R_{f,t+1}) &= -\frac{cov_t(m_{t+1},\ R_{i,t+1} - R_{f,t+1})}{E_t(m_{t+1})} \\ &= -R_{f,t+1} cov_t(m_{t+1},\ R_{i,t+1} - R_{f,t+1}) \end{aligned} \quad (2\text{-}16)$$

or, equivalently, $E_t[m_{t+1}(R_{t+1} - R_{f,t+1})] = 0$. The conditional volatility of the SDF is related to the $\mathcal{L}^2$ metric distance between the risk-neutral and the objective probability measures (Bakshi *et al.*, 2005). In the absence of arbitrage opportunities, Bakshi *et al.* (2005) prove that the one-to-one correspondence is

$$\begin{aligned} \sigma_t(m_{t+1}) = \delta_2[\pi - f_{t+1}] &= \frac{\|\pi - f_{t+1}\|}{\|\tau\|} \\ &= \sqrt{\int_{\omega\in\Omega} (\pi(\omega) - f_{t+1}(\omega))^2 f_{t+1}(\omega) d\omega} \end{aligned} \quad (2\text{-}17)$$

As just seen, the fact that m_{t+1} prices excess returns gives that $E_t[m_{t+1}(R_{i,t+1} - R_{f,t+1})]$ equals zero. Also the absolute value of correlation between SDF m_{t+1} and $(R_{i,t+1} - R_{f,t+1})$ is bounded from above by 1, and this implies that, under no-arbitrage and for a given level of correlation between the SDF and asset returns, spreads in expected excess returns across assets are proportional to SDF volatility,

$$\left|E_t(R_{i,t+1} - R_{f,t+1})\right| \le R_{f,t+1}\sigma_t(R_{i,t+1} - R_{f,t+1})\sigma_t(m_{t+1}). \quad (2\text{-}18)$$

Noting that (2-18) must hold for any SDF that prices the assets, including the minimum variance SDF m_{t+1}^*, which is the projection of m_{t+1} on the payoff set, we can rearrange (2-18) to obtain the well-known Hansen and Jagannathan (1991) lower bound on SDF volatility

$$\begin{aligned} \sigma_t(m_{t+1}) \geq \sigma_t(m_{t+1}^*) &= \frac{\left|E_t(R_{i,t+1} - R_{f,t+1})\right|}{R_{f,t+1}\sigma_t(R_{i,t+1} - R_{f,t+1})} \\ .or. \quad &= \left|\frac{1 - E_t(m_{t+1})E_t(R_{i,t+1})}{\sigma_t(R_{i,t+1} - R_{f,t+1})}\right|, \end{aligned} \tag{2-19}$$

where, $\sigma(m_{t+1}^*) = \inf\{\sigma(m_{t+1})\}$ indictates the miminum variance of m_{t+1} over the support of admissible kernels (not a singleton, in incomplete markets). Obviously, if the payoff space contains N securities, condition (2-19) satisfies

$$\sigma_t^2(m_{t+1}) \geq [\tau - E_t(m_{t+1})E_t(R_{t+1})]'\Sigma^{-1}[\tau - E_t(m_{t+1})E_t(R_{t+1})], \tag{2-20}$$

where $\Sigma = cov_t(R_{t+1})$ denotes the covariance matrix of asset returns and τ is the $N \times 1$ unit vector.

Moreover, Ross (2005) points out that risk-averse investors prefer volatile SDFs. Specifically, he shows that, in a complete market and given a family of SDFs with the same mean but different volatility, the expected utility of a risk-averse investor is uniformly larger if the payoffs are priced by a more volatile SDF. This implies that, if a risk-free asset with a given return identifies the SDF mean, the expectation of a concave utility function increases in the volatility of the unique SDF (i.e. the state-price density) that prices the investment opportunity set. In incomplete markets, more than one SDF price the marketed payoffs and thus expected utility increases in the volatility of the projection of all admissible pricing kernels on the payoff space, i.e. it increases in the volatility of the minimum-variance SDF. Similar arguments can be traced to Bakshi *et al.* (2005) and Chabi-Yo *et al.* (2008). More precisely, Bakshi *et al.* (2005) show that $\sigma_t(m_{t+1})$ is related to both the level of risk aversion, γ, and the variance of the consumption-growth $\sigma_t(c_{t+1})$

$$\begin{aligned} \sigma_t(m_{t+1}) &= \sqrt{e^{\gamma^2\sigma_t^2(c_{t+1})}-1} \\ &\approx \gamma\sigma_t(c_{t+1}) \end{aligned}. \tag{2-21}$$

where $c_{t+1} = log(C_{t+1}/C_t)$, represents the growth of consumption. The argument of Chabi-Yo, *et al.*, (2008) is more general than Bakshi, *et al.*, (2004) without assuming a specific utility function, and they show that the growth of the SDF is proportional to the Arrow-Pratt index of absolute-risk aversion (ARA),

$$\frac{\partial log(m_{t+1})}{\partial C_{t+1}} = \frac{U''(C_{t+1})}{U'(C_{t+1})} = -ARA\,. \tag{2-22}$$

Of course, a higher volatility of the SDF requires a higher ARA. In other words, if the ARA is restricted within a reasonable range, $\sup\{\sigma_t(m_{t+1})\}$ is bounded.

Cochrane and Saa-Requejo (1996) and Cochrane (2001) introduced the term ‘good deals’ to denote desirable investment opportunities, i.e. arbitrage opportunities and investment opportunities that offer a large reward for risk. Similarly, Cerný and Hodges

(2000) define good deals as desirable investment opportunities that have zero or negative cost. They also formalize the conditions under which good deals can be ruled out in complete and incomplete markets, given the absence of good deals in the space of payoffs spanned by a subset of the traded assets. Again, the absence of good deals is connected to the volatility restriction of the SDF. Cochrane and Saa-Requejo (1996) suggest that,

$$\sigma_t(m^*_{t+1}) \leq \frac{h}{R_{f,t+1}}. \tag{2-23}$$

where h is the pre-specified volatility bound. The results in Ross (2005) shows that good deals offer expected utility improving opportunities to the risk-averse investors who prefer a high SDF volatility. Potì (2007) and Potì and Wang (2010), in order to restrict the volatility of the minimum-variance SDF, extend this argument by restricting, by means of an upper bound, the volatility of the representative investor's inter-temporal marginal rate of substitution (IMRS) between present and future consumption, which is specified as in the typical three-moment CAPM, i.e. it coincides with the representative investor's IMRS and is a linear function of the market return and its square.

2.5. The SDF Approach and Linear Pricing Models

The classical linear Capital Asset Pricing Model (CAPM) has been established in the work of Sharpe (1964), Lintner (1965), Black (1972), Merton (1973) and Breeden (1979). The main prediction of the model is that the expected return on a financial asset is a function of its covariance with some systematic risk factors or, equivalently, linearly related to the coefficients, known as betas, of the regression of the excess-return on the factors. Cochrane (1996) subsequently formalizes the relation between the SDF approach and linear beta-pricing models and reformulates the CAPM in a SDF setting. Generally, inference in the SDF approach requires the specification of fewer assumptions about the distribution of asset returns than when estimating linear pricing models.

Sharpe (1964) and Lintner (1965) assume that all investors are single-period mean-variance optimizers to show that the market portfolio is mean-variance efficient and the expected return of each asset is linearly related with the market portfolio through the asset risk exposure beta. Ross (1976) points out that this conclusion can also be reached using an asymptotic no-arbitrage argument and the assumption that the market portfolio is the only source of common, undiversifiable risk. In other words, if there are several common factors that generate undiversifiable risk, then a multifactor model holds. Within the SDF framework, these same conclusions can be reached directly from the assumption that the SDF is a linear combination of K common factors $f_{k,t+1}$, $k = 1...K$. For expositional simplicity the factors are assumed to have conditional mean zero and are orthogonal to one another. If

$$m_{t+1} = a_t + \sum_{k=1}^{K} b_{k,t} f_{k,t+1} . \tag{2-24}$$

then the negative of the covariance of any excess return with the SDF can be written as

$$-\text{cov}_t(m_{t+1},\ R_{i,t+1} - R_{f,t+1}) = -\sum_{k=1}^{K} b_{k,t}\sigma_{ik,t} = -\sum_{k=1}^{K} b_{k,t}\sigma_{k,t}^2\left(\frac{\sigma_{ik,t}}{\sigma_{k,t}^2}\right) = \sum_{k=1}^{K}\lambda_{k,t}\beta_{ik,t}\ . \quad \text{(2-25)}$$

Here $\sigma_{ik,t}$ is the conditional covariance of asset return i with the k^{th} factor, $\sigma_{k,t}^2$ is the conditional variance of the k^{th} factor, $\lambda_{k,t} = -b_{k,t}\sigma_{k,t}^2$ is the price of risk with the respect to the k^{th} factor, and $\beta_{ik,t} = \sigma_{ik,t}/\sigma_{k,t}^2$ is the regression coefficient of asset return i on that factor, or the formerly mentioned beta . Then equation (2-25) together with the equation (2-16) implies that the risk premium on any asset can be rewritten as a sum of the asset's betas with common factors times the risk prices of those factors. More particularly, if we set the asset i as the market portfolio, i.e. $R_{i,t+1} = R_{m,t+1}$, $f_{k,t+1} = R_{m,t+1} - E_t(R_{m,t+1})$ and $m_{t+1} = b_t[R_{m,t+1} - E_t(R_{m,t+1})]$, then $\beta_{ik,t} = 1$ and

$$\lambda_{k,t} = E_t(R_{m,t+1} - R_{f,t+1})/R_{f,t+1}\ . \quad \text{(2-26)}$$

In other words, if $b_{kt} = -E_t(R_{m,t+1} - R_{f,t+1})/(R_{f,t+1}\sigma_{R_{m,t}}^2)$, then m_{t+1} will exactly generate the CAPM. Here, it is noticed that m_{t+1} is negatively correlated with $R_{m,t+1}$.

2.6. Conditional Liner Models

The unconditional or static asset pricing model was derived by considering the rational behaviour of investors living for only one period. In extending this model to a multiperiod setting, one of the commonly made assumptions is that asset betas remain constant over time. According to Jagannathan and Wang (1996), this is not a particularly reasonable assumption since the relative risk of a firm's cash flow is likely to vary over the business cycle. During a recession, for example, financial leverage of firms in relatively poor shape may increase sharply relative to other firms, causing their stock betas to rise. Also, to the extent that the business cycle is induced by technology or taste shocks, the relative share of different sectors in the economy fluctuates, inducing fluctuations in the betas of firms in these sectors. Hence, betas and expected returns will in general depend on the nature of the information available at any given point in time and vary over time. Therefore, they specify a relation between conditional expected return and conditional betas, i.e.

$$E_t(R_{t+1} - R_{f,t+1}) = \sum_{k=1}^{K} \lambda_{k,t}\beta_{k,t} \ . \tag{2-27}$$

We can take the unconditional expectation of both sides of equation (2-27) to get

$$\begin{aligned} E[E_t(R_{t+1} - R_{f,t+1})] = E(R_{t+1} - R_{f,t+1}) &= \sum_{k=1}^{K} E(\lambda_{k,t})E(\beta_{k,t}) + \sum_{k=1}^{K} cov(\lambda_{k,t}, \beta_{k,t}) \\ &= \sum_{k=1}^{K} \bar{\lambda}_k \bar{\beta}_k + \sum_{k=1}^{K} var(\lambda_{k,t})\vartheta_k \end{aligned} . \tag{2-28}$$

where, $\bar{\lambda}_k = E(\lambda_{k,t})$ and $\bar{\beta}_k = E(\beta_{k,t})$. ϑ_k is the regression coefficient of $\beta_{k,t}$ on $\lambda_{k,t}$, i.e.,

$$\vartheta_k = cov(\beta_{k,t}, \lambda_{k,t}) / var(\lambda_{k,t}) \ . \tag{2-29}$$

Jagannathan and Wang (1996) defined two types of unconditional betas:

$$\beta_k = cov(R_{t+1}, \lambda_{k,t+1}) / var(\lambda_{k,t+1}) . \quad \text{(2-30)}$$

$$\beta_k^z = cov[R_{t+1}, E_t(\lambda_{k,t+1})] / var[E_t(\lambda_{k,t+1})] . \quad \text{(2-31)}$$

where β_k and β_k^z are all constant. Moreover they proved that,

$$\begin{pmatrix} \beta_k \\ \beta_k^z \end{pmatrix} = \Gamma_0 + \Gamma_1 \begin{pmatrix} \overline{\beta}_k \\ \vartheta_k \end{pmatrix}, \quad \text{(2-32)}$$

Here, Γ_0 and Γ_1 are constant coefficient vectors. Then, the unconditional expectation of the conditional expected return is

$$E(R_{t+1} - R_{f,t+1}) = \sum_{k=1}^{K} \lambda_{k,t+1}\beta_k + \sum_{k=1}^{K} E_t(\lambda_{k,t+1})\beta_k^z . \quad \text{(2-33)}$$

The betas in equation (2-33) are not time-varying as the ones in equation (2-27), yet they follow from the same asset pricing model. In fact, equation (2-33) is the unconditional implication of equation (2-27).

According to Cochrane (2001), letting z_t represent a vector of variables that summarize the relevant conditioning information, we can reformulate conditional linear models according to a SDF representation by writing the parameters in (2-24) as functions of the conditioning variables, i.e. we can write $a_t = a_t(z_t)$ and $b_t = b_t(z_t)$ in (2-24). The simplest way to model conditional time-variation in these parameters is to specify them as a linear function of the set of conditioning variable:

$$
\begin{aligned}
a_t &= a^0 + a^{1'} z_t \\
b_t &= b^0 + b^{1'} z_t
\end{aligned}
\quad . \tag{2-34}
$$

Here, a_t is a scalar, b_t, b^0 are $N \times 1$ vectors, z_t is a $k \times 1$ vector of conditioning variables, a^1 is a $k \times 1$ vector and b^1 is an $k \times N$ matrix. Using, (2-34) we can rewrite (2-24) as follows:

$$
\begin{aligned}
m_{t+1} &= a^0 + a^{1'} z_t + (b^0 + b^{1'} z_t)' f_{t+1} \\
&= a^0 + a^{1'} z_t + b^{0'} f_{t+1} + z_t{}' b^1 f_{t+1} \\
&= a^0 + a^{1'} z_t + b^{0'} f_{t+1} + b^{2'} (f_{t+1} \otimes z_t)
\end{aligned}
\quad . \tag{2-35}
$$

Here, b^2 is a $[(k \times N) \times 1]$ vector obtained stacking the N columns of b^1, i.e

$$
b^2 = vec(b^1) \,. \tag{2-36}
$$

The specification in (2-35) and (2-36) can be seen as an unconditional model, i.e. a model with time-invariant parameters, in the new set of factors

$$
F_{t+1} = \begin{bmatrix} z_t \\ f_{t+1} \\ f_{t+1} \otimes z_t \end{bmatrix} . \tag{2-37}
$$

For convenience, we can also rewrite (2-35) folding the unconditional mean of these factors in the constant and write the SDF as a linear function of a new set of unconditionally de-meaned factors:

$$
\begin{aligned}
m_{t+1} &= [a^0 + a^{1'} E(z_t) + b^{0'} E(f_{t+1}) + b^{2'} E(f_{t+1} \otimes z_t)] + \\
&\quad + a^{1'} [z_t - E(z_t)] + b^{0'} [f_{t+1} - E(z_t)] \\
&\quad + b^{2'} [f_{t+1} \otimes z_t - E(f_{t+1} \otimes z_t)] \\
&= \tilde{a} + a^{1'} \tilde{z}_t + b^{0'} \tilde{f}_{t+1} + b^{2'} (f_{t+1} \tilde{\otimes} z_t) \\
&= \tilde{a} + b^{3'} \tilde{F}_{t+1}
\end{aligned} \quad . \qquad (2\text{-}38)
$$

Here,

$$\tilde{a} = a^0 + a^{1'} E(z_t) + b^{0'} E(f_{t+1}) + b^{2'} E(f_{t+1} \otimes z_t) . \qquad (2\text{-}39)$$

$$\tilde{F}_{t+1} = \begin{bmatrix} z_t - E(z_t) \\ f_{t+1} - E(f_{t+1}) \\ f_{t+1} \otimes z_t - E(f_{t+1} \otimes z_t) \end{bmatrix} . \qquad (2\text{-}40)$$

$$b^3 = \begin{bmatrix} a^1 \\ b^0 \\ b^2 \end{bmatrix} . \qquad (2\text{-}41)$$

Since the parameters of the factors in (2-38) are by definition constant over time, the conditional and unconditional implications of the model are the same. In particular, we can derive the unconditional implications without worrying about co-variation between the parameters of the SDF and the factors. To this end, we may take the unconditional expectation of (2-6) with the SDF specified as in (2-38). The covariance and beta-pricing representations of the implication of this unconditional expectation are the following:

$$E(R_{i,t+1}) = -cov\left(R_{i,t+1}, \tilde{F}_{t+1}\right) R_{f,t+1} b^3 \cong -cov\left(R_{i,t+1}, F_{t+1}\right) b^3$$
$$E\left(R_{i,t+1}\right) = \beta_i' \lambda \quad (2\text{-}42)$$

where,

$$\beta_i = var\left(\tilde{F}_{t+1}\right)^{-1} cov\left(\tilde{F}_{t+1}, R_{i,t+1}\right) = var\left(F_{t+1}\right)^{-1} cov\left(F_{t+1}, R_{i,t+1}\right). \quad (2\text{-}43)$$

$$\lambda = -R_{f,t+1} var\left(\tilde{F}_{t+1}\right) b^3 \cong -var\left(F_{t+1}\right) b^3. \quad (2\text{-}44)$$

2.7. Asset Pricing Paradigms and EMH

The efficient market hypothesis (henceforth, EMH), as formulated by Fama (1970, 1976), requires that conditional expectations of future cash-flows and conditional moments of the multivariate return distribution be formed using all the available *relevant* information[6] and, ultimately, that returns in excess of the rate of return on riskless assets do not deviate in any systematic (i.e. exploitable) way from their conditional expectation[7]. This implies that returns deviations from their possibly time-varying equilibrium conditional expectations follow a *fair game* process (see, for a simple taxonomy, Copeland *et al.* (2004)) with zero conditional and unconditional mean but with possibly time-varying higher order moments. On average, then, returns equal conditional expected returns or, equivalently, expected returns conditional on the available information set are unbiased estimates[8] of actual future returns. The key

---

[6] This is a definition of market efficiency implied by Fama (1970) discussion and reported in Fama (1976).

7 Recall that, if the distribution of the relevant conditioning variables is known, unconditional moments can be derived from the conditional ones. Therefore, if either asset cash-flows conditional expectations or conditional return moments are not formed using all the available relevant information, superior forecasts of asset prices could be formed by using conditioning variables that convey the relevant information neglected by market prices. These forecasts would be exploitable to earn above-average risk-adjusted returns. Clearly, this does not need to apply to conditional asset cash-flows expectations and return moments formed using subsets of the available information set, such as the data available to the econometrician.

[8] This condition can be formulated as follows: $E(r_{i,t}) = E[r_{i,t} - E(r_{i,t} |\Omega)]$, where Ω denotes the conditioning information set.

difference between rational asset pricing under the old and new paradigm[9] is that conditional expected returns and higher order moments of the returns deviations from their conditional means are fixed in the former and possibly time-varying in the latter.

The old paradigm implies the EMH, but the reverse is not true. In particular, within the new paradigm of asset returns, it is possible to recognize explanations for asset pricing phenomena based either on asset pricing models with investors that process information and decide upon it rationally, and thus consistently with the EMH, or on models that allow for some degree of investors' irrationality. I refer to the former as *rational asset pricing models* and to the as latter *behavioural asset pricing models*.

2.8. Rational Asset Pricing Models

Rational asset pricing models can be interpreted as specifications of a unified theoretical framework, the neoclassical rational economic model (Constantinides, 2002), that views expected excess returns as the reward demanded by risk averse, expected utility optimizing investors for bearing non diversifiable risk. These investors have unambiguously defined preferences over consumption. If we add the assumptions that investors' expectations are rational and investors' beliefs consistent, in the sense implied by Sargent (1996) discussion of the rational expectation equilibrium, this framework implies the EMH. Versions of this theory allow for market incompleteness, market imperfections, informational asymmetries, and learning. The theory also allows

[9] The key difference is therefore that the former relies on a random walk whereas the latter in based on a fair game view of conditionally unexpected returns. Also recall that conditional and unconditional moments are the same only in the random walk case.

for differences among assets for liquidity, transaction costs, tax status, and other institutional factors.

2.8.1 Consumption Asset Pricing

For inter-temporal utility maximizing investors, the SDF depends on their impatience and on the marginal utility of whatever they must give up in order to acquire additional units of the payoff x_{t+1}. To see this, suppose that investors extract utility from consumption, and that they have the following 2-period utility function:

$$U(C_t^j, C_{t+1}^j) = U(C_t^j) + \beta E\left[U(C_{t+1}^j)\right]. \tag{2-45}$$

Here, C_t^j denotes investor *j*'s consumption and β is the subjective discount factor that represents the investors' impatience[10], that is by how much, under *any* circumstance, *any* payoff is worth less if it is paid at a later date. Subjective discount factors should be always less than unity for impatient investors. Desirable properties of investors' utility function, as argued by Arrow (1971), are positive and decreasing marginal utility of wealth and non-increasing absolute risk aversion. Positive marginal utility of wealth, or $U' > 0$, implies investors' non satiation (NS), whereas decreasing marginal utility, $U'' < 0$, implies risk aversion (RA). Non increasing absolute risk aversion (NIARA), $\frac{d(-U''/U')}{dC^j} \leq 0$, implies that risky assets are not inferior goods and, as shown in Arditti (1967), it is a sufficient condition for $U''' \geq 0$. Hence $U''' \geq 0$ implies NIARA and aversion to negative skewness. NIARA, for a utility maximizing,

[10] It has nothing to do with the CAPM asset beta but I keep this notation because it is almost standard in the literature.

risk-averse individual, and hence with positive marginal utility and RA, is also related to prudence as defined by Kimball (1990). Included in the set of utility functions that display these desirable attributes are the logarithmic, power and negative exponential utility function. It should be noted that the popular quadratic utility function does not satisfy NIARA.

The investor in t must decide how much to consume and how much to invest in the asset that offers the payoff x_{t+1}. Subject to his inter-temporal budget constraint, the more of the asset he purchases, the less his consumption today but the more he will be able to consume in the future. The problem of a rational investor, therefore, is to find the level of investment that maximizes his expected utility. Assuming that the utility function is concave, denoting by $U'(C_t^j)$ the marginal utility of consumption and setting $m_{t+1}^j = \beta \frac{U'(C_{t+1}^j)}{U'(C_t^j)}$, (2-7) can be seen as the first order condition for the maximization of the investor *j*'s expected utility, i.e. the expectation of (2-45), given the price p_t of the pay-off x_{t+1}. In this setup, m_{t+1}^j is known as the investor's inter-temporal marginal rate of substitution and acts as the SDF that prices the assets faced by such investor. Treating the subjective discount factor as an inter-temporal constant, we thus have that the SDF m_{t+1}^j is proportional to the investor's marginal utility growth and (2-16) implies that the investor is willing to pay more for assets that are expected to pay off handsomely when her marginal utility of consumption is high.

The economy SDF, i.e. the process m_{t+1} that prices all pay-offs (that is, the payoffs faced by all investors), depends in general on the circumstances (factors) that determine the extent to which investors' aggregate marginal utility in $t+1$ is high relative to the previous period. The shape of investors' utility functions and the extent to which they can freely form portfolios has also implications for the shape of the SDF that prices the assets. For example, NS imply no-arbitrage and therefore a positive SDF. Furthermore, if the utility function is concave, marginal utility is high when resources to purchase additional units of consumption are scarce and therefore consumption is low. A payoff that made additional resources available when these were needed the most would be particularly welcome and the investors would value it more (m_{t+1} would be high). This implies a SDF decreasing in wealth. At a more technical level, the shape of investors' utility also has implications for how closely the SDF that prices all assets resembles the shape of individual investors' marginal utility growth. In other words, whether aggregation of individual investors' marginal utility growth results in a SDF defined over aggregate wealth with the same shape as the individual investors' SDF depends, in general, on the shape of the utility function. In empirical applications, the assumption that prices are set by a representative investor allows to bypass this issue (essentially, leaving it in the background for asset pricing theorists).

2.8.2 Representative Investor

Under the representative investor assumption, $\{C_i^j = C_i, \quad i = t, t+1\}$ and the SDF m_{t+1} can be expressed in terms of aggregate consumption as the growth of such investor's marginal utility:

$$m_{t+1} = \beta \frac{U'(C_{t+1})}{U'(C_t)}. \tag{2-46}$$

For example, in LeRoy (1973), Lucas (1978) or Breeden (1979), whose specifications represent instances of the consumption-based CAPM (henceforth C-CAPM), one can identify the stochastic discount factor with the IMRS of a representative agent. The asset pricing implications of the representative investor assumption and of the assumption that capital markets are complete are the same. This is because, in complete capital markets, as in Lucas (1978), investors can exchange contingent claims on any future state of the world. Full risk-sharing and diversification are therefore optimal for all investors, who then hold portfolios with risky assets in identical proportions. In these circumstances, pricing assets with respect to individual investors' consumption or with respect to aggregate consumption is equivalent because marginal utility growth is the same for all investors.

In a 2-period setting, investors must consume at the end of the second period all their wealth. Thus, in the SDF in (2-46), we can substitute out the representative investor's consumption with wealth. In a multi-period setting, consumption and wealth are equivalent only if either returns are unpredictable, as in the old paradigm, or predictability has no effect on inter-temporal optimal consumption-investment and portfolio choices. Strictly, the latter condition requires the assumption of logarithmic[11] utility. The empirical literature, e.g. Jagannathan and Wang (1996) and Lettau and Ludvigson (2001b), however, often assume that the SDF pricing equation holds conditionally period by period even under other types of utility functions. This corresponds to the assumption that predictability is at most a second order effect relative

[11] As explained in Cochrane (2005), for this type of utility function substitution effects (higher expected returns imply an higher opportunity cost of current consumption and therefore tend to decrease it) and income effects (higher expected returns imply higher next period wealth and therefore tends to increase consumption) exactly offset each other.

to the variability in consumption and wealth. Under these conditions, the inter-temporal marginal rate of substitution in (2-46) can be expressed as a function of aggregate wealth:

$$m_{t+1} = \beta \frac{U'(W_{t+1})}{U'(W_t)}. \tag{2-47}$$

The SDF of a representative investor with preferences defined over wealth that display NS, RA and NIARA is positive, decreasing and concave in wealth.

2.8.3 CAPM

The CAPM is a special single factor model. In its original version, it is a static equilibrium model. Under investors' NS, it can be derived either by assuming a representative investor with quadratic utility, thus excluding preference for moments of the multivariate distribution of asset returns higher than the second, or allowing for preference for higher moments (as under a power utility function) but assuming that returns are multivariate normal, and that investors, rational and risk averse, can freely diversify and have access to the same information. The latter assumption, even when a subset of investors is imperfectly rational, can be replaced by the assumption that the most informed marginal investor is rational and can borrow and lend without limits at the risk-free rate (this, essentially, requires a frictionless capital market). Quadratic utility assures that the U''' and U'''' term in (A-3) in Appendix A is zero. Under a multivariate normal distribution, the covariance with the squared rate of return on investor's wealth is zero (because of the symmetry of the normal distribution). In either case, the SDF depends linearly only on the return on the mean-variance efficient portfolio of risky assets, i.e. (2-24) becomes:

$$m_{t+1} = a + bR_{m,t+1} . \qquad (2\text{-}48)$$

From (2-42), then, the expected excess return on any asset is proportional to the coefficient of the regression of the asset excess returns on the portfolio excess return (that captures the asset systematic risk exposure). The proportionality coefficient, i.e. the risk-premium, is the market expected excess return. This is because, by construction, the regression coefficient $\beta_{m,m}$ of the market excess return on itself equals 1 and therefore, from (2-42), $\lambda_m = E(R_{m,t+1})$.

2.8.4 Conditional CAPM

We might extend the CAPM to an inter-temporal setting, where returns are not i.i.d. and moments are allowed to be time-varying, by letting the CAPM hold conditionally, period by period. This is clearly an approximation, as a rational mean-variance investor would anticipate the possibility of variation in the first moment of the return distribution and thus would seek to hedge against adverse (negative) changes in expected returns, i.e. a demand for hedging against reinvestment risk would arise and a corresponding risk premium would enter the equilibrium expected return determination equation. For the general derivation, please see Section 2.6. Since in the one-period CAPM expected returns are proportional to expected market variance, the latter would show up in the multi-period CAPM as an additional risk factor with a positive risk price in the

SDF of the representative investor, as in Merton (1973) Inter-temporal CAPM (henceforth, ICAPM), i.e.

$$E_t(R_{i,t+1}) = -b_{1,t}cov_t(R_{m,t+1}, R_{i,t+1}) - b_{2,t}cov_t(z_{t+1}, R_{i,t+1}). \qquad (2\text{-}49)$$

Here z_{t+1} is a state variable that describes the state of the investment opportunity set, i.e. it captures reinvestment risk. Merton (1980), however, points out that the hedging motive is likely not very important. Following Merton (1980) advice, Jagannathan and Wang (1996) set the price of reinvestment risk to zero and approximate the SDF as a linear function of the return on the market portfolio with time-varying parameters, i.e. $m_{t+1} = a_t + b_t R_{m,t+1}$. Such a SDF summarizes the asset pricing implication of the conditional CAPM, henceforth (C)CAPM. Alternatively, we could just treat this specification of the SDF as a reduced form representation of the true inter-temporal SDF. In any case, letting the SDF parameters depend linearly on the conditioning variable z_t, as in (2-34), and setting $f_{t+1} = R_{m,t+1}$ in (2-24), we have:

$$\begin{aligned} m_{t+1} &= a_0 + a_1 z_t + (b_0 + b_1 z_t) R_{m,t+1} \\ &= a_0 + a_1 z_t + b_0 R_{m,t+1} + b_1 z_1 R_{m,t+1} \end{aligned}. \qquad (2\text{-}50)$$

Using (2-50) in (2-42), the beta-pricing representation of the conditional excess return pricing implications of the (C)CAPM is the following:

$$E_t(R_{i,t+1}) = \beta_{m,i,t}\lambda_{m,t}. \qquad (2\text{-}51)$$

Here, $\beta_{m,i,t}$ is a time-varying coefficient of the regression of $R_{i,t+1}$ on $R_{m,t+1}$ and $\lambda_{m,t}$ is the conditional market risk premium, given by (2-44):

$$\lambda_{m,t} = -var_t(R_{m,t+1})b_t . \tag{2-52}$$

Also, since by definition $\beta_{m,m,t} = 1$, we have from (2-51) that $\lambda_{m,t} = E_t(R_{m,t+1})$. Hence, the conditional market risk premium is equal to the conditional market expected excess return. Since, as shown in Appendix A, $b_{1,t} = -RRA_t$, the market risk premium in (2-52) can be rewritten as follows:

$$\lambda_{m,t} = E_t(R_{m,t+1}) = RRA_t var_t(R_{m,t+1}) . \tag{2-53}$$

RRA_t can be interpreted as the representative investor's relative risk aversion parameter for reasons that become clear by examining the derivation of the stylized risk-return relation reported in Appendix (B-12) and (B-16).

To derive the unconditional implications of the conditional SDF model in (2-50), I apply (2-51) and take unconditional expectations of both sides:

$$E(R_{i,t+1}) = E(\beta_{m,i,t}\lambda_{m,t}) = \beta_{i,t}\lambda_t = \beta_{z,i}\lambda_z + \beta_{m,i}\lambda_m + \beta_{zm,i}\lambda_{zm} . \tag{2-54}$$

Here, $\beta_{z,i}$, $\beta_{m,i}$ and $\beta_{zm,i}$ are regression coefficients of $R_{i,t+1}$ on, respectively, z_t, $R_{m,t+1}$ and $z_t R_{m,t+1}$. Equivalently, the SDF in (2-50) can be seen as a linear function of

z_t, $R_{m,t+1}$ and $z_t R_{m,t+1}$. Hence, (2-54) can be derived applying (2-42) to (2-50) with the elements of F_t in (2-37) given by z_t, $R_{m,t+1}$ and $z_t R_{m,t+1}$. Notice that, if the parameters of the SDF are fixed, $a_t = a_0 = a$ and $b_t = b_0 = b$, the preceding equations simplify to the unconditional CAPM (CAPM), i.e. $m_{t+1} = a + bR_{m,t+1}$, and $E(R_{i,t+1}) = \beta_{m,i}\lambda_m$.

2.8.5 Conditioning Variables

A critical consideration in estimating the (C)CAPM, or any conditional asset pricing model, is the choice of the conditioning variable z_t. The conditioning variable should capture the time variation in the parameters of the SDF. There are two main theoretical reasons why the parameters of a SDF conditionally defined over market wealth might change over time.

One relates to non-market sources of risk and the impact of economy-wide shocks on the marginal utility of stock market wealth. From this perspective, we seek conditioning variables that proxy for the state of the economy and, in particular, for sources of systematic variation in non-market wealth, such as labor income shocks and real estate returns. These are labeled by Cochrane (2001) “distress risk” factors or recession variables and should capture sources of systematic risk different from the stock market. During a recession unemployment is high, labor income is low and more volatile and property prices falter. If investors’ marginal utility of stock market wealth is higher under these circumstances than in good times, variables that capture the state of the economy should show up as priced risk factors alongside the stock market factor. This ultimately implies that investors’ utility is not defined only over stock market wealth but also over other forms of wealth. In turn, this implies that the stock market is not a good proxy for overall wealth. The recession state variables do not need to predict anything (either the stock market or the future state of the economy) but they should be highly correlated with the wider economy or particular (sizeable) portions of it unrelated to the stock market. In other words, they should represent good instruments for the state of portions of the economy, unrelated to the stock market but relevant in determining investors' marginal utility. High correlation implies that the conditioning variable should be either highly pro-cyclical or anti-cyclical relative to these portions of the economy. If they were pro-cyclical, they would command a positive risk premium. If they were anti-cyclical, they would command a negative risk premium (exposure to them would represent an insurance against a non-stock market source of systematic risk).

The other theoretical reason why the parameters of the SDF might change over time relates, in Merton (1973) ICAPM framework, to inter-temporal risk and to the impact of changes to the future investment opportunity set on marginal utility of wealth. Thus we seek conditioning variables that summarize the predictable evolution of the investment opportunity set and hence provide a summary measure of expected excess returns. These variables should, in other words, predict excess returns. In particular, in a world where only systematic risk matters to investors, the conditioning variable should help forecast market returns. The empirical literature has proposed a number of variables that help predict future returns. The most successful are the stochastically de-trended short term interest rate, employed among others by Scruggs (1998), the book to market value ratio, the dividend-price ratio, used by Campbell and Shiller (1988), and the observable proxy for the consumption-wealth ratio proposed by Lettau and Ludvigson (2001a).

Theoretical arguments that suggest that the consumption-wealth ratio and the dividend-price ratio should predict future returns are especially compelling.

