DOWNSIDE RISK AND EMPIRICAL ASSET PRICING

ERIM Ph.D. Series Research in Management, 49

Erasmus Research Institute of Management (ERIM)
Erasmus University Rotterdam
Internet: http://www.erim.eur.nl

ISBN 90-5892-075-5

DOWNSIDE RISK AND EMPIRICAL ASSET PRICING

De rol van neerwaarts risico bij de prijsvorming van effecten:
een empirisch onderzoek

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus

Prof.dr. S.W.J. Lamberts

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 16 december 2004 om 13:30 uur

door

WILLEM NICOLAAS VAN VLIET
geboren te Bodegraven

Promotiecommissie

Promotor:
Prof.dr. G.T. Post

Overige leden:
Prof.dr. W.F.M. de Bondt
Prof.dr. J.T.J. Smit
Prof.dr. J. Spronk

Preface

In October 2000, I started my PhD project in Finance at ERIM. During the past four years I have learnt a lot from family, friends, colleagues and supervisors. This is a good time to write down some words of thanks for those who have had faith in me.

I would like to thank Han Smit for his enthusiasm which stimulated me to start a PhD and explore the impact of growth options on the performance of stocks. Both your guidance in the early years and your passion for financial research inspired me. You also improved my presentation skills and level of academic writing. I cherish our conversations, especially the ones we had at the camp fire during the retraite. I am especially indebted to my thesis supervisor Thierry Post for helping me to push my research to a higher level. Our cooperation started during a meeting of the EWGFM on the beautiful island of Capri. Together with Martijn and our girlfriends, we had great fun during this conference and I especially remember the speedboat, champagne and cigars. Since then, my research focused on downside risk and empirical asset pricing. I appreciate your efficiency, brilliant ideas and work-ethic and I feel very privileged working with you for such a long period. You always had time to help me out and responded to questions very quickly, irrespective of time or place. You are the co-author of most of my papers and taught me to be critical and to structure ideas in a logical way. I also remember the work-outs in the gym and the "appelflappen" on Saturday morning.

I am also grateful to Werner de Bondt who explained me some 'mores' in empirical finance on sunny terraces in Antwerp and Maastricht. Your laughter and Belgium humor cheered up the days we had contact. I would also like to thank Jaap Spronk. You have been a stable factor during the past four years, gave good advice and showed genuine personal interest. Next I would also like to thank the other members of the committee; Haim Levy, Marno Verbeek and Casper de Vries for evaluating this thesis.

During my PhD track I shared a room with Kevin Pak. We had lots of fun, did some ERIM courses together, shared the fridge and listened to music, preferably of Shirley Bassey. Further I would like to thank my fellow promovendi, most of them from the 16th floor: Martijn van den Assem 'ik ben geciteerd!', Guido Baltussen 'lekker knallen', Reimer Beneder 'je verdient zo een ton', Klaas Beniers 'is dat mijn koffiemok?', Ward van den Berg 'liberaal is niet conservatief', Bram van Dijk 'links conservatief', Dennis Fok 'dat is toch makkelijk?', Hans Haanappel 'ik wil applyen', Wilco van der Heuvel 'IJsselvogels', Patrick Houweling 'ik ga long in vastgoed', Joop Huij 'het zijn de flows', Jos van Iwaarden 'goeiesmeurgens', Erik Kole 'Berlijn is fijn',

Daina Konter ‘vrouwen zijn anders’, Joost Loef ‘heb je je Liebherr verkocht?’, Ivo Nobel ‘Bodegraven bravo!’, Klaas Staal ‘reizen is leuk’, Rutger van Oest ‘iedereen promoveert op yoghurt data’, Sandra Philippen ‘in een kerk geboren’, Remco Prins ‘ja’, Jan-Frederik Slijkerman ‘ik spreek je straks’, Mariëlle Sonnenberg ‘waar is mijn portemonnee?’, Philippe Versijp ‘de staatsidee’, Björn Vroomen ‘lunch!’ and Amy Wong ‘echt waar?’. Further I would like to thank Dick van Dijk, Mathijs van Dijk, Anna Gutkovska, Winfried Hallerbach, Ben Jacobsen, Jedid-Jah Jonker, Roy Kouwenberg, Roger Lord, Wouter de Maeseneire, Ronald Mahieu, Martin Martens, Gerard Moerman, Erjen van Nierop, Mariëlle Non, Haikun Ning, Richard Paap, Michiel de Pooter, Igor Pouchkarev, Nico van der Sar, Marc Schauten, Stephen Smit, Onno Steenbeek, Rob Stevens and Yulia Veld. Many thanks to Marien de Gelder and Suzanne van der Tang for programming assistence to sort out the CRSP database. I am indebted to the staff of ERIM and the Finance department for excellent support: Trudy van Florestein, Wilfred Mijnhardt, Helene Molenaar, Betty Rietveld, Tineke van der Vhee and Daphne Willems. Thank you for letting me organize a rainy poldersport adventure and listening to my speeches.

During my life, friends and family have always been a great support. I would like to thank Arian, Boy, Corne, Cornelis, Christoph, Dennis, Erik, Erik, Foppe, Hadrian, Hans, Jan, Joan, Johan, Jon, Koen, Leon, Marcel, Maarten, Martin, Martijn, Martijn, Navin, Patrick, Peter, Pieter, Ralph, Robert, Roel, Remon, Tammo, Timon, Wilco, Wouter-Jan for their friendship. I especially give thanks to my parents for their unconditional love during my whole life. Thank you father for your inspired wisdom and thank you mother for all your support and prayers. Further, I would like to thank my three brothers Bert, Arjen, Hans, my special sister Margreet, and Wendy, Elsien, Cees, Julia and their children. Thank you for accepting me as a family member and listening to my attempts to explain my research in accesable terms. I thank my parents in law, Linetta and Marjolein for giving me a second family. Thank you sweet Sabine for sharing your life with me. You complete me in a beautiful way and I highly appreciate your committed love for me. Your encouraging faith and hope motivated me to go the extra mile. You have taught me many things such as windsurfing, and I consider you as a special gift from heaven.