To show that the consumption–aggregate wealth ratio summarizes agents' expectations of future returns, Lettau and Ludvigson (2001a), using a log-linear approximation to a representative investor's inter-temporal budget constraint $W_{t+1} = R_{m,t+1}(W_t - C_t)$, express the log consumption-wealth ratio in terms of future returns to the market portfolio and future consumption growth. Because this approximation is based on the agent's inter-temporal budget constraint, it holds both ex post and ex ante. Accordingly, the log consumption–wealth ratio may be expressed in terms of expected returns to the market portfolio and expected consumption growth as:

$$c_t - w_t \cong E_t \sum_{j=1}^{\infty} \rho_w^j (r_{m,t+j} - \Delta c_{t+j}) . \qquad (2\text{-}55)$$

Here, lower case letters denote logarithms of (per capita) consumption and wealth and ρ_w is the steady-state ratio of invested to total wealth. This essentially means that, given the representative investor's wealth, the amount of consumption today depends on the amount he wishes to be able to afford to consume tomorrow and, therefore, on his expected future consumption. Under Muth (1961) rational expectations (henceforth RE), the above equation implies that, if consumption growth is not too volatile (something that appears to be true empirically), the variation in the log consumption-wealth ratio must be driven by variation in expected returns. It therefore summarizes expectations of future returns on the market portfolio. Intuitively, if the consumption-wealth ratio is high, then the agent must be expecting either high returns on wealth in the future or low consumption growth rates (boosting in both cases current consumption). Since consumption growth rates are fairly stable, however, swings in the consumption-wealth ratio should be related to changing agents' expectations about aggregate returns and therefore, under RE, they should predict aggregate returns.

Of course, the log consumption-aggregate wealth ratio is not observable because human capital is not observable. To overcome this obstacle, Lettau and Ludvigson (2001a) construct a proxy based on *observable* quantities. Denote non-human or asset wealth by A_t and its logarithm as $a_t = log(A_t)$. Also, assume that human capital H_t is on average a constant multiple of labor Y_t income, i.e. $H_t = KY_t$. Its logarithm then can be written as $h_t = k + y_t + v_t$, where k is a constant and v_t is a mean zero stationary random variable with $E(v_t) = 0$. Lettau and Ludvigson (2001a) reformulate the bivariate cointegrating relation between c_t and w_t in the consumption-wealth ratio equation (c_t and w_t are both integrated but their linear combination on the right hand side is stationary) as a trivariate co-integrating relation involving the three *observable* variables log consumption c_t, log nonhuman or asset wealth a_t, and log labor earnings y_t. Since c_t and a_t are both $I(1)$, such a reformulation is possible, by Engle and Granger (1987) representation theorem, under the condition that labor income is integrated and the rate of return to human capital is stationary. Aggregate wealth is $W_t \equiv A_t + H_t$ and log aggregate wealth may be approximated as $w_t \cong \omega a_t + (1-\omega)h_t$ where ω equals the average share of nonhuman wealth in total wealth, A_t/W_t. The left-hand side of (2-55) may then be expressed as follows:

$$
\begin{aligned}
c_t - w_t &\cong c_t - \omega a_t - (1-\omega) h_t \\
&= c_t - \omega a_t - (1-\omega)(k+y+v)_t \\
&= c_t - \omega a_t - (1-\omega) y_t - (1-\omega)(k+v)_t \\
&= cay_t - (1-\omega) k_t - (1-\omega) v_t
\end{aligned}
\quad . \qquad (2\text{-}56)
$$

Here, $cay_t \equiv c_t - \omega a_t - (1-\omega) y_t$ is the difference between log consumption and a weighted average of log asset wealth and log labor income. Solving (2-56) for cay_t and using (2-55), we can write:

$$
\begin{aligned}
cay_t &\equiv (1-\omega)k + E_t \sum_{j=1}^{\infty} \rho_w^j (r_{m,t+j} - \Delta c_{t+j}) + (1-\omega) v_t \\
&\equiv const. + E_t \sum_{j=1}^{\infty} \rho_w^j (r_{m,t+j} - \Delta c_{t+j}) + (1-\omega) v_t
\end{aligned}
\quad . \qquad (2\text{-}57)
$$

Because all the variables on the right-hand side of the above equation are stationary, the model implies that cay_t is stationary and hence that consumption, asset wealth, and labor income share a common stochastic trend (they are cointegrated), with ω and $1-\omega$ parameters of this shared trend. If the cointegrating parameter ω can be consistently estimated, cay_t can be treated as observable. As long as the error term v_t on the right-hand side is not too variable, this equation also implies that cay_t should be a good proxy for the unobservable quantities on the right hand side of (2-57) and therefore for variation in the log consumption–aggregate wealth ratio and expected returns. An important issue in using the left-hand side of this equation as a conditioning variable is the estimation of the parameters in cay_t. Lettau and Ludvigson (2001a) discuss how the cointegrating parameter ω can be estimated consistently. As suggested by Lettau and Ludvigson (2001a) , it is the cay_t time-series constructed using the estimated ω parameter and the observed log consumption c_t, log asset wealth a_t and log labor earnings y_t that can be employed as a scaling variable in a conditional asset pricing model.

The specification of the consumption-wealth ratio equation reported above is analogous to the linearized formula for the log dividend–price ratio (Campbell & Shiller, 1988), where consumption enters in place of dividends and wealth enters in place of price:

$$p_t - d_t \equiv const. + E_t \sum_{j=1}^{\infty} \rho^{j-1} (\Delta d_{t+j} - r_{m,t+j}) . \qquad (2\text{-}58)$$

Here, d_t denotes log-dividends, r_t denotes returns, and ρ can be seen as the steady state dividend yield. Because all the variables on the right-hand side of the above equation are stationary, the model implies that $p_t - d_t$ is stationary and hence that prices and dividends share a common stochastic trend (they are cointegrated), with 1 and –1 parameters of their cointegrating relation. If the dividend-price ratio is high, investors must be expecting either high returns on the stock market portfolio in the future or low dividend growth rates. Since both consumption and dividends are not very volatile and their growth rates are relatively unpredictable, high wealth and high stock market prices relative to, respectively, consumption and dividends (but also relative to the book value and other metrics) must predict low future returns. The key difference between the consumption-wealth ratio and the dividend-price ratio is what is on the right-hand side: in the equation for the consumption-wealth ratio it is the return to the *entire* market portfolio and consumption growth, whereas in the dividend-price ratio equation it is the return to the stock market component of wealth and dividend growth.

Lettau and Ludvigson (2001a) and Guo and Savickas (2003) present evidence that cay_t is a good predictor of excess returns on aggregate stock market indices. Evidence that the dividend-price ratio or the dividend yield is a good predictor of returns is given, among others, by Campbell and Shiller (1988), Campbell (1991), more recently, Cochrane (1999a, 2001); Fama and French (1996). It is worth stressing that predictability is a long-horizon effect. The updated predictability of 1 and 5 year returns using the dividend-price ratio as a forecasting regression variable is reported in Table 2-1, reproduced from Cochrane (2011). The dividend-price ratio predicts 9 percent of the variation in 1 year returns and its explanatory power rises steadily as the horizon increases. It predicts up to 28 percent of the variation in 5 year returns. Even though the explanatory power, R^2, of the regression is inflated by an overlapping observations problem, the results at different horizons are reflections of a single underlying phenomenon. Even a small short run predictive power or non-zero contemporaneous correlation build up to yield substantial returns predictability at longer horizon if the forecasting variable is persistent. For example, if daily returns are very slightly predictable by a slow-moving (i.e., persistent) variable, that predictability adds up over long horizons. As argued by Cochrane (1999a) in a very illuminating way, you can predict that the temperature in Chicago will rise about one-third of a degree per day in spring. This forecast explains very little of the day to day variation in temperature but, because temperature changes are persistent (within each season), it tracks almost all the rise in temperature from January to July. Thus, the R^2 rises with horizon. Precisely, suppose that we forecast excess returns with a forecasting variable x:

$$r_{t+1} = a + bx_t + e_{t+1}$$
$$x_{t+1} = c + \rho x_t + \varepsilon_{t+1} \tag{2-59}$$

Even for small values of short-horizon b and R^2 in the first equation above, a large coefficient ρ in the second equation implies that the long-horizon regression has a large regression coefficient b and a large R^2. This regression has a powerful implication: stocks are in many ways like bonds. Any bond investor understands that a string of good past returns that pushes the price up is bad news for subsequent returns. Many stock investors see a string of good past returns and interpret this as a sign of a bull market, concluding that future stock returns will be good as well. The regression reveals the opposite: a string of good past returns which drives up stock prices is bad news for subsequent stock returns, as it is for bonds.

Table 2-1

Return Predictability

Horizon *k*	***b***	**t-statistic**	**R^2**
1 year	3.8	(2.6)	0.09
5 years	20.6	(3.4)	0.28

Notes: This table reports OLS regressions of excess-returns (value-weighted NYSE - Treasury bill rate) on value-weighted dividend/price ratio (reproduced from Cochrane (2011)):

$$r_{t\to t+k} = a + b(D_t / P_t) + e_t$$

$r_{t\to t+k}$ denotes the *k*th year return. Standard errors use Hansen–Hodrick correction for heteroskedasticity and error autocorrelation.

The co-integrating relation between consumption, asset wealth, and labor income and between consumption and dividends imply that asset prices are set as predicted by rational asset pricing models, i.e. prices and wealth equal the present value of the rational expectation of future cash flows, either consumption (real cash flows) or dividends, discounted at the equilibrium expected rate of return. The valuation implied by any such model is the solution to a stochastic differential equation where prices equal the present value of the rational expectation of next period dividend or consumption flows and capital gains discounted at the equilibrium expected rate of return. For this differential equation to have a determinate solution, a boundary condition that rules out bubbles must hold. Without this condition (equivalent to requiring that sooner or later any bubble bursts), any self-fulfilling expectation of capital gains would imply a different yet legitimate solution. In turn, the lack of this boundary condition would imply that the right hand side of (2-58) and (2-59) contains a non-stationary bubble component (in addition to the stationary terms in the rate of future returns and of consumption or dividend growth) and the left hand side would be non-stationary.

The intimate relation between stationarity of the left-hand side of (2-58) and (2-59) and rational valuation has generated intense interest in tests of the co-integrating relation between variables such as prices and dividends or consumption, asset wealth and labor income. While Lettau and Ludvigson (2001a) find that consumption, asset wealth and labor income are co-integrated and a large body of evidence suggests that the dividend-price ratio is stationary, see for example Cochrane (1999a, 2001), the evidence that prices and dividends are co-integrated is at best weaker. In particular, tests based on the Engle and Granger (1987) methodology find limited evidence of cointegration between dividend and prices, see for example Campbell and Shiller (1988), Diba and Grossman (1988), Froot and Obstfeld (1991), Balke and Wohar (2001). Since prices are much more volatile than dividends, see for example Campbell and Shiller (1988) and Campbell (1987), it is possible that these tests fail to detect cointegration because the parameters of the cointegrating relation are time varying and, in particular, they display heteroskedastic variability characterized by clustering over time. Heteroskedastic time-variation in the parameters of the cointegrating relation in turn might help explain heteroskedastic excess volatility of prices over fundamentals. Harris *et al.* (2002) introduce a test for stochastic cointegration, where the parameters of the cointegrating relation are allowed to be time varying. This test encompasses the test for cointegration with fixed parameters of the cointegrating relation, defined stationary cointegration. Harris *et al.* (2002) find mixed evidence in favour of stochastic cointegration between stock and dividends but this evidence is stronger than the evidence in favour of stationary cointegration.

2.8.6 The Role of Systematic Skewness

Non-normal return distributions cannot be entirely described by first and second moments. Unless investors display a special type of preferences (quadratic), they care about higher moments. In particular, while NIARA rules out preference for negative portfolio skewness, decreasing absolute risk aversion (DARA) implies preference for positive skewness. As argued by Richter (1960), Levy (1969) and Kraus and Litzenberger (1976), an exact preference ordering for risky portfolios using the first three moments of the portfolio return is possible, in general, only for an investor with a cubic utility function of wealth. Unfortunately, as shown by Levy (1969) and Tsiang (1972), this third degree polynomial utility function is unsuitable to model the preferences of a risk adverse investor. Duly restricted third order Taylor expansions of admissible non-polynomial utility functions can be used instead. Under (NS, RA and) DARA and hence if the investor has a preference for positive portfolio skewness, he should be willing to accept a somewhat lower expected return to hold assets with positive coskewness.

2.8.7 *The 3M-CAPM*

Kraus and Litzenberger (1976) consider the optimal portfolio choice of a representative investor that lives in a 1-period economy. His utility is defined over end of period wealth W, i.e. $U = U(W)$, and it is not restricted to any particular functional form. The only requirement is that it be continuous and three times continuously differentiable over the range of wealth. In this very simple 1-period setting, where the investor does not have to solve the usual optimal consumption-investment decision problem that arises in multi-period (2 or more periods) models, the Euler equation for the maximization of his expected utility is:

$$E_t\left[U'(W_{m,t+1})r_{i,t+1}\right] = 0\,. \tag{2-60}$$

As shown in Section 8.1 Appendix A, a third order Taylor expansion of a standardized utility function around the point $W_0 = E(W) = 1$ yields:

$$\begin{aligned} U(W_{m,t+1}) &\cong \left[R_{m,t+1} - E(R_{m,t})\right] + \theta_{1,t}\left[R_{m,t+1} - E(R_{m,t})\right]^2 + \theta_{2,t}\left[R_{m,t+1} - E(R_{m,t})\right]^3 \\ &\cong \left[r_{m,t+1} - E(r_{m,t})\right] + \theta_{1,t}\left[r_{m,t+1} - E(r_{m,t})\right]^2 + \theta_{2,t}\left[r_{m,t+1} - E(r_{m,t})\right]^3 \end{aligned}\,.$$

(2-61)

Here, $\theta_{1,t} = \dfrac{1}{2}U''(1)$ and $\theta_{2,t} = \dfrac{1}{6}U'''(1)$. In the second line of (2-61), I use excess returns instead of returns because, in this simple 1-period setting where the distinction between unconditional and conditional moments is irrelevant, the risk free rate is known with certainty (also conditionally) and, therefore, $R_{m,t+1} - E_t(R_{m,t}) = r_{m,t+1} - E_t(r_{m,t})$. Differentiating (2-61) once with respect to wealth, marginal utility can be approximated as follows:

$$U'(r_{m,t+1}) \cong 1 + 2\theta_{1,t}\left[r_{m,t+1} - E(r_{m,t})\right] + 3\theta_{2,t}\left[r_{m,t+1} - E(r_{m,t})\right]^2. \qquad (2\text{-}62)$$

Using (2-62) in (2-60) yields Kraus and Litzenberger (1976) 3M-CAPM. Interpreting marginal utility $U'(W_{m,t+1})$ as a SDF, (2-60) can be seen as a version of (2-16) where $r_{i,t+1} = R_{i,t+1} - R_{f,t+1}$ and, because $r_{i,t+1}$ is an excess return, $p_t = 0$. In (2-62), the SDF is approximated as a linear function of the market excess return and its square and thus it can be seen as an instance of (2-24) with $f_{1,t+1} = r_{m,t+1} - E(r_{m,t})$ and $f_{2,t+1} = \left[r_{m,t+1} - E(r_{m,t})\right]^2$, $a_t = 1$, $b_{1,t} = 2\theta_1$, $b_{2,t} = 3\theta_2$. Applying (2-16), and dropping time-subscripts for notational simplicity, (2-60) and (2-62) imply:

$$E(r_i) = -\frac{cov[U'(r_m \mid \theta), r_i]}{E[U'(r_m \mid \theta)]} = -\frac{E\{[U'(r_m \mid \theta) - E(U'(r_m \mid \theta))][r_i - E(r_i)]\}}{E[U'(r_m \mid \theta)]}. \qquad (2\text{-}63)$$

Here, differentiating (2-62) once, $U''(r_m \mid \theta) = 2\theta_1 + 6\theta_2 r_m$ and, differentiating it once more, $U'''(r_m \mid \theta) = 6\theta_2$. Finally, multiplying and dividing the first and second term on the right-hand side of this equation by, respectively, $E[r_m - E(r_m)]^2$ and $E[r_m - E(r_m)]^3$ and re-arranging, we can write[12]:

$$E(r_i) \cong \delta_1 \beta_i + \delta_2 \gamma_i \,. \tag{2-64}$$

where,

$$\delta_1 = -\frac{E[U''(r_m \mid \theta)]E[r_m - E(r_m)]^2}{E[U'(r_m \mid \theta)]} \,. \tag{2-65}$$

$$\delta_2 = -\frac{E[U'''(r_m \mid \theta)]E[r_m - E(r_m)]^3}{2E[U'(r_m \mid \theta)]} \,. \tag{2-66}$$

$$\beta_i = \frac{E[(r_i - E(r_i))(r_m - E(r_m))]}{E[r_m - E(r_m)]^2} \,. \tag{2-67}$$

$$\gamma_i = \frac{E[(r_i - E(r_i))(r_m - E(r_m))^2]}{E[r_m - E(r_m)]^3} \,. \tag{2-68}$$

Here, the coefficient δ_1 is the beta premium and the coefficient δ_2 is the gamma premium. The assumption of greed implies $E[U'(r_m \mid \theta)] > 0$ and, under RA, $E[U''(r_m \mid \theta)] \leq 0$. Thus, since $E[r_m - E(r_m)]^2 \geq 0$, the beta coefficient δ_1 is *positive* for risk-averse, greedy investors. If the market portfolio skewness is negative (as it is often the case empirically) and if there is a market reward for holding assets with negative

---

[12] I also assume that the second derivative of the utility function does not depend on the interaction between market and asset unexpected returns, $Cov\left\{u''(r_m \mid \theta), \left[r_i - E(r_i)\right]\left[r_m - E(r_m)\right]\right\} = 0$, and that the third derivative does not depend on the interaction between squared market unexpected returns and asset unexpected returns $Cov\left\{u'''(r_m \mid \theta), \left[r_i - E(r_i)\right]\left[r_m - E(r_m)\right]^2\right\}$. These are very useful and reasonable simplifications that, intuitively, correspond to the requirement that absolute risk aversion and preference towards skewness do not depend on the relation between a single asset and the market portfolio or its square (rather, they should depend only on the latter, i.e. the market return and its square). Essentially, only changes in overall wealth and in its volatility should determine moves along the utility function and, therefore, changes in the point at which its derivatives are evaluated.

systematic asset coskewness, the gamma coefficient δ_2 is *positive*. This can also be seen by noting that, under the assumption of greed and NIARA, $E[U'(r_m \mid \theta)] > 0$ and $E[U'''(r_m \mid \theta)] \geq 0$ respectively. Since empirical market portfolio skewness is usually found to be negative, i.e. $E[r_m - E(r_m)]^3 \leq 0$, then $\delta_2 \geq 0$. While δ_1 represents investors' reward for systematic variance, i.e. for holding assets that increase the volatility of the overall market portfolio, δ_2 compensates investors for systematic negative skewness, i.e. for holding assets that decrease the skewness of the overall market portfolio (that cause the distribution of portfolio returns to be skewed to the left).

2.8.8 The 3M-(C)CAPM

While allowing utility to contain a cubic term in wealth, its parameters could be allowed to be time varying. For example, the elements of θ in (2-60) could be specified as a function of conditioning information. A particularly interesting possibility is that they vary with the business cycle or that they are a function of conditioning variables that represent investors' expectations about future returns. This would yield a conditional version of Kraus and Litzenberger (1976) 3M-CAPM. Following a similar approach, Harvey and Siddique (2000) propose a conditional asset pricing equation where expected asset excess returns are a function of their conditional covariance and coskewness with the market portfolio and the prices of covariance and coskewness risk also vary over time:

$$\begin{aligned} E_t(r_{i,t+1}) &= \lambda_t cov_t(r_{i,t+1}, m_{t+1}) \\ &= \lambda_{1,t} cov_t(r_{i,t+1}, r_{m,t+1}) + \lambda_{2,t} cov_t(r_{i,t+1}, r_{m,t+1}^2) \end{aligned} . \tag{2-69}$$

where,

$$
\begin{aligned}
\lambda_{1,t} &= \frac{var_t(r_{m,t+1}^2)E_t(r_{m,t+1}) - skew_t(r_{m,t+1})E_t(r_{m,t+1}^2)}{var_t(r_{m,t+1})var_t(r_{m,t+1}^2) - [skew_t(r_{m,t+1})]^2} \\
\lambda_{2,t} &= \frac{var_t(r_{m,t+1})E_t(r_{m,t+1}^2) - skew_t(r_{m,t+1})E_t(r_{m,t+1})}{var_t(r_{m,t+1})var_t(r_{m,t+1}^2) - [skew_t(r_{m,t+1})]^2}
\end{aligned} . \qquad (2\text{-}70)
$$

Here, the symbol $skew_t(r_{m,t+1}) = E_t\{[r_{m,t+1} - E_t(r_{m,t+1})],[r_{m,t+1}^2 - E_t(r_{m,t+1}^2)]\}$ represents the skewness of the market portfolio in $t+1$ conditional on information available at time t , while the other symbols (e.g. conditional expectation and variance operators) have the usual meaning. To prove this, we only assume:

$$
E_t(r_{i,t+1}) = \beta_{1,t}E_t(r_{m,t+1}) + \beta_{2,t}E_t(r_{m,t+1}^2) . \qquad (2\text{-}71)
$$

Since,

$$
\begin{aligned}
\beta_{1,t} &= \frac{cov_t(r_{i,t+1}, r_{m,t+1})var_t(r_{m,t+1}^2) - skew_t(r_{m,t+1})cov_t(r_{i,t+1}, r_{m,t+1}^2)}{var_t(r_{m,t+1})var_t(r_{m,t+1}^2) - [skew_t(r_{m,t+1})]^2} \\
\beta_{2,t} &= \frac{cov_t(r_{i,t+1}, r_{m,t+1}^2)var_t(r_{m,t+1}) - skew_t(r_{m,t+1})cov_t(r_{i,t+1}, r_{m,t+1})}{var_t(r_{m,t+1})var_t(r_{m,t+1}^2) - [skew_t(r_{m,t+1})]^2}
\end{aligned} , \qquad (2\text{-}72)
$$

then, we combine (2-71) and (2-72) by taking the terms $cov_t(r_{i,t+1}, r_{m,t+1})$ and $cov_t(r_{i,t+1}, r_{m,t+1}^2)$ out, we will have the equations in (2-70), i.e.

$$
\begin{aligned}
E_t(r_{i,t+1}) = {} & \frac{var_t(r_{m,t+1}^2)E_t(r_{m,t+1}) - skew_t(r_{m,t+1})E_t(r_{m,t+1}^2)}{var_t(r_{m,t+1})var_t(r_{m,t+1}^2) - [skew_t(r_{m,t+1})]^2} cov_t(r_{i,t+1}, r_{m,t+1}) \\
& + \frac{var_t(r_{m,t+1})E_t(r_{m,t+1}^2) - skew_t(r_{m,t+1})E_t(r_{m,t+1})}{var_t(r_{m,t+1})var_t(r_{m,t+1}^2) - [skew_t(r_{m,t+1})]^2} cov_t(r_{i,t+1}, r_{m,t+1}^2)
\end{aligned} . \qquad (2\text{-}73)
$$

The pricing equation in (2-71) can be derived from (2-25) specifying the SDF m_{t+1} as a quadratic polynomial in the market excess return $r_{m,t+1}$ with parameters a_t, $b_{1,t}$ and $b_{2,t}$ that are allowed to vary over time, $R_{f,t+1} \cong 1$, and therefore[13], $a_t + b_t' E_t(f_{t+1}) \cong 1$. That is

$$m_{t+1} = a_t + b_{1,t} r_{m,t+1} + b_{2,t} r_{m,t+1}^2 . \tag{2-74}$$

Interpreting m_{t+1} in (2.60) as the SDF implied by a third order Taylor expansion of the representative investor's utility function, the pricing equation in (2.59) can be seen as the cross-sectional implication of a conditional version of the 3 moment CAPM (henceforth, 3M-(C)CAPM). Under this model, if investors like positive portfolio skewness, they should accept a negative risk premium to hold assets with positive coskewness because these assets contribute to increase the skewness of the overall market portfolio. In other words, the portfolio with positive coskewness is a ideal hedging instruments when the market is volatile, and the investors who hold such portfolio should pay extra price by bearing the lower premium. The price of coskewness risk $\beta_{2,t}$, therefore, should be negative. It is worth at this point highlighting the difference with the 3-moment model derived by Kraus and Litzenberger (1976) where, if market portfolio skewness is negative, positive asset coskewness implies a negative gamma and a positive δ_2. In other words, the specification of the systematic third

---

[13] Recall that, as shown in Section 2.2, in (2.12), and 2.3, in (2.19) and (2.22), without this approximation and the resulting restriction on the relation between the intercept and the mean of the factors the risk free return would show up in the equations for the risk prices.

moment premium used by Harvey and Siddique (2000) and by Kraus and Litzenberger (1976) are not equivalent.

2.8.9 Tests of the 3M-(C)CAPM

Harvey and Siddique (2000) test the 3M-CAPM on Centre for Research on Security Prices (CRSP) NYSE, AMEX and NASDAQ stock data over the period 1963-1993. They find that the 3M-CAPM significantly improves on a two-moment CAPM specification. They report that coskewness helps explain the cross-section of average excess-returns on 32 industry portfolios and 25 size and book-to-market value sorted portfolios. Moreover, they find that coskewness retains a significant explanatory power even after the inclusion of factors related to size and book to market value that have been found by Fama and French (1992, 1993, 1995) to empirically explain a large portion of the cross-sectional variation in average asset returns. In particular, they find that systematic skewness is important and commands on average a risk premium of 3.6 percent per annum.

Dittmar (2002) specifies a conditional model by expressing the parameters of a quadratic and cubic SDF as linear functions of a set of conditioning variables. The quadratic SDF implies the 3M-(C)CAPM whereas the cubic SDF implies a 4 moment CAPM where preference for co-kurtosis (the systematic fourth moment), is allowed. The conditioning variables include one lag of the market excess-return, of the dividend yield, the spread of the 3 Month T-Bill over the 1 Month T-Bill rate and the 1 Month T-Bill rate itself. He finds evidence of substantial non-linearity in the pricing kernel and that both the quadratic and cubic SDF fit well the cross-section of US industry equity indices average returns over the period 1963-1995. After imposing the regularity conditions on the shape of the utility functions that correspond to *standard* risk aversion, i.e. positive marginal utility, RA and NIARA over all values of wealth, the estimated gamma premium remains statistically and economically significant but it becomes much smaller, thus considerably reducing the ability of the estimated 3 and 4 moment conditional specifications to explain the cross-section of average returns.

Post *et al.* (2008) criticise previous empirical tests of the 3M-CAPM, such as Harvey and Siddique (2000), on the grounds that they fail to check whether the decreasing marginal risk-aversion requirement is satisfied by the estimated pricing model. Consistently with Dittmar (2002), they show that the *gamma* (standardised asset co-skewness) premium turns out very small when the appropriate regularity conditions (risk aversion) are imposed on the shape of the investor utility function. In fact, fitting a cubic utility to data on the Fama and French (1995) market portfolio and on 10 size-ranked portfolios for the period 1963-2001, their estimated expected utility function does not satisfy the concavity requirement over the relevant wealth interval and thus the market portfolio is not guaranteed to be efficient for the representative investor. Moreover, they find that the market portfolio is likely to minimize the sample expected utility, rather than maximize it as predicted by the 3M-CAPM.

2.8.10 3M-CAPM vs. (C)CAPM

Even though the (C)CAPM uses the assumption that investors have a quadratic utility function and its pricing kernel does not incorporate 3^{rd} order terms, the unconditional implications of both the 3M-CAPM and the (C)CAPM contain a premium for a cross third moment of asset returns. The 3M-CAPM contains a premium for the cross third moment between asset return and the square of the market return, i.e. a premium for

$cov(r_{i,t}, r_{m,t}^2)$. The (C)CAPM contains a cross third moment between the asset return, the market return and a conditioning variable that influences marginal utility of market wealth, i.e. a premium for $cov(r_{i,t}, z_{t-1}r_{m,t})$. Equivalently, asset coskewness can be seen as the covariance between the asset return and market volatility[14], whereas in the (C)CAPM the expression $cov(r_{i,t}, z_{t-1}r_{m,t})$ can be interpreted as the covariance between the asset return and the sensitivity of the market return to the conditioning variable. In other words, in the 3M-CAPM investors are rewarded for holding assets that perform poorly at times of high market volatility, whereas in the (C)CAPM they are rewarded for holding assets that do poorly when the return on the stock market portfolio is very reactive to the conditioning variable, and hence when it is very reactive to either returns on non-market wealth or expected stock market returns. There are a number of circumstances under which the (C)CAPM expression $cov(r_{i,t}, z_{t-1}r_{m,t})$ could proxy for asset coskewness and vice versa. This would be the case if z_{t-1} was a good proxy for $r_{m,t}$ and hence when the former forecasts the latter.

2.8.11 Higher Moment CAPM and the Shape of the Utility Function

Consistent with the 3M-CAPM, Dittmar (2002) finds that SDF specifications quadratic and cubic in the market return provide a much better fit to the observed cross-section of stock returns than linear models. This author also emphasises, however, that the best fit obtains with a U-shaped SDF, whereas the superior performance of the 3M and 4M-CAPM is significantly reduced when the SDF is restricted to be decreasing in the market return. Post *et al.* (2008) report similar findings and reach similar conclusions.

[14] More accurately, coskewness should be seen as the covariance with the realization of the market second moment.

This evidence implies that marginal utility of wealth increases above a threshold and thus that the utility function takes an “inverse *S*” shape. Such a shape, in turn, implies risk-seeking over a range of sample wealth.

It is not uncommon, in the extant literature, to encounter studies that find evidence of risk-seeking and specifications of investors’ preferences that admit this type of behaviour. Psychologists, led by Kahneman and Tversky (1979), find experimental evidence of local risk seeking. Friedman and Savage (1948) and Markowitz (1952) suggest that the willingness to purchase both insurance and lottery tickets implies that marginal utility is increasing over a range of investors’ wealth. Golec and Tamarkin (1998) and Garrett and Sobel (1999) provide evidence that risk-averse individuals take actuarially unfair “long shot” gambles, i.e. low-probability but high-variance bets, and argue that this behaviour can be rationalized by postulating skewness preference and an inverse-S shaped utility function. Post and Van Vliet (2006) and Post *et al.* (2008) argue, however, that non-concave utility is problematic from the viewpoint of the 3M-CAPM. In essence, they point out that, if the representative investor’s utility function is not concave, the market portfolio is not guaranteed to maximize her expected utility function and the 3M-CAPM does not necessarily hold even if the estimated SDF is quadratic in the return on such portfolio. Their remark extends in a straightforward manner to models of the SDF cubic in market wealth, implying a similar difficulty in interpreting the explanatory power of such models as evidence in favour of the 4M-CAPM.

2.9. Behavioural Models

The behavioural finance explanation of the stylized features of the distribution of asset returns also belongs to the new paradigm. While it does not rule out time varying risk and risk premia, it allows for investors' irrationality and market inefficiency. Under this approach, it is admissible that asset prices and expected returns are not the solution to a general equilibrium model with fully rational, risk averse economic agents and competitive financial markets. See, for a review, Barberis and Thaler (2003)). The literature on limits to arbitrage clarified that, in the presence of noise trader risk, risk-averse market participants with short horizons (finitely lived) might not have the incentive to trade quickly as to exploit all available information even though financial markets are competitive and hence investors are price takers. This is the perspective advocated, among others, by De Long *et al.* (1990a, 1990b) and Shleifer and Vishny (1997). Noise trader risk is the risk that mispricing caused by the net demand of irrational (and hence uninformed) noise traders might worsen in the short run before trades by rational (and hence informed) traders manage to correct it. The relevant notion of rationality is, in this context, the definition embedded in the rational expectation hypothesis of Muth (1961).

The behavioural perspective also allows for non-standard utility functions where investors either do not have unambiguously defined preferences over consumption or they display risk seeking over certain portions of the utility function domain. For example, Prospect Theory and Cumulative Prospect Theory, formulated by Kahneman and Tversky (1979) and Tversky and Kahneman (1992), respectively, imply framing and S-shaped utility functions defined over gains and losses instead of over consumption and wealth as in the standard Expected Utility framework. In particular, these utility functions display risk aversion over gains and risk seeking over losses below a threshold. Behavioural Portfolio Theory, advocated by Shefrin and Statman (2000), predicts instead risk aversion over losses and risk seeking over gains and thus an inverse S-shaped utility function. These non-standard utility functions rationalize evidence that investors, under certain circumstances, display risk seeking behaviour. Active stock traders appear to play negative-sum games and their behavior can sometimes be interpreted as "gambling" (Statman, 2002). In addition, psychologists led by Kahneman and Tversky (1979) find experimental evidence for local risk seeking behavior. More specifically, Post and Levy (2005) argue that a number of celebrated asset pricing anomalies, such as the low average yield on stocks with large capitalization, growth stocks and past winners, could be explained by risk aversion over losses and risk seeking over gains.

Numerous contributions from the literature on non-standard utility theory and behavioural asset pricing (see for a review Shefrin (2010)), thus, admit a non-linear pricing kernel that implies non concavity of the utility function over certain ranges of wealth, and thus an increasing SDF and a violation of RA. Friedman and Savage (1948) and Markowitz (1952) argue that the willingness to purchase both insurance and lottery tickets implies that marginal utility is increasing over a range. See Hartley and Farrell (2002) and Post and Levy (2005) for a recent discussion. *S*-shaped utility functions, such as the function implied by the prospect theory of Kahneman and Tversky (1979), do not satisfy either RA or NIARA. Inverse *S*-shaped utility functions, such as the specification implied by the behavioural portfolio theory of Shefrin and Statman (2002), violate RA but satisfy NIARA at every point in the domain where the function is differentiable.

2.10. Absolute vs. Relative Pricing

Financial theory has extensively addressed the issue of how to model the mean behaviour of asset returns and link it to other variables. In particular, asset pricing models relate mean returns to higher moments. The latter are typically cross-moments formed between the asset return and other variables. These variables can be either economic fundamentals and other non-asset variables, or returns on other assets such as the market portfolio. The first approach is known as absolute pricing, e.g. Lucas' (1978) Consumption CAPM, whereas the second is known as relative pricing, e.g. the Sharpe (1964) and Lintner (1965) CAPM and especially the Asset Pricing Theory (APT), proposed by Ross (1976). The APT requires for its derivation less restrictive assumptions than the CAPM, such as that investors are greedy, that markets are frictionless (or, at least, that diversification is not too costly) and that the returns variance-covariance matrix has a well defined factor structure. The latter requirement guarantees that diversified portfolios can be closely replicated by portfolios that mimic the exposure to single factors. It does not, however, require market completeness (or, equivalently, the representative investor assumption). It provides a no-arbitrage pricing relation between diversified portfolios of assets based on their sensitivity to a set of pervasive risk factors and on the equilibrium risk premium for the exposure to each factor. Chen *et al.* (1986) proxy for the factors using macro-economic variables deemed to drive the variation in stock returns. Within a multi-factor model for asset returns derived from the APT, Koutoulas and Kryzanowski (1996) estimated conditional time-varying risk premia and conditional volatilities associated with each pervasive risk factor. They found that five pervasive risk factors, namely the lag of industrial production, the Canadian Index of 10 Leading Indicators, the US Composite Index of 12 leading Indicators, the exchange rate and the residual market factor, have priced risk premia, including the residual market factor.

For large and diversified portfolios, the implications of the CAPM and of the APT are the same when there is only one pervasive risk factor, the market portfolio. In this case, the expected excess return on the market portfolio would be the only risk premium priced in equilibrium for any diversified portfolio. Neither the CAPM nor the APT admit any risk premium related to idiosyncratic risk (exposure to asset-specific, non-pervasive risk factors), which is expected to be diversified away. It should be noted however that the CAPM, contrary to a popular misinterpretation, is not a special case of the APT. The latter imposes an assumption, namely that the idiosyncratic residuals are uncorrelated, that the CAPM does not require. In the CAPM, idiosyncratic residuals are uncorrelated only on average (with capitalization weights). This is not an assumption, but an implication that follows by construction from the CAPM prediction that these residuals are the error terms of the regression of a set of asset excess returns on their own capitalization-weighted average, namely the market excess returns.

2.11. Transaction Costs

Despite the formal elegance and analytical simplicity of the C-CAPM, the empirical performance of the model has been, at best, mixed. Since the early studies by Hansen and Singleton (1982, 1983), it has been clear that observed asset returns are inconsistent with the dynamics of consumption choices, at least as observed in aggregate data. This evidence was reinforced and confirmed in a large number of subsequent studies (see, Mehra and Prescott (1985)). Some studies, such as Hansen and Jagannathan (1991), suggested that one of the reasons for the poor empirical performance of the model was the low level of variability of aggregate consumption growth. More recently, there have been several attempts at rationalizing this discouraging evidence and several studies have explored the possibility that limited participation in financial markets might explain the disparity between theoretical predictions and empirical evidence. More precisely, since the first-order conditions of representative-agent rational asset pricing models hold as equalities only for households that own diversified portfolios, the models should be tested for this subset of households only and not for the whole population. As a consequence, since in practice relatively few households hold shares directly and even fewer hold a fully diversified portfolio, the use of aggregate consumption data in evaluating asset pricing models can be misleading, This is arguably the case even if we abstract from standard aggregation issues arising from the nonlinearity of the marginal rate of substitution

These points have been stressed by Mankiw and Zeldes (1991), Attanasio *et al.* (2002), Vissing-Jørgensen (2002) and Paiella (2004), among others, who propose limited financial market participation as a unified framework for rationalizing the empirical rejection of the C-CAPM. These authors show that accounting for portfolio heterogeneity and in particular for non-participation in financial markets helps reconcile the predictions of the theory with the empirical evidence. Attanasio *et al.* (2002), for instance, show that focusing on the consumption of stockholders not only yields estimates of preference parameters that are in line with the theory, but also one does not reject the over identifying restrictions implied by the model and, relatedly, the moments of the marginal rate of substitution are within the Hansen-Jagannathan bounds. However, as pointed by Attanasio and Paiella (2011), limited participation is itself a puzzle for the intertemporal consumption model, just like the observed substantial differences in portfolio composition across agents and over the life cycle. Merton (1969) and Samuelson (1969) have illustrated how such behaviour is inconsistent with the maximization of expected lifetime utility, which predicts that rational agents should invest an arbitrarily small amount in all assets with positive expected return, including risky ones, unless there are nonlinearities in the budget constraint.

One possible and obvious way to reinstate consumption risk as the key determinant of equilibrium asset prices is by invoking other market conditions, such as trading costs, which may explain the tenuous link between aggregate consumption risk and asset volatility. This approach has been pursued by He and Modest (1995), Luttmer (1996, 1999) and, more recently, Attanasio and Paiella (2011). Investors trading directly on the New York Stock Exchange (NYSE) face a bid-ask spread of at least one eighth of a dollar per share (one "tick"). Also, taxes can create a gap between the interest rates at which consumers can borrow and lend. More indirectly, about 2.5 percent of the labour force in the U.S. is employed in the financial sector, presumably providing some costly intermediating service between buyers and sellers of assets (Luttmer, 1996). Empirical studies of the capital asset pricing model, the arbitrage pricing theory, and of consumption-based asset pricing models are typically premised on the assumption that the empirical implications of these models are not very much affected by the presence of trading frictions of commonly observed magnitudes. Using aggregate data, Luttmer (1999) provides a lower bound on the transaction costs that would rationalize the model in the face of available data. Paiella (2007), using micro data, provides evidence in support of the participation cost hypothesis by bounding from below the costs of participating in some financial markets and showing that such lower bound is realistic. Her bounds for the stock market participation costs are as small as $130 per year.

2.12. Summary and Conclusions

In this section I have summarized the important developments in asset pricing theory along with the transition from the old to the new paradigm of asset returns, and I have shown how the various asset pricing models can be seen as specializations of the general SDF model. I have then reviewed a number of specifications of this model. Whenever possible, I highlighted the connections between the implications of the various asset pricing models and interesting patterns in equity returns and their moments.

The SDF representation of the asset pricing problem is surprisingly flexible, yet it allows explanations for the observed patterns in asset returns to be generated in a rigorous and testable manner. The only requirement is that the SDF be linear in the factors. Since this approach allows for considerable flexibility in specifying the functional form of the SDF, it can capture non-linearity in the behaviour of marginal utility and time variation in the parameters of the utility function. It therefore serves as a useful framework to specify alternative asset pricing models that allow for a variety of factors to be priced in the time series and cross section of asset returns under alternative assumptions about the multivariate return distribution, investors' preferences and market completeness.

All the asset pricing models considered in the empirical part of the thesis, i.e. Chapter 3, 4, 5 and 6, can be seen as specializations of the SDF model. Chapter 3 proposes a novel approach to testing non-linear SDF specifications that arise in rational representative investor models. Chapter 4 will check whether adding the sentiment factor, and thus allowing for a behavioural influence on asset pricing, allows to significantly improve the explanatory power of the (C)CAPM and 3M-CAPM. In Chapter 5, I will allow for market frictions, and investigate the performance of the conditional higher-moment CAPM under realistic transaction cost assumptions. The attribution model used in Chapter 6 to assess the performance of Chinese mutual fund managers is based on a SDF representation of a multi-factor pricing model specified under the assumption that the shares of the funds are held at the margin by a representative investor that faces an investment opportunity set made up of Chinese stocks.

3. The Coskewness Puzzle

3.1. Introduction

In this chapter, we propose a novel approach to testing non-linear SDF specifications that arise in rational representative investor models. Our approach does not require overly-restrictive assumptions about the shape of investors' preferences, typically imposed by the extant literature, and is based instead on restrictions that rule out "good deals", i.e. arbitrage opportunities as well as unduly large Sharpe ratios. We apply this framework to test the empirical admissibility of 3 and 4-moment versions of the CAPM in explaining differences in average returns across a number of stock strategies and portfolios, including static strategies based on a fine industry-level diversification, momentum strategies and portfolios managed on the basis of available information.

Our study aims to shed light on the puzzling conundrum arising from the problematic shape of the representative investor's utility function implied by unrestricted estimates of the 3M and 4M-CAPM. On the one hand, the findings of Dittmar (2002) and the critique put forth by Post *et al.* (2008) suggest that omitting appropriate restrictions on the shape of the candidate SDF might lead to over-fitting the cross-section of the test asset returns and thus to spurious estimates of the parameters of the 3M and 4M-CAPM. On the other hand, concavity of utility is a sufficient but not necessary condition for the market portfolio to maximize expected utility. This implies that it is formally impossible to make any conclusive inference on the empirical validity of higher moment versions of the CAPM if the latter are rejected when this condition is imposed in estimation, as in the tests performed by Dittmar (2002) and Post *et al.* (2008). Such tests, in fact, while suggestive, amount to tests of the joint hypothesis that the higher moment CAPM holds and utility is concave.

To mitigate this difficulty, we pursue an alternative approach. We only assume that marginal utility is positive and the representative investor's relative risk aversion (RRA) does not exceed an upper bound motivated by survey and experimental evidence (secondary data) on economic agents' risk aversion, as well as by introspection. These restrictions, and especially the bound on RRA, have the effect of limiting the volatility of the estimated SDF and, together, they rule out both arbitrage opportunities and unduly high Sharpe ratios (henceforth, SR). Imposing such restrictions boils down to limiting the admissible curvature of the representative investor's utility function and, ultimately, mitigates the danger of over-fitting the cross-section of the test asset returns. Because this approach does not rely on overly-restrictive assumptions about the shape of utility, it permits a more direct appraisal of the empirical admissibility of higher-moments versions of the CAPM than the approach followed by Dittmar (2002) and Post *et al.* (2008).

Our empirical results paint a complex picture. On the one hand, somewhat surprisingly, we find that coskewness and cokurtosis risk matter only for the pricing of strategies characterized by relatively high SRs, such as dynamic portfolios managed on the basis of available conditioning information or, at least until the late 90s, payoffs spanned by a fine industry-level diversification, as well as momentum strategies. On the other hand, our results conclusively demonstrate that, as suspected by Post and Van Vliet (2006), the CAPM and its higher-moment versions cannot provide an exhaustive description of average stock returns. This is also true of specifications, akin to those estimated by Dittmar (2002), consistent with 3 and 4-moment conditional versions of this model. We conclude that, while preferences towards higher moments are important in explaining average returns, especially on certain strategies, the CAPM and its higher-moment versions provide at best a partial account of the stock price determination mechanism.

In the next Section, we discuss the problem of bounding from above SDF volatility in the context of the higher-moment CAPM. In Section 3.3, we outline the estimation methodology that underlies our tests. In Section 3.4, we describe our dataset. In Section 3.5, we present our empirical results for the case of unconditional models. In Section 3.6, to facilitate comparisons with the extant literature, we report estimates of the unconditional models obtained using different sample periods and test asset payoffs. In Section 3.7, we report our empirical results for the case of conditional models. In Section 3.8, we discuss the implications of our findings and offer our conclusions.

3.2. The Risk Aversion and the SDF Volatility

We specify the candidate SDF as the following third order polynomial defined over the market portfolio return $R_{m,t+1}$:

$$m_{t+1} = a_t + b_{1,t} R_{m,t+1} + b_{2,t} R_{m,t+1}^2 + b_{3,t} R_{m,t+1}^3 . \qquad (3\text{-}1)$$

The specification in (3-1) can be interpreted as a factor model with factors given by the first three integer powers of the return on the market portfolio of risky assets. In view of the covariance-pricing representation of the implications of (3-1) outlined in Appendix

A, $-b_{1,t}$, $-b_{2,t}$, and $-b_{3,t}$ is related with the prices of market covariance, coskewness and cokurtosis risk, respectively. It is well known that the CAPM imposes a number of restrictions on (3-1). In this model, the market portfolio is the inter-temporal optimal allocation that maximizes the representative investor's multi-period expected utility, subject to the economy's inter-temporal budget constraint. Denoting by $U(W_{m,t+1})$ the investor's time-separable utility function of wealth and by $U'(W_{m,t+1})$ its first derivative and according to the analysis in the last chapter, the first order condition (FOC) for the optimum is that, for every payoff x_{t+1}, the following equality holds:

$$U'(W_{m,t})p_t = \gamma E_t\left[U'(W_{m,t+1})x_{t+1}\right]. \tag{3-2}$$

This equation also can be interpreted as

$$p_t = E_t\left[\gamma \frac{U'(W_{m,t+1})}{U'(W_{m,t})} x_{t+1}\right] = E_t(m_{t+1}x_{t+1}). \tag{3-3}$$

The key prediction of the CAPM is that the market portfolio is efficient. This requires that the market portfolio maximizes the representative investor's expected utility, subject to the inter-temporal budget constraint. That is, the market portfolio must represent the solution to a constrained maximization problem. The equalities in (3-2) represent FOCs of this problem. A necessary and sufficient condition for their solution to coincide with the solution to the representative investor's problem is that the expectation of the utility function (the maximand) be quasi-concave in wealth. Deriving the implications of this condition for the parameters of (3-1) is, however, extremely complex. This difficulty explains why, in tests of the higher-moment CAPM, researchers usually impose global concavity of utility and thus require that its second derivative U'' be negative, by either using a power specification or suitably restricting the parameters of a polynomial expansion of the utility function, as in Dittmar (2002). This condition is relatively easy to enforce and it is sufficient to guarantee that the expectation of utility be quasi-concave in wealth, but it is not necessary for this to be the case. Therefore, if the aim is to test higher moment versions of the CAPM, such condition is too restrictive.

We do not wish to impose the overly restrictive assumption that utility is concave, nor we are able to directly impose the requirement that its expectation be quasi-concave. We do wish, however, to restrict the curvature of the utility function in a meaningful fashion. To achieve this aim, we exploit the link between the curvature and volatility of the representative investor's IMRS. Referring to Appendix A, we have:

$$\begin{aligned} E_t(m_{t+1}) \cong \gamma + \gamma \frac{U''(W_t)}{U'(W_t)} W_t E_t(r_{m,t+1}) + \cdots \\ + \frac{1}{2}\gamma \frac{U'''(W_t)}{U'(W_t)} W_t^2 E_t(r_{m,t+1}^2) + \frac{1}{6}\gamma \frac{U''''(W_t)}{U'(W_t)} W_t^3 E_t(r_{m,t+1}^3) \end{aligned} . \qquad (3\text{-}4)$$

As a first step, we obtain an expression for the variance of the IMRS by applying the conditional variance operator to both sides of (3-4). In doing so, we neglect all terms in the Taylor expansion of the IMRS that contain second and higher order powers of the market return. This yields the following locally valid approximation of the IMRS conditional variance:

$$\sigma_t^2(m_{t+1}) \cong \gamma^2 \sigma_t^2 \left[\frac{U''(W_{m,t})}{U'(W_{m,t})} W_{m,t} r_{m,t+1} \right] \equiv \gamma^2 \sigma_t^2 (RRA_{m,t} r_{m,t+1}) . \quad (3\text{-}5)$$

The above result illustrates the relation between RRA and the volatility of the IMRS in a neighbourhood of an initial wealth value $W_{m,t}$. Under an upper bound on the (absolute value of) RRA, i.e. letting $|RRA_{m,t}| \leq RRA_V$, we have the following inequality:

$$\sigma_t^2 (RRA_{m,t} r_{m,t+1}) \leq \sigma_t^2 (RRA_V r_{m,t+1}) = RRA_V^2 \sigma_t^2 (r_{m,t+1}) . \quad (3\text{-}6)$$

Using (3-6) to bound from above the right-hand side of (3-5), and recalling that the coefficient γ which is often called the subjective inter-temporal rate of substitution should satisfies $0 < \gamma \leq 1$, we obtain the following upper bound to the volatility of the IMRS:

$$\sigma_t^2(m_{t+1}) \leq \gamma^2 RRA_V^2 \sigma_t^2 (r_{m,t+1}) \leq RRA_V^2 \sigma_t^2 (r_{m,t+1}) . \quad (3\text{-}7)$$

Since $\sigma_{t+1}^2(r_{m,t+1}) = \sigma_{t+1}^2(1 + r_{m,t+1}) = \sigma_{t+1}^2(R_{m,t+1})$, (3-7) could be re-written as,

$$\sigma_t^2(m_{t+1}) \leq \gamma^2 RRA_V^2 \sigma_t^2 (R_{m,t+1}) \leq RRA_V^2 \sigma_t^2 (R_{m,t+1}) . \quad (3\text{-}8)$$

This inequality, ultimately a restriction on the curvature of utility, implies that the volatility of the IMRS is bounded from above by a quantity that depends on market volatility and a bound on the admissible values of RRA. It should be noted that, while (3-5) only holds in a neighbourhood of the initial wealth value $W_{m,t}$, the upper bound in (3-7) and (3-8) holds in the neighbourhood of every value of wealth as long as RRA, itself a function of wealth, evaluates to a quantity that is within the RRA bound. This makes the bound relevant for restricting the volatility of the IMRS of investors who exhibit non constant RRA and thus care about moments of third and higher orders. Therefore, (3-7) and (3-8) can be seen as a generalization to a possibly non mean-variance world of a result already formulated by Ross (2005) under the assumption of either log-normally distributed returns or exponential utility. Importantly, Ross (2005) also demonstrates[15] that the volatility of the representative investor's IMRS provides a "no good deal" upper bound to the economy maximal SR or, equivalently, to the volatility of the minimum-variance SDF m_{t+1}^{*} that prices all assets traded in the economy, i.e. $\sigma_t^2(m_{t+1}^{*}) \leq \sigma_t^2(m_{t+1})$. Thus, using (3-8), we obtain the following bound on the variance of m_{t+1}^{*}:

$$\sigma_t^2(m_{t+1}^{*}) \leq \sigma_t^2(m_{t+1}) \leq RRA_V^2 \sigma_t^2(R_{m,t+1}) . \tag{3-9}$$

Taking unconditional expectations of both sides of (3-9) gives the following upper bound on the unconditional variance of the candidate SDF:

$$\sigma^2(m_{t+1}^{*}) \leq RRA_V^2 \sigma^2(R_{m,t+1}) . \tag{3-10}$$

To give empirical content to this inequality, one needs to identify a suitable value for the RRA bound. To this end, we follow the advice of Ross (2005) and seek guidance from survey and experimental evidence on investors' RRA provided by the extant literature. Ross (2005) suggests imposing an upper bound of 5 on the RRA of the marginal investor, i.e. $RRA_V = 5$. Among the motivations advanced by Ross (2005) to do so, the one that most easily applies to a world with possibly non-normally distributed returns and non quadratic utility is the simple observation that RRA higher than 5 would imply that the investor is willing to pay more than 10 percent per annum to avoid a 20 percent volatility of his wealth (i.e., about the unconditional volatility of the S&P in the last 75 years), which seems a rather large amount. Similar indications can be drawn from the analysis offered by Meyer and Meyer (2005), who provide a comprehensive

[15] See Proposition 1 in Ross (2005).

re-evaluation of the hitherto scattered empirical evidence on economic agents' risk aversion. They show that RRA estimates reported by the extant literature are less heterogeneous and extreme if one takes into account measurement issues and the outcome variable with respect to which each study defines risk aversion. Using returns on stock investments as the outcome variable, calculations by Meyer and Meyer (2005) show that the RRA coefficient in the classical Friend and Blume (1975) study of household asset allocation choices ranges between 6.4 and 2.0, and decreases in investors' wealth. Using returns on the investors' overall wealth, including real estate and a measure of human capital, the RRA estimate ranges between 3.0 and 2.4. The same calculations show[16] that the RRA implied by the findings of Barsky *et al.* (1997) ranges between 0.8 and 1.6.

These considerations allow us to identify a plausible upper bound on RRA. To be able to compute the SDF volatility upper bound in (3-10), all that remains to be done is to obtain an estimate of the market portfolio volatility, or at least an upper bound on such a quantity. To this end, we note that, even though the market portfolio of risky assets includes both a traded and a non traded portion, the former is likely the most volatile. Thus, we use the S&P500 index to proxy for the traded portion and neglect the non-traded one, which is hard to measure. The sample estimate of the S&P500 index volatility over the period 1965-2005 is about 15.5 percent per annum. Therefore, using (3-10), the SDF volatility bound for this period is about 78 percent per annum if, as

[16] Meyer and Meyer (2005) calculate somewhat higher values based on estimates provided by studies of the equity premium puzzle. Since these estimates are backed out parametrically from estimates of a particular asset pricing model, often based on a narrow definition of the market portfolio, they are of no interest for the purpose of computing the SDF volatility bound. Moreover, their use would imply a circular argument.

suggested by Ross (2005), the RRA upper bound is set to RRA_V = 5.0 and about 99 percent per annum if the RRA upper bound is set to match the RRA of the most risk averse households' cohort in Friend and Blume (1975) study, i.e. RRA_V = 6.4.

3.3. Factor Model Estimation under SDF Sign and Volatility Restrictions

Given a set of n basis test asset payoffs x_{t+1}, a NA restriction and an upper bound on the volatility of the candidate SDF m_{t+1}, the latter can be estimated by solving the following problem:

$$\begin{aligned} &\min_{\{m\}} \ g_t' W_{n\times n} g_t \\ &s.t.\ \ m_{t+1} \geq 0,\ \ \sigma_t^2(m_{t+1}) \leq A_t \end{aligned} . \tag{3-11}$$

with

$$g_t = E_t(m_{t+1}x_{t+1}) - p_t(x_{t+1}) . \tag{3-12}$$

The elements of the vector g_t correspond to the moment conditions implied by the restriction in (3-3) and can be interpreted as pricing errors, while W is a suitable $n\times n$ weighting matrix. The efficient choice for the latter is Hansen (1982) optimal weighting matrix but in this study, following Dittmar (2002), we mainly use Hansen and Jagannathan (1997) second moment matrix because it does not reward spurious variability of the candidate SDF. When using this weighting matrix, the minimized pricing error metric in (12) becomes T times the square of Hansen and Jagannathan (1997) distance, where T denotes the length of the sample period. To make inferences about the empirical admissibility of the candidate SDF, we use the asymptotic distribution of this statistic, under the null of zero pricing errors, provided by Jagannathan and Wang (1996). To identify the mean of the SDF, as done by Dittmar

(2002) and recommended by Dahlquist and Söderlind (1999) and Farnsworth *et al.* (2002a), we include the risk-free asset amongst the basis test asset payoffs x_{t+1}. In estimation, under the usual ergodicity assumptions, we use sample averages $E_T()$ instead of unconditional means and we expand the set of orthogonality conditions by imposing the pricing errors to be unpredictable using information carried by a vector of instruments z_t:

$$\begin{aligned} &\min_{\{m\}} \; g_T' W_{nk\times nk} g_T \\ s.t. \quad &m_{t+1} \geq 0, \;\; \sigma_T^2(m_{t+1}) \leq A \end{aligned} . \tag{3-13}$$

with

$$g_T = E_T\left[(m_{t+1}x_{t+1} - p_t)\otimes z_t\right]. \tag{3-14}$$

Here, z_t is a vector of k conditioning variables. This yields a total of $N = n\times k$ orthogonality conditions for the pricing of as many test assets with payoffs represented by the cross-products $x_{t+1}\otimes z_t$. The latter can be seen as payoffs of dynamic portfolios made up of basis assets having payoffs x_{t+1} and managed using information carried by the conditioning variables z_t. The candidate SDF m_{t+1} is the polynomial described by (3-1). We call the unrestricted version of the latter the (conditional) cubic market factor model, and denote it by the shorthand notation CMFM. Letting $b_{2,t}$ equal zero yields a quadratic market factor model (QMFM). Setting both $b_{2,t}$ and $b_{3,t}$ equal to zero yields a linear market factor model (LMFM). Given the seminal role played by the study of Dittmar (2002) in the literature on tests of the higher moment-CAPM under preference restrictions, we adopt an empirical specification of the candidate SDF that is as close as possible to the one used by this author. We thus model the market portfolio as a linear combination of a value-weighted portfolio of traded assets, with rate of return denoted by $R_{vw,t+1}$, and human capital, with rate of return $R_{l,t+1}$. That is, we let:

$$R_{m,t+1} = \theta_t R_{vw,t+1} + (1-\theta_t) R_{l,t+1}. \qquad (3\text{-}15)$$

Here, $0 \le \theta_t \le 1$ represents the fraction of market wealth allocated to traded financial assets. Thus, using (3-15) in (3-1), we have the following empirical specification of the candidate SDF:

$$\begin{aligned} m_{t+1} = a_{0,t} &+ b_{vw,1,t} R_{vw,t+1} + b_{vw,2,t} R^2_{vw,t+1} + b_{vw,3,t} R^3_{vw,t+1} + \cdots \\ &+ b_{l,1,t} R_{l,t+1} + b_{l,2,t} R^2_{l,t+1} + b_{l,3,t} R^3_{l,t+1} \end{aligned}. \qquad (3\text{-}16)$$

Here, the cross product between $R^i_{vw,t+1}$ and $R^i_{l,t+1}$ is ignored. This specification is the same as the one used by Dittmar (2002). We note, however, that the weights θ_t and $1-\theta_t$ are time-varying. For this reason, imposing that the candidate SDF in (3-16) be decreasing in $R_{vw,t+1}$ and $R_{l,t+1}$, as done by Dittmar (2002), is not the same as imposing it to be decreasing in $R_{m,t+1}$. In fact, the former requirement is much more restrictive than the latter and it is overly restrictive if the aim is to impose concavity of the utility function, which only requires that the candidate SDF be decreasing in the overall market portfolio wealth rather than in each one of its components[17]. These considerations provide further motivation for imposing positivity and volatility restrictions on the estimated SDF in place of global concavity in tests of the higher moment CAPM. Finally, following again Dittmar (2002), we model conditional time-variation in the shape of the utility function by specifying the parameters of the

---

[17] Admittedly, imposing that the estimated SDF be decreasing in $R_{m,t+1}$, rather than in $R_{vw,t+1}$ and $R_{l,t+1}$ separately, can be quite challenging because of obvious programming difficulties and econometric software limitations as well as because the weights with which traded and non traded wealth enter the market portfolio are difficult to observe and measured with error.

candidate SDF as linear functions of the set of conditioning information variables z_t, i.e. we let,

$$a_t = \eta' z_t \qquad b_{vw,j,t} = \varphi_{vw,j}' z_t \qquad b_{l,j,t} = \varphi_{l,j}' z_t . \tag{3-17}$$

Here, recalling that z_t is a vector of k conditioning variables, $\varphi_{vw,j}$ and $\varphi_{l,j}$, with $j = [1,\ 2,\ 3]$, are conformable coefficients vectors.

3.4. Empirical Specification Details and Data

We use monthly data, from 1963 to 2005, on the rate of returns on an overall value-weighted portfolio of stocks included in the Centre for Research on Security Prices (CRSP) database and on portfolios of such stocks sorted into 17 and 30 industries as well as 10 momentum deciles[18]. As customary, the value-weighted portfolio represents the proxy for the return $R_{vw,t+1}$ on the traded portion of the market portfolio of risky assets. The returns on the 17 and 30 industry-sorted stock portfolios represent payoff spaces similar to those spanned by the returns on the 20 and 27 industry-sorted portfolios used by Dittmar (2002) and Harvey and Siddique (2000), respectively. The 10 momentum portfolios are also similar to those used by the latter authors. We use NIPA data to construct, as in Dittmar (2002), a proxy for the return on human capital $R_{l,t+1}$ and we employ the yield of the 1-month US Government Treasury Bill as a proxy for the risk-free rate. We use the rates of return on the test asset payoffs to compute gross returns (i.e., one plus the rate of return or the yield). The conditioning information variables z_t are the unit series and variables drawn from a set that includes the

[18] We thank K. French for making this data publicly available for download.

excess-return on the market portfolio proxy $r_{m,t}$, its dividend-yield dy_t, the yield spread ys_t of the 3-month T-Bill in excess of the 1-month T-Bill, the return tb_t on the T-Bill closest to one-month maturity, i.e, $z_t = [r_{m,t},\ dy_t,\ ys_t,\ tb_t]'$. These variables are known to predict market-wide returns, as noted by Dittmar (2002) and references therein.

Figure 3-1 plots the Hansen and Jagannathan (1991) mean-variance frontier of the SDFs that price by construction the payoff space for the case when the payoffs x_{t+1} are the gross returns either on one of the two sets of industry portfolios or on the 10 momentum portfolios, augmented by the gross return on the risk-free asset proxy. The figure also plots the SDF mean-variance frontier for the case when the payoffs are the cross-products $x_{t+1} \otimes z_t$ between the gross returns on the 17 industry-sorted portfolios, augmented by the gross return on the risk-free asset proxy, and the conditioning variables z_t, represented by the unit series, $r_{m,t}$, dy_t, ys_t and tb_t. Panel A of the Figure 3-1 plots the frontiers for the sub-sample period 1965-1993, which ends at the same time as the sample period considered by Harvey and Siddique (2000), while Panel B plots the frontiers for our benchmark sample period 1965-2005. By construction, the 30 industry-sorted portfolios span all strategies spanned by the 17 industry-sorted ones, but not the other way round. As a result, the frontier for the former set of portfolios lies above the frontier for the latter. Similarly, the managed portfolios span all strategies spanned by the static portfolios, but not the other way round, which is reflected in a tighter Hansen and Jagannathan (1991) bound for the former and thus in a higher frontier. One interesting feature of all these frontiers is that they are noticeably lower in the longer sample period.

3.5. Unconditional Estimates

We first estimated unconditional versions of (3-16), setting z_t equal to the unit series and using gross returns as test asset payoffs. The price of these payoffs, by construction, equals one. In (3-14), we therefore set $p_t = 1$. In Panel A of Table 3-1, we summarize the unconstrained estimates of the LMFM, QMFM and CMFM. The panel reports the point estimates of the coefficients of the candidate SDFs, along with the *p*-values of the corresponding *t*-statistics computed using heteroskedasticity and autocorrelation adjusted (HAC) Newey and West (1987) standard errors, the Hansen and Richard (1987) distance and the associated *p*-value, based on the asymptotic distribution under the null of zero pricing errors provided by Jagannathan and Wang (1996), as well as the annualized volatility of the estimated SDF and the implied RRA bound. The sample period is 1965-2005.

The estimated b_2 coefficients are statistically insignificant for all models. While the non-linear models are not rejected, neither are the linear ones. As shown in Panel B of the Table 3-1, imposing the positivity constraint on the estimated SDF does not significantly diminish the explanatory power of the models under consideration and the restriction is in many cases not binding. In Table 3-2, we report estimates of the candidate SDFs obtained under additional restrictions. To enforce RA and NIARA in estimation, we imposed $(-1)^j b_{i,j,t} > 0$, with $i \in [vw,\ l]$ and $j = [1,\ 2,\ 3]$. We also restricted the estimated SDF to satisfy a strict positivity constraint and imposed an upper bound on its volatility, as specified by (3-13). The volatility bound was set to 78 percent per annum, corresponding to a RRA upper bound of 5. Again, the Hansen and Jagannathan (1997) test fails to reject any of the models under consideration. Overall, these results suggest that third and fourth systematic moments play at most a limited role in the pricing of the test asset payoffs used in estimation, i.e. the gross returns on the 17 industry-sorted portfolios and the risk-free asset.

3.6. The Role of Coskewness Risk over Time and Strategies

On balance, our findings thus far suggest that the price of coskewness (and cokurtosis) risk does not greatly matter in the pricing of unconditional (i.e., static) portfolios of stocks. This is consistent with the results reported by Friend and Westerfield (1980) and Fang and Lai (1997) but is in sharp contrast with the findings of other authors, most notably Harvey and Siddique (2000). To double-check on these seemingly conflicting results, we estimate the QMFM under alternative choices for the weighting matrix of the moment conditions, including the identity matrix and the optimal weighting matrix, and using alternative sets of test asset payoffs and sample periods.

To save space, we report in Table 3-3 only the estimation results obtained using a sample counterpart of Hansen (1982) optimal weighting matrix, based on a continuously updating (CUE) estimate of the spectral density matrix, and a set of test asset payoffs represented by excess-returns (in place of gross returns) on the 17 and 30 industry-sorted portfolios over the sample period 1963-2005 and portions thereof. For comparison, we also report estimates obtained using excess-returns on the 10 momentum portfolios as the test asset payoffs. The weighting matrix choice yields

CUE-GMM, which has nice statistical properties that allow for more efficient estimates. The use of excess-returns as test asset payoffs implies exact pricing of the risk free asset. By construction, excess-returns have a zero price. Therefore, in (3-14), we set $p_t = 0$.

Table 3-3 reports two sets of results. The first comprises point estimates of the b_1 and b_2 parameters of the QMFM and the *p*-values of the corresponding *t*-statistics constructed using standard errors based on our estimate of the spectral density matrix. The second set of results include the stock market covariance and coskewness annualized risk-premia $\lambda_{vw,1}$ and $\lambda_{vw,2}$, respectively, implied by the estimated $b_{vw,1}$ and $b_{vw,2}$ according to the beta-pricing representation, described by (A-3), (A-4) and (A-5) in section 8.1 Appendix A, of the QMFM. The Table also reports *p*-values of the bootstrapped risk-premia distribution. The latter is generated using the so called "estimation-based" bootstrap introduced by Freedman and Peters (1984) and Peters and Freedman (1984)[19]. A comparison of Panel A and B of Table 3-3 shows that the role of coskewness risk in explaining pricing phenomena is emphasized when the 30 industry-sorted portfolios are used in place of the 17 industry-sorted ones. As shown by Figure 1, the frontier for the 30 industry-sorted portfolios lays noticeably above the frontier for the 17-industry sorted portfolios. This suggests that a positive price of

[19] We first estimated time-series regressions of the test assets excess-returns on the factors (the stock market return proxy $R_{vw,t+1}$ and its square) and stored the residuals. We then re-sampled 5,000 times, with replacement, blocks of 5 consecutive realizations from the stored residuals time-series, i.e. we employed 'block re-sampling' to capture any residual serial correlation not explained by the estimated time-series regression. Using the time-series of the re-sampled residuals and the point estimates of the time-series regression parameters, we generated 5,000 separate bootstrapped return series for each test asset, for which we then re-estimated the QMFM model and recorded the $\lambda_{vw,1}$ and $\lambda_{vw,2}$ coefficients, calculated using (A5). This generates a bootstrapped distribution of the $\lambda_{vw,1}$ and $\lambda_{vw,2}$ coefficients.

coskewness risk helps generate the additional SR attainable by a finer industry-level diversification.

As shown in Panel C of Table 3-3, the quadratic term coefficient is also highly significant when the momentum portfolios are used as the test asset payoffs. This is consistent with the findings of Harvey and Siddique (2000) and also in agreement with the conclusions reached by Rachev *et al.* (2007), who argue that momentum strategies entail considerable tail risk. The coskewness risk premium, however, is not statistically significant. In the terminology used by Cochrane (2001), this implies that the quadratic factor helps "price" the momentum portfolios but it is not "priced" in their cross-section. Interestingly, this suggests that a non-zero price of coskewness risk might help explain the momentum effect uncovered by Jegadeesh and Titman (1993) rather than the other way round. Noticeably, however, in spite of the extremely large amount of volatility of the estimated SDF, the quadratic model is rejected in all sample periods at the 5 percent level and at the 10 percent level in the period 1963-1993. In un-tabulated results, we find a similarly disappointing empirical performance for 4-moment versions of the model and for specifications that include human capital.

Another, and perhaps most noteworthy, fact highlighted by Table 3-3 is that the pricing ability of coskewness risk has diminished in recent times, as implied by the considerably reduced significance of both the price of coskewness risk and the coskewness risk premium in the 1963-2005 sample period relatively to the benchmark sample period used in Harvey and Siddique (2000) study, i.e. 1963-1993[20]. The less important role played by coskewness risk after the early 90s, while somewhat puzzling, might be explained by a number of circumstances. For example, the discovery by Harvey and Siddique (2000) of the strength of the coskewness 'effect' might have led investors to engage in strategies designed to exploit it. The rise of hedge funds springs to mind as a possible key ingredient for such development. The analysis conducted by Zakamouline and Koekebakker (2009) suggests that hedge funds managers pursue strategies that exhibit negative skewness (and excess-kurtosis), for example by shorting options or engaging in dynamic portfolio insurance, in a quest for higher Sharpe Ratios. This behaviour can, in principle, drive down both the price of coskewness risk and the maximal SR attainable from the investment opportunity set, producing the drop in the SDF frontier that we observe in Figure 3-1.

On balance, these results clarify that our earlier findings (those reported in Table 3-1 and Table 3-2) do not contradict those of the extant literature but rather all such findings are sample-specific, in the sense that the estimated price of coskewness risk and the associated risk premium depend on the test asset payoffs used in estimation as well as on the sample period. Overall, coskewness appears to play a more important role in the pricing of strategies with relatively high SRs, such as those spanned by the finer industry level diversification allowed by the 30 industry portfolios, and in explaining

[20] Our estimated market coskewness risk premium for this period, when using the 30 industry-sorted portfolios as test assets, is –2.90 per cent per annum. It is thus very close to the –3.60 per cent per annum estimated by Harvey and Siddique (2000) and it is larger, in absolute value, than the market covariance risk-premium, which is just 2.70 per cent per annum. By contrast, in the extended sample period 1993-2005, the coskewness premium is –0.48 per cent per annum, and thus the coskewness discount is a mere 0.48 percent, whereas the covariance premium is a much larger 6.68 per cent per annum.

returns on hard-to-price strategies, such as those spanned by the 10 momentum portfolios. The investigation of the determinants of the changing role of coskewness in asset pricing, as well as the difficult task of exploring the relations between momentum and coskewness, are outside the scope of this chapter. In order to focus on the main task at hand, i.e. the testing of the CAPM and its higher moment versions, we leave these intriguing tasks to future research.