Finally, I would like to thank my Father who gives me life through Jesus and inspires me through his Spirit. I am grateful for the talents and abilities I received from God my Creator. This dissertation is in honor of Him who deserves all worship and glory.

Soli Deo Gloria
Pim van Vliet
Rotterdam, October 2004

Contents

Chapter 1

Introduction

1.1 Purpose

The last decades have witnessed some major developments in the field of *asset pricing*. These have contributed to a better understanding of stock, bond and other asset prices and have influenced other disciplines such as corporate finance and macro economics. Currently (2004), the Nobel prize winning *Capital Asset Pricing Model* (CAPM) of Sharpe (1964) and Lintner (1965) celebrates its 40th birthday. This seminal model is the most widely applied model in asset pricing. During the past decades, much progress has been made in the theoretical, methodological and empirical fields of asset pricing. First, several generalizations of the CAPM exist, each based on relaxing one of the main *assumptions* of the CAPM. Second, advanced *econometric techniques* and an increase in computer power have made it possible to handle large datasets and to control for several statistical issues. Third, high-quality *financial databases* have become available and the richness and precision of these databases is unmatched by other economic disciplines.

Still, the CAPM faces some severe empirical difficulties. Specifically, Basu (1977), Banz (1981), Reinganum (1981), DeBondt and Thaler (1985), and Jegadeesh and Titman (1993) among others show that the CAPM fails to explain the returns of several equity investment strategies (e.g. based on accounting data or past returns). Some authors explain the failures of the CAPM with *nonrisk-based* explanations such as biases in the empirical methodology (e.g. Lo and MacKinlay (1990), MacKinlay (1995) and Kothari, Shanken and Sloan (1995)), or investor irrationality (e.g. DeBondt and Thaler (1987), Lakonishok, Shleifer and Vishny (1994) and Daniel and Titman (1997)), while others take a *rational view* and explain differences in return with differences in risk (e.g. Fama and French (1996), Cochrane (1996) and Lettau and Ludvigson (2001)).

This research specially considers several empirical and methodological issues typical for asset pricing. We take a rational view and question CAPM's use of *variance* as the relevant risk measure. Today there is substantial evidence that returns are typically not normally distributed and that investors are more sensitive to downside than to upside price movements. Alternative *downside risk* measures, such as semi-variance, may better describe *investor preferences*. Surprisingly, despite the theoretical appeal of downside risk and the empirical problems of the CAPM, this line of research has not been thoroughly analyzed yet.

The purpose of this thesis is to document the empirical performance of the CAPM and to examine if relaxing the preference assumptions of the CAPM can help to explain stock prices. This thesis has five distinguishing features. First, we pay special attention to adhering to first principles such as nonsatiation and risk aversion. Second, we are the first to apply recently developed *non-parametric techniques* on large scale datasets covering most of the 20th century. Due to methodological advances it is now possible to study a wide class of (downside) risk measures without having to parameterize the model in advance. Third, to the best of our knowledge, we are the first to model and empirically test *time-varying downside risk aversion*. We consider a simple model in which investors care more about downside deviations during economic recession periods than during economic expansion periods. Fourth, we *construct benchmark portfolios* that are specially designed for the analysis of downside risk. For statistical testing we use the *Generalized Method of Moments* (GMM), a unifying framework which embeds various regression techniques such as OLS and GLS as special cases. Fifth, the empirical analysis considers a large number of datasets and carries out many robustness checks.

1.2 Variance as a risk measure

> Lorie (1966, page 108) "*I believe we will ultimately find an objective measure of sensitivity to decline which avoids the inherent absurdity of calling a stock risky because in the past it has gone up much faster than the market and only as fast in others, whereas we call a security that never varies in price not risky at all.*"

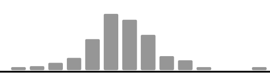
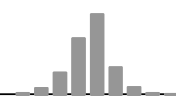
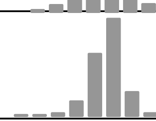
The weak empirical performance of the CAPM may be explained by the limitation of the MV criterion. Roughly speaking, this criterion can only be applied if either investors have quadratic utility, or if asset returns are *normally distributed* (Tobin (1958) and Berk (1997)). In this section we shortly discuss that financial asset returns generally are not normally distributed and preferences of investors are generally not captured by the mean-variance criterion.

1.2.1 Non-normal returns

Table 1.1 shows the empirical monthly return distribution for a single stock, an industry stock portfolio and the value-weighted stock market index covering the January 1983 – December 2002 period. Clearly, the return series display *asymmetry* and *fat tails* and mean and variance alone fail to describe asset returns. At the individual stock level we observe high variance combined with positive skewness and excess kurtosis. Variance is lower at higher aggregation levels, but the return distributions are still characterized by non-normalities. Interestingly, the return distributions become negatively skewed and maintain fat tailed. This means that large negative returns occur more often than expected under normality. Under normality, an observation more than 5 standard deviations from the mean (such as the stock market crash of October 1987) should be observed about once every 1 million years. In fact such observations have happened more than once in the 20th century (e.g. during the 1930s). Not surprisingly, the Jarque-Bera (JB) test, which tests for normality, has to be rejected for all three series; all p-values are below conventional levels of significance.