3.7. Conditional Estimates

Next, we turn to conditional versions of (3-16). Initially, in (3-14) and (3-17), we use the full set of conditioning variables. In Table 3-4, we report the unconstrained estimation results, obtained using gross returns on the 17 industry portfolios and the risk-free asset as the basis test asset payoffs. The reported SDF coefficients are the values of a_t, $b_{vw,j,t}$ and $b_{l,j,t}$, with $j = [1,\ 2,\ 3]$, evaluated at the sample average values of the conditioning variables z_t. The associated *p*-values are significance levels of Wald tests of the restriction that the coefficient under consideration is zero. Consistently with Dittmar (2002) findings, the non-linear SDF specifications achieve considerable empirical success. As shown in the lower panel of Figure 3-2, the estimated CMFM expected returns line up reasonably closely with sample average returns. Both the QMFM and the CMFM, but not the LMFM, pass the Hansen and Jagannathan (1997) distance test when the market portfolio proxy includes human capital. The results imply that coskewness and cokurtosis risk are significantly priced in the cross-section of stock portfolios managed on the basis of available information. This contrasts with the insignificance of the coskewness and cokurtosis risk coefficients in unconditional estimates, suggesting that the quadratic and cubic terms generate the additional SDF volatility that, as shown in Figure 3-1, is required to price the managed portfolios used in conditional estimates.

The unrestricted estimates do not satisfy, however, a number of requirements of Kimball (1993) standard risk aversion. For example, as shown in the top panel of Figure 3-2, the estimated SDF is not always positive and, as shown in the lower right-hand corner of the same figure, it is not decreasing in the return on the stock market proxy. It may be argued that the estimated SDF takes negative values only very seldom, essentially only around times of exceptional market turmoil, namely the bear market that followed the 1974 oil shock, the 1987 stock market crash, the so-called 1997 Asian crisis and the crisis that followed the default of the Russian Federation on its domestic debt in 1998, as well as the bust of the "dot com" bubble in March 2000. Admittedly, these may be viewed as exceptional one-off circumstances, akin to outliers. The estimated SDF, however, is also very volatile, especially in the case of non-linear models. For example, the annualized volatility of the estimated SDF is 273 percent for the QMFM and 299 percent for the CMFM. Given a 15.5 percent annualized volatility of the stock market portfolio, this would imply that RRA takes values as large as 17.56 and 19.23, respectively. These values are puzzling when compared with the upper bound on RRA suggested by Ross (2005), i.e. $RRA_V = 5$, and survey evidence, as summarized by Meyer and Meyer (2005), which shows that, even for the most risk-averse households cohorts, RRA does not exceed 6.4 when the wealth portfolio includes only financial assets and 3.0 when it includes also non-financial assets.

To disentangle the pricing implications of violations of the NA positivity requirement from those of excessive SDF volatility, we estimate the candidate SDFs imposing each one of the two restrictions in turn. In

Table **3-5**, we report estimates obtained under a positivity restriction on the SDF. All candidate SDFs fail the Hansen and Jagannathan (1997) distance test when estimated under this restriction. This is also the case of non-linear specifications with human capital included in the market portfolio. Next, we estimate the candidate SDFs without the positivity restriction but under various upper bounds on their volatility. In Figure 3-3, we summarize the results of this exercise for the CMFM with human capital. In the Figure, we plot the model Hansen and Jagannathan (1997) distance against the annualized volatility upper bound used in estimation and the corresponding RRA upper bound. All models estimated under a volatility upper bound set to less than 188 percent per annum are rejected at the 5 percent level. That is, it takes a RRA upper bound as loose as about 12.04 for the 4M-CAPM with human capital to be empirically admissible. These results are consistent with those reported by Dittmar (2002), who also reject the 4M-CAPM as well as more restricted versions of this model. Our findings, however, clarify that the model is rejected not because of attempts to enforce overly restrictive requirements on the shape of the estimated utility function but rather, and more conclusively, because of the binding implications of sensible 'no good-deal' restrictions. In unreported estimates, we find this to be the case also in the earlier sample period used by Dittmar (2002), i.e. in the period 1963-1997.

These results imply that, while the prices of coskewness and cokurtosis risk are statistically significant, 3 and 4-moment versions of the CAPM provide an inadequate account of asset pricing patterns. In interpreting these findings, however, a caveat is in order. The test asset payoffs second moment matrix used to compute the weighting matrix of the moment conditions is close to singular. This implies that the pricing error metric being minimized in estimation is the pricing error of a portfolio that contains extreme long and short positions[21]. This is a widespread problem in empirical asset pricing and occurs also with the second moment matrix estimated using data for the period 1963-1997, i.e. the sample period used by Dittmar (2002). To mitigate this problem, we re-estimate using one conditioning variable at a time, i.e. including in z_t only one of the conditioning variables $r_{m,t}$, dy_t, ys_t, and tb_t, as well as the unit-series. This way, only 36 test asset payoffs at a time are used in estimation[22]. As shown in Table 3-6, the empirical performance of the CMFM with human capital improves somewhat and it passes the Hansen and Jagannathan (1997) distance test at the five percent significant level when the conditioning information variables set includes the lag of either the dividend yield, a term spread variable or the 3-month T-Bill return. The model, however, fails the test at the 10 percent in all cases except when the conditioning variables are the unit series and the lagged term spread and no restriction is placed on SDF volatility. When the conditioning variable is represented by the unit series and the lagged market excess return, the CMFM is rejected at the 0.1 percent level. These findings are important in that they confirm that our rejection of the 4M-CAPM is not a

---

[21] We verified this by computing the weights assigned by this matrix to the asset payoffs in the definition of the pricing error metric. These weights can be calculated by taking the sum by column of the elements of the upper triangular matrix obtained from a Cholesky decomposition of the weighting matrix, i.e. of the inverse of the second moment matrix.

[22] Reassuringly, the resulting second moment matrix is much less close to singular.

spurious by-product of a near-singular inverse weighting matrix. Interestingly, the fact that the model fares worst when the conditioning variable is represented by the lagged market excess return might be related to the well-known difficulty, also noted earlier, of explaining abnormal returns of momentum strategies.

3.8. Final Remarks and Conclusions

In this chapter, we acknowledge the importance of Dittmar (2002) findings and of the critique put forth by Post *et al.* (2008), in that they highlight the danger of spurious estimates of higher-moment versions of the CAPM. We emphasize, however, that a decreasing SDF, albeit sufficient, is not a necessary condition for the CAPM. Moreover, even if one is willing to restrict focus on the joint hypothesis of the higher moment CAPM and concave utility, the specification of the market portfolio proxy may seriously affect inferences on the empirical admissibility of such hypothesis, much in the same way in which, as pointed out by Roll (1977), the circumstance that the market portfolio is *de facto* un-observable affects all tests of the CAPM. Recognising this difficulty amounts to acknowledging the ramifications of Roll (1977) critique for tests of non-linear versions of the CAPM. In estimating the model, we thus impose alternative restrictions on the shape of the candidate SDF, namely a positivity requirement and a volatility upper bound. These restrictions boil down to ruling out arbitrage opportunities and SRs that, at least to the marginal investor, would resemble obvious near arbitrage opportunities[23]. This way, we limit the danger of over-fitting the

[23] Beside, since coskewness is an asset characteristic that explains a considerable portion of the cross-section of asset returns, such an approach is also consistent with a multi-factor, no-arbitrage perspective, along the lines of Ross'

cross-section of asset returns without the need to resort to the overly-restrictive assumption of concave utility.

Our findings confirm that the quadratic and cubic market factors help explain observed stock returns. They play an important role in the pricing of certain payoffs, including strategies characterized by relatively high SRs, such as those spanned by a fine industry-level diversification, most notably until the late 90s, or by dynamic portfolios managed on the basis of available conditioning information, as well as momentum portfolios. They do so, however, by generating high levels of SDF volatility. To rationalize this evidence within a higher moment CAPM framework, we would need to postulate implausibly high levels of investors' risk aversion. We conclude, therefore, that the 3M and 4M-CAPM provide at best a partial explanation of the differences in average returns on stocks and stock strategies. This gives rise to a coskewness (and cokurtosis) puzzle. The solution of the latter requires an explanation, different from the 3M and 4M-CAPM, for why the quadratic and cubic market factors are priced in the cross-section of stock returns.

One obvious possibility is that these factors proxy for other priced but omitted factors. Vanden (2006) suggests that powers of the market returns proxy for omitted option-related factors. This possibility, while intriguing, requires further scrutiny because Vanden (2006) sample period is relatively short and, more importantly, it

---

(1976) APT. From this perspective, it is similar to the approach followed by Cochrane and Saà-Requejo (2000) and Cochrane (2005) to extend, in incomplete markets, the pricing implications of the factor prices and of (1) to a non-redundant security.

remains to be established whether the SDF estimated by this author satisfies a sensible volatility upper bound. Another fruitful extension of our study would be a check of whether the more flexible specification used by Smith (2007)[24] is empirically admissible under appropriate positivity and volatility restrictions on the estimated SDF. An alternative approach would be to allow for a wedge between volatility of the candidate SDF and IMRS volatility by letting sentiment and variables related to investors' errors generate, as suggested by Shefrin (2010), the extra SDF volatility required to price stocks. There is also the possibility that some of the managed portfolios used in Dittmar (2002) study and in ours correspond to unfeasible strategies, i.e. strategies with unfeasibly high SRs. Luttmer (1996), for example, shows how even modest proportional transaction costs, short sales restrictions and margin requirements considerably lower the mean-variance SDF frontier.[25] We leave the investigation of these possible explanations of the coskewness puzzle for the remaining parts of the thesis.

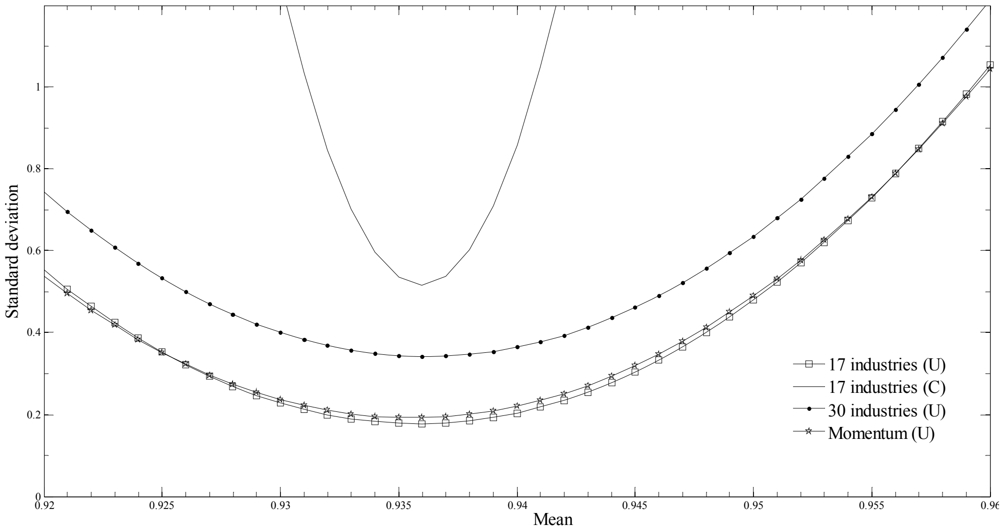
---

[24] As shown by Smith (2007), his specification nests Dittmar's (2002) specification and hence the specification used in this paper.
[25] While we took this problem partially into account by re-estimating using one conditioning variable at a time, we feel that additional and more direct checks on the implications of transaction costs and market microstructure frictions might be worthwhile.

Figure 3-1
SDF Mean-Variance frontier

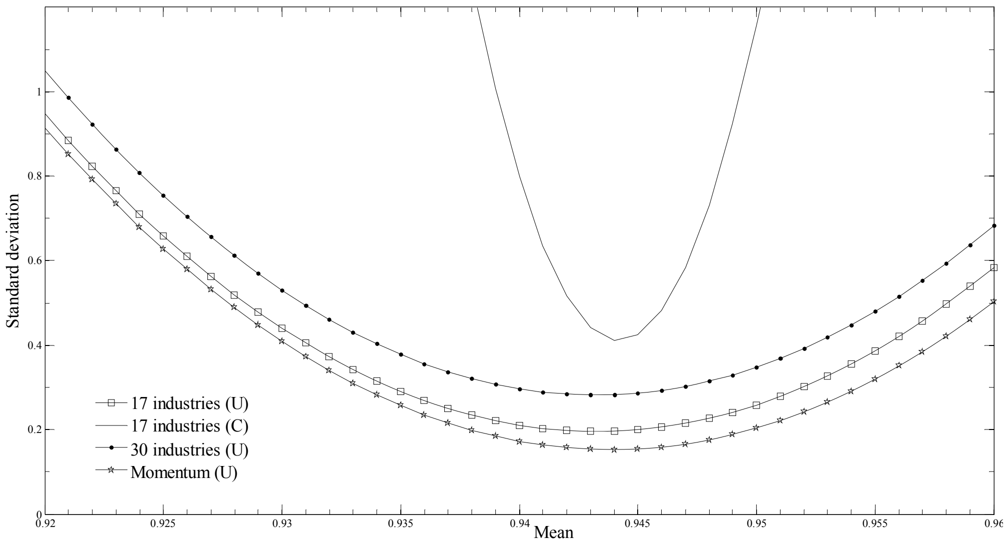
Panel A

Sample period: 1965-1993



Panel B

Sample period: 1965-2005



Notes. This Figure plots the Hansen and Jagannathan (1991) SDF mean-variance frontiers for sets of portfolios of basis assets managed using information conveyed by a set of conditioning variables. The basis assets are either industry-sorted (17 or 30 industries) or momentum portfolios augmented by the risk-free asset proxy. The data is monthly, but results are converted to an annual basis, for the indicated sample periods. The symbol U (C) in the legend means that the set of conditioning variables used in forming the managed portfolios includes only the unit series (the full set of conditioning variables).

Table 3-1
GMM estimates of unconditional models
(1965-2005)

Model	a	$b_{vw,1}$	$b_{vw,2}$	$b_{vw,3}$	$b_{l,1}$	$b_{l,2}$	$b_{l,3}$	HJ_T (*p.-value*)	$\sigma(m)$	RRA_v
					Panel A (Unconstrained estimates)					
					Market portfolio without human capital					
LMFM	1.02	-2.55						0.18	0.40	2.55
	(0.000)	(0.006)						(0.522)		
QMFM	0.99	-2.57	1.25					0.17	0.40	2.57
	(0.000)	(0.018)	(0.978)					(0.511)		
CMFM	1.04	4.48	-34.48	-743.79				0.16	1.28	8.29
	(0.000)	(0.138)	(0.241)	(0.039)				(0.677)		
					Market portfolio with human capital					
LMFM	0.985	-2.32			47.82			0.17	0.76	4.86
	(0.000)	(0.040)			(0.321)			(0.567)		
QMFM	1.01	-2.22	22.66		-3536.1	1783.12		0.17	0.88	5.67
	(0.000)	(0.096)	(0.687)		(0.736)	(0.732)		(0.525)		
CMFM	0.87	2.36	-14.60	-486.23	32.16	360.07	-516,15*	0.12	1.86	11.99
	(0.001)	(0.303)	(0.395)	(0.154)	(0.262)	(0.373)	(0.033)	(0.798)		
					Panel B (Estimates under NA)					
					Market portfolio without human capital					
LMFM	1.02	-2.55						0.18	0.40	2.55
	(0.000)	(0.006)						(0.522)		
QMFM	0.99	-2.57	1.25					0.17	0.40	2.57
	(0.000)	(0.018)	(0.978)					(0.511)		
CMFM	1.03	3.68	-34.48	-743.80				0.17	1.12	7.22
	(0.000)	(0.186)	(0.241)	(0.039)				(0.454)		
					Market portfolio with human capital					
LMFM	0.99	-3.21			-53.53			0.19	0.86	5.53
	(0.000)	(0.000)			(0.000)			(0.292)		
QMFM	1.26	-2.23	3.64		-46.32	-334.81		0.17	1.18	7.62
	(0.000)	(0.017)	(0.470)		(0.110)	(0.361)		(0.420)		
CMFM	0.86	1.77	-14.59	-486.23	32.17	360.07	-516,15*	0.15	1.15	7.43
	(0.001)	(0.349)	(0.395)	(0.154)	(0.262)	(0.373)	(0.033)	(0.428)		

Notes. This Table reports unrestricted GMM estimates of the unconditional LMFM, QMFM and CMFM, obtained using Hansen and Jagannathan's (1997) second moment weighting matrix. Panel A reports the unconstrained estimates of SDF, while Panel B reports the estimates under no-arbitrage condition. For each included factor, we report the corresponding SDF coefficient point estimate and its p-value in brackets. We also report the Hansen and Jagannathan (1997) distance and its p-value in brackets. The second last column reports the annualized volatility of the stochastic discount factor in percentage. All the variables are defined as in the text. The test asset payoffs are gross returns on the 17 industry-sorted portfolios augmented by the gross return on the risk-free asset. The sample period is 1965-2005.

Table 3-2
GMM estimates of unconditional models under NA, Local RA, NIARA and SDF volatility restrictions
(1965-2005)

Model	*a*	*bvw,1*	*bvw,2*	*bvw,3*	*bl,1*	*bl,2*	*bl,3*	*HJT (p.-value)*	*σ(m)*	*RRAv*
					σ(m) ≤ 0.78					
					Market portfolio without human capital					
LMFM	1.02	-2.55						0.18	0.40	2.55
	(0.000)	(0.006)						(0.519)		
QMFM	0.99	-2.57	1.25					0.17	0.40	2.57
	(0.000)	(0.018)	(0.978)					(0.511)		
CMFM	0.93	-2.61	45.64	-0.08				018	0.78	5.00
	(0.000)	(0.012)	(0.087)	(0.974)				(0.376)		
					Market portfolio with human capital					
LMFM	1.01	-2.44			-0.01			0.17	0.37	2.39
	(0.000)	(0.010)			(0.975)			(0.432)		
QMFM	1.07	-0.71	1.39		-53.75	4795.55		0.18	0.78	5.00
	(0.000)	(0.093)	(0.925)		(0.000)	(0.000)		(0.160)		
CMFM	0.91	-0.37	1.91	-0.01	-20.47	4739.12	-73215.94	0.18	0.78	5.00
	(0.000)	(0.218)	(0.878)	(0.925)	(0.007)	(0.000)	(0.893)	(0.140)		

Notes. This Table reports GMM estimates of the unconditional LMFM, QMFM and CMFM for the period 1965-2005, obtained using Hansens and Jagannathan's (1997) second moment weighting matrix, under positivity and volatility restrictions on the estimated SDF. For each included factor, we report the corresponding SDF coefficient point estimate and its p-value in brackets. We also report the Hansen and Jagannathan (1997) distance and its p-value in brackets. The second last column reports the annualized volatility of the stochastic discount factor in percentage. All the variables are defined as in the text.

Table 3-3
QMFM without human capital CUE-GMM estimates

Sample Period	$m > 0$ (yes/no)	R^2	$b_{vw,1}$ (*p.-value*) $\lambda_{vw,1}$ (*p.-value*)	$b_{vw,2}$ (*p.-value*) $\lambda_{vw,2}$ (*p.-value*)	TJ_T (*p.-value*)	$\sigma(m)$	RRA_v
			Panel A (**17** ***industry portfolios***)				
1963-2005	yes	0.01	-2.63	-8.28	0.18	0.43	2.77
			(0.010)	(0.840)	(0.710)		
			6.40	0.20			
			(0.000)	(0.500)			
1963-1993	yes	0.17	-2.59	87.48	0.14	1.35	8.85
			(0.258)	(0.083)	(0.934)		
			4.32	-1.96			
			(0.000)	(0.150)			
			Panel B (**30** ***industry portfolios***)				
1963-2005	yes	0.05	-2.82	22.42	0.21	0.54	3.50
			(0.027)	(0.572)	(0.761)		
			6.68	-0.48			
			(0.000)	(0.270)			
1963-2000	yes	1.08	-3.94	78.49	0.22	1.26	8.29
			(0.038)	(0.092)	(0.922)		
			8.26	1.61			
			(0.000)	(0.250)			
1963-1997	yes	0.15	-4.18	116.95	0.21	1.73	11.68
			(0.060)	(0.010)	(0.981)		
			6.95	-2.33			
			(0.000)	(0.070)			
1963-1993	no	0.15	-2.25	127.96	0.21	1.94	12.70
			(0.336)	(0.004)	(0.956)		
			2.71	-2.90			
			(0.000)	(0.090)			
1963-1993	yes	0.12	-2.34	105.24	0.22	1.61	10.51
			(0.278)	(0.000)	(0.930)		
			3.39	-2.37			
			(0.000)	(0.050)			
			Panel C (**10** ***momentum portfolios***)				
1963-2005	no	21.52	-2.67	-69.95	0.32	1.10	7.23
			(0.006)	(0.145)	(0.000)		
			6.66	1.48			
			(0.000)	(0.210)			

1963-2005	yes	0.03	-6.79	106.10	0.27	1.82	11.98
			(0.000)	(0.000)	(0.002)		
			15.14	-2.17			
			(0.000)	(0.150)			
1963-1993	no	0.22	-17.99	383.30	0.20	6.21	40.66
			(0.003)	(0.004)	(0.057)		
			34.45	-8.47			
			(0.050)	(0.300)			
1963-1993	yes	0.10	-8.02	108.55	0.35	1.96	12.87
			(0.000)	(0.000)	(0.000)		
			16.60	-2.34			
			(0.050)	(0.220)			

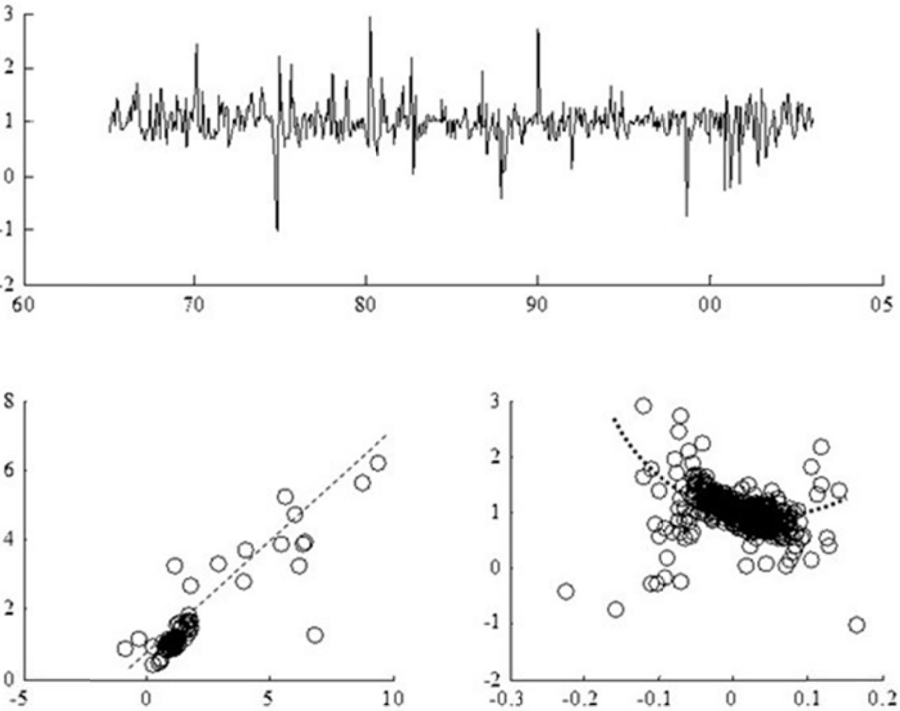
Notes. This Table reports CUE-GMM estimates of the unconditional QMFM for various sample periods, specified in the first column. The second column indicates whether the estimated SDF satisfies the positivity requirement. The third column reports the coefficient of determination R^2, i.e. the squared correlation coefficient between sample average excess returns and their model estimate. Next, we report the SDF coefficient point estimates with, in brackets, the p-values of their t-statistics computed using HAC standard errors. We also report the corresponding implied annualized percentage risk-premia and, in brackets, the p-values of their bootstrapped distribution. In the last 3 columns, we report Hansen's (1982) TJ_T statistic (and the corresponding p-value, in brackets), the estimated SDF volatility and the associated RRA bound. The test asset payoffs are monthly excess returns on the 17 and 30 industry-sorted and momentum portfolios.

Table 3-4
Unconstrained GMM estimates of conditional models
(1965-2005)

Model	a	$b_{vw,1}$	$b_{vw,2}$	$b_{vw,3}$	$b_{l,1}$	$b_{l,2}$	$b_{l,3}$	HJ_T (p.-value)	$\sigma(m)$	RRA_v
					Market portfolio without human capital					
LMFM	1.05	-3.67						0.53	0.73	4.69
	(0.000)	(0.001)						(0.000)		
QMFM	1.06	-5.80	24.25					0.50	1.24	7.96
	(0.000)	(0.000)	(0.002)					(0.011)		
CMFM	1.07	-2.03	5.73	-613.23				0.49	1.39	8.94
	(0.000)	(0.024)	(0.128)	(0.605)				(0.012)		
					Market portfolio with human capital					
LMFM	1.04	-3.09			-1.96			0.50	1.60	10.27
	(0.000)	(0.005)			(0.009)			(0.012)		
QMFM	1.01	-4.14	38.46		-20.70	1,098.81		0.44	2.73	17.56
	(0.000)	(0.001)	(0.002)		(0.039)	(0.041)		(0.228)		
CMFM	0.96	-0.24	13.78	-548.68	-12.62	2,157.97	-75,667.48	0.42	2.99	19.23
	(0.000)	(0.030)	(0.544)	(0.335)	(0.106)	(0.249)	(0.483)	(0.170)		

Notes. This Table reports unrestricted GMM estimates of the LMFM, QMFM and CMFM with conditioning variables for the period 1965-2005, obtained using Hansen and Jagannathan's (1997) second moment weighting matrix. For each included factor, we report the corresponding SDF coefficient point estimate and its p-value in brackets. We also report the Hansen and Jagannathan (1997) distance and its p-value in brackets. The second last column reports the annualized volatility of the stochastic discount factor in percentage. All the variables are defined as in the text.

Figure 3-2
SDF implied by estimates of the unconstrained CMFM
(1965-2005)



Notes. This Figure reports, in the top panel, the SDF time-series implied by estimates of the unconstrained CMFM and, in the bottom panel, the plot of the estimated expected return on the 17 industry sorted portfolios against their sample average return and the plot of the estimated SDF against the realizations of the market return sorted from lowest to highest. The data is monthly for the period 1965-2005.

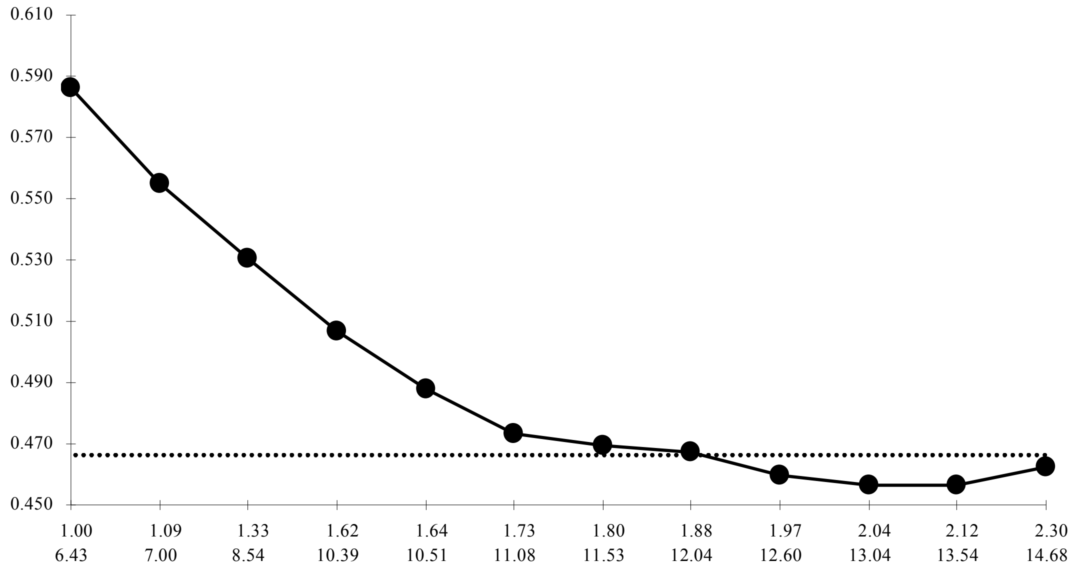
Table 3-5
GMM estimates of conditional models under NA
(1965-2005)

Model	a	$b_{vw,1}$	$b_{vw,2}$	$b_{vw,3}$	$b_{l,1}$	$b_{l,2}$	$b_{l,3}$	HJ_T (p.-*value*)	$\sigma(m)$	RRA_v
					Market portfolio without human capital					
LMFM	1.05	-3.67						0.53	0.72	4.66
	(0.000)	(0.001)						(0.000)		
QMFM	1.06	-5.78	24.25					0.50	1.15	7.40
	(0.000)	(0.000)	(0.005)					(0.007)		
CMFM	1.06	-2.05	5.63	-613.24				0.50	1.21	7.81
	(0.000)	(0.064)	(0.178)	(0.773)				(0.006)		
					Market portfolio with human capital					
LMFM	1.03	-3.26			0.65			0.52	1.17	7.54
	(0.000)	(0.001)			(0.100)			(0.002)		
QMFM	1.00	-4.86	38.00		-18.32	1,098.82		0.47	1.87	12.03
	(0.000)	(0.001)	(0.003)		(0.199)	(0.521)		(0.024)		
CMFM	0.92	-0.25	13.78	-548.68	-12.62	2,157.97	-75,667.48	0.46	2.30	14.81
	(0.000)	(0.049)	(0.804)	(0.781)	(0.873)	(0.596)	(0.999)	(0.015)		

Notes. This Table reports GMM estimates of the LMFM, QMFM and CMFM for the period 1965-2005, obtained using Hansens and Jagannathan's (1997) second moment weighting matrix under a positivity restriction on the estimated SDF. For each included factor, we report the corresponding SDF coefficient point estimate and its p-value in brackets. We also report the Hansen and Jagannathan (1997) distance and its p-value in brackets. The second last column reports the annualized volatility of the stochastic discount factor in percentage. All the variables are defined as in the text. The conditioning variables are $r_{m,t}$, dy_t, ys_t, tb_t.

Figure 3-3
Pricing errors vs. SDF volatility/RRA bound

Conditional CMFM with human capital



(1965-2005)

Notes. This figure plots the Hansen and Jagannathan (1997) distance of GMM estimates of the conditional CMFM with human capital under a varying volatility upper bound on the estimated SDF. The annualized volatility of the estimated SDFs and the corresponding *RRA* upper bound are reported on the horizontal axis. The sample period is 1965-2005. Values of the Hansen and Jagannathan (1997) distance that plot above the dotted line are significant at the 5 percent level. The conditioning variables are $r_{m,t}$, dy_t, ys_t, tb_t. All variables are defined as in the text.

Table 3-6
GMM estimates of conditional CMFM under NA (1965-2005)

z_t	a	$b_{vw,1}$	$b_{vw,2}$	$b_{vw,3}$	$b_{l,1}$	$b_{l,2}$	$b_{l,3}$	HJ_T (*p.-value*)	$\sigma(m)$	RRA_v
$r_{m,t}$	1.00	-2.67	0.00	-5.10	-5.58	1458.78	-45873.47	0.34	0.46	2.97
	(0.000)	(0.009)	(0.955)	(0.759)	(0.882)	(0.504)	(0.842)	(0.001)		
dy_t	1.21	-2.62	6.32	-25.35	-39.24	1313.89	-73356.49	0.29	0.74	4.77
	(0.000)	(0.020)	(0.758)	(0.772)	(0.040)	(0.050)	(0.094)	(0.061)		
ys_t	0.84	-2.23	0.13	-2.00	-10.43	4550.70	-5.09	0.27	0.94	6.07
	(0.000)	(0.028)	(0.954)	(0.907)	(0.751)	(0.011)	(0.995)	(0.192)		
ys_t	0.91	-0.58	39.34	-40.75	-1.85	1006.42	-77507.20	0.32	0.78	5.00
	(0.000)	(0.178)	(0.114)	(0.228)	(0.495)	(0.049)	(0.045)	(0.005)		
tb_t	0.86	-2.19	33.82	-2.88	-8.38	2739.05	-10212.62	0.30	0.78	5.00
	(0.000)	(0.032)	(0.283)	(0.854)	(0.686)	(0.091)	(0.759)	(0.069)		

Notes. This Table reports GMM estimates of the CMFM for the period 1965-2005, obtained using Hansens and Jagannathan's (1997) second moment weighting matrix under a positivity restriction on the estimated SDF. The first column reports the conditioning variable used in estimation. For each included factor, we report the corresponding SDF coefficient point estimate and its p-value in brackets. We also report the Hansen and Jagannathan (1997) distance and its p-value in brackets. The second last column reports the annualized volatility of the stochastic discount factor in percentage. All the variables are defined as in the text. Since the RRA_V implied by ys_t is over 5, then we also report the SDF estimates under the restricted condition $RRA_V \leq 5$.

4. The Sentiment in SDF: Behaviour Approach

4.1. Introduction

Habit persistence, as in the model of Campbell and Cochrane (1999), and recursive non-(time) separable preferences, as those employed by Epstein and Zin (1989), go some way towards explaining the equity premium puzzle. These models represent at least partially successful attempts to overcome some of the empirical shortcomings of the consumption-based CAPM without abandoning the tenets of rational asset pricing theory and thus, essentially, the view that asset prices are set by an expected utility optimizing representative investor endowed with rational expectations. Rational asset pricing theory, however, has not yet succeeded in the task of explaining the cross-section of stock returns. For example, it is difficult to reproduce the explanatory power of a number of stock characteristics[26] using empirical specifications of available rational asset pricing models, most notably the CAPM. Well known examples of such characteristics are firms' size and their book-to-market ratio, as in Fama and French (1992, 1993) studies, Jegadeesh and Titman (1993) momentum and, more recently, idiosyncratic volatility, as in Ang *et al.* (2006); Ang *et al.* (2009), (Peterson & Smedema, 2011), and (Fousseni, 2011).

Conditional specifications, that allow for time-variation in the parameters of the representative investors IMRS, and 3 and 4-moment versions, are amongst the most empirically successful extensions of the CAPM. For example, Lettau and Ludvigson (2001b) demonstrate that a multifactor model inspired by the (C)CAPM performs

[26] Other examples of relatively successful characteristics are the momentum effect documented by Jegadeesh and Titman (1993) and the liquidity effect documented by Pastor and Stambaugh (2003).

much better than the unconditional CAPM and almost as well as the Fama and French (1992, 1993, 1995) 3-factor model, while Harvey and Siddique (2000) argue that a conditional version of the 3M-CAPM captures a large portion of the cross-sectional variation in average stock returns[27].