Table 1.1
Financial asset return distributions

This table shows some examples of return series of financial assets at an individual level, industry level and market level. The monthly return series cover the January 1983-December 2002 period (240 months). Excess returns are computed from the raw return observations by subtracting the return on the one-month US Treasury bill. Besides descriptive statistics the p-value of the Jarque-Bera test statistic (JB-test) is displayed. The JB-test tests the null hypothesis of normal returns.

Series	Avg	Stdev	Skew	Kurt	JB-test	Distribution
Individual Level:						
AT&T Corp	0.53	8.33	0.27	1.81	0.00	
Industry Level:						
Telecom	0.65	5.77	-0.11	0.96	0.01	
Market Level:						
All-share index	0.58	4.53	-0.92	3.06	0.00	
						-30% 40%

Source: CRSP monthly US stock return database (2002)

Actually, the choice for monthly stock returns does not represent the most strongest case against normality, because other return intervals and other asset types also show violations of normality. For example, it is well-known that daily stock returns are leptokurtic (see Fama (1965)). Further, the returns of other asset types, such as corporate bonds (Ibbotson (2002)), derivates (Coval and Shumway (2001)) and foreign stocks/bonds (Dimson and Marsh (2001)), are characterized by asymmetry and fat tails. For these series, mean and variance also fail to completely describe the return distribution.

Interestingly, individual stocks are on average positively skewed, while portfolio returns are negatively skewed (for example Table I and Ang and Chen (2002)). The change in skewness sign is caused by an increase in correlation of stock returns during bear markets (Campbell, Koedijk and Kofman (2002)). Due to this asymmetric correlation, it is more difficult to diversify away downside risk than it is to destroy upside potential. An open research question is why individual stock returns are positively skewed and why correlations are stronger in downside markets. For example, positive skewness in individual stock returns may be caused by real option leverage (see Myers (1977), Knez and Ready (1997) and Smit and van Vliet (2002))and negative portfolio skewness may be caused by differences in opinion and short-sales constraints (Hong and Stein (2003)). However, in this thesis we take the empirical return distributions as given.

1.2.2 Non-quadratic utility

Since returns are non-normal, another way to justify the mean-variance criterion is by assuming a *quadratic utility* function. A quadratic utility function implies that investors care about mean and variance only, even if returns exhibit asymmetry or fat tails.

Hanoch and Levy (1970) argue that quadratic utility is subject to some serious limitations. First, a quadratic function will ultimately become negative over some return interval. Put differently, a quadratic utility function could violate the nonsatiation condition. Nonsatiation is a very important property and underlies virtually all models of decision-making. Second, a quadratic utility function is inconsistent with *decreasing absolute risk aversion* (DARA). This implies that, all other things equal, individuals should prefer distributions that are right-skewed to distributions that are left-skewed (see Arditti (1967)). In fact, several financial studies suggest that investors display skewness preference, see Cooley (1977, Simkowitz and Beedles (1978), Scott and Horvath (1980), Kane (1982), Harvey and Siddique (2000) and Dittmar (2002). Further, a quadratic utility function fails to capture *loss aversion* (LA). Loss aversion means that individuals care much more about losses than gains. For example, Benartzi and Thaler (1995) and Berkelaar, Kouwenberg and Post (2005) demonstrate that loss aversion helps to better understand asset prices.

In sum, a quadratic utility has several important limitations and may fail to describe investor preferences. Therefore, this study loosens this classical preference assumption and considers a wider range of alternative investor preferences.

1.3 Asset pricing theory

1.3.1 Assumptions

> William Sharpe in Burton (1998): "*The CAPM was a very simple, very strong set of assumptions that got a nice, clean, pretty result. And then almost immediately, we all said, let's bring more complexity into it to try to get closer to the real world.*"

The CAPM is a *single-period*, *representative investor, portfolio-oriented* model of a *perfect capital market.* The literature proposes at least five ways to generalize the CAPM. We briefly discuss these generalizations and motivate our choice to stay close to CAPM's assumptions and to adjust the investor preference assumption only.

First, the MV criterion is obtained either if one makes *distributional assumptions* (normal distribution), or if one makes *investor preferences assumptions* (quadratic utility). In this thesis, we take a preference-based perspective and interpret the MV-criterion in terms of a quadratic utility function. We do not make further distributional assumptions, because theory does not guide us to select the correct return distribution or the relevant moment(s) of this alternative return distribution. By contrast, investor preferences can be modeled in such a way that they obey the basic *regularity conditions*. These conditions impose that the utility function of the investor should be increasing and concave to reflect *nonsatiation* (NS) and *risk aversion* (RA). We will relax the quadratic utility assumption and consider a broader set of *well-behaved* utility functions. Relaxation of the preferences assumption has received relatively little attention compared to the other generalizations of the CAPM.

Second, we use a *single-period* model. In reality, the portfolio-choice problem of investors may be better described by a multi-period model. The *multi-period* problem generally is more complex than the single-period problem, because the mean-variance efficient frontier changes through time and investors revise their portfolios accordingly. Merton's (1973) *intertemporal capital asset pricing model (ICAPM)* captures this multi-period aspect of financial market equilibrium. The main insight of the ICAPM is that the optimal portfolio must also provide the best hedge against unfavorable changes in the set of future investment opportunities. However, economic theory does not guide us which factors should be used to proxy for the future opportunity set. As a result the ICAPM is sometimes misused as a "factor fishing license" (see Fama (1991)). Further, even if we know the relevant factor, then we still do not know how to specify the relationship. For this reason, we maintain the single-period assumption.

Third, we use *portfolio-oriented* models, which means that we assume that investors derive utility directly from portfolio returns. In reality however, investors derive utility from *consumption* of goods and services. Future consumption can be financed through cash-flows obtained from other sources than investment returns, such as return on human capital. The consumption-based CAPM (CCAPM, see Breeden (1979)) defines the utility function over future consumption. However, several empirical problems arise when using *consumption data*. For example, observed consumption data measure *expenditures*, not actual consumption. This may explain the poor empirical performance of consumption based asset pricing models (Breeden, Gibbons and Litzenberger (1989)). Further, aggregate expenditure data are not available at higher frequencies such as monthly or daily intervals. Although theoretically less appealing, we favor the use of portfolio-oriented models because of these empirical issues.

Fourth, we assume the existence of a *perfect capital market.* This means that capital markets are frictionless (no transaction costs and securities are infinitely divisible) and market participants are price-takers, have costless access to information and face no taxes. Arguably, security markets come closest to the definition of a perfect capital market because there are many market participants who face low transaction costs, low absolute prices, few entrance barriers and have access to low-cost, high-quality information. Still, we are aware that some financial assets are illiquid (usually small caps) and some investment strategies involve high trading costs (such as momentum). Therefore, we mainly consider low-cost buy and hold investment strategies and put less emphasis on the small cap and/or momentum strategies.

Fifth, we consider representative investor models. In these models all investors act in such a way that their cumulative actions might as well be the actions of one investor maximizing its expected utility function. Market equilibrium can be described as an optimization problem of a single investor who chooses an efficient portfolio. Alternatively, one may assume *heterogeneity* in investors borrowing constraints, income or preferences. We motivate our use of a representative investor model in three ways. First, by the assumption of *complete markets* which means that the complete set of future cash flows in future states-of-the-world can be constructed with existing assets. With complete markets a Pareto optimal outcome results, which means that all investors share risk perfectly (for a in-depth treatment of complete markets see Huang and Litzenberger (1988, Chapter 5)). Second, by the assumption that investors have *sufficient similar preferences* which means that investors may differ in several respects, but obey the same set of optimality conditions. For example, Rubinstein (1974)) shows that for a broad class of logarithmic and quadratic utility functions this result is found. Third, one can motivate the use of a representative investor by the *revealed preferences* of (some) investors. Even though individual preferences of investors may differ substantially, many investors are interested in value-weighted market indexes. This is apparent from the popularity of passive mutual funds that track broad well-diversified portfolios (for example the S&P500 index). Hence it can also be seen as an attempt to rationalize the choice of the investors who hold the market portfolio.

Figure 1.1 illustrates the five main generalizations of the CAPM which result from relaxation of the five different assumptions. Note that these relaxations can also be combined, for example alternative preferences with market imperfections.

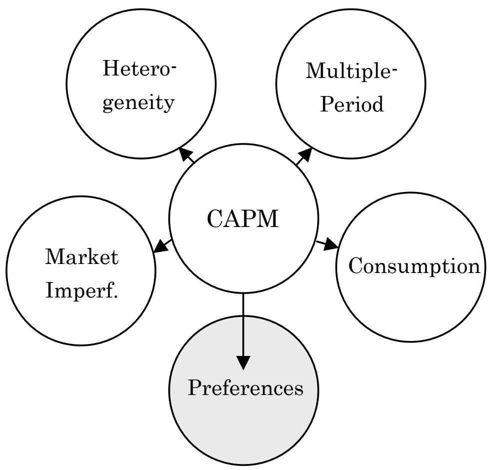


Figure 1.1: Generalizations of the CAPM

1.3.2 Risk-return relation

Let us consider a non-satiable and risk averse representative investor who maximizes an increasing and concave utility function, $u(R)$, by choosing how much to invest in several assets. The investment universe consists of a number of risky assets and a risk-free asset. The first-order condition, also known as the Euler equation, gives the necessary and sufficient optimality condition for this portfolio choice problem:

$$E\left(R_i m\right) = 0 \tag{1.1}$$

where R_i is the excess return on asset i and m is the marginal utility function of the representative investor also known as the *pricing kernel* or stochastic discount factor (SDF). The marginal utility loss of buying a little more of one asset should equal the marginal utility gain of buying another asset. Thus the market portfolio is the efficient or *optimal portfolio*. The pricing kernel is an aggregate of the marginal utility functions of all individual investors and is evaluated at the return of the market portfolio.

We can re-write the Euler equation (1.1) into a more familiar beta representation. Applying the covariance decomposition gives:

$$E(R_i) = -\frac{\operatorname{cov}(R_i, m)}{E(m)} \tag{1.2}$$

we may multiply the right-hand side of this equation with $\operatorname{cov}(R_M, m)[\operatorname{cov}(R_M, m)]^{-1}$ and arranging terms gives a generalized version of the *Security Market Line* (SML):

$$E(R_i) = \frac{\operatorname{cov}(R_i, m)}{\operatorname{cov}(R_M, m)} E(R_M) = \beta_{m,i} E(R_M) \tag{1.3}$$

where $E(R_i)$ is the expected excess return of asset i, $\beta_{m,i}$ is the *kernel beta* and $E(R_M)$ is the market risk premium. This equation simply states that assets that make the market portfolio (R_M) more risky must promise a higher expected return to lure the representative investor to hold it. On the other hand, assets that reduce the risk of the market portfolio have a lower expected return. In equilibrium all assets should be priced according to their risk, i.e. they should all lie on the SML. The kernel beta is a generalization of the traditional CAPM market beta. Even complex multi-factor nonlinear kernels can be expressed in this single kernel beta representation. Interestingly, although the pricing kernel may be nonlinear, the risk-return relation (Equation 1.3) will always be linear.