Lewellen and Nagel (2006); Lewellen *et al.* (2010) however, warn that the sign of the risk premia estimated by Lettau and Ludvigson (2001b) is problematic from the perspective of the conditional consumption CAPM. More specifically, the risk premia point estimates reported by Lettau and Ludvigson (2001a) imply that (conditional) relative risk-aversion takes negative values for certain sample realizations of the conditioning variable. Moreover, these estimates imply a SDF, or pricing kernel, that might take negative values. This implies the existence of arbitrage opportunities, i.e. non-negative payoffs with a negative price. It is also inconsistent with the assumption that the representative investor's preferences display non-satiation, as in the (C)CAPM the SDF and the representative investor's IMRS coincide. There is therefore a puzzling contrast, in the specification of the (C)CAPM estimated by Lettau and Ludvigson (2001b), between the high cross-sectional explanatory power of the factors implied by the model and the inconsistence between the parameter estimates and fundamental assumptions underlying the model itself. We might label this problem as the "(C)CAPM puzzle".

On a related note, Dittmar (2002) and Post *et al.* (2008) point out that covariance and coskewness risk prices estimated in empirical tests of the 3 and 4 moment CAPM

[27] Adesi, Gagliardini and Urga (2004) and Potì (2005), using a quadratic market model consistent with the 3M-CAPM, add to the evidence that models that allow for both covariance and coskewness premium fit the cross-section of stock returns well.

imply a non-concave utility function, to an extent that might be inconsistent with the models being tested. More worryingly, these authors show that the empirical fit of these models is greatly reduced when the shape of the representative investor's utility function is restricted to display non satiation, risk aversion and non increasing absolute risk aversion (henceforth, NS, RA and NIARA, respectively). Yet, the evidence on the cross-sectional explanatory power of coskewness is compelling. This evidence, coupled with the critique put forth by Post *et al.* (2008) and Potì and Wang (2010), gives rise to a "coskewness puzzle" in the complete market-representative investor setting of the 3M-CAPM.

There is the concrete possibility that an omitted factor problem might be behind both puzzles. In this case, OLS estimates of the model parameters, i.e. factor risk premia and risk prices, would be inconsistent. In this chapter, we explicitly explore this possibility and, in particular, we investigate whether the omitted factors are related to systematic investor error. We thus propose a multi-factor specification that augments 2 and 3 moment versions of the (C)CAPM with a sentiment factor. This specification is based on the central result of Shefrin (2010), namely that the pricing kernel can be decomposed into two terms, one being sentiment and the other being an expression that depends only on economic fundamentals. The former term can be seen as the behavioural component of the kernel, while the second can be seen as its rational component. We check whether adding the sentiment factor, and thus allowing for a behavioural influence on asset pricing, allows to retain or improve the explanatory power of the (C)CAPM and 3M-CAPM while admitting risk price point estimates consistent with investors risk aversion (RA) and non increasing absolute risk aversion (NIARA), and thus with the underlying economic theory. Baker and Wurgler (2006) carefully constructed a proxy for investors' sentiment and demonstrated that exposure to this variable can explain a significant portion of the cross-section of stock returns. We thus employ this variable to proxy for sentiment in our specification. These authors, however, did not specify the representative investor's problem and her role in setting prices and therefore their results, while suggestive, cannot be directly used to characterize the relative importance of the rational and behavioural component of asset prices. On the contrary, Cecchetti *et al.* (2000) and Abel (2002) did model the impact of sentiment in a representative investor setting, but their empirical focus was on the equity premium rather then on cross-sectional differences in stocks average returns.

The remaining of this chapter is organized as follows. In the next Section, we introduce Shefrin (2010) SDF decomposition into a rational and sentiment related component and we map this decomposition into a linear factor model. In Section 4.3, we outline our estimation strategy. In Section 4.4, we present our dataset. In Section 4.5, we report our empirical results. Section 4.6 concludes.

4.2. Sentiment and the Pricing Kernel

As in Shefrin (2010), we may model the stochastic process for the pricing kernel as follows:

$$m(x_i) = \delta_R \Lambda(x_i) \frac{U'(C(x_i))}{U'(C_0)}. \tag{4-1}$$

Here, δ_R is the representative investor's subjective rate of time preference, $\Lambda(x_i) \equiv \dfrac{P_R(x_i)}{\Pi(x_i)}$ is the ratio of the probability weight $P_R(x_i)$ assigned to state x_i by the representative trader and the objective probability $\Pi(x_i)$ of occurrence of state x_i, C_0 represents current aggregate consumption and $C(x_i)$ represents aggregate consumption in state x_i. The quantity $\Lambda(x_i)$ can be interpreted as sentiment and it plays a role similar to investors' pessimism and optimism in the models of the equity premium developed by Cecchetti *et al.* (2000) and Abel (2002). Letting for simplicity $\delta_R \cong 1$ and exploiting the fact[28] that $\Lambda(x_i) \cong 1 + \ln \Lambda(x_i)$, the kernel process can be approximated as follows:

$$m(x_i) \cong \frac{U'(C(x_i))}{U'(C_0)} + \frac{U'(C(x_i))}{U'(C_0)} \ln \Lambda(x_i) = m_c + m_s . \tag{4-2}$$

In (4-2), $m_c \equiv \frac{U'(C(x_i))}{U'(C_0)}$ is aggregate marginal utility growth, while $m_s \equiv \frac{U'(C(x_i))}{U'(C_0)} \ln \Lambda(x_i)$ is the product of marginal utility growth and the log of the ratio between the representative investor's probabilities and the correct probabilities. The above equation thus states that the pricing kernel is the sum of two distinct processes, one based on the kernel of a rational expected utility optimizing representative investor and the other based on sentiment. The sentiment component is zero only when the likelihood ratio $\Lambda(x_i)$ equals one and thus only when prices are set in such a way that the representative investor holds objectively correct belief.

Recalling from the last chapter, Arrow (1971) argued that investors' utility functions should display non satiation (NS), risk aversion (RA) and non-increasing absolute risk aversion (NIARA). The latter is related to the notion of prudence, see Kimball (1990,

[28] This approximation implicitly assumes that $\Lambda(x_i)$ is never much different from one.

1993). With utility functions $U(W)$ defined over wealth, NS implies positive marginal utility of wealth, i.e. $U'(W) > 0$, RA implies decreasing marginal utility, i.e. $U''(W) < 0$, whereas NIARA, i.e. $\frac{d(-U''/U')}{dC} \leq 0$, implies that the rate of decrease of marginal utility does not increase in wealth. A necessary condition for NIARA, as shown in Arditti (1967), is $U''' \geq 0$. Hence NIARA implies $U''' \geq 0$ and aversion to negative skewness. We may approximate the marginal utility growth of the representative investor with preferences defined over aggregate consumption using a Taylor's expansion,

$$\frac{U'(C(x_i))}{U'(C_0)} = 1 + \frac{1}{2}\frac{U''(C_0)}{U'(C_0)}C_0 r_c(x_i) + \frac{1}{6}\frac{U'''(C_0)}{U'(C_0)}C_0^2 r_c(x_i)^2 + \cdots . \tag{4-3}$$

Here, $r_c(x_i) \equiv \frac{\Delta C(x_i)}{C_0}$ is aggregate consumption growth. Differentiating (4-3) twice with respect to wealth, it becomes clear that a necessary and sufficient condition for $U''' \geq 0$ is $b_2 \geq 0$ and thus this is also a necessary condition for NIARA. When this condition holds, a necessary condition for RA is $b_{1,t} < 0$. Following Dittmar (2002), we might allow for a multi-period setting with predictability due to a possibly time-varying investment opportunity set, and we might thus generalize (4-3) as follows:

$$m_{c,t+1} = 1 + h_{1,t}\frac{U''(C_t)}{U'(C_t)}r_{c,t+1} + h_{2,t}\frac{U'''(C_t)}{U'(C_t)}r_{c,t+1}^2 + h_{3,t}\frac{U''''(C_t)}{U'(C_t)}r_{c,t+1}^3 + \cdots . \tag{4-4}$$

Here, the h_i terms, i = 1, 2 3…, are non-negative expansion parameters. Dropping terms of order higher than the second, we might rewrite $m_{c,t+1}$ more compactly as follows:

$$m_{c,t+1} = 1 + b_{1,t} r_{c,t+1} + b_{2,t} r_{c,t+1}^2 \ldots. \tag{4-5}$$

Since $b_{1,t} \equiv h_{1,t} \frac{U''(C_t)}{U'(C_t)}$ and $b_{2,t} \equiv h_{2,t} \frac{U'''(C_t)}{U'(C_t)}$, RA and NIARA imply $b_{1,t} < 0$ and $b_{2,t} \geq 0$. With power utility with risk aversion parameter γ, we have that $b_{1,t} = -\gamma$ while the coefficients of the higher order terms are zero. This is the constant relative risk aversion (CRRA) case considered by Shefrin (2010). The component m_s of the kernel in (4-2) is the cross-product of m_c and the log-likelihood ratio $\ln \Lambda(x_i)$. Letting, for notational simplicity, $\ln \Lambda(x_i) = s(x_i)$, using (4-2) and (4-5), and dropping terms of third and higher orders, we can thus rewrite the kernel as follows:

$$m_{t+1} \cong 1 + b_{1,t} r_{c,t+1} + b_{2,t} r_{c,t+1}^2 + b_{3,t} s_{t+1} r_{c,t+1}. \tag{4-6}$$

4.3. Estimation Strategy

Consider the general conditional factor model, where the SDF is a (conditionally) linear function of a set of factors f_{t+1},

$$m_{t+1} = a_t + b_t' f_{t+1}. \tag{4-7}$$

The model in (4-6) is a specification of the general conditional factor model in (4-7), with $a_t = 1$ and factors f_{t+1} given by a constant, the growth rate of aggregate consumption, $r_{c,t+1}$, the square of the latter, $r_{c,t+1}^2$, and the cross-product between the sentiment proxy, s_{t+1}, and the consumption growth rate, $s_{t+1} r_{c,t+1}$. Since the price of excess returns $r_{i,t+1}$ is by definition zero, the prcing of $r_{i,t+1}$ can be rewritten as $0 = E_t\left(m_{t+1} r_{i,t+1}\right)$. The unconditional implications of the latter are

$$E(m_{t+1}r_{i,t+1}) = 0\ . \tag{4-8}$$

As in Dittmar (2002), we model variation in b_t as a linear function of the first lag of a vector of conditioning variables z_t (which typically include a constant), i.e.

$$b_t = b_i' z_t \qquad i = 1, \ldots, k\ . \tag{4-9}$$

The unconditional factor model implied by (4-7), (4-8) and (4-9) can then be written as follows,

$$m_{t+1} = a + b' F_{t+1}\ . \tag{4-10}$$

Here, $a \cong 1$ and b is a vector that stacks all the b_i vectors in a column. The specification in (11) is an unconditional model, i.e. a model with time-invariant parameters, in the new set of (unconditionally) de-meaned factors $F_{t+1} = (f_{t+1} \otimes z_t) - E(f_{t+1} \otimes z_t)$. When pricing excess returns, the means of excess returns do not identify the mean of the risk free rate[29], see for example Cochrane (2001). Thus, for simplicity, we will set the mean of the kernel in (4-10) equal to one,

$$E\left(m_{t+1}\right) = a + b' E\left(F_{t+1}\right) \cong 1\ . \tag{4-11}$$

Based on (4-10) and (4-11), we can rewrite the restrictions that (4-8) impose on the cross-section of expected returns as follows:

$$E(r_{i,t+1}) = \beta_i' \lambda\ . \tag{4-12}$$

Here, $\beta_i = var\left(F_{t+1}\right)^{-1} cov\left(F_{t+1}, r_{i,t+1}\right)$ is a vector of factor loadings of the regression of asset i on the factors. The elements of λ are factor risk-premia (of the unconditional model). Following a widely used terminology, we will refer to (4-12) as the beta-pricing representation of the restrictions that (4-8) and (4-10) impose on the

---

[29] This, however, is strictly true only as long as the risk free rate is not unrealistically high.

cross-section of expected returns. The risk-premia and the parameters of the unconditional SDF model are linked as follows:

$$\lambda = -var(F_{t+1})b \,. \tag{4-13}$$

We estimate (4-12) using a robust 2-pass regression without intercept in the second pass. In the first pass of this procedure, we regress the test asset payoff excess returns on the factors of the unconditional model and in the second pass we regress average excess returns on the factor loadings estimated in the first pass. This yields factor risk premia λ estimates and we then retrieve the parameters of the SDF (of the unconditional model) using (4-13). It can be shown that this approach is equivalent to first-stage GMM. In a GMM setting, in fact, the parameters of the kernel that prices a vector of n test asset payoff x_{t+1} can be estimated by solving the following problem:

$$\min_{\{m\}} \; g_T' W_{n(1+k)\times n(1+k)} g_T \,. \tag{4-14}$$

$$g_T = E_T(g_t) \qquad g_t = \left[m_{t+1}(b)x_{t+1} - p_t\right] \otimes z_t \,. \tag{4-15}$$

Here, $E_T(\cdot)$ denotes a sample average, i.e. an arithmetic average over a sample of T observations, and z_t is a vector of k instruments that coincide with the conditioning variables in (4-9). Based on (4-8), the elements of the $n \times 1$ vector g_t can be interpreted as pricing errors and the moment conditions g_T as pricing errors sample averages. Under the usual ergodicity assumption, the latter are consistent estimates of the unconditional expectations of (4-8). The $n \times n$ matrix W is a weighting matrix for the moment conditions. The 2-pass OLS regression amounts to minimizing (4-14) using the identity matrix to weight the moment conditions, i.e. setting $W = I$. Efficient second-stage and iterated GMM estimates are obtained instead by setting W equal to the optimal weighting matrix of Hansen (1982). We also experiment with the latter approach but we do not report the results as these are qualitatively the same as those obtained using the 2-pass regression. Finally, with the parameters of the unconditional model in hand, we back out the parameters of the conditional model using (4-9).

4.4. Data

The sample period starts in the first quarter of 1966 (to avoid the impact of dividend taxation reform in the early part of the previous decade) and ends in the last quarter of 2005. The test asset payoffs are 25 size and book-to market sorted portfolio and 30 value-weighted sorted portfolios of the NYSE, AMEX, and NASDAQ stocks[30]. We use returns on the three-month Treasury Bill as the risk-free rate and aggregate consumption expenditure data constructed as in Lettau and Ludvigson (2001b)[31]. We set the elements of s_t equal to the first difference of a proxy for investors' sentiment constructed as described by Baker and Wurgler (2006)[32], namely the updated version of their $\Delta SENT^{\perp}$ variable. More specifically, the latter is the first difference of a sentiment index[33] based on the first principal component of six (standardized) sentiment proxies over the period 1966-2005, where each of the proxies has first been orthogonalized with respect to a set of macroeconomic conditions. Finally, we use the consumption-wealth ratio estimates cay_t supplied by Lettau and Ludvigson (2001a) as the conditioning variable that drives, in (4-9), the variation of the parameters of the representative investor's IMRS. Notice that the much debated "look-ahead" bias of

---

[30] Data on this portfolios was downloaded from Kenneth French website.
[31] We thank Martine Lettau and Sidney Ludvigson for making this data available on their web-sites.
[32] We thank these authors for making this data available for download from the AFA-Journal of Finance web-site.
[33] This index is given by Equation (3) in their paper.

this variable, if present, is actually a desirable feature in our context. This is because rational expectations themselves should indeed display look ahead bias, as rational investors know the return data generating process, see Muth (1961) for a seminal reference and, more recently, Sargent (1996).

4.5. Empirical Results

Table 4-1 and Table 4-2 report estimates of an empirical specification of the (C)CAPM and 3M(C)CAPM, with an without sentiment. In particular, Table 4-1 reports risk premia estimates and associated *t*-statistics. We construct *t*-statistics using standard errors adjusted, based on Shanken (1992) correction, for sampling error that arises because the regressors β_t are estimated in a first-stage time-series regression. Following Lettau and Ludvigson (2001b), we also construct conventional OLS *t*-statistics. This is motivated by the argument put forth by Jagannathan and Wang (1996), who show that OLS standard errors from a 2-pass procedure do not necessarily overstate the precision of the standard errors, even if conditional heteroskedasticity is present. Table 4-2 reports the estimates of the SDF parameters (of the unconditional model), their *t*-statistics, without and with Shanken's (1992) correction, and a first-stage GMM T_J test statistic of the null that the pricing errors are jointly zero. This statistic is constructed as a quadratic form of the pricing errors $g_T' W g_T$, weighted by the first-stage GMM weighting matrix, $W = Cov(g_T(\hat{b}))^{-1}$. The latter is defined as follows

$$Tcov(g_T(\hat{b})) = \left[I_n - d(d'd)^{-1}d'\right]' cov_T(g_t)\left[I_n - d(d'd)^{-1}d'\right] \qquad (4\text{-}16)$$

Here, I_n is the $n \times n$ identity matrix, $cov_T(\cdot)$ denotes the sample variance-covariance matrix and $d = \frac{\partial g_T(b)}{\partial b} = E_T(F_{t+1}r'_{t+1})$ is the vector of the first derivatives of the moment conditions in (4-15) with respect to the model parameters.

Concerning the (C)CAPM estimates, all the factor risk premia are statistically significant when *t*-statistics are computed using OLS standard errors. With Shanken's (1992) standard error adjustment, however, only the risk premium of $cay_t r_{c,t+1}$ is statistically significant. This should not surprise as it simply means that there is considerable sampling error in the estimation of the factor loadings β_t. The coefficient of determination is, however, not very high. In particular, it is much lower than the coefficient of determination reported by Lettau and Ludvigson (2001b). This is due to the fact that they include an intercept in the second pass regression, thus reducing pricing errors. More importantly, the estimates of b imply that both the SDF and $b_{1,t}$ display the wrong sign for prolonged portions of the sample period. As shown in Panel A of Figure 4-1, the estimated SDF often takes negative values, thus assigning negative prices to non-negative payoffs almost surely. This implies the existence of arbitrage opportunities and violation of the assumption that investors preferences display non-satiation. Moreover, as shown in Panel B, the estimated $b_{1,t}$ does not always take a negative value, thus implying a violation of the assumption that investors are risk averse. Interestingly, $b_{1,t}$ becomes positive roughly when anecdotal evidence suggests that the market assessment of perspective earnings growth might have been affected by over-optimism, i.e. in the second part of the

1990s, end it peaks on June 2000, roughly at the beginning of the stock market correction. The estimates of the 3M(C)CAPM suffer from the same problems, i.e. the SDF is not always positive and b_{1t} is not always negative (neither however plotted to save space), alongside a non-positive point estimate of b_2, in contrast with the assumption of non-increasing absolute risk aversion and thus with standard risk aversion.

The Sentiment-(C)CAPM and Sentiment-3M(C)CAPM are empirically more successful, and most of the increase in the explanatory power is due to the inclusion of sentiment. As shown in Table 4-1, the R^2 of Sentiment-(C)CAPM jumps to 58.9 (50.7 adjusted). This value is comparable to the coefficient of determination obtained by Lettau and Ludvigson (2001b) allowing for an intercept in the second pass regression. The R^2 of the 3-moment (C)CAPM with sentiment is not much larger (it is actually lower, when adjusted for degrees of freedom) than the R^2 of its 2-moment version but it is much larger than the 3-moment specification without sentiment.

The T_J test statistic reported in Table 4-2 is never statically significant, thus implying that pricing errors are jointly insignificant. Moreover, while not all risk premia are statistically significant at conventional levels (merely implying that the price of the factors themselves might be close to zero), both aggregate consumption and sentiment are priced in the cross-section of average returns, as demonstrated by the significance of the corresponding elements of the b vector reported in Table 4-2. More importantly, the point estimates of the elements of b imply a much better behaved behaviour of the rational component of the SDF compared to the specifications without sentiment. As shown in Panel A of Figure 4-2, while the SDF still often takes negative values, its rational component (the ticker line) is rarely negative. Perhaps more remarkably, $b_{1,t}$ is now always negative, both in the case of the 2-moment model, as shown in Panel B of Figure 4-2, and in the case of the 3-moment model (not reported to save space). This is important as it implies that, including sentiment as a factor, the sign of the price of systematic conditional covariance is consistent with the risk-aversion assumption for any realization of the conditioning variable. In Figure 4-3, we report the sentiment component of the SDF. This is the component that helps the SDF fit the cross-section of average excess returns while allowing for a representative investor's IMRS much more consistent with the tenets of rational optimizing behaviour. As it is evident from the Figure, this component of the SDF peaks at times of high market valuations, such as the so called 'dot-com' boom of the late 1990s, and its troughs coincide with the bottom of market corrections, when investors' judgment might have been clouded by pessimism, as in the aftermath of the Latin American debt crisis in 1982. At these times, it is necessary to allow for a substantial amount of systematic investors' information processing error, to be able to justify stock valuations and still retain the assumption that, by and large, investors are greedy, risk averse, expected utility maximizing individuals.

We also report estimates that add an extra sentiment factor, namely s_t alongside $s_t cay_{t+1}$, and, to facilitate comparison of our estimates with those reported by Lettau and Ludvigson (2001b), we include cay_{t+1} among the factors. The estimates of these augmented specifications are reported in Table 4-3 and Table 4-4. The only qualitatively important difference is that now the point estimate of the b_2 parameter of the 3M-(C)CAPM is positive, in accordance with NIARA. This suggests that there might be additional factors with which both cay_{t+1} and the square of the consumption growth rate are correlated. We leave for future research the identification of the extra factor (factors) and, more importantly, of whether it is should enter the rational or the sentiment-related component of the SDF.

4.6. Conclusions

This chapter shows that augmenting the (C)CAPM with sentiment, and thus allowing for systematic investor error in forming beliefs about the distribution of returns, permits to largely reconcile investors' optimizing behaviour with the cross-section of average returns. This implies that investors must either commit systematic errors in assessing the joint distribution of stock returns and aggregate consumption or behave in a way that, at the aggregate level, is inconsistent with expected utility maximization and with standard risk aversion assumptions, as formulated for example by Kimball (1993). We leave for future research to ascertain whether these systematic (ex-post) errors might have been avoided making full use of available information, thus implying a violation of the Efficient Market Hypothesis, as formulated by Fama (1970, 1976).

Table 4-1
Second Pass Regressions (1966-2005) Estimates
25 Size and Book-to-Market Portfolios
Beta-Pricing Representation

Model	$R_{c,t+1}$	$cay_t R_{c,t+1}$	$s_{t+1} R_{c,t+1}$	$R^2_{c,t+1}$	R^2 *(Adj.* R^2*)*
(C)CAPM	0.48	0.02			26.1
	(3.21)	(5.97)			(22.8)
	(0.99)	(1.83)			
3M(C)CAPM	0.21	0.01		0.01	33.7
	(1.36)	(4.71)		(4.63)	(27.6)
	(0.45)	(1.53)		(1.57)	
Sentiment-(C)CAPM	0.84	0.01	-1.35		56.5
	(5.59)	(3.70)	(-6.22)		(52.6)
	(1.76)	(1.21)	(-1.96)		
Sentiment-3M(C)CAPM	0.67	0.01	-1.22	0.01	56.5
	(4.61)	(3.25)	(-5.74)	(5.00)	(50.3)
	(1.59)	(1.16)	(-1.99)	(1.78)	

Notes. This Table reports 2-step regression estimates of the beta-pricing representation of various factor models for the period 1966-2005. The second pass regressions are estimated without an intercept term. The top row indicates the factors included in each model. For each included factor, we report the risk premia point estimates in percentage and two sets of t-statistics in brackets. These statistics are computed using OLS standard errors that account for correlated errors across test portfolios while the second set of *t*-statistics also uses Shanken' (1992) correction for the fact that the beta coefficients are estimated. The third and second last columns report the percent coefficient of determination R^2, both unadjusted and adjusted for the degrees of freedom. All the variables are defined as in the text. The data frequency is quarterly.

Table 4-2
Second Pass Regressions (1966-2005) Estimates
25 Size and Book-to-Market Portfolios
Implied SDF Parameters

Model	$R_{c,t+1}$	$cay_t R_{c,t+1}$	$s_{t+1} R_{c,t+1}$	$R^2_{c,t+1}$	***TJShanken***	σ_m	σ_{mc} σ_{ms}
(C)CAPM	-418.00	-37,904.00			17.00	629.8	629.8
	(0.000)	(0.000)			(0.763)		629.8
3M(C)CAPM	284.00	-30,117.00		-70,956.00	24.28	632.2	632.2
	(0.399)	(0.003)		(0.038)	(0.230)		632.2
Sentiment-(C)CAPM	-526.00	-13,219.00	305.00		15.08	617.5	446.8
	(0.000)	(0.309)	(0.007)		(0.772)		383.5
Sentiment-3M(C)CAPM	-209.00	-14,153.00	252.00	-30,205.00	24.25	433.84	239.25
	(0.987)	(0.319)	(0.056)	(0.450)	(0.232)		316.76

Notes. This Table reports the elements of the *b* vector, the negative of the risk prices, implied by 2-pass regression estimates (without intercept in the second pass regression) and, in brackets, associated p-values. These are computed using standard errors based on a weighting matrix equal to the inverse of the sample (first stage) estimate of the pricing errors variance-covariance matrix. The same weighting matrix is used in the computation of the average pricing errors *TJ* reported in the third last column. The last two columns report the annualized volatility of the estimated stochastic discount factor (in percentage) and its decomposition in a rational and sentiment-related component. The market Sharpe ratio is about 40 percent. All the variables are defined as in the text. The data frequency is quarterly.

Table 4-3
Augmentation by cay_t and s_t
Beta-Pricing Representation

Model	cay_t	$R_{c,t+1}$	$cay_t R_{c,t+1}$	s_{t+1}	$s_{t+1} R_{c,t+1}$	$R^2_{c,t+1}$	R^2 $(Adj.\ R^2)$
1	2.34	0.41	0.02				22.8
	(4.86)	(2.6)	(6.41)				(15.8)
	(1.49)	(0.79)	(1.93)				
2	0.86	0.64	0.01			0.003	34.3
	(2.23)	(3.87)	(4.22)			(2.13)	(24.9)
	(0.74)	(1.27)	(1.39)			(0.71)	
3	1.16	0.71	0.01		-1.46		54.7
	(3.18)	(4.54)	(4.2)		(-6.43)		(48.2)
	(1.34)	(1.71)	(1.52)		(-3.10)		
4	1.09	0.65	0.01		-1.39	0.01	55.16
	(3.32)	(4.44)	(3.59)		(-7.91)	(4.2)	(46.2)
	(1.02)	(1.47)	(1.44)		(-2.11)	(1.43)	
5	0.84	0.6	0.01	-80.61	-1.12		58.9
	(2.59)	(3.93)	(4.36)	(-3.36)	(-6.45)		(50.7)
	(0.97)	(1.43)	(1.62)	(-1.25)	(-2.38)		
6	0.86	0.64	0.01	-74.86	-1.13	0.003	59.2
	(2.61)	(4.33)	(4.41)	(-3.19)	(-6.42)	(2.13)	(48.5)
	(0.92)	(1.49)	(1.52)	(-1.12)	(-2.24)	(0.73)	

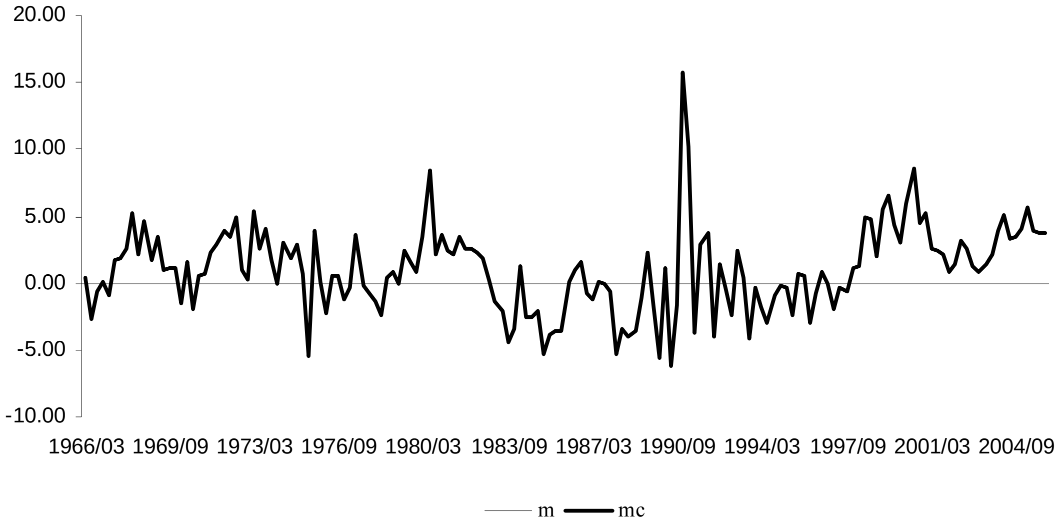
Notes. This Table reports 2-step regression estimates of the beta-pricing representation of various factor models for the period 1966-2005. The second pass regressions are estimated without an intercept term. The top row indicates the factors included in each model. For each included factor, we report the risk premia point estimates in percentage and two sets of t-statistics in brackets. These statistics are computed using OLS standard errors that account for correlated errors across test portfolios while the second set of t-statistics also uses Shanken' (1992) correction for the fact that the beta coefficients are estimated. The third and second last columns report the percent coefficient of determination R^2, both unadjusted and adjusted for the degrees of freedom. All the variables are defined as in the text. The data frequency is quarterly.

Table 4-4
Augmentation by cay_t and s_t
Implied SDF Parameters

Model	cay_t	$R_{c,t+1}$	$cay_t R_{c,t+1}$	s_{t+1}	$s_{t+1} R_{c,t+1}$	$R^2_{c,t+1}$	***TJShanken***	$\sigma(m)$	σ_{mc} σ_{ms}
1	-21.67	-391.61	-36,754.60				24.14	636.6	636.6
	(0.655)	(0.000)	(0.000)				(0.340)		636.6
2	-31.00	-510.00	-10,945.00			14,812.00	25.34	364.2	364.2
	(0.536)	(0.132)	(0.284)			(0.672)	(0.189)		364.2
3	-55.40	-468.80	-7,893.20		334.00		16.24	616.8	409.2
	(0.270)	(0.000)	(0.568)		(0.004)		(0.756)		419.9
4	-48.00	-342.00	-8,977.00		308.00	-12,772.00	16.23	544.2	332.4
	(0.393)	(0.448)	(0.531)		(0.037)	(0.775)	(0.702)		387.2
5	-27.96	-381.17	-11,457.46	-1.22	387.55		20.46	524.9	350.3
	(0.633)	(0.002)	(0.426)	(0.368)	(0.003)		(0.429)		340.3
6	-31.00	-510.00	-10,945.00	-1.00	429.00	14,812.00	12.56	556.6	298.91
	(0.600)	(0.297)	(0.451)	(0.618)	(0.033)	(0.785)	(0.895)		410.82

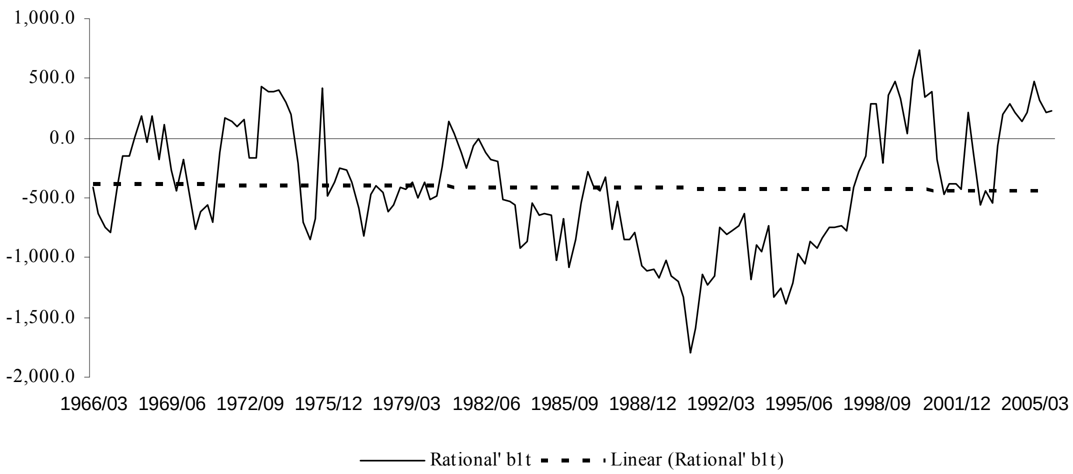
Notes. This Table reports the elements of the *b* vector, the negative of the risk prices, implied by 2-pass regression estimates (without intercept in the second pass regression) and, in brackets, associated p-values. These are computed using standard errors based on a weighting matrix equal to the inverse of the sample (first stage) estimate of the pricing errors variance-covariance matrix. The same weighting matrix is used in the computation of the average pricing errors *TJ* reported in the third last column. The last two columns report the annualized volatility of the estimated stochastic discount factor (in percentage) and its decomposition in a rational and sentiment-related component. The market Sharpe ratio is about 40 percent. All the variables are defined as in the text. The data frequency is quarterly.

Figure 4-1
(C)CAPM
Panel A
(Estimated SDF)



Panel B

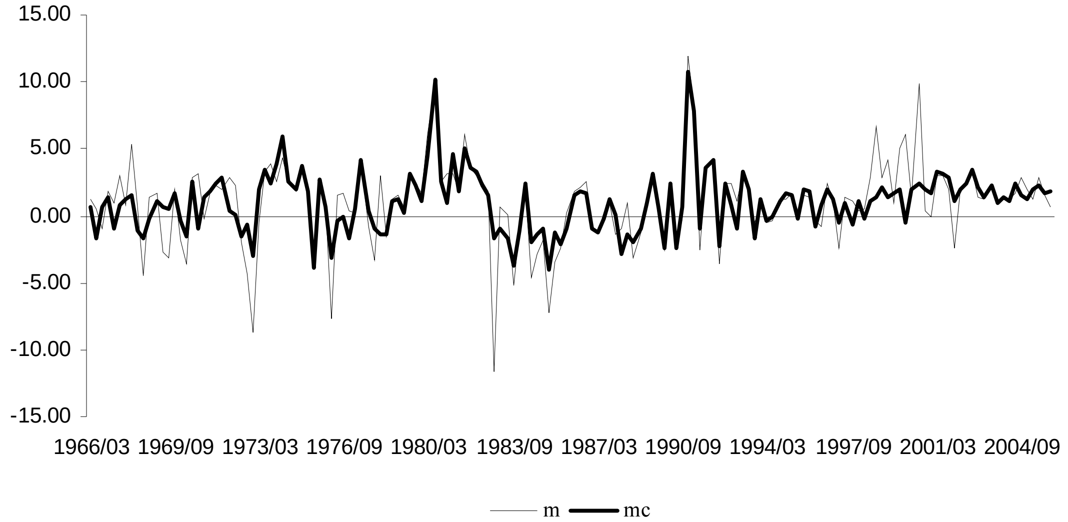
(Estimated $b_{1,t}$)



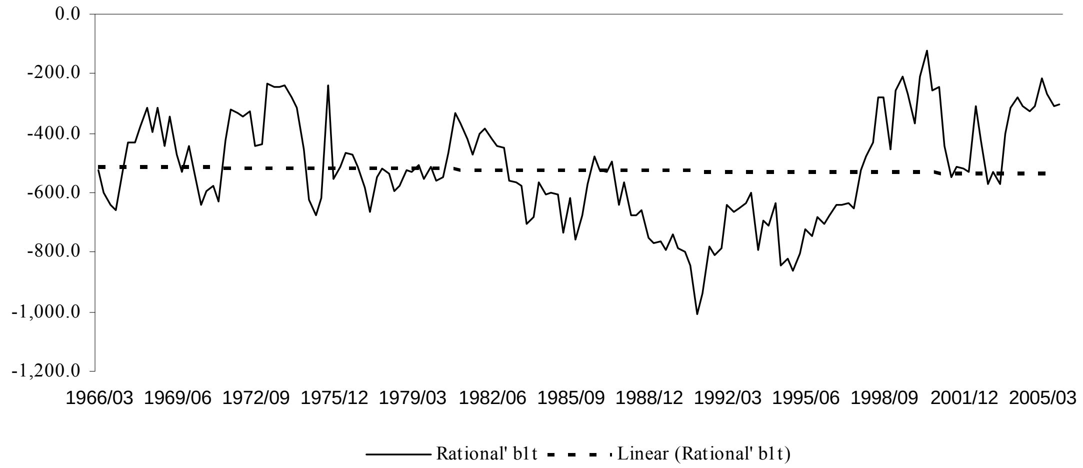
Notes. Panel A of this figure reports the SDF implied by the 2-step regression estimates of the (C)CAPM. Panel B reports the estimate of the conditional b_{1t} parameter for the same model and its linear interpolation (denoted as “linear”). The sample period is 1966-2005.