The pricing kernel representation is currently very popular because a broad range of asset pricing models can be analyzed in this unifying framework. With a correctly specified pricing kernel, we know 'the price of everything'. The CAPM specifies the pricing kernel as a *linear function* of market portfolio return: $m=a+bR_M$. In this special case, the kernel beta ($\beta_{m,}$) is identical to the traditional market beta (β_{CAPM}). The linearity assumption of the pricing kernel is unnecessary and Cochrane's quote nicely illustrates this point. Still, theory does not impose this exact *specification* of the functional form of the pricing kernel, but only gives some basic regularity conditions.

John Cochrane (2001, page 169) '*Why bother linearizing a model? …The tricks were developed when it was hard to estimate nonlinear models. It is clear how to estimate a beta and a market risk premium by regressions, but estimating nonlinear models used to be a big headache. Now, GMM has made it easy to estimate and evaluate nonlinear*

1.3.3 Regularity conditions

In this thesis we emphasize the need for imposing basic *regularity conditions* in order to guarantee economically meaningful results. Specifically we will consider three different regularity conditions: *nonsatiation* (NS), *risk aversion* (RA) and *decreasing absolute risk aversion* (DARA). More wealth makes the investor happier, but the increase in happiness declines at a decreasing rate with the wealth level. This means that the marginal utility function is a positive and decreasing and convex function of wealth. There are economical and mathematical arguments why these regularity conditions should be imposed.

Risk aversion (RA) is needed to ensure that the market portfolio is the optimal portfolio. In general, we observe risk aversion given the large insurance market (insurance premium) and the higher average returns on risky assets such as stocks (equity premium). Still, risk aversion is not a law of nature and there are also arguments to support (local) risk seeking. For example, Markowitz (1952) argues that the willingness to purchase both insurance and lottery tickets (the Friedman-Savage puzzle) implies that marginal utility is increasing for gains. However, a representative risk-seeking investor will not fully diversify because he likes risk. Assuming such preferences will not generate economically meaningful results because this investor will not hold the market portfolio in equilibrium.

From a mathematical perspective, a decreasing marginal utility function is required in order to justify the approach of checking the first-order condition (Euler-equation or SML). For a non-concave utility function the first-order condition is a necessary but nut sufficient condition for establishing a global maximum for expected utility; the first-order condition also applies for possible local optima and a possible global minimum. This important result is sometimes ignored when using non-concave utility functions.

The regularity conditions are of great help to impose structure on the functional form of the marginal utility function of the representative investor (the pricing kernel). In the portfolio-oriented model, we know that the pricing kernel is some function of market return: $m=f(R_M)$. With the regularity conditions we further know that $f>0$, $f'<0$, $f''>0$, which means that the pricing kernel is a globally positive, decreasing and convex function of market return.

Figure 1.2 shows a nonlinear pricing kernel specification as an alternative to the linear pricing kernel. This figure provides further insight in the regularity conditions and the possible violations. The linear kernel corresponds to the CAPM specification. We recall that the CAPM assumes that the pricing kernel is a linear function of market return. The curved kernel corresponds to a generalized version of the CAPM. Interestingly, the nonlinear pricing kernel obeys the three regularity conditions: NS, RA and DARA. By contrast, the linear pricing kernel violates the NS condition for large market values. In the same spirit, Dybvig and Ingersoll (1982) demonstrate that the CAPM allows for arbitrage opportunities. Further, the linear pricing kernel does not display skewness preference and thus violates the DARA condition. We note that the nonlinear kernel is more flexible than the linear kernel to attach more weight to low probability events, such as market crashes. The next section, kernel specification, further discusses how alternative investor preferences can be modeled through a nonlinear pricing kernel.

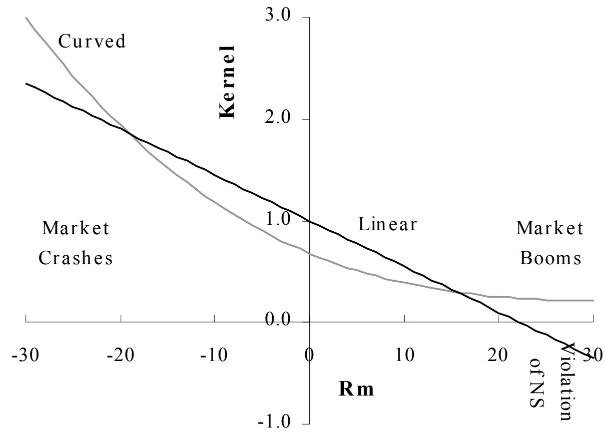


Figure 1.2: Linear and nonlinear pricing kernel. This figure shows a linear pricing kernel (black line) and a nonlinear curved pricing kernel (grey line). The linear kernel (CAPM) violates the nonsatiation (NS) condition during market booms and does not display decreasing absolute risk aversion (DARA). The curved kernel (grey line) obeys all three regularity conditions: NS, risk aversion and DARA.

1.4 Methodology

1.4.1 Kernel specification

In the previous section we discussed how nonlinear pricing kernels could better describe investor preferences, while adhering to the regularity conditions. In the asset pricing literature several nonlinear pricing kernel specifications have been proposed.