Figure 4-2
Sentiment-(C)CAPM

Panel A
(Estimated SDF)

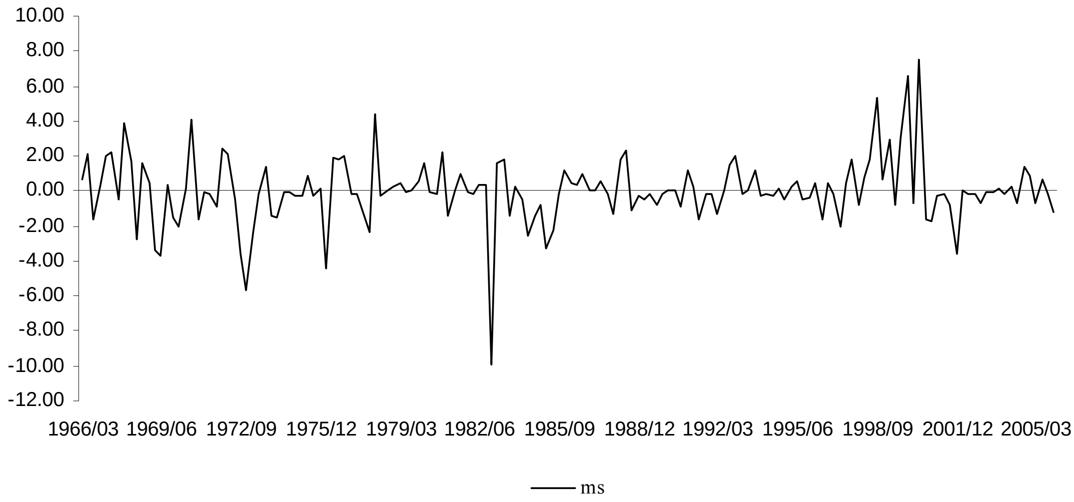


Panel B
(Estimated $b_{1,t}$)



Notes. Panel A of this figure reports the SDF implied by the 2-step regression estimates of the Sentiment-(C)CAPM, and its rational component. Panel B reports the estimate of the conditional b_{1t} parameter for the same model and its linear interpolation (denoted as "linear"). The sample period is 1966-2005.

Figure 4-3
The Sentiment Adjustment



Notes. This figure reports the sentiment component of the SDF implied by the 2-step regression estimates of the Sentiment-(C)CAPM. The sample period is 1966-2005.

5. SDF and Transaction Costs

5.1. Introduction

In this chapter, we compare several competing pricing kernels using a modified version of Hansen-Jagannathan distance (Hansen & Jagannathan, 1997), which not only accounts for conditional information but also allows for transaction costs. The competing pricing kernels we consider here include linear and quadratic models, together with the relevant conditional version. The assessment of candidate pricing kernels in this chapter is preceded by the discussion of extensions of the Hansen-Jagannathan (henceforth HJ) bound and distance measure allowing for both conditional information and transactions costs. On the one hand, it is shown how the payoff space is extended, by including scaled asset returns, when the investment opportunity set considered in deriving the bound includes dynamic investment strategies that exploit conditioning information. This modification leads to a sharper bound. On the other hand, we discuss the implications for the bound of allowing for the presence of transaction costs, which are assumed to be absent by a large part of the earlier literature. This modification leads to a loosening of the bound. Although former literature has already delt with these two issues separately (see Gallant *et al.* (1990), Bekaert and Liu (2004), and He and Modest (1995)), no study, to our knowledge, addresses both. .

The remaining of this chapter is organized as follows. Section 5.2 reviews the tradidtional Hansen-Jagannathan volatility bounds as well as the version that allows for transaction costs. Section 5.3 discusses the HJ Distance and the testing procedure followed in this chapter. Section 5.4 presents the dataset and the empirical results. Finally, Section 5.5 concludes the chapter.

5.2. Volatility Bounds and Transaction Costs

He and Modest (1995) generalized the equilibrium condition for the optimal intertemporal consumption-investment problem by allowing for transaction costs, which include both bid-ask spreads and commissions. Over a one period horizon, each unit investment leads to the end of period return $(1-\rho_i)R_{i,t+1} / (1+\rho_i)$. Here ρ_i denotes the rate of transaction costs in the price of asset i. Similarly, a one unit borrowing generates the end of period payment $(1+\rho_i)R_{i,t+1} / (1-\rho_i)$. For a portfolio, long and short positions must therefore satisfy the following pricing condition,

$$\frac{1-\rho_i}{1+\rho_i} \le E_t[\beta \frac{U'(C_{t+1})}{U'(C_t)} R_{i,t+1}] \le \frac{1+\rho_i}{1-\rho_i}. \tag{5-1}$$

Assuming transaction costs are the same across all assets and equal to ρ, Equation (4-2) becomes

$$\lambda_i^L \le E_t[m_{t+1} R_{i,t+1}] \le \lambda_i^U. \tag{5-2}$$

where $\lambda_i^L = (1-\rho)\tau / 1+\rho$ and $\lambda_i^U = (1+\rho)\tau / 1-\rho$ represent the lower and upper bounds, respectively, on the initial outlay for future payoffs, R_{t+1}. And τ is a vector of $N \times 1$ ones.

Hansen and Jagannathan (1991) suggested a non-parametric technique to find the projection of the admissible pricing kernels on the payoff space, based on the fact that

$$m_{t+1} = E_t(m_{t+1}) + [\tau - E_t(m_{t+1})E_t(R_{t+1})]\Sigma^{-1}[R_{t+1} - E_t(R_{t+1})] + \varepsilon \,. \qquad (5\text{-}3)$$

where $\Sigma = cov_t(R_{t+1},\ R'_{t+1})$ represents the covariance matrix of asset returns. This results derives directly from the Law of One Price. By a well known property of even powers, $\sigma_t^2(\varepsilon)$ is non-negative and therefore

$$\sigma_t^2(m_{t+1}) \geq [\tau - E_t(m_{t+1})E_t(R_{t+1})]'\Sigma^{-1}[\tau - E_t(m_{t+1})E_t(R_{t+1})] \,. \qquad (5\text{-}4)$$

This inequality has been used as one of the most popular diagnostic tools for inspecting unconditional pricing models. As a relatively week selection requirement, however, it fails to capture the implications of conditional information, and it does not take into account market frictions such as transaction costs, including bid-ask spreads, commission fees, taxation and so on. Many authors have examined the implications for the admissibility of candidate pricing models of either the use of conditioning information in dynamic investment strategies (Lettau and Ludvigson (2001b), Bekaert and Liu (2004), Ferson and Siegel (2003), and Dittmar (2002)) or of frictions such as the transaction costs (He & Modest, 1995). To our knowledge, however, no empirical study has considered the two issues at the same time.

We consider a payoff space augmented by the inclusion of dynamic strategies based on conditioning information. Lettau and Ludvigson (2001b) pointed out that the consumption-wealth ratio summarizes the return expectations of the representative agent and, as a result, it must be strongly linked to the expected stock market return, i.e.

$$cay_t \approx E_t \sum_{i=1}^{\infty} \rho_w^i r_{m,t+i} \,, \tag{5-5}$$

where cay_t denotes deviations from the long run value of the consumption-wealth ratio. We therefore consider conditioning information summarized by this variable and an augmented payoff space given by $(R'_{i,t+1},\ cay_t R'_{i,t+1})'$. If we set R_{t+1}^{new} equals $(R'_{i,t+1},\ cay_t R'_{i,t+1})'$, and the price vector τ^{new} equals $(\tau',\ cay_t \tau')'$, (5-2) then can be rewritten as follows,

$$\begin{aligned} \lambda_i^L \le E_t[m_{t+1} R_{i,t+1}] \le \lambda_i^U \\ cay_t \lambda_i^L \le E_t[m_{t+1}(cay_t R_{i,t+1})] \le cay_t \lambda_i^U \end{aligned} . \tag{5-6}$$

The conditional expectation can be dropped, since the extended payoff space includes the dynamic strategies in the basis of the conditional expectations. The corresponding pricing kernel volatility bound, accounting for transaction costs and conditional information, is as follows,

$$\begin{aligned} &\sigma^2(m_{t+1}) \ge \inf[\lambda^{new} - E(m_{t+1})E(R_{t+1}^{new})]'\Sigma_{new}^{-1}[\lambda^{new} - E(m_{t+1})E(R_{t+1}^{new})] \\ &s.\,t. \qquad \frac{1-\rho}{1+\rho}\tau^{new} \le \lambda^{new} \le \frac{1+\rho}{1-\rho}\tau^{new} \end{aligned} . \tag{5-7}$$

where $\Sigma_{new} = cov(R_{t+1}^{new},\ R_{t+1}^{new\prime})$.

5.3. Estimation and Diagnostic Tests

Hansen and Jagannathan (1997) proposed the so-called Hansen–Jagannathan (HJ) distance to measure the pricing error for any particular candidate pricing model. Setting families of pricing kernels that price assets correctly as the benchmark, the HJ distance shows the maximum pricing error per unit norm on a set of portfolio with n risky assets and takes the form

$$HJ_T(\theta) = \sqrt{E[g_T(\theta)]' G^{-1} E[g_T(\theta)]}\,, \tag{5-8}$$

where the weighting matrix $G = E(R_{t+1}^{new} R_{t+1}^{new\prime})$ is assumed to be nonsingular and

$$\begin{gathered} g_T(\theta) = E[(m_{t+1}(\theta) R_{t+1}^{new} - \tau^{new}) \otimes I_t] \\ s.\,t. \quad m_{t+1}(\theta) = \theta' f \end{gathered}\,, \tag{5-9}$$

here I_t stands for the conditioning variables cay_t and z_t in this chapter. This distance measure is equivalent to $\|\tilde{p}\|$, where $\tilde{p}$ is the correction to the proxy stochastic discount factor necessary to make it consistent with the data. As such, the HJ-distance is typically interpreted as the least-square distance between the given candidate pricing kernel and the true pricing kernel. The first order condition for its minimization criteria is

$$D_T' G_T^{-1} g_T(\theta) = 0\,, \tag{5-10}$$

which gives,

$$\hat{\theta}_T = (\hat{D}_T' \hat{G}_T^{-1} \hat{D}_T')^{-1} \hat{D}_T' \hat{G}_T^{-1} \tau^{new}\,, \tag{5-11}$$

where $D_T = E(R_{t+1} f')$. The asymptotic distribution of HJ Distance, as shown by Jagannathan & Wang (1996) differs from the χ^2 distribution of the pricing errors under the optimal weighting matrix Hansen (1982). It follows a combination of χ^2 distributions weighted by the nonzero eigenvalues of a particular matrix. Mathematically,

$$T[HJ_T(\theta)]^2 \xrightarrow{d} \sum_{j=1}^{N-K} \lambda_j v_j\,, \tag{5-12}$$

where v_j is random variable drawn from a $\chi^2(1)$ distribution, and λ_j are the nonzero eigenvalues of the following matrix:

$$\Theta = S^{\frac{1}{2}} G^{-\frac{1}{2}} [I - (G^{-\frac{1}{2}})' D (D' G^{-1} D)^{-1} D' G^{-\frac{1}{2}}] (G^{-\frac{1}{2}})' (S^{\frac{1}{2}})', \tag{5-13}$$

where S is the variance matrix of pricing errors, with sample counterpart $\hat{S} = \frac{1}{T} \sum_{1}^{T} [g_T(\hat{\theta}) g_T'(\hat{\theta})]$. $S^{\frac{1}{2}}$ and $G^{\frac{1}{2}}$ are the upper-triangle matrices obtained from a cholesky decomposition of matrices S and G, respectively. The testing procedure involves the simulation on M sets of $N-K$ random $\chi^2(1)$ variables. Then the *p*-value is

$$p = \sum_{i=1}^{M} I\{ \sum_{j=1}^{N-K} \lambda_j v_{ij} \geq T[HJ_T(\theta)]^2 \} \Big/ M, \tag{5-14}$$

where $I(\cdot)$ is the discrete choice function, which is equal to one when the underlying condition holds, and zero otherwise. This provides a diagnostic test based on the HJ-distance, which should be not statistically significant for the null that the model is correctly specified not to be rejected. This approach differs from the conventional two-stage GMM estimation by Hansen (1982). The HJ distance, as a measure of the departure between the candidate pricing kernels and the set of true pricing kernels, has a more intuitive interpretation than the spefification error statistic based on the GMM estimator that uses the optimal weighting matrix or it two-stage approximation..Perhaps more importantly, since the distance measure is formed on a weighting matrix that is invariant across all models tested, it can be used to directly compare the performance not only of nested models, but nonnested models as well.

Another advantage of the Hansen-Jagannathan approach is that it largely avoids the pitfall of favoring pricing models that produce volatile pricing errors. The Hansen-Jagannathan criterion is a function of the inverse of the second moment matrix of returns rather than the inverse of the second moment matrix of pricing errors.

Consequently, the HJ distance will fall only if the least-square distance to an admissible pricing kernel is reduced, and not if the proxy pricing kernel generates volatile pricing errors. Thus, the distance rewards models exclusively for improving pricing and not for adding noise.

One caveat is in order. The distribution of the Hansen-Jagannathan test statistic is a function of the optimal GMM weighting matrix. Consequently, when testing the significance of the Hansen-Jagannathan distance, one may find a high *p*-value because the parameters imply a “small” optimal GMM weighting matrix; that is, a weighting matrix characterized by highly volatile pricing errors. One potential safeguard against failing to reject a model due simply to noise in the pricing kernel is to analyze the significance of the parameter estimates. Whereas the distribution of the distance measure is rewarded for a small GMM weighting matrix, the distribution of the parameter estimates is penalized by a small GMM weighting matrix. That is, although a model may be accepted due to volatile pricing errors, the volatility will tend to reduce the significance of the parameter estimates. Consequently, we perform Wald tests to assess the significance of adding each marginal term in the pricing kernel. These tests provide some surety not only that a pricing kernel is not rewarded simply for being noisy, but also provides evidence as to the importance of adding polynomial terms, potentially alleviating concerns about overfitting.

A final advantage to the Hansen-Jagannathan distance measure is that the results may be more robust than in standard GMM estimates (Cochrane, 2001). Since the weighting matrix is not a function of the parameters, the results should be more stable.

Despite this advantage, Ahn and Gadarowski (2004) suggest that the size of the test statistic is poor in finite samples; the distance measure rejects correctly specified models too often. These results suggest the possibility that using the HJ GMM estimator rather than the optimal or two-stage GMM estimators may trade size for power. To gauge the possible impact of this trade-off, we also estimate the models using two-stage GMM.

5.4. Data and Empirical Results

5.4.1 Data

The data used in this chapter includes the logarithmic quarterly returns of 30 value-weighted industry-sorted portfolios of the stocks listed on the NYSE, AMEX, and NASDAQ for the period from the first quarter of 1954 to the third quarter of 2002, the quarterly return on the three-month Treasury Bill to proxy for the risk-free rate, the individual quarterly non-durable and services consumption expenditure from DataStream to proxy for the representative agent's intertemporal consumption stream. Finally, we use the consumption-wealth ratio series, cay_t as the proxy of conditional information to scale the factors. Empirically, the choice of conditioning variables has been discussed in many studies. The variables should reflect investors' expectations about future market conditions and predict asset returns. To check the robustness of our results, we also replicate the work of Dittmar (2002), i.e. we include the set of instruments $z_t = \{1,\ r_{m,t},\ dy_t,\ ys_t,\ dp_t\}$ as another information set differing from cay_t. In detail, the $r_{m,t}$ represents the market excess return, dy_t is the aggregate dividend yield, ys_t measures the yield term spread between three-month treasury bill return and one-month treasury bill return, and dp_t captures the industrial production growth. Table 5-1 is the statistic summary of conditioning factors.

To test the predictive ability of z_t, we project all returns R_{t+1} on the conditioning variable set z_t through the following regression,

$$R_{i,t+1} = b'z_t + e_{t+1}, \tag{5-15}$$

Table 5-2 provides Wald statistics, computed using the Newey and West heteroskedasticity and autocorrelation consistent covariance matrix, and the corresponding *p*-values. The null hypothesis for the test is that the lagged conditioning variables set has no predictive power for asset returns. Statistically, the coefficients b of the predictive regressions are jointly zero. The null is rejected at 1 percent critical level for all 30 industry portfolios.

5.4.2 Empirical Results

5.4.2.1 Estimation of pricing kernels

A traditional GMM procedure is employed to estimate the coefficients of the candidate pricing kernels, which include the linear, quadratic and cubic specifications with time-varying coefficients, with and without human capital $r_{c,t+1}$ as a component of the return on aggregate wealth. Moreover, two alternative sets of conditioning variables, i.e. $\{cay_t\}$ and z_t, are employed in estimation.

Table 5-3 presents results of specification tests when the measure of aggregate wealth does not include human capital $r_{c,t+1}$. The table presents average values of the coefficients $\overline{a}_t$ and $\overline{b}_{i,t}$, $i = 1, 2, 3$ corresponding to the i^{th} order term of the return on the market portfolio polynomial. The table also presents the TJ_T measure, i.e. T times the weighted average pricing error with the weighting given by the inverse of the error variance-covariance matrix $S = E[u(\hat{\theta})u'(\hat{\theta})]$, and associated p-values under the null that the candidate model is correctly specified. Panel A reports the coefficients of the linear conditional pricing models, with SDF $m_{t+1} = a_t + b_{1,t} r_{m,t+1}$, where the $r_{m,t+1}$ stands for the market return proxy, i.e. the market index return, and a_t and $b_{i,t}$ are time-varing as linear functions of the conditioning variable cay_t, as defined in Lettau and Ludvigson (2001b). Both these two coefficients are statistically significant, and the coefficient of maket return is negative, in accordance with the CAPM. The mispecification test statistic is not significant at conventional levels. As shown in Panel B, the linear market return term is significant in the quadratic model too and dominates the quadratic term. The p-values of the TJ_T statistic and of the individual coefficients suggest marginal improvement in the model fit in moving from a linear specification to a nonlinear one. The quadratic pricing kernel reduces the TJ_T measure from 75.25 to 74.99, and furthermore, the additional cubic term reduces TJ_T by another 0.65 drop. However, the addition of either the quadratic or cubic term does not materially improve the performance of the conditional pricing kernel. If we choose 10 percent as the critical p-value, none of the above models can generate the admissible pricing kernel which prices the cross section of returns on the 30 value-weighted industry-sorted portfolios.

We next analyse the impact of incorporating a measure of human capital in the return on aggregate wealth. These results are displayed in Table 5-4. The outcomes of the specification tests are markedly different from those in Table 5-3. The fit of all three pricing kernels improves relative to the case in which human capital is not included in the measure of aggregate wealth. The TJ_T measure implied by the linear pricing kernel falls to 74.81. This result is consistent with the findings of Jagannathan and Wang (1996), who find that incorporating human capital improves the performance of the conditional CAPM. However, the linear pricing kernel and the quadratic pricing kernel are rejected at the 10 percent significance level.

We subsequently re-estimated by mimimizing the HJ-distance and using the set of information variables z_t . Estimation results are qualitatively unchanged. Consistent with the two-stage GMM estimation, the distance measures and *p*-values for the tests of significance of the coefficients suggest marginal improvement in moving from a linear specification to a nonlinear specification. As shown in Panel B of Table 5-5, a cubic specification of the pricing kernel with human capital results in a decrease in the distance measure of 4.27 percent relative to the linear kernel without human capital. The same measure in Table 5-6 improves by 16.35 percent, which is in line with the improvement of 12.50 percent reported by Dittmar (2002). In either case, the cubic pricing kernel cannot be rejected at the 10 level. These results suggest that the performance of a pricing kernel grounded in preference theory, i.e. estimated imposing preference restrictions implied by decreasing absolute risk aversion and decreasing absolute prudence, can capture cross-sectional variation in returns. The results of Table 5-5 and Table 5-6 suggest that incorporating only nonlinear functions of the return on the value-weighted index or a linear function of the return on labour is insufficient to generate an admissible pricing kernel. However, consistently with Dittmar (2002), by utilizing both the return on labour and the nonlinearities implied by the series expansion, we are able to generate an admissible pricing kernel.

5.4.2.2 Volatility of pricing kernels and the HJ Bounds with Transaction Costs

Dittmar (2002) examines the relation of the estimated pricing kernels to the volatility bounds of Hansen and Jagannathan (1991), and shows that the cubic pricing kernel with human capital is able to generate sufficient volatility that meets the Hansen-Jagannathan bound for a certain range of values of the mean, but the estimated value of the latter is slightly too high for the pricing kernel to actually lie within the bound. This finding essentially implies that the cubic pricing kernel, even with human capital, is not volatile enough to be admissible though it is very close.

In order to cope with this problem, and gather some insight into the role of market frictions in asset pricing, we compare the estimated pricing kernels with the HJ bounds in the presence of transaction costs. As indicated in Section 5.2, the bounds represent the minimum volatility that a pricing kernel must exhibit, given its mean, to be admissible. In this respect, the bounds depict the set of admissible pricing kernels in mean-standard deviation space. Since the pricing kernel approach relates the first moment of returns to the second moment of the discount factor, this provides further insight into the specification of the model. The analysis differs from the specification test of the Hansen-Jagannathan distance measure, which asks whether there is some specific admissible pricing kernel that is statistically indistinguishable from that of the model.

The Hansen-Jagannathan bounds for the 30 value-weighted industry-sorted portfolios augmented by the three-month T-bill return are presented in Figure 5-1 panel A and panel B. The assumption of 0.80 percent transaction costs per quarter is similar to the one made by He and Modest (1995), which is reasonable given the existence of stamp duty, brokage commissions (or the bid-ask spread) and slippage costs. As suggested by Table 5-5 and Table 5-6, the conditional cubic pricing kernels with human capital dominate the others in terms of mean-volatility pairs, although neither of them reaches the HJ bounds in the absence of transaction costs. In contrast, the conditional cubic and quadratic pricing kernels lie insider the HJ bounds when transaction costs are taken into account. In detail, Panel A of Figure 5-1 shows that the conditional cubic pricing kernel with human capital is admissible when the conditioning information set is cay_t and transaction costs are 0.8 percent, and Panel B shows that the conditional cubic and quadratic pricing kernels either with or without human capital are all admissible when the conditioning information set is z_t and transaction costs are 0.8 percent. These findings are novel relative to those reported by Dittmar (2002).

The Hansen-Jagannathan bounds plot, together with the decomposition of the distance measure, indicate that incorporating human capital and transaction costs substantially improves the nonlinear pricing kernels' ability to match the volatility of the set of pricing kernels that are admissible for the industry portfolios. That is, it substantially lowers the standard deviation of the adjustment necessary to make the nonlinear pricing kernels admissible.

5.5. Conclusion

In this chapter, we follow the approach of He and Modest (1995) and Luttmer (1996), who showed how to adapt the the Hansen-Jagannathan volatility bounds to economies with the kinked budget constraints that arise in the presence of proportional transaction costs, to assess whether allowing for transaction costs can help solve the coskewness puzzle. Since these bounds do not depend on a particular model for the stochastic discount factor, but only on the form of the budget constraint, this provides a robust way to quantify the extent to which market frictions affect inferences about important features of asset pricing models.

Like in the study of Dittmar (2002), we consider nonlinear pricing kernels that can be seen as providing a link between nonparametric and parametric approaches to describing cross sectional variation in equity returns. The common element in these pricing kernels and those of nonparametric models is nonlinearity in priced risk factors. In contrast to these nonparametric approaches, and in common with parametric approaches, the pricing kernels are defined over an endogenous risk factor, and preference restrictions govern the sign of the relationship between returns and the terms in the pricing kernel. The risk factor is the return on aggregate wealth, and the nonlinearity arises from an expansion of a representative investor’s Euler equations for portfolio and consumption choice. Adding the additional assumption that the agent’s preferences exhibit decreasing absolute prudence allows us to restrict the sign of the first three terms of this expansion. We show that this framework is consistent with a setting in which agents are averse to kurtosis, and consequently asset returns are affected by covariance, coskewness, and cokurtosis with the return on aggregate wealth.

Tests of the model show that incorporating nonlinearity substantially improves upon the pricing kernel's ability to describe the cross section of returns. In particular, when human capital is incorporated into the measure of aggregate wealth, a cubic pricing kernel is able to fit the cross section of industry-sorted portfolio returns. Moreover, the conditioning information set is the crucial factor to the effect of determining the volatility of pricing kernel. Although the information set cay_t and the information set z_t generate similar admissible cubic pricing kernels, they perform differently under the HJ bounds framework. The success of former depends on the inclusion of human capital, while for the later it does not.

The main finding is that the conditional quadratic pricing kernel, given by the quadratic polynomial approximation of the representative investor's IMRS that arises in the 3M-CAPM, dominates the competing kernels under both the sharpened HJ bound and the sharpened HJ distance and it is admissible under relatively high, though not unrealistically so, proportional transaction costs. The pricing errors, as summarized by the HJ distance, are not significant for the conditional quadratic pricing model for a quarterly transaction cost rate as low as 2 percent when the risk factor is the return on the stock market portfolio. This break-even rate increases to 3 percent per quarter when the risk factor is the growth rate of aggregate consumption In contrast, if the market is assumed frictionless (i.e., under the no transaction cost assumption), the volatility of the admissible nonparametric pricing kernels is so high that cannot be matched even by any of the models we consider. This suggests that market friction consisting of proportional transaction costs can explain, at least in part, the coskewness puzzle. .

Table 5-1
Statistic Summary: conditioning factors

variables	mean	std.	Correlation coefficients $r_{m,t+1}$	$r_{c,t+1}$	cay_t	r_{mt}	dy_t	ys_t	dp_t
$r_{m,t+1}$	0.030	0.083	1						
$r_{c,t+1}$	0.005	0.005	0.128	1					
cay_t	0.000	0.013	0.023	-0.035	1				
r_{mt}	0.017	0.084	0.045	0.230	0.300	1			
dy_t	0.044	0.020	0.144	-0.194	0.345	0.076	1		
ys_t	0.016	0.007	0.102	-0.026	-0.080	0.109	0.486	1	
dp_t	0.008	0.019	-0.112	0.221	0.044	-0.057	-0.221	-0.315	1

Note. Quarterly means and standard deviations of market return $r_{m,t+1}$, consumption growth $r_{c,t+1}$, and conditioning variable cay_t and Z_t are reported. And the correlation coefficients among all variables are provided as well. The market return and consumptioin return are used as the proxy of investors' wealth growth rate. The set of information variables includes cay_t and z_t, where z_t is defined as $\{1, r_{mt}, dy_t, ys_t, dp_t\}$. The r_{mt} represents the lag market excess return, dy_t is the aggregate dividend yield, ys_t measures the yield term spread between three-month treasury bill return and one-month treasury bill return, and dp_t captures the industrial production growth. The data samples are in the period from 1954:Q3 to 2002:Q3.

Table 5-2
Predictive power test of the conditioning factors

Industry	χ^2_4	p-value	Industry	χ^2_4	p-value
1 Food Products	56.828	0.000	16 Transportation equipment	53.906	0.000
2 Beer & Liquor	22.460	0.000	17 Mines	18.010	0.000
3 Tobacco Products	8.753	0.068	18 Coal	19.449	0.000
4 Games Recreation	84.721	0.000	19 Oil	27.138	0.000
5 Books Publishing	78.158	0.000	20 Utilities	29.731	0.000
6 Hshld Consumer Goods	64.656	0.000	21 Communication	45.174	0.000
7 Cloths Apparel	37.544	0.000	22 Personal and Business Services	103.537	0.000
8 Healthcare	51.978	0.000	23 Business Equipment	108.226	0.000
9 Chemicals	76.353	0.000	24 Business Supplies	131.079	0.000
10 Textiles	52.241	0.000	25 Transportation	85.891	0.000
11 Construction	112.450	0.000	26 Wholesale	95.787	0.000
12 Steel Works Etc	44.530	0.000	27 Retail	101.969	0.000
13 Fabricated Products	160.892	0.000	28 Restaraunts, Hotels, Motels	58.510	0.000
14 Electrical Equipment	72.824	0.000	29 Finance	156.387	0.000
15 Automobiles and Trucks	55.995	0.000	30 Others	66.668	0.000

Note. This table reports a summary of the predictive power of the conditioning variables z_t used in this chapter. z_t is defined as $\{1, r_{mt}, dy_t, ys_t, dp_t\}$. The r_{mt} represents the market excess return, dy_t is the aggregate dividend yield, ys_t measures the yield term spread between three-month treasury bill return and one-month treasury bill return, and dp_t captures the industrial production growth. The data samples are in the period from 1954:Q3 to 2002:Q3. The predictive power test of z_t is accessed by the linear regression

$$R_{i,t+1} = b_0 + b_1 r_{mt} + b_2 dy_t + b_3 ys_t + b_4 dp_t + e_{t+1}.$$

The reported χ^2_4 and the p-values presents the results in terms of Newey and West Wald tests, which employs heteroskedasticity and autocorrelation-consistent covariance matrix in computing statistics. The null hypothesis is

$$H_0 : b_1 = b_2 = b_3 = b_4 = 0.$$

Table 5-3

GMM Estimation of Pricing Kernels (Conditioing Variables: cay_t)

Human Capital Excluded

Parameters ($\widehat{\theta}$)	$\overline{a}_t$	$\overline{b}_{1,t}$	$\overline{b}_{2,t}$	$\overline{b}_{3,t}$	$TJ_{T-Stat.}$
			Panel A: Conditional Linear Pricing Kernel		
Estimate	1.56	-3.07			75.25
p-value	(0.000)	(0.000)			(0.089)
			Panel B: Conditional Quadratic Pricing Kernel		
Estimate	1.48	-2.82	5.44		74.99
p-value	(0.000)	(0.002)	(0.572)		(0.078)
			Panel C: Conditional Cubic Pricing Kernel		
Estimate	1.51	-2.41	4.37	-43.86	74.34
p-value	(0.000)	(0.177)	(0.700)	(0.502)	(0.062)

Note. GMM estimation is employed to estimate the coefficients in these various specifications of pricing kernel. The conditioning information cay_t captures the time-varying coefficients in conditional pricing kernels. The moment conditions here are

$$g_T(\widehat{\theta}) = E[(m_{t+1}(\widehat{\theta})R_{t+1} - \tau) \otimes cay_t],$$

where different forms of $m_{t+1}(\widehat{\theta})$ imply different pricing models. And the scaled factor S^{-1} for minimizing the qudratic GMM criterion is the optimal weighting matrix $S = E[u(\widehat{\theta})u'(\widehat{\theta})]$, where $u(\widehat{\theta}) = (m_{t+1}(\widehat{\theta})R_{t+1} - \tau) \otimes cay_t$. The p-values for each individual coefficient are based on the Wald χ^2 test, and the p-value for $J_{T-Stat.}$ is based on χ^2 over-identifying restrictions. The data samples are in the period from 1954:Q3 to 2002:Q3. Moreover, the human capital $r_{c,t+1}$ is excluded from the measure of aggregate wealth.

Table 5-4

GMM Estimation of Pricing Kernels (Conditioing Variables: cay_t)

Human Capital Included

Parameters ($\widehat{\theta}$)	$\overline{a}_t$	$\overline{b^m}_{1,t}$	$\overline{b^c}_{1,t}$	$\overline{b^m}_{2,t}$	$\overline{b^c}_{2,t}$	$\overline{b^m}_{3,t}$	$\overline{b^c}_{3,t}$	$TJ_{T-Stat.}$
Panel A: Conditional Linear Pricing Kernel								
Estimate	1.79	-3.47	-20.55					74.81
p-value	(0.000)	(0.000)	(0.020)					(0.060)
Panel B: Conditional Quadratic Pricing Kernel								
Estimate	1.21	-4.41	-15.35	4.00	5,906.70			74.59
p-value	(0.000)	(0.000)	(0.691)	(0.726)	(0.062)			(0.059)
Panel C: Conditional Cubic Pricing Kernel								
Estimate	0.82	-4.21	-10.55	4.04	14,774.21	-40.92	-35,992.59	73.88
p-value	(0.000)	(0.054)	(0.842)	(0.785)	(0.049)	(0.587)	(0.002)	(0.133)

Note. GMM estimation is employed to estimate the coefficients in these various specifications of pricing kernel. The conditioning information cay_t captures the time-varying coefficients in conditional pricing kernels. The moment conditions here are

$$g_T(\widehat{\theta}) = E[(m_{t+1}(\widehat{\theta})R_{t+1} - \tau) \otimes cay_t],$$

where different forms of $m_{t+1}(\widehat{\theta})$ imply different pricing models. And the scaled factor S^{-1} for minimizing the qudratic GMM criterion is the optimal weighting matrix $S = E[u(\widehat{\theta})u'(\widehat{\theta})]$, where $u(\widehat{\theta}) = (m_{t+1}(\widehat{\theta})R_{t+1} - \tau) \otimes cay_t$. The p-values for each individual coefficient are based on the Wald χ^2 test, and the p-value for $J_{T-Stat.}$ is based on χ^2 over-identifying restrictions. The data samples are in the period from 1954:Q3 to 2002:Q3. Moreover, the human capital $r_{c,t+1}$ is included within the measure of aggregate wealth.

Table 5-5

HJ Distance Estimation of Pricing Kernels (Conditioing Variables: cay_t)

Parameters $(\widehat{\theta})$	$\overline{a}_t$	$\overline{b^m}_{1,t}$	$\overline{b^c}_{1,t}$	$\overline{b^m}_{2,t}$	$\overline{b^c}_{2,t}$	$\overline{b^m}_{3,t}$	$\overline{b^c}_{3,t}$	$HJ-Dist.$
Panel A: Conditional Linear Pricing Kernel (Human Capital Excluded)								
Estimate	1.14	-4.36						0.77
p-value	(0.000)	(0.000)						(0.000)
Panel B: Conditional Quadratic Pricing Kernel (Human Capital Excluded)								
Estimate	1.11	-4.93		6.41				0.77
p-value	(0.000)	(0.000)		(0.568)				(0.000)
Panel C: Conditional Cubic Pricing Kernel (Human Capital Excluded)								
Estimate	1.09	-3.93		4.20		-59.03		0.76
p-value	(0.000)	(0.035)		(0.714)		(0.377)		(0.002)
Panel D: Conditional Linear Pricing Kernel (Human Capital Included)								
Estimate	1.28	-5.10	-24.31					0.76
p-value	(0.000)	(0.000)	(0.038)					(0.000)
Panel E: Conditional Quadratic Pricing Kernel (Human Capital Included)								
Estimate	1.01	-5.73	-17.43	4.02	5,285.78			0.75
p-value	(0.000)	(0.000)	(0.629)	(0.729)	(0.047)			(0.001)
Panel F: Conditional Cubic Pricing Kernel (Human Capital Included)								
Estimate	0.79	-5.48	-12.54	4.29	13,175.35	-34.44	-490,248.86	0.74
p-value	(0.000)	(0.005)	(0.755)	(0.721)	(0.015)	(0.612)	(0.002)	(0.171)

Note. Hansen-Jagannathan Distance estimation is employed to estimate the coefficients in these various specifications of pricing kernel. The conditioning information cay_t captures the time-varying coefficients in conditional pricing kernels. The moment conditions here are

$$g_T(\widehat{\theta}) = E[(m_{t+1}(\widehat{\theta})R_{t+1} - \tau) \otimes cay_t],$$

where different forms of $m_{t+1}(\widehat{\theta})$ imply different pricing models. And the scaled factor S^{-1} for minimizing the qudratic GMM criterion is the Hansen-Jagannathan weighting matrix $S = E[(R_{t+1} \otimes cay_t)(R_{t+1} \otimes cay_t)']$. The p-values for each individual coefficient are based on the Wald χ^2 test, and the p-value for $HJ - Dist.$ is based on the asymptotic distribution of HJ Distance (Jagannthan & Wang, 1996). The data samples are in the period from 1954:Q3 to 2002:Q3.

Table 5-6

HJ Distance Estimation of Pricing Kernels (Conditioing Variables: z_t)

Parameters ($\widehat{\theta}$)	$\overline{a}_t$	$\overline{b^m}_{1,t}$	$\overline{b^c}_{1,t}$	$\overline{b^m}_{2,t}$	$\overline{b^c}_{2,t}$	$\overline{b^m}_{3,t}$	$\overline{b^c}_{3,t}$	$HJ-Dist.$
Panel A: Conditional Linear Pricing Kernel (Human Capital Excluded)								
Estimate	1.18	-4.39						0.77
p-value	(0.000)	(0.000)						(0.006)
Panel B: Conditional Quadratic Pricing Kernel (Human Capital Excluded)								
Estimate	1.11	-4.26		1.89				0.76
p-value	(0.000)	(0.000)		(0.849)				(0.011)
Panel C: Conditional Cubic Pricing Kernel (Human Capital Excluded)								
Estimate	1.09	-3.68		2.82		-23.95		0.75
p-value	(0.000)	(0.064)		(0.783)		(0.740)		(0.009)
Panel D: Conditional Linear Pricing Kernel (Human Capital Included)								
Estimate	1.15	-3.98	-0.51					0.77
p-value	(0.000)	(0.000)	(0.985)					(0.000)
Panel E: Conditional Quadratic Pricing Kernel (Human Capital Included)								
Estimate	1.14	-4.33	-13.93	0.20	1,058.32			0.75
p-value	(0.000)	(0.000)	(0.682)	(0.984)	(0.075)			(0.003)
Panel F: Conditional Cubic Pricing Kernel (Human Capital Included)								
Estimate	1.19	-4.27	-22.95	1.22	2,274.54	-40.92	-120,557.95	0.64
p-value	(0.000)	(0.034)	(0.588)	(0.907)	(0.060)	(0.588)	(0.006)	(0.233)

Note. Hansen-Jagannathan Distance estimation is employed to estimate the coefficients in these various specifications of pricing kernel. The conditioning information z_t captures the time-varying coefficients in conditional pricing kernels. The moment conditions here are

$$g_T(\widehat{\theta}) = E[(m_{t+1}(\widehat{\theta})R_{t+1} - \tau) \otimes z_t],$$

where different forms of $m_{t+1}(\widehat{\theta})$ imply different pricing models. And the scaled factor S^{-1} for minimizing the qudratic GMM criterion is the Hansen-Jagannathan weighting matrix $S = E[(R_{t+1} \otimes z_t)(R_{t+1} \otimes z_t)']$. The p-values for each individual coefficient are based on the Wald χ^2 test, and the p-value for $HJ - Dist.$ is based on the asymptotic distribution of HJ Distance (Jagannthan & Wang, 1996). The data samples are in the period from 1954:Q3 to 2002:Q3.

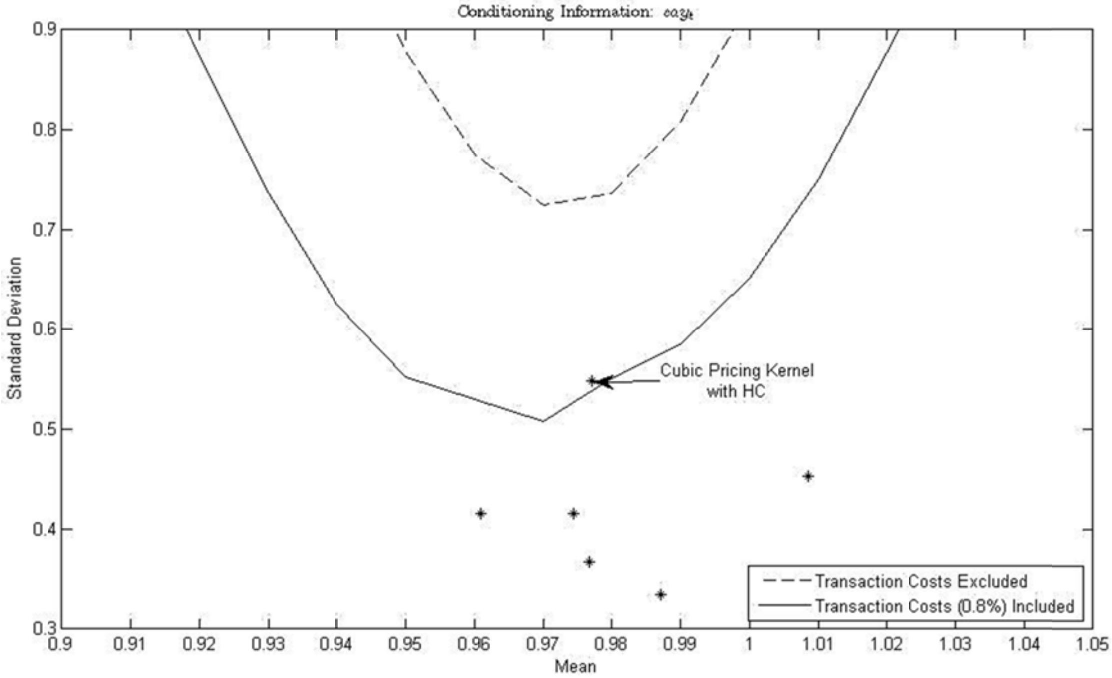
Table 5-7
Mean and Standard Deviation of Conditional Pricing Kernels

Panel A: Conditioning Information Set: cay_t			
Model	Mean	Std.	Dist.
	Human Capital Excluded		
Linear:	0.9872	0.3345	0.7735
Quadratic:	0.9767	0.3678	0.7663
Cubic:	0.9609	0.4148	0.7602
	Human Capital Included		
Linear:	0.9745	0.4149	0.7641
Quadratic:	1.0086	0.4532	0.7531
Cubic:	0.9770	0.5482	0.7405
Panel B: Conditioning Information Set: z_t			
Model	Mean	Std.	Dist.
	Human Capital Excluded		
Linear:	1.0498	0.3470	0.7683
Quadratic:	0.9980	0.3510	0.7556
Cubic:	0.9910	0.3530	0.7512
	Human Capital Included		
Linear:	1.0302	0.3293	0.7658
Quadratic:	0.9890	0.3588	0.7650
Cubic:	0.9999	0.5557	0.6427

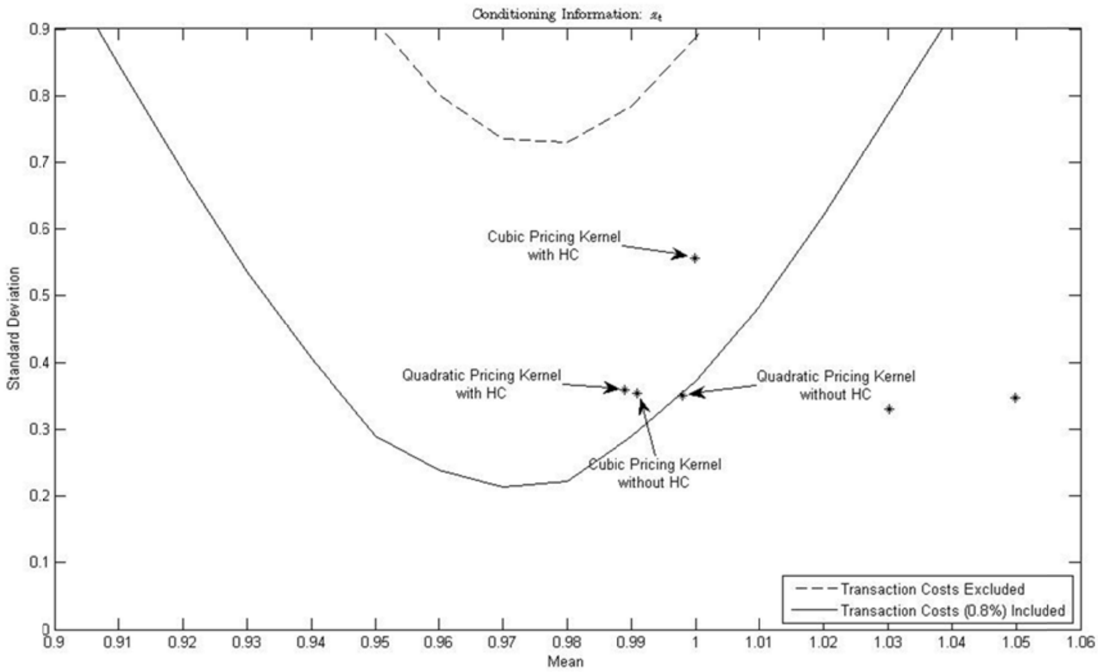
Note. This table presents the mean and the standard deviation of the pricing kernels estimated via the Hansen-Jagannathan Distance approach. The column labeled "Mean" represents the quarterly average of the estimated pricing kernel m_{t+1}, the column labeled "Std." represents its quarterly standard deviation, and the column labeled "Distance" represents the HansenCJagannathan distance. The rows labeled under "Human Capital Excluded" represent the decomposition for the polynomial pricing kernels, omitting human capital. The rows labeled under "Human Capital Excluded" represent the decomposition for the polynomial pricing kernels, including human capital. In Panel A and B, the coefficients of the pricing kernel are modeled with two conditioning variables cay_t and z_t, respectively, and the sign restrictions for each coefficient are imposed. The data samples are in the period from 1954:Q3 to 2002:Q3.

Figure 5-1
Conditional Hansen-Jagannathan Bounds
with Transaction Costs

Panel A



Panel B



6. Fund Evaluation

6.1. Introduction

The evaluation of fund performance has attracted lots of attention from both practitioners and scholars. The essential concern of this part of the thesis, as in Ahn and Shivdasani (1999), is to provide an empirical assessment of fund performance in the Chinese market. Although research on fund performance evaluation has seen countless contributions over the years, applications are somewhat limited geographically (Leite & Cortez, 2009). This is especially the case of Chinese funds. The reason may be traced to the lesser development of the Chinese security exchange market[34]. The lack of an efficient regulatory framework during the early stage after the birth of the Chinese stock exchange inevitably discouraged the emerging of open-end funds during 1990s. But since 2000, the open-end mutual fund industry has become one of the fastest growing segments of the financial sector in China. The first fund, called Hua'an Innovation, was introduced with the assistance of J.P. Morgan in September, 2001. According to the WIND financial Database[35], the biggest financial data service provider in China, by the end of 2011, the number of listed open-end funds had risen to 992, and more than 400 mutual funds are heavily invested in equities.

[34] The Shanghai stock exchange rebegan to operate on December 19th, 1990, after the founding of the People's Republic of China. And the Shenzhen stock exchange is founded on December 1st, 1990.
[35] Please check http://www.wind.com.cn/En/Default.aspx for more details.

In analysis of the trading practices in China, Kang *et al.* (2002) pointed out that the trading decisions of most individual investors in the Chinese market are mainly driven by sentiment and market rumours. Then the question is whether fund managers in the Chinese market construct their portfolio more efficiently than the ordinary individual investors. Most studies of mutual funds rely on Jensen's alpha as a measures of risk adjusted performance. As noted by Grinblatt and Titman (1989), however, Jensen's alpha may assigns erroneously negative performance to a market timer. The reason is, as shwon by Grant (1977), that the estimate of the CAPM beta coefficient is biased upwards, and therefore the estimation of performance is downward-biased. In order to cope with this weakness, recent studies of mutual fund returns have moved beyond performance measures based on Jensen's alpha derived from the ordinary CAPM. As emphasized by Lee and Rahman (1990), the basis for a mutual fund manager to generate superior performance consists of two components: microforecasting (security analysis ability) and macroforecasting (market timing ability). In their paper, the authors allow for the risk coefficient to be time-varying in order to reflect market-timing ability. In this chapter, we apply the approach of Chen and Knez (1996) to analyse the open-end fund performance in the Chinese market. Following Dahlquist and Söderlind (1999), we examine the hypothesis that the open-end fund managers in Chinese market process information more professionally than individual investors to generate significant superior returns and also try to establish whether their performance could be replicated by employing some simple strategies that make use of public information. This analysis makes several contributions to the fund performance literature. Firstly, we extend our understanding of the trade-off between risk and return to the Chinese open-end fund industry. Instead of using, as in many studies, the the Sharpe ratio[36], our approach echoes the newly emerging literature (see Dahlquist and Söderlind (1999), Ahn *et al.* (2003) and Ferson *et al.* (2006)) and defines performance relative to a pre-specified SDF. As discussed by Dahlquist and Söderlind (1999), given an pre-derived SDF, a measure of abnormal return (SDF alpha) is easily obtained. Secondly, by adding conditioning information, we test the fund performance conditional on ex-ante economic states (Ferson *et al.*, 2006). In fact, due to the short history of the open-end fund in Chinese market, the conditional performance of this industry is still largely unexplored. A few studies challenge the pertinence of the use of conditioning information in fund performance evaluation outside the US market (see, Bauer *et al.* (2006), Blake *et al.* (2002), Otten and Bams (2002), and Sawicki and Ong (2000)), then the question is whether this argument applies to the Chinese market as well.

6.2. Theoretical Framework

For the identification of superior return, Chen and Knez (1996) require two conditions: a) the existence of a reference portfolio set x_o, which is available to both fund managers and retail investors, b) the existence of an admissible pricing kernel m^+ which assigns an identical performance measure for portfolios constructed through simple replication strategies from the reference portfolio set by using public

[36] Chen and Knez (1996) proposed four basic conditions for the legitimacy of evaluation measure. In fact, the Sharpe measure, and the RAROC (risk adjusted rate of return) index violate at least one of these conditions. For instance, the RAROC does not satisfy the condition II in Chen and Knez (1996): the measure function λ must be linear.

information. If the law of one price (LOP) holds, the price for a portfolio equals the value-weighted price of each individual stock. Mathematically,

$$p(\sum_{i=1}^{N}\alpha_i x_i) = \sum_{i=1}^{N}\alpha_i p(x_i) . \tag{6-1}$$

where $p(\cdot)$ denotes the linear pricing function which is a projection from payoff space to price space $p: x_o \to \mathbb{R}$, α_i represents the investment weight of each individual stock, and x_i is the payoff of each individual stock, i.e., $x_i \in x_o$. If we define the future return on the asset with payoff x_i as $R_i = x_i / p(x_i)$, where R_i forms the return space R_o, then for any two R_i and R_j, we have,

$$p(R_i) - p(R_j) = p[\frac{x_i}{p(x_i)}] - p[\frac{x_j}{p(x_j)}] = 0 . \tag{6-2}$$

For a fund manager, whose performance R^* lies inside the reference return space R_o, the assessment of superior investment ability is zero, i.e. $p(R^* - R_o) = 0$. If the generated performance R^* lies inside the managed superior return space R_s, which indicates the existence of $R^*(\omega) - R_0(\omega) > 0$ for at least one state of the world ω, then under the assumption of NA, $p(R^* - R_o) > 0$. In fact, Chen and Knez (1996) point out that the only way for fund managers to achieve superior performance is to exploit superior information. Chen and Knez (1996) define the superior return affine hull as

$$R_s \equiv \{R_s = \frac{x_s}{p(x_s)} \mid x_s = \sum_{i \in N} w_i(s) x_i \text{ and } \sum_{i \in N} w_i(s) = 1\} . \tag{6-3}$$

Here, the fund manager can generate dynamic trading strategies $w(s)$ by using the superior information set s. By contrast, the reference return affine hull is

$$R_o \equiv \{R_o = \frac{x_o}{p(x_o)} \mid x_o = \sum_{i \in N} w_i(o) x_i \text{ and } \sum_{i \in N} w_i(o) = 1\}, \tag{6-4}$$

where the information set o is publicly available to the retail investors. Inevitably, some fund managers may only follow a passive constant strategy or the dynamic trading strategies on the basis of public information. In this case, the superior return space R_s should contain the reference return affine hull R_o and the passive return affine hull R_c. More precisely, $R_c \subseteq R_o \subseteq R_s$ where R_c is the constant composition return by a passive strategy. Chen and Knez (1996) propose a generalized performance measure m^+ which satisfies

$$E(m^+ R_o) = 1 \qquad m^+ > 0, \tag{6-5}$$

According to (6-5), we have:

$$\begin{aligned} E[m^+(R_s - R_o)] \geq 0 \qquad & \text{if } (R_s - R_0) \geq 0 \\ \text{and } E[m^+(R_s - R_o)] > 0 \qquad & \text{if } (R_s - R_0) > 0 \end{aligned}. \tag{6-6}$$

6.3. Conditional Performance Evaluation

Traditional or unconditional alphas compare returns and risks measured as averages over an evaluation period, and these averages are taken "unconditionally" or without regard to variations in the state of financial markets or the broader economy. In the conditional performance evaluation approach, the state of the economy is measured using predetermined, public information variables. This takes the view that a managed portfolio strategy that can be replicated using readily available public information should not be judged as having superior performance.

Similarly, when the fund manager predicts the price movement, ideally he will adjust his portfolio in an attempt to fully explore beneficial opportunities. Literally, if the market has a high probability to generate positive (negative) return over the next investment period, the manager will reconstruct his portfolio by including more (less) high-risk stocks and less (more) low-risk stocks. The analysis of mean-variance efficient sets with respect to conditioning information is developed by Hansen and Richard (1987). In their paper, they state that unconditional efficiency can be treated as a special case of conditionally efficiency, but not the converse. In fact, a number of papers have since extended Jensen's (1968) work by including conditioning information to evaluate the timing ability. Grinblatt and Titman (1989) suggest that the excess return of the manager's portfolio should be expressed as

$$r_{p,t+1} = J\alpha_t + \beta_{p,t+1} r_{m,t+1} + \varepsilon_{p,t+1}, \qquad \text{(6-7)}$$

where $J\alpha$ denotes the Jansen's alpha. Ferson and Harvey (1999) provide an empirical study on the ability of the Fama-French model to capture common dynamic patterns in returns by using a set of lagged, economy-wide predictor variables. The betas in the asset pricing model are allowed to be time-varying and depend linearly on the predetermined instruments,

$$\begin{aligned} r_{p,t+1} &= J\alpha_{p,t} + b'_{p,t+1} f_{t+1} + \varepsilon_{p,t+1} \\ &= J\alpha_{p,t} + (b_{0,p} + b'_{1,p} z_t) f_{t+1} + \varepsilon_{p,t+1} \end{aligned}, \qquad \text{(6-8)}$$

Here, f_{t+1} is a vector of excess returns on the risk factor-mimicking portfolios over the risk-free rate $R_{f,t}$, and $f_{t+1} + R_{f,t} \subseteq R_o$. Assume $R_{p,t+1} = R_{f,t} + r_{p,t+1}$ is in the affine hull R_o, then $J\alpha_{p,t}$ is statistically insignificant. Otherwise, $R_{p,t+1}$ outperforms the basis assets R_o and lies inside the superior return affine hull R_s, if $J\alpha_{p,t}$ is significant positive in measure. According to Ferson and Siegel (2009), both of the above cases (6-7) and (6-8) can be represented conditionally, in a similar way as in (6-6), In fact,

$$\begin{aligned} E_t(m_{t+1}^{+} r_{p,t+1}) &= E_t(m_{t+1}^{+} J\alpha_{p,t}) + E_t[m_{t+1}^{+}(\beta_{p,t+1} f_{p,t+1} + \varepsilon_{p,t+1})] \\ \frac{1}{R_{f,t+1}} J\alpha_{p,t} &= E_t[m_{t+1}^{+}(R_{p,t+1} - R_{o,t+1})] \\ J\alpha_{p,t} &= R_{f,t+1} E_t[m_{t+1}^{+}(R_{p,t+1} - R_{o,t+1})] \end{aligned} \quad , \tag{6-9}$$

or equivalently, we could write (6-9) as,

$$J\alpha_{p,t} = R_{f,t+1} E_t(m_{t+1}^{+} R_{p,t+1}) - R_{f,t+1} \qquad \forall R_{p,t+1} \in R_s , \tag{6-10}$$

Therefore, a natural hypothesis to test is whether $J\alpha_{p,t}$ is equal to zero.

Through equation (6-3), we know that an investment manager forms a portfolio of the primitive assets with gross return $R_{p,t+1} = \sum_{i \in N} w_i(s) R_{i,t+1}$, here, $\sum_{i \in N} w_i(s) = 1$ and $R_{i,t+1} \in R_o$. To see how a manager with superior information can generate alpha, substitute portfolio $R_{p,t+1}$ into (6-10) and use the definition of covariance and equation (6-5) to obtain

$$\begin{aligned} J\alpha_t &= R_{f,t+1} E_t(m_{t+1}^{+} R_{i,t+1}) E_t[w_i(s)] + R_{f,t+1} cov_t[m_{t+1}^{+} R_{i,t+1}, w_i(s)] - R_{f,t+1} \\ &= R_{f,t+1} cov_t[m_{t+1}^{+} R_{i,t+1}, w_i(s)] \end{aligned} \quad . \tag{6-11}$$

If the trading strategy $w_i(s)$ is based on the public information set o, the trading strategy is predetermined at time t and $J\alpha_t$ is zero. Otherwise, if the fund manager has the access to the superior information set s, then an abnormal return can be generated.

6.4. Estimation of Alpha

Similarly to Ahn *et al.* (2003), the method that we employ follows Chen and Knez (1996) in that, to measure risk-adjusted performance, we use a stochastic discount factor retrieved from a set of basis assets. It is important to note the differences between this method and other parametric approaches in assessing the abnormal performance of superior strategies employed by fund managers. Nonparametric performance measures attempt to recover a set of admissible stochastic discount factors based on minimal conditions such as the law of one price or no arbitrage conditions. Parametric approaches put forth instead a particular parametric pricing model, which is assumed to price all basis assets, and then test whether the strategy of interest outperforms the efficient portfolio relative to this model.

As discussed by Ahn *et al.* (2003), the nonparametric approach has several advantages. First, as mentioned above, estimating a stochastic discount factor from a set of basis assets imposes equilibrium pricing conditions without the need to specify a parametric benchmark. This can help avoid the problem that the success of superior performance might be only conditional on a particular parametric pricing model assumed to be true. Second, the estimation of a discount factor leads to natural measures of risk-adjusted abnormal performance. If the funds considered outperform when measured in this manner, then it is more likely that their performance is due to investor irrationality. However, if these funds cannot outperform the benchmark, their success may be consistent with rational asset pricing. Third, the nonparametric measures I use can be easily extended to conditional measures which incorporate the possibility that risk premiums are time varying.

Hansen and Jagannathan (1991) investigate how to retrieve the stochastic discount factors m_{t+1} from a given set of tradable, or basis, assets. The key underlying assumption therein is that there is no pricing inconsistency among the basis assets: that is, the SDFs are admissible. Hansen and Jagannathan (1991) suggest two particular solutions for m_{t+1} which are the minimum-norm discount factors defined in different metrics. The first is defined as the m_{t+1} that is in the space R_o. That is,

$$m_{t+1} = \theta^T R_{i,t+1} \quad \forall R_{i,t+1} \in R_o . \tag{6-12}$$

Here $R_{i,t+1}$ is the basis assets in R_o, and θ usually equals $\tau' E_t(R_{t+1}R'_{t+1})^{-1}$. We follow Chen and Knez (1996) and term this solution for the SDF the law of one price (LOP) discount factor, since its existence necessitates only that the law of one price holds. The second SDF satisfies equation (6-5), and the further requirement that m_{t+1} is strictly positive,

$$m^+_{t+1} = \max[\theta^T R_{i,t+1}, 0] \quad \forall R_{i,t+1} \in R_o . \tag{6-13}$$

This SDF satisfies the stronger condition of no arbitrage, thereby ruling out investment opportunities with positive payoffs and non-positive prices. We refer to this discount factor as the no-arbitrage SDF.

The method for estimating $J\alpha_t$ in a GMM setting is taken directly from Farnsworth *et al.* (2002a), and it is based on the system of moment conditions listed below,

$$\begin{aligned} g_o(\theta) &= E[m^+(\theta)R_{o,t+1} \otimes z_t] - \tau \\ g_s(\theta) &= J\alpha_t = R_{f,t+1}E[m^+(\theta)R_{s,t+1}] - R_{f,t+1} \end{aligned} . \tag{6-14}$$

The parameters can be estimated by minimizing the criterion function $G(\theta)$,

$$\theta = \arg\min_{\theta} \; G(\theta)$$
$$= \arg\min_{\theta} \; [g_o'(\theta), g_s(\theta)] W \begin{bmatrix} g_o(\theta) \\ g_s(\theta) \end{bmatrix}. \tag{6-15}$$

where the weighting matrix W is,

$$W = E\{\begin{bmatrix} g_o(\theta) \\ g_s(\theta) \end{bmatrix} [g_o'(\theta), g_s(\theta)]\}^{-1}. \tag{6-16}$$

As shown by Hansen (1982), under the null of no model mispecification $TG(\theta) \sim \chi^2_{n-k}$, where T is the number of the observation in the sample period, n is the number of moment conditions and k is the number of parameters. Each of the n assets is associated with one moment condition. We incorporate the return on a risk-free asset among the test asset payoffs, since Dahlquist and Söderlind (1999) emphasize the importance of doing this in order to fix the mean of the stochastic discount factor at a reasonable level. The fund performance return adds an additional moment condition so that $TG(\theta) \sim \chi^2_1$.

6.5. Data

All the data we use here are provided by WIND[37], which is the most famous and reliable financial data provider in the Chinese market. Due to the short period during which the open-end fund industry in China has been in existence, we only focus on the relatively high weekly frequency in order to have a sufficiency large number of observations in our sample period. The latter runs from January 1, 2006 to December

[37] The official website of WIND is: http://www.wind.com.cn/En/Default.aspx

31, 2011, for a total of 306 observations. We choose the fund performances and the basis portfolios as follows.

6.5.1 The Fund Performance

We only focus on the funds whose main investment holdings are stocks, and meanwhile we exclude the funds with passive trading strategies, such as the ETFs, and index-tracking funds. In fact, the trading purpose of the ETFs or the index-tracking funds is to replicate the targeted index. The fund managers follow a predetermined passive strategy, and rebalance the investment portfolio according to the change in the index components. It seems relatively pointless to include such funds in an assessment of investment ability of active fund managers. After these filters are applied, we are left with only 12 open-end funds. Summay statistics on these funds are given in Table 6-1.

6.5.2 The Basis Portfolios

The choice of the basis portfolios effectively determines the measure of abnormal performance. In complete markets, the SDF is unique. However, when markets are incomplete, there exists a multiplicity of SDFs that correctly price the assets in the economy (Harrison & Kreps, 1979). If the reference set from which the SDF is formed spans the payoff opportunity set which is available to investors, then measuring abnormal performance relative to this reference set will provide a correct (and unique) inference. However, if the reference set does not span the payoffs, it is possible to incorrectly reject the null hypothesis of zero abnormal performance (Ahn & Shivdasani, 1999).

Ideally, to prevent an incorrect rejection of the null, the reference assets should mimic the entire opportunity set from which the trading strategies are chosen. However, this approach is not implementable in practice. Therefore we must choose a more parsimonious set of reference assets that capture as much of the investment opportunity set as possible; that is, we wish to group securities in a manner that maximizes intragroup correlation and minimizes intergroup correlation. King (1966) demonstrates that industry groupings do precisely what we need. In an exhaustive analysis of factors important in the determination of stock returns, he concludes that market and industry factors capture most, if not all, of the common variation in stock returns. For example, he demonstrates that “large” positive covariance in returns cluster strongly within industry groupings, and negative covariances are observed exclusively across industry groupings. Therefore we form the reference set by forming portfolios on the basis of industry. In this part, we choose Shen & Wan 23 Industry portfolio groups, which are published everyday and treated as the standard industry portfolios in Chinese market.

6.6. Empirical Findings

6.6.1 Unconditional Performance Measures

We examine unconditional performance based on no-arbitrage first. The purpose of the unconditional estimation is to investigate whether the fund manager is able to achieve superior performance compared with the passive strategy. In this case, we are testing the fund performance $R_{i,t+1} \in R_c$. As discussed previously, the basis assets used for the tests consist of the 23 industry-sorted portfolios plus the return on the riskless asset. Results of this test are shown in Table 4-1.

The results in Table 6-1 suggest that the abnormal performances of all the listed funds are trivial when the unconditional performance measure is used, as none of the $J\alpha_t$ is significant at any conventional level. These results suggest that the fund managers of the selected funds add no superior performance to their managed investment.

6.6.2 Conditional Performance Measures

In order to further assess performance, we now apply the conditional measure used by Ahn *et al.* (2003) and check whether our results are affected if we allow investors' expectations to vary conditional on three publicly known information variables: the one-week lag market excess return over risk-free rate, the return on the Treasury bill with maturity closest to one month, and the term spread, measured as the difference in yields on three-year maturity Treasury bonds over one-year maturity Treasury bills, i.e., $Z_t = \{rm_t,\ tb_t,\ ys_t\}$. These results are presented in Table 6-3. All the funds appear to display significant negative performances. These results, combined with those in Table 6-1, show that the fund performances $R_{i,t+1}$ only belong to R_c, but are outside R_o and even R_s. These results suggest that the open-end fund purchaser in Chinese market is largely wasting money by buying these funds, since their performance can be simply replicated by the individual investor through passive strategies.

6.7. Conclusion

In this part, we followed the approach of Chen and Knez (1996), and in a similar spirit to Dahlquist and Söderlind (1999), to analyse the open-end fund performance in the Chinese market. The results of our analysis show that the fund managers of active open-ended funds in fact add no superior performance to the managed portfolios. Even the naive buy-and-hold trading strategy is able to replicate their performance easily. Then the question is why we assist to the phenomenon of the rapid expansion of the open-fund industry in spite of the poor fund performance. Tight restrictions on admissible investment policies apply to all fund managers in China. Frequent rebalancing, short selling and leverage trading are banned by the CSRC (China Securities Regulatory Commission) for the open-end funds. These restrictiction might help explain the poor performce of the mutual fund industry. Other possible explanations include high transaction costs, and in this respect a comparison with non active mutual funds would be useful, and investors inexperience and limited monitoring ability. We leave the investigation of these possibilities for future research.

Table 6-1

Statistical Summary of the Fund Performances

Quotation Codes	Mean	Stdev	Quotation Codes	Mean	Stdev
160106.OF	0.456	4.570 percent	161706.OF	0.495	4.220 percent
160211.OF	0.604	3.892 percent	161810.OF	0.440	3.988 percent
160314.OF	0.542	3.435 percent	161903.OF	0.181	5.008 percent
160505.OF	0.517	3.757 percent	162006.OF	0.496	3.825 percent
161607.OF	0.364	4.466 percent	162605.OF	0.445	4.581 percent
161610.OF	0.430	4.044 percent	162703.OF	0.502	4.496 percent

Notes. This table presents weekly means and standard deviation of the returns to the 12 target funds. In fact, only the funds whose main investment holdings are stocks are included, and meanwhile the funds with passive trading strategies, such as the ETFs, and index-tracing funds are excluded. The sample period is from January 1, 2006 to December 31, 2011. The ".OF" in quotation codes stands for "open-end fund".

Table 6-2

NA-based Performance Measures (Unconditional)

Quotation Codes	$J\alpha_t$	*p-value*	Quotation Codes	$J\alpha_t$	*p-value*
160106.OF	0.0011	0.9009	161706.OF	0.0017	0.8231
160211.OF	0.0032	0.6359	161810.OF	0.0016	0.8127
160314.OF	0.0037	0.5496	161903.OF	0.0001	0.9906
160505.OF	0.0033	0.6160	162006.OF	0.0034	0.6140
161607.OF	0.0012	0.8806	162605.OF	0.0009	0.9141
161610.OF	0.0024	0.7434	162703.OF	0.0022	0.7762

Notes. This table presents results from the unconditional no arbitrage estimation of performance measures. $J\alpha_t$ represents average weekly excess performance over the portfolio of basis assets for the passive strategy. The *p*-value represent a chi-squared test of the hypothesis $H_0 : J\alpha_t = 0$. The basis assets in this sample consist of 23 industry sorted portfolios plus the risk-free rate.

Table 6-3

NA-based Performance Measures (Conditional)

Quotation Codes	$J\alpha_t$	*p-value*	Quotation Codes	$J\alpha_t$	*p-value*
160106.OF	-0.0336	0.0077	161706.OF	-0.0322	0.0086
160211.OF	-0.0309	0.0076	161810.OF	-0.0322	0.0087
160314.OF	-0.0305	0.0051	161903.OF	-0.0336	0.0077
160505.OF	-0.0319	0.0087	162006.OF	-0.0301	0.0051
161607.OF	-0.0338	0.0072	162605.OF	-0.0337	0.0076
161610.OF	-0.0308	0.0060	162703.OF	-0.0329	0.0084

Notes. This table presents results from the conditional no arbitrage estimation of performance measures. $J\alpha_t$ represents average weekly excess performance over the portfolio of basis assets for the active strategies. The *p*-value represent a chi-squared test of the hypothesis $H_0 : J\alpha_t = 0$. The basis assets in this sample consist of 23 industry sorted portfolios plus the risk-free rate, augmented by managed portfolios based on the instrumental variables $Z_t = \{rm_t, tb_t, ys_t\}$.

7. Implication, Limitations and Future Work

7.1. Introduction

In this Chapter I summarize my main findings and discuss their implications for asset pricing. I start, in the next Section, by reviewing the main findings and the important analytical results. Section 7.3 discusses their implications. In Section 7.4, I then outline the main limitation of this study, I suggest possible extensions and I highlight opportunities for future research. The final Section presents some final remarks and draws together the main conclusions.

7.2. The Main Findings Restated

My thesis contributes to depict a representation of the multivariate distribution of stock returns where the relations between moments and their dynamics are important in explaining their cross-sectional differences. More innovatively, we explore whether the SDF implied by the 3 and 4-moment CAPM is plausible under restrictions that are weaker than those considered by Dittmar (2002) yet sufficient to rule out implausible curvature of the representative investor's utility functions. We find that, even under these weaker restrictions, the 3 and 4-moment CAPM cannot solve well known puzzles which plague the empirical performance of extant rational asset pricing models, even though the higher order terms do generate considerable additional explanatory power. In chapter 3, our findings confirm that the quadratic and cubic market factors help explain observed stock returns. They play an important role in the pricing of certain payoffs, including strategies characterized by relatively high SRs, such as those spanned by a fine industry-level diversification, most notably until the late 90s, or by dynamic portfolios managed on the basis of available conditioning information, as well as momentum portfolios. They do so, however, by generating high levels of SDF volatility. To rationalize this evidence within a higher moment CAPM framework, we would need to postulate implausibly high levels of investors' risk aversion. We conclude, therefore, that the 3M and 4M-CAPM provide at best a partial explanation of the differences in average returns on stocks and stock strategies. This gives rise to a coskewness (and cokurtosis) puzzle. The solution of the latter requires an explanation, different from the 3M and 4M-CAPM, for why the quadratic and cubic market factors are priced in the cross-section of stock returns.