First, some models add *higher-order central moments* such as skewness and kurtosis to the mean-variance framework. Kraus and Litzenberger (1976) propose a quadratic kernel and Bansal, Hsieh and Viswanathan (1993) consider higher order polynomials. The higher order kernels seem especially suited for modeling *risk-seeking* ($f'>0$) rather than downside risk aversion. For example, Dittmar (2002) shows that a cubic increasing pricing kernel, which allows for risk seeking ($f'>0$), gives the best empirical fit. However, as discussed in section 1.3.3, regularity conditions, the pricing kernel should be positive and decreasing. In this spirit, Levy, Post and van Vliet (2003) show that the first order conditions (Equation 1.1) do not necessarily hold for a (locally) increasing quadratic pricing kernel. Stated differently, a risk loving investor does not fully diversify because this disproportionally destroys upside potential.

Second, Bawa and Lindenberg (1977) propose the general class of *lower partial moment* (LPM) kernels. A special case of the LPM nonlinear kernels is the *Mean-Semivariance (MS) CAPM* (Hogan and Warren (1974)). The MS CAPM replaces variance by semi-variance and regular beta by downside beta as the relevant risk measure. Interestingly, contrary to the MV CAPM, the MS CAPM has received much attention. An attractive feature of the LPM kernels is that they by definition obey the basic regularity conditions of nonsatiation and risk aversion. This makes this class of nonlinear kernels especially suited for analyzing downside risk.

In general, there is always the chance of kernel misspecification because economic theory only gives minimal guidance on how investor preferences should be modeled. An alternative approach is to use the rules of *Stochastic Dominance* (SD). With SD rules the researcher puts minimal structure on the investor preferences in order to let 'the data speak for themselves'. The strength of the non-parametric approach is that it considers all kernels that obey the regularity conditions. A positive kernel corresponds to (FSD), a positive and decreasing pricing kernel corresponds to second-order SD (SSD) and a positive, decreasing and convex kernel to third-order SD (TSD). Recently, Post (2003) derived computationally tractable empirical tests for SSD efficiency of a given portfolio. This test cleared an important hurdle for practical implementations of SD rules (see quote Haim Levy). With this new test, Post and van Vliet (2004a) find that the value effect can be rationalized in the 1978-1998 period. The weakness of the SD approach is that it does not directly test a particular asset pricing model, but rather a class of asset pricing models. In my opinion, the recently developed SD methodology serves as a helpful and rigorous 'screening device' and indicates whether well-behaved nonlinear kernels exist that better explain asset prices than linear kernels.

> Haim Levy (1992, p.583) '*Ironically, the main drawback of the SD framework is found in the area of finance where it is most intensively used, namely, in choosing the efficient diversification strategies. This is because as yet there is no way to find the SD efficient set of diversification strategies as prevailed by the MV-framework. Therefore the next important contribution will be in this area.*'

1.4.2 Statistical inference

In this section we will discuss how the central pricing equation can be empirically tested. Different approaches exist for asset pricing model evaluation. In general, the pricing kernel is specified in such a way to minimize the *pricing errors* and then the model is evaluated by examining how large those errors are. The pricing errors, or alphas are the empirical deviations from the central pricing equation.

Since the early tests of the CAPM, the asset pricing methodology has evolved substantially (Fama and French (2003)). In the early 1970s, *cross-sectional* OLS regressions (e.g. Fama and MacBeth (1973, FM)) and *time-series* OLS regressions (e.g. Black, Jensen and Scholes (1972)) where used to test the CAPM. Soon afterwards, it was recognized that various statistical problems such as heteroskedasticity and dependency in the errors should be resolved. An important refinement of the time-series methodology was made by Gibbons, Ross and Shanken (1989, GRS) who proposed a multi-variate test statistic. Currently, the Generalized Method of Moments (GMM) framework (for a discussion see textbooks of Campbell, Lo and MacKinlay (1997) and Cochrane (2001)) is popular, because it embeds a broad range of statistical estimation methods, including cross-sectional regressions (FM) and multi-variate time-series regressions (GRS). Additional attractive features of GMM framework are (1) sufficiently flexible to handle a broad range of linear and nonlinear kernels (2) place restrictions on the pricing kernel to obey the regularity conditions and (3) solve the "errors in variables" problem by estimating the risk premium and kernel beta simultaneously.

While we cast our parametric asset pricing tests in the GMM framework, the non-parametric asset pricing methodology is very young and still under development. In general, empirical SD-efficiency tests are somewhat less powerful, that is, they have a lower probability of correctly rejecting the null-hypothesis. This thesis uses the recently developed Post (2003) test. At the time of writing (October 2004), Post and Versijp (2004) cast the SD test in a more powerful GMM-framework. We expect further developments in this area with possible applications for a broader base of representative agent models (e.g. FSD tests and conditional SD tests).

1.5 Empirics

> Eugene Fama (by Roger Ibbotson, 2000): *'I bill it [finance] as the most successful area of economics. The development of the theory, empirical work and computers all joined together to make an explosion of work in this area. We are advantaged relative to other areas of economics because data on markets are much easier to get.'*

The financial researcher faces some empirical issues before the central pricing equation can be tested in practice. This section motivates a number of empirical choices made in this thesis.

1.5.1 Return data

We consider monthly returns of U.S stocks as a proxy for the investment universe and use the 1-month T-bill rate as a proxy for the risk free rate. The stock returns are obtained from the Center for Research in Security Prices (CRSP) at the University of Chicago and the T-bills are obtained from Ibbotson.

The choice for a monthly return interval reflects the trade-off between (1) the sampling error of a sufficiently large sample and (2) a realistic evaluation horizon. Increasing the return interval (e.g. yearly) would lead to small dataset, while decreasing it (e.g. daily) to an unrealistically short evaluation horizon. Further, the use of high-frequency data introduces several micro-structure problems such as the bid-ask bounce which tend to distort the empirical results. Therefore we adhere to the common approach of using monthly returns.

Today, we have about forty years of out-of-sample data available since the first empirical tests of the CAPM. The quality of the CRSP database has improved and the number of assets covered has also substantially increased. We employ common U.S stocks because the CRSP database covers a long time period (1926-2002), includes market capitalization information and is very clean. It is free of the survivorship bias and delisting bias and contains a minimum of typos and missing values due to frequent backfilling updates. We study an extended sample period including the bear markets of the 1930s (see quote de Bondt) 1970s and early 2000s. We use the value-weighted average of all stocks as a proxy for the market portfolio. Unfortunately, to the best of our knowledge, no comparable databases exist which include price information for other asset classes and/or non-US financial assets, yet. Construction of new bias-free databases, including non-equity and non-US assets which cover a long sample period, will certainly be helpful to the field of empirical asset pricing.

> Werner de Bondt (Antwerp 2004): *'Some financial researchers exclude the observations of the 1930s from their analysis. This is similar to throwing out all the divorces from the sample when studying happiness in marriages… '*

1.5.2 Benchmark portfolios

In this thesis, we group stocks into a small set of *benchmark portfolios* based on one or more characteristics, for example historical beta or size (market capitalization).

The number of benchmark portfolios varies from 10 for a single sort and 100 for a double sort on two characteristics.

There are a few reasons why we group assets into portfolios. First, a frequent rebalancing procedure generates portfolios with stable characteristics. For example, an individual stock may wander through time from low beta to high beta, while beta-sorted portfolios always contain firms with similar betas and as a result have more constant betas. Second, a smaller set of portfolios gives a more reliable estimate for the joint return distribution, because with a large number of assets an extremely large number of time-series observations is needed. Third, a multiple-sorting routine proves particularly helpful to disentangle correlated characteristics. For example, small stocks tend to have higher downside betas, thus by sorting on size first and subsequently on downside beta helps to unravel these two effects.

In our analyses, we put more emphasis on portfolio sorts based on stock characteristics that are unrelated to the outcomes of prior empirical research such as regular beta, downside beta or industry classification. Therefore our results are less sensitive to the datasnooping issues as discussed by Lo and MacKinlay (1990).

1.5.3 Time-varying risk

There is substantial evidence that investor preferences and return distributions are *time-varying* and not constant in time. In the spirit of the habit utility models (Campbell and Cochrane (1999)) the CAPM can be specified and tested in a *conditional* way to capture the time-varying nature of risk. The failure of the CAPM may be caused because this dynamic aspect of risk is ignored in an unconditional empirical specification. In fact, there seems to be empirical evidence which suggests that (1) the relative risk of assets and (2) the risk premia change through time. To some extent the risk premia tend to vary with the business cycle, reflecting a higher risk aversion and therefore higher risk premiums in recession periods (e.g. Jagannathan and Wang (1996) and Lettau and Ludvigson (2001)).

However, the use of a conditional model introduces a serious risk of specification error. In order to describe the time varying relation one needs to choose a specific variable to proxy for the state-of-the-world (e.g. dividend yield) and to specify a particular relation (e.g. linear). Ghysels (1998) argues that this relation may also change over time and shows that this mis-specification often results in larger pricing errors. To address this issue we will consider several conditional variables and cross-validate our results with split-sample analyses.

Another serious risk of conditioning is that the regularity conditions (section 1.3.3) are violated. Wang and Zhang (2004)) show that conditional kernels often take large negative values; this means a violation of nonsatiation. Further, the analysis of Lewellen and Nagel (2004)) indicates that in typical empirical tests of conditional asset pricing models, implausible risk premiums imply violations of risk aversion. Therefore we use a new approach, pioneered in this thesis, which puts additional restrictions on the pricing kernel, which ensures us that the pricing kernel obeys the regularity conditions. We explain this in more detail in Chapter 4.

1.6 Outline

The second chapter investigates the empirical performance of the CAPM compared to some well-known multi-factor extensions. We cast the asset pricing tests in a GMM framework to control for the statistical problems and special attention to the portfolio formation procedures. We construct a large number of benchmark sets covering a long sample period and control for market capitalization. The third chapter investigates a large number of well-behaved kernels in a non-parametric fashion. In this chapter we apply the recently developed empirical tests for SD efficiency on some well-known datasets. We test if the high returns on small, value, winner stocks can be explained by other return moments than variance alone. The fourth chapter tests the mean semivariance CAPM, where regular beta is replaced by downside beta. Surprisingly, this model has not been rigorously tested due to methodological and empirical problems. The MS CAPM is evaluated in a GMM framework and we impose the regularity conditions. To better understand downside risk, we construct portfolios based on downside beta and we also carry out conditional tests to capture the time-varying aspect of risk. The final chapter offers a conclusion.

Chapter 2

Do multiple factors help or hurt?[1]

ABSTRACT: This chapter compares the single-factor CAPM with the Fama and French three-factor model and the Carhart four-factor model using a broad cross-section and long time-series of U.S stock portfolios and controlling for market capitalization. Confirming known results, multiple factors help for value and momentum portfolios in the post-1963 period, most notably for the small cap market segment. However, multiple factors generally do not help or even hurt (1) in the pre-1963 period, (2) for size, beta, reversal, and industry portfolios and (3) within the large cap market segment. These empirical findings support the data snooping hypothesis or other non-risk based explanations such as high transaction costs and low market liquidity for small caps.