Faced with this difficulty, we then explore whether the failure to fully account for cross-sectional differences in average returns can be explained by the presence of either transaction costs or a behavioural component of the SDF, reflecting investors' systematic mistakes in processing information. In chapter 4, we show that augmenting the (C)CAPM with sentiment, and thus allowing for systematic investor error in forming beliefs about the distribution of returns, permits to largely reconcile investors' optimizing behaviour with the cross-section of average returns. In fact, The Sentiment-(C)CAPM and Sentiment-3M(C)CAPM are empirically more successful, and most of the increase in the explanatory power is due to the inclusion of sentiment.

This implies that investors must either commit systematic errors, at least ex-post, in assessing the joint distribution of stock returns and aggregate consumption or they must behave in a way that, at the aggregate level, is inconsistent with expected utility maximization and with standard risk aversion assumptions.

In chapter 5, we compare several competing pricing kernels using a modified version of Hansen-Jagannathan distance (Hansen & Jagannathan, 1997), which not only accounts for the conditional information but also recognizes the existence of transaction costs. We follow the approach of He and Modest (1995) and Luttmer (1996), who show how the Hansen-Jagannathan volatility bounds can be derived for economies with the kinked budget constraints that arise from proportional transaction costs. Since these bounds do not depend on a particular model for the stochastic discount factor, but only on the form of the budget constraint, this provides a robust way to quantify the extent to which market frictions affect inferences about important features of asset pricing models. In fact, if the market is frictionless (i.e., there are no transaction costs), the volatility of the admissible nonparametric pricing kernels is so high that even the candidate kernels with nonlinear terms cannot match it, however, a number of the estimated pricing kernels reach the minimum volatility requirement when transaction costs are assumed to be 0.8 percent per quarter.

In chapter 6, we follow the approach of Chen and Knez (1996) to analyse the open-ended fund performance in the Chinese market. In a similar spirit to Dahlquist and Söderlind (1999), we examine the validity of the hypothesis that the open-ended fund managers in the Chinese market process information more efficiently than individual investors do, so as to generate significantly superior returns, and also try to observe whether their performance can be replicated by employing some simple strategies that make use of publicly available information. Our results show that the fund managers' performance is significantly negative and is outperformed even by a naive buy-and-hold trading strategy. While these findings supports the view that the stock market is efficient, they beg the question why there has been such a rapid expansion of the open-fund industry in spite of such a poor performance of the mutual funds.

7.3. Implications

The main implications of our results are in the field of asset pricing. The candidate pricing kernel specification in (3.1), estimated using the 30 Fama and French US Industry portfolios and the CRSP index as a proxy for the market portfolio, implies a 'coskewness puzzle'. The puzzle arises because, while the $b_{i,t}$ parameter estimates fit the cross section of industry returns relatively well, they imply risk seeking over gains and thus a non-concave utility function. Given the shape of the utility function implied by these estimates, the market portfolio is not necessarily efficient for the representative investor. In turn, if the market portfolio is inefficient, the 3M-CAPM and 4M-CAPM do not hold.

Pending the investigation into the theoretical explanation of these findings, the interesting question is then whether we should price assets based on expected returns that reflect a coskewness premium and, in particular, the large coskewness premium implied by the unrestricted quadratic SDF specification. It is clear that, as long as we do not have an equilibrium asset pricing model that can account for this large coskewness premium, we cannot strictly consider coskewness a risk measure. However, since $b_{2,t}$ and $b_{3,t}$ are the factor loadings of a multifactor model that explains the cross-section of industry returns relatively well, we might draw pricing implications for other assets based on no-arbitrage arguments.

Alternatively, in performance attribution, we may follow the approach of Chen and Knez (1996) adopted in Chapter 6, based on constructing a non parametric kernel from a set of traded asset payoffs.

7.4. Limitations of the Analysis and Future Work

Predictability and time varying risk premia likely reflects a premium for holding macroeconomic risk associated with the business cycle, for holding assets that do poorly in times of high volatility and financial distress and for holding assets that do poorly when the market portfolio is negatively skewed. Therefore, they seem to be closely related to the issue of asymmetry and thick tails in the multivariate return distribution and hence to asset coskewness. The exploration of the link between aggregate idiosyncratic risk, higher moments and asymmetries of the multivariate distribution of asset returns opens fascinating yet challenging possibilities for future research. For example, further research might suitably expand the set of conditioning variables to better model variation in the utility function parameters and might use a more meaningful proxy for the market portfolio of all risky assets. This, beside improving the fit of the model, might lead to a 3M-(C)CAPM specification with parameter estimates that do not violate RA and NIARA. I leave these developments, however, for future research.

Turning to the 'coskewness puzzle', its solution requires a theory that predicts a stochastic discount factor quadratic in the market return without implying that the market portfolio is efficient. We might appeal to Harrison and Krepp's (1979) theorem to motivate the stochastic discount factor representation of the asset pricing problem without requiring that the market portfolio maximizes investors' expected utility. Recall that this theorem states that, given free portfolio formation and under the law of one price, there exists an m_{t+1} such that, for every payoff x_{t+1}, $p_t = E_t(m_{t+1}x_{t+1})$. This approach, however, rises the problem of motivating why (2.74) specifies m_{t+1} as a function of the market excess return and its square. Alternatively, we might specify individual utility functions that exhibit DIARA and then determine equilibrium prices without imposing restrictive assumptions such as investors' homogeneity and market completeness or the equivalent representative investor assumption. The interesting question then becomes why the market return and its square should be good proxies for aggregate marginal utility growth even though the market portfolio is not necessarily efficient for the representative investor. I leave the investigation of these issues for further research.

7.5. Final Comments and Conclusion

This Chapter reviewed and summarized the main findings reported by this thesis and their implications for theoretical asset pricing. The unifying theme of this thesis is about asset pricing from a cross-sectional perspective. In the cross-section of average returns, assets with negative coskewness, and therefore with exposure to volatility risk, command a risk premium on top of the reward for market risk. This relation, as shown in Chapter 3, is empirically strong and bears puzzling implications for the shape of the stochastic discount factor, and thus for the possibility that prices are set by a representative investor. A deeper understanding of the relation between 'good deal' opportunities and the volatility of the SDFs contributes to the formulation of a richer investment advice and more meaningful fund evaluation system. From this perspective, Chapter 4 and 5 have explored whether the failure to fully account for cross-sectional differences in average returns can be explained by the presence of either transaction costs or a behavioural component of the SDF, reflecting investors' systematic mistakes in processing information. We find evidence of both problems, though our analysis is not conclusive in this respect. Finally, in a more applied exercise, we employ the SDF-framework to test whether Chinese fund managers generate superior investment performance, and find that Chinese fund managers generate significantly negative performance under both the unconditional and the conditional measure.

8. Appendix

8.1. Appendix A: Higher Co-moments in IMRS

Consider the local variation in a marginal utility function defined over wealth W_t, $U' = U'(W_t)$, given by a third order Taylor's expansion:

$$\begin{aligned} U'(W_{t+1}) &\cong U'(W_t) + U''(W_t)(W_{t+1} - W_t) + \frac{1}{2}U'''(W_t)(W_{t+1} - W_t)^2 + \cdots \\ &+ \frac{1}{6}U''''(W_t)(W_{t+1} - W_t)^3 \end{aligned} . \quad \text{(A-1)}$$

Thus,

$$\begin{aligned} \frac{U'(W_{t+1})}{U'(W_t)} &\cong 1 + \frac{U''(W_t)}{U'(W_t)}(W_{t+1} - W_t) + \cdots \\ &+ \frac{1}{2}\frac{U'''(W_t)}{U'(W_t)}(W_{t+1} - W_t)^2 + \frac{1}{6}\frac{U''''(W_t)}{U'(W_t)}(W_{t+1} - W_t)^3 \end{aligned} . \quad \text{(A-2)}$$

Let $r_{m,t+1} = \frac{W_{t+1} - W_t}{W_t}$, then for the admissible SDF or the representative investor's IMRS $E_t(m_{t+1}) = \gamma \frac{E_t(U'_{W_{t+1}})}{U_{W_t}}$, we will have

$$\begin{aligned} E_t(m_{t+1}) &\cong \gamma + \gamma\frac{U''(W_t)}{U'(W_t)}W_t E_t(r_{m,t+1}) + \cdots \\ &+ \frac{1}{2}\gamma\frac{U'''(W_t)}{U'(W_t)}W_t^2 E_t(r_{m,t+1}^2) + \frac{1}{6}\gamma\frac{U''''(W_t)}{U'(W_t)}W_t^3 E_t(r_{m,t+1}^3) \end{aligned} . \quad \text{(A-3)}$$

It is easy to have the following restrictions on the parameters of the candidate SDF by identification of (A-3):

$$a_t = \gamma \qquad \text{and} \qquad b_{i,t} \equiv \gamma \frac{U^{(i)}(W_{m,t})}{U'(W_{m,t})} W_{m,t}^i \quad i \in [1,2,3,4]. \quad \text{(A-4)}$$

Then the admissible SDF should be rewritten as

$$E_t(m_{t+1}) = a_t + \sum_{i=1}^{4} b_{i,t} E_t(r_{m,t+1}^i) \, . \tag{A-5}$$

In (A-4), $U^{(i)}$ denotes the *i*-th derivative of the utility function $U(W_{m,t})$. Arrow (1971) argues that desirable properties of this function are non satiation (NS), risk aversion (RA) and non-increasing absolute risk aversion (NIARA). NS implies $U' > 0$, RA implies decreasing MU, i.e. $U'' < 0$, whereas NIARA implies that the rate of decrease of MU does not increase in wealth and thus $U''' \geq 0$. Hence, NIARA implies aversion to negative skewness of the distribution of the return on wealth. Kimball (1993) adds non-increasing absolute prudence (NIAP) to the set of desirable properties of a 'well behaved' utility function. Under NIARA, NIAP implies $U'''' \leq 0$ and thus aversion to kurtosis. Together, NS, RA, NIARA and NIAP yield standard risk aversion, as defined by Kimball (1993). Preferences of rational expected utility maximizers will necessarily display the first of these properties, namely NS. This implies a no-arbitrage (NA) positivity restriction on the sign of the admissible SDFs, and hence $m_{t+1} > 0$, to avoid assigning a zero or negative price to strictly positive payoffs and thus to rule out unexploited arbitrage opportunities. NIAP, NIARA and RA rule out counter-intuitive behaviour and can be used to further restrict m_{t+1}, but they are not necessary conditions for the CAPM. When $U' > 0$, and thus under NS, NIAP implies $b_3 \leq 0$. Under NS, a necessary condition for $U''' \geq 0$, and thus for NIARA, is $b_2 \geq 0$. Finally, a necessary condition for $U'' < 0$, and thus for RA, is $b_1 < 0$. The latter condition guarantees local risk aversion, a milder requirement than concave utility (i.e., global risk aversion).

Since utility functions are equivalent up to a linear transformation, we might let $U(W_t)=0$ and $U'(W_t)=1$. This standardization is often very useful when working with utility functions in that it simplifies their manipulation. From (A-2), the third order Taylor expansion of this standardized marginal utility function around an initial level of wealth $W_t=1$ is therefore:

$$\begin{aligned} U'(W_{t+1}) &\cong 1+U''(W_t)(W_{t+1}-W_t)+\cdots \\ &+\frac{1}{2}U'''(W_t)(W_{t+1}-W_t)^2+\frac{1}{6}U''''(W_t)(W_{t+1}-W_t)^3 \\ &=1+\theta_1 r_{m,t+1}+\theta_2 r_{m,t+1}^2+\theta_3 r_{m,t+1}^3 \end{aligned} \quad \text{(A-6)}$$

According to the above discussion, we can easily find $\theta_1<0$, $\theta_2\geq 0$, and $\theta_3\leq 0$.

8.2. Appendix B: The Pratt-Arrow Risk Premium

In this Appendix, I provide a derivation of the equilibrium relation between expected variance and expected return. Define the simple gamble (an actuarially neutral gamble) as follows:

$$Z\sim\Phi(0,\sigma_z^2). \quad \text{(B-1)}$$

with:

Z: the random variable

$\Phi(0,\sigma_z^2)$: a probability distribution with zero mean and σ_z^2 variance.

I then assume that investors' utility is a function of wealth W only:

$$U=U(W). \quad \text{(B-2)}$$

Now we can define the condition for the investor to accept the gamble according to the following equation:

$$E[U(W+Z)] = U[W + E(Z) - \pi(Z)] \\ = U[W - \pi(Z)] \quad \text{(B-3)}$$

In equation (B-3) the expression $\pi(Z)$ represents the risk premium that makes the investor indifferent between accepting the actuarially neutral gamble Z and not accepting. It is assumed to be a function solely of wealth and of the gamble itself. Now, writing out the Taylor expansion of the left-hand side and right-hand side of equation (B-3):

$$E[U(W+Z)] = E[U(W) + U'_W Z + \tfrac{1}{2} U''_W Z^2 + \cdots] \\ U[W - \pi(Z)] = U(W) - U'_W \pi(Z) - \tfrac{1}{2} U''_W \pi^2(Z) - \cdots \quad \text{(B-4)}$$

Here, the terms U'_W and U''_W are the first and second total derivatives of the utility function. Now, since $E[U(W)] = U(W)$ and $E[Z] = 0$, equating and simplifying (B-4) we get:

$$\frac{1}{2} U''_W E(Z^2) = -U'_W \pi(Z) - \frac{1}{2} U''_W \pi^2(Z) . \quad \text{(B-5)}$$

If we further assume the term $-U''_W \pi^2(Z) / 2$ to be negligible, we can solve for π, the risk premium:

$$\pi(Z) = -\frac{U''_W}{2U'_W} E(Z^2) = -\frac{1}{2} \sigma_Z^2 \frac{U''_W}{U'_W} . \quad \text{(B-12)}$$

The above expression $-U''_W / U'_W$ is analogous to the Pratt-Arrow absolute risk-aversion (*ARA*) coefficient. Therefore, using equation (B-12) we can write:

$$\pi = \frac{1}{2} \sigma_Z^2 ARA . \quad \text{(B-7)}$$

Denoting by $r = Z / W$ the return on the gamble given the risk premium, we can write:

$$\sigma_Z^2 = E(Z^2) = E\{W^2[r - E(r)]^2\} = W^2 var(r). \tag{B-8}$$

Therefore, using (B-8) in (B-7) we can write:

$$\pi = \frac{1}{2} W^2 var(r) ARA. \tag{B-9}$$

Then,

$$-r_\pi = \frac{\pi}{W} = \frac{1}{2} var(r)(W * ARA) = \frac{1}{2} var(r) RRA, \tag{B-16}$$

where:

$$RRA = ARA * W = -\frac{U''_W}{U'_W} W. \tag{B-11}$$

Here, the quantity $\frac{\pi}{W}$ denotes negative expected return that the investor is willing to accept to remove the risk of an otherwise actuarially neutral gamble. The coefficient *RRA* in equation (B-11) denotes relative risk aversion. As long as the 'local shape' of the utility function does not change, it should be constant against changes in wealth. Equation (B-11) displays a linear relation between risk and expected excess-return that is valid only locally since the relation has been derived on the basis of a second order Taylor expansion of a potentially non-linear equation.

9. Bibliography

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric Estimation of State‐Price Densities Implicit in Financial Asset Prices. *The Journal of finance, 53*(2), 499-547.

Abadir, K., & Rockinger, M. (1997). Density-embedding Functions, with Applications to the Pricing of Financial Options: Working Paper, HEC.

Abel, A. B. (2002). An exploration of the effects of pessimism and doubt on asset returns. *Journal of Economic dynamics and control, 26*(7-8), 1075-1092.

Abken, P. A., Madan, D. B., & Ramamurtie, S. (1996a). Estimation of risk-neutral and statistical densities by Hermite polynomial approximation: with an application to Eurodollar futures options. *Federal Reserve Bank of Atlanta Working Paper, 96*(5).

Abken, P. A., Madan, D. B., & Ramamurtie, S. (1996b). *Pricing S & P 500 Index Options Using a Hilbert Space Basis*: Federal Reserve Bank of Atlanta.

Ahn, D. H., Conrad, J., & Dittmar, R. F. (2003). Risk adjustment and trading strategies. *Review of Financial studies, 16*(2), 459-485.

Ahn, D. H., & Shivdasani, A. (1999). Long-term returns following seasoned equity issues: Bad performance or bad model. *University of North Carolina working paper*.

Ahn, S. C., & Gadarowski, C. (2004). Small sample properties of the GMM specification test based on the Hansen-Jagannathan distance. *Journal of Empirical finance, 11*(1), 109-132.

Ang, A., & Bekaert, G. (1999). International asset allocation with time-varying correlations: National Bureau of Economic Research.

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial studies, 15*(4), 1137-1187.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics, 63*(3), 443-494.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross‐section of volatility and expected returns. *The Journal of finance, 61*(1), 259-299.

Ang, A., Hodrick, R. J., Xing, Y. H., & Zhang, X. Y. (2009). High idiosyncratic volatility and low returns: International and further US evidence. *Journal of Financial Economics, 91*(1), 1-23. doi: DOI 10.1016/j.jfineco.2007.12.005

Aparicio, S. D., & Hodges, S. D. (1998). Implied risk-neutral distribution: A comparison of estimation methods. *FORC Preprint, University of Warwick*.

Arditti, F. D. (1967). Risk and the required return on equity. *The Journal of finance, 22*(1), 19-36.

Arrow, K. J. (1971). Essays in the theory of risk-bearing.

Attanasio, O., Berloffa, G., Blundell, R., & Preston, I. (2002). From earnings inequality to consumption inequality. *The Economic Journal, 112*(478), C52-C59.

Attanasio, O. P., & Paiella, M. (2011). Intertemporal consumption choices, transaction costs and limited participation in financial markets: reconciling data and theory. *Journal of Applied Econometrics, 26*(2), 322-343.

Bae, K. H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *Review of Financial studies, 16*(3), 717-763.

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross - Section of Stock Returns. *The Journal of finance, 61*(4), 1645-1680.

Bakshi, G., Chen, Z., & Hjalmarsson, E. (2005). Volatility of the Stochastic Discount Factor, and the Distinction between Risk-Neutral and Objective Probability Measures. *Working Papers in Economics*.

Balke, N., & Wohar, M. (2001). Explaining stock price movements: is there a case for fundamentals? *Economic and Financial Policy Review*(Q III), 22-34.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance, 1*, 1053-1128.

Barsky, R. B., Kimball, M. S., Juster, F. T., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement survey: National Bureau of Economic Research.

Bauer, R., Otten, R., & Rad, A. T. (2006). New Zealand mutual funds: measuring performance and persistence in performance. *Accounting & finance, 46*(3), 347-363.

Bekaert, G., & Liu, J. (2004). Conditioning information and variance bounds on pricing kernels. *Review of Financial studies, 17*(2), 339-378.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business, 45*(3), 444-455.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 637-654.

Blake, D., Lehmann, B. N., & Timmermann, A. (2002). Performance clustering and incentives in the UK pension fund industry. *Journal of Asset Management, 3*(2), 173-194.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics, 7*(3), 265-296.

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, 621-651.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics, 18*(2), 373-399.

Campbell, J. Y. (1991). A variance decomposition for stock returns: National Bureau of Economic Research.

Campbell, J. Y. (2000). Asset pricing at the millennium. *The Journal of finance, 55*(4), 1515-1567.

Campbell, J. Y., & Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy, 107*(2).

Campbell, J. Y., & Cochrane, J. H. (2000). Explaining the poor performance of consumption - based asset pricing models. *The Journal of finance, 55*(6), 2863-2878.
Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets* (Vol. 1): princeton University press Princeton, NJ.
Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial studies, 1*(3), 195-228.
Campbell, R., Koedijk, K., & Kofman, P. (2002). Increased correlation in bear markets. *Financial Analysts Journal*, 87-94.
Cecchetti, S. G., Lam, P., & Mark, N. C. (2000). Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True? *American Economic Review*, 787-805.
Cerný, A., & Hodges, S. (2000). *The theory of good-deal pricing in financial markets*.
Chabi-Yo, F., Garcia, R., & Renault, E. (2008). State dependence can explain the risk aversion puzzle. *Review of Financial studies, 21*(2), 973-1011.
Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, 383-403.
Chen, Z., & Knez, P. J. (1996). Portfolio performance measurement: Theory and applications. *Review of Financial studies, 9*(2), 511-555.
Cochrane, J. H. (1996). A Cross-Sectional Test of an Investment-Based Asset Pricing Model. *The journal of political economy, 104*(3), 572-621.
Cochrane, J. H. (1999a). New facts in finance: National Bureau of Economic Research.
Cochrane, J. H. (1999b). Portfolio advice for a multifactor world: National Bureau of Economic Research.
Cochrane, J. H. (2001). Asset pricing.
Cochrane, J. H., & Saa-Requejo, J. (1996). Beyond Arbitrage:" Good-Deal" Asset Price Bounds in Incomplete Markets: National Bureau of Economic Research.
Constantinides, G. M. (2002). Rational asset prices: National Bureau of Economic Research.
Cont, R. (1997). Scaling and correlation in financial data. *Arxiv preprint cond-mat/9705075*.
Copeland, T., Weston, J. F., & SHASHTRI, K. (2004). Financial theory and corporate policy 4th ed (International edition). *Recherche, 67*, 02.
Dahlquist, M., & Söderlind, P. (1999). Evaluating Portfolio Performance with Stochastic Discount Factors. *The Journal of Business, 72*(3), 347-383.
De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990a). Noise Trader Risk in Financial Markets. *The journal of political economy, 98*(4), 703-738.
De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990b). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of finance, 45*(2), 379-395.

Diba, B. T., & Grossman, H. I. (1988). Explosive rational bubbles in stock prices? *The American Economic Review, 78*(3), 520-530.

Dittmar, R. F. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of finance, 57*(1), 369-403.

Drew, M. E., Naughton, T., & Veeraraghavan, M. (2003). Firm size, book-to-market equity and security returns: Evidence from the Shanghai Stock Exchange. *Australian Journal of Management, 28*(2), 119-139.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251-276.

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

Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1994). Forecasting international equity correlations. *Financial Analysts Journal*, 32-45.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of finance, 25*(2), 383-417.

Fama, E. F. (1976). *Foundations of finance: portfolio decisions and securities prices*: Basic Books New York.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*(1), 23-49.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance*, 131-155.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 55-84.

Fang, H., & Lai, T. Y. (1997). Co-Kurtosis and Capital Asset Pricing. *Financial Review, 32*(2), 293-307.

Farnsworth, H., Ferson, W., Jackson, D., & Todd, S. (2002a). Performance Evaluation with Stochastic Discount Factors. *The Journal of Business, 75*(3), 473-583.

Farnsworth, H., Ferson, W., Jackson, D., & Todd, S. (2002b). Performance Evaluation with Stochastic Discount Factors*. *The Journal of Business, 75*(3), 473-583.

Ferson, W., Henry, T. R., & Kisgen, D. J. (2006). Evaluating government bond fund performance with stochastic discount factors. *Review of Financial studies, 19*(2), 423-455.

Ferson, W. E., & Harvey, C. R. (1999). Conditioning variables and the cross section of stock returns. *The Journal of finance, 54*(4), 1325-1360.

Ferson, W. E., & Siegel, A. F. (2003). Stochastic discount factor bounds with conditioning information. *Review of Financial studies, 16*(2), 567-595.

Ferson, W. E., & Siegel, A. F. (2009). Testing Portfolio Efficiency with Conditioning Information. *Review of Financial studies, 22*(7), 2735-2758. doi: 10.1093/rfs/hhn112

Fousseni, C.-Y. (2011). Explaining the idiosyncratic volatility puzzle using Stochastic Discount Factors. *Journal of Banking & Finance, 35*(8), 1971-1983. doi: 10.1016/j.jbankfin.2011.01.002

Freedman, D. A., & Peters, S. C. (1984). Bootstrapping an econometric model: Some empirical results. *Journal of Business & Economic Statistics*, 150-158.

Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. *The journal of political economy, 56*(4), 279-304.

Friend, I., & Blume, M. E. (1975). The demand for risky assets. *The American Economic Review, 65*(5), 900-922.

Friend, I., & Westerfield, R. (1980). Co-skewness and capital asset pricing. *The Journal of finance, 35*(4), 897-913.

Froot, K. A., & Obstfeld, M. (1991). Exchange-rate dynamics under stochastic regime shifts:: A unified approach. *Journal of International Economics, 31*(3-4), 203-229.

Gallant, A. R., Hansen, L. P., & Tauchen, G. (1990). Using conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution. *Journal of econometrics, 45*(1-2), 141-179.

Garrett, T. A., & Sobel, R. S. (1999). Gamblers favor skewness, not risk: Further evidence from United States' lottery games. *Economics Letters, 63*(1), 85-90.

Golec, J., & Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, 205-225.

Grant, D. (1977). Portfolio performance and the" cost" of timing decisions. *The Journal of finance, 32*(3), 837-846.

Grinblatt, M., & Titman, S. (1989). Mutual fund performance: An analysis of quarterly portfolio holdings. *Journal of business*, 393-416.

Guo, H., & Savickas, R. (2003). Does idiosyncratic risk matter: Another look. *Working Paper Series*.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029-1054.

Hansen, L. P., & Jagannathan, R. (1991). Restrictions on intertemporal marginal rates of substitution implied by asset returns. *Journal of Political Economy, 99*(2), 225.

Hansen, L. P., & Jagannathan, R. (1997). Assessing Specification Errors in Stochastic Discount Factor Models. *Journal of Finance*, 557-590.

Hansen, L. P., & Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica: Journal of the Econometric Society*, 587-613.

Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, 1269-1286.

Hansen, L. P., & Singleton, K. J. (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *The Journal of political economy*, 249-265.

Harris, D., McCabe, B., & Leybourne, S. (2002). Stochastic cointegration: estimation and inference. *Journal of econometrics, 111*(2), 363-384.

Harrison, J. M., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory, 20*(3), 381-408.

Hartley, R., & Farrell, L. (2002). Can expected utility theory explain gambling? *The American Economic Review, 92*(3), 613-624.

Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of finance, 55*(3), 1263-1295.

He, H., & Modest, D. M. (1995). Market frictions and consumption-based asset pricing. *Journal of Political Economy*, 94-117.

Ingersoll, J. E. (1987). *Theory of financial decision making*: Rowman & Littlefield Pub Inc.

Jackwerth, J. C. (1999). Option-implied risk-neutral distributions and implied binomial trees: A literature review. *Journal of Derivatives, 7*(2), 66-82.

Jagannathan, R., & Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, 3-53.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 65-91.

Jegadeesh, N., & Titman, S. (1999). Profitability of momentum strategies: An evaluation of alternative explanations: National Bureau of Economic Research.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kang, J., Liu, M. H., & Ni, S. X. (2002). Contrarian and momentum strategies in the China stock market: 1993-2000. *Pacific-Basin Finance Journal, 10*(3), 243-265.

Kimball, M. S. (1990). Precautionary Saving in the Small and in the Large. *Econometrica: Journal of the Econometric Society*, 53-73.

Kimball, M. S. (1993). Standard risk aversion. *Econometrica: Journal of the Econometric Society*, 589-611.

King, B. F. (1966). Market and industry factors in stock price behavior. *The Journal of Business, 39*(1), 139-190.

Koutoulas, G., & Kryzanowski, L. (1996). Macrofactor Conditional Volatilities, Time‐Varying Risk Premia and Stock Return Behavior. *Financial Review, 31*(1), 169-195.

Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of finance, 31*(4), 1085-1100.

Lee, C. F., & Rahman, S. (1990). Market timing, selectivity, and mutual fund performance: An empirical investigation. *Journal of business*, 261-278.

Leite, P. A., & Cortez, M. C. (2009). Conditioning information in mutual fund performance evaluation: Portuguese evidence. *The European Journal of Finance, 15*(5-6), 585-605.

LeRoy, S. F. (1973). Risk aversion and the martingale property of stock prices. *International Economic Review, 14*(2), 436-446.

Lettau, M., & Ludvigson, S. (2001a). Consumption, aggregate wealth, and expected stock returns. *The Journal of finance, 56*(3), 815-849.

Lettau, M., & Ludvigson, S. (2001b). Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 1238-1287.

Levy, H. (1969). A utility function depending on the first three moments. *The Journal of finance, 24*(4), 715-719.

Lewellen, J., & Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics, 82*(2), 289-314.

Lewellen, J., Nagel, S., & Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics, 96*(2), 175-194.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*(1), 13-37.

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of finance, 56*(2), 649-676.

Loretan, M., & Phillips, P. C. B. (1994). Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial datasets. *Journal of Empirical finance, 1*(2), 211-248.

Loughran, T., & Ritter, J. R. (1997). The operating performance of firms conducting seasoned equity offerings. *Journal of Finance*, 1823-1850.

Lucas, J. R. E. (1978). Asset prices in an exchange economy. *Econometrica: Journal of the Econometric Society*, 1429-1445.

Luttmer, E. G. J. (1996). Asset pricing in economies with frictions. *Econometrica: Journal of the Econometric Society*, 1439-1467.

Luttmer, E. G. J. (1999). What level of fixed costs can reconcile consumption and stock returns? *Journal of Political Economy, 107*(5), 969-997.

Malkiel, B. G. (1999). *A random walk down Wall Street: including a life-cycle guide to personal investing*: WW Norton & Company.

Mankiw, N. G., & Zeldes, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics, 29*(1), 97-112.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77-91.

Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics, 15*(2), 145-161.

Melick, W. R., & Thomas, C. P. (1997). Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis. *Journal of Financial and Quantitative Analysis, 32*(1).

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of economics and statistics, 51*(3), 247-257.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 141-183.

Merton, R. C. (1980). On estimating the expected return on the market* 1:: An exploratory investigation. *Journal of Financial Economics, 8*(4), 323-361.

Meyer, D. J., & Meyer, J. (2005). Relative risk aversion: What do we know? *Journal of Risk and Uncertainty, 31*(3), 243-262.

Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum? *The Journal of finance, 54*(4), 1249-1290.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, 315-335.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica: Journal of the Econometric Society*, 703-708.

Otten, R., & Bams, D. (2002). European mutual fund performance. *European Financial Management, 8*(1), 75-101.

Pagan, A. (1996). The econometrics of financial markets. *Journal of Empirical finance, 3*(1), 15-102.

Paiella, M. (2004). Heterogeneity in financial market participation: appraising its implications for the C-CAPM. *Review of Finance, 8*(3), 445-480.

Paiella, M. (2007). The forgone gains of incomplete portfolios. *Review of Financial studies, 20*(5), 1623-1646.

Peters, S. C., & Freedman, D. A. (1984). Some notes on the bootstrap in regression problems. *Journal of Business & Economic Statistics, 2*(4), 406-409.

Pliska, S. R. (1997). *Introduction to mathematical finance: discrete time models*: Wiley.

Posner, S., & Milevsky, M. (1998). Valuing exotic options by approximating the SPD with higher moments.

Post, T., & Levy, H. (2005). Does risk seeking drive stock prices? A stochastic dominance analysis of aggregate investor preferences and beliefs. *Review of Financial Studies, 18*(3), 925-953.

Post, T., & Van Vliet, P. (2006). Downside risk and asset pricing. *Journal of Banking & Finance, 30*(3), 823-849.

Post, T., Van Vliet, P., & Levy, H. (2008). Risk aversion and skewness preference. *Journal of Banking & Finance, 32*(7), 1178-1187.

Potì, V. (2007). Arbitrage, Good Deals and Stochastic Discount Factor Restrictions on Multi-Factor Models With Coskewness. *SSRN eLibrary*.

Potì, V., & Wang, D. L. (2010). The coskewness puzzle. *Journal of Banking & Finance, 34*(8), 1827-1838.

Potters, M., Cont, R., & Bouchaud, J. P. (1998). Financial markets as adaptive systems. *EPL (Europhysics Letters), 41*, 239.

Pritsker, M. (1998). Nonparametric density estimation and tests of continuous time interest rate models. *Review of Financial studies, 11*(3), 449-487.

Rachev, S., Jasic, T., Stoyanov, S., & Fabozzi, F. J. (2007). Momentum strategies based on reward-risk stock selection criteria. *Journal of Banking & Finance, 31*(8), 2325-2346.

Richter, M. K. (1960). Cardinal utility, portfolio selection and taxation. *The Review of Economic Studies, 27*(3), 152-166.

Ritchey, R. J. (1990). Call option valuation for discrete normal mixtures. *Journal of Financial Research, 13*(4), 285-296.

Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics, 4*(2), 129-176.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341-360.

Ross, S. A. (2005). *Neoclassical finance*: Princeton Univ Pr.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *The review of economics and statistics, 51*(3), 239-246.

Sargent, T. J. (1996). Bounded rationality in macroeconomics: The Arne Ryde memorial lectures. *OUP Catalogue*.

Sawicki, J., & Ong, F. (2000). Evaluating managed fund performance using conditional measures: Australian evidence. *Pacific Basin Finance Journal, 8*, 505-528.

Scruggs, J. T. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two‐factor approach. *The Journal of finance, 53*(2), 575-603.

Shanken, J. (1992). On the estimation of beta-pricing models. *Review of Financial studies, 5*(1), 1-55.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425-442.

Shefrin, H. (2010). *Behavioralizing Finance* (Vol. 4): Now Pub.

Shefrin, H., & Statman, M. (2000). Behavioral portfolio theory. *Journal of Financial and Quantitative Analysis, 35*(02), 127-151.

Shefrin, H., & Statman, M. (2002). The style of investor expectations. *The Handbook of Equity Style Management, forthcoming*.

Shleifer, A., & Vishny, R. W. (1997). The Limits of Arbitrage. *The Journal of finance, 52*(1), 35-55.

Simkowitz, M. A., & Beedles, W. L. (1978). Diversification in a three-moment world. *Journal of Financial and Quantitative Analysis, 13*(05), 927-941.

Singleton, J. C., & Wingender, J. (1986). Skewness persistence in common stock returns. *Journal of Financial and Quantitative Analysis, 21*(03), 335-341.

Smith, D. R. (2007). Conditional coskewness and asset pricing. *Journal of Empirical finance, 14*(1), 91-119.

Smith, P., & Wickens, M. (2002). Asset pricing with observable stochastic discount factors. *Journal of Economic Surveys, 16*(3), 397-446.

Statman, M. (2002). Lottery players/stock traders. *Financial Analysts Journal*, 14-21.

Tsiang, S. C. (1972). The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money. *The American Economic Review, 62*(3), 354-371.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.

Vanden, J. M. (2006). Option coskewness and capital asset pricing. *Review of Financial studies, 19*(4), 1279-1320.

Vissing-Jørgensen, A. (2002). Limited Asset Market Participation and the Elasticity of Intertemporal Substitution. *Journal of Political Economy, 110*(4).

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of monetary Economics, 24*(3), 401-421.

Zakamouline, V., & Koekebakker, S. (2009). Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *Journal of Banking & Finance, 33*(7), 1242-1254.