2.1 Introduction

CONCERNS ABOUT THE EMPIRICAL VALIDITY OF THE SINGLE-FACTOR CAPM explain much of the current popularity of multifactor asset pricing models. Various empirical studies suggest that the stock market portfolio is highly and significantly mean-variance inefficient relative to portfolios formed on stock characteristics such as market capitalization (size), book-to-market-equity ratio (BE/ME) and price momentum. In response to these findings, several multifactor models have been developed. Currently, the most popular models are the three-factor model (3FM; Fama and French (1993) and Fama and French (1995)) and the four-factor model (4FM; Carhart et. al (1996) and Carhart (1997)).

The economic rationale behind the multifactor models is not entirely clear. The models can be motivated and interpreted with various different theoretical alternatives to the CAPM, ranging from Merton's (1973) ICAPM to Ross' (1976) APT. Nevertheless, from an instrumentalist perspective, the empirical performance of these models may be regarded as more important than the precise theoretical underpinnings. Surprisingly, the existing empirical studies that analyze the single-

1 This chapter is based on the paper titled: 'Do multiple factors help or hurt?' by Post and van Vliet (2004d). This paper can also be found at: http://ssrn.com/abstract=582101

factor or multifactor models are far from decisive on the added value of multiple factors.

Existing studies typically consider a relatively narrow cross-section of size, BE/ME (or other related multiples) and momentum based portfolios and a short time-series of post-1963 data. This approach introduces the risk of data snooping, because the empirical problems of the CAPM for this type of data were known prior to the formulation of the multifactor models (see also Lo and MacKinlay (1990)). For this reason, it would be useful to cross-validate the existing results with other cross-sectional effects, including the pre-1963 period.

Further, existing studies do not always control in a satisfactory way for the size of the stocks under evaluation. The small caps generally involve relatively low market liquidity and high transaction costs. Transaction costs decrease with firm size due to more favorable (1) bid-ask spreads, (2) commission fees and (3) price impact of trade for large caps. With a new measure which incorporates all three effects, Lesmond, Ogden and Trzcinka (1999) estimate the round-trip transaction costs to be 1.2 percent for largest decile stocks and 10.3 percent for smallest decile stocks (NYSE/AMEX stocks, 1963-1990). Similarly, Pastor and Stambaugh (2003) show that stock liquidity increases with size. Due to these microstructure problems, the results for small caps may have less economic significance and practical relevance to investors than the results for large caps. For this reason, it is interesting to analyze the interaction of size with the other cross-sectional effects, again including the pre-1963 period.

Several studies suggest that the sample period and size segment are important for judging the empirical validity of the CAPM. For example, Loughran (1997) shows that a substantial portion of the BE/ME effect is driven by the low returns of small newly-listed growth stocks. Further, Ang and Chen (2003) Ang and Chen (2003)show that the BE/ME effect disappears in the pre-1963 period. However, these studies do not consider the combined effect of the benchmark set, the sample period and the size segment. Also, these studies focus on the CAPM and do not directly address the issue of the added value of multiple factors. Similarly, studies that focus on multiple factor models without analyzing the CAPM, such as Davis, Fama and French (2000) also cannot determine the added value.

The purpose of this chapter is to systematically analyze the empirical performances of the CAPM and the multifactor models (3FM and 4FM) using a broad cross-section and a long time-series of stock portfolios and controlling for size. We employ benchmark portfolios based on beta, reversal and industry in addition to the typical size, BE/ME and momentum portfolios. Further, we analyze the pre-1963 period in addition to the typical post-1963 period. To control for size, we construct size-based double-sorted portfolios and analyze multifactor efficiency within each size segment separately.

Our main findings (summarized in Table 2.1) are remarkable given the current popularity of multifactor models. Most notably, multiple factors do not significantly improve the fit for size, beta and reversal portfolios. Also, there are no improvements for BE/ME portfolios in the pre-1963 period and in the large cap segment, roughly 80% of the total stock market capitalization. Further, multiple factors significantly worsen the fit for industry portfolios and for large cap momentum portfolios. Combined with the weak theoretical underpinnings, these findings lead us to seriously question the use of multifactor models.

Table 2.1
Summary of test results

This table summarizes the outcomes of our GMM test for multifactor efficiency. Specifically, it shows the p-value associated with the null that the alphas are jointly equal to zero. Section I and II give further details about the data and methodology use to arrive at these results. P-values below the significance level of 10% are highlighted.

Formatted: Font: 9 pt

Characteristic	Period	Size Segment	CAPM	3FM	4FM	Help or hurt?
Size	1931-1962	All	0.95	0.96	0.80	-
BE/ME	1931-1962	All	0.34	0.27	0.65	-
Momentum	1931-1962	All	**0.01**	**0.00**	0.36	Help
Beta	1963-2002	All	0.21	0.21	0.15	-
Reversal	1963-2002	All	0.73	0.99	0.75	-
Industry	1963-2002	All	0.54	**0.02**	**0.03**	Hurt
BE/ME	1963-2002	Large	0.46	0.31	0.37	-
Momentum	1963-2002	Large	0.12	**0.01**	**0.01**	Hurt
Beta	1963-2002	Large	0.37	0.93	0.98	-
Reversal	1963-2002	Large	0.44	0.95	0.39	-
Industry	1963-2002	Large	0.16	**0.00**	**0.01**	Hurt

Our empirical results complement the theoretical results of MacKinlay (1995). He demonstrates that omitted risk factors are unlikely to cause large deviations from the CAPM and that non-risk based explanations, including data-snooping, are more likely to yield large deviations. In contrast to MacKinlay, we use a purely empirical approach, showing that the violations of CAPM and the improvements from using multiple factors are not robust with respect to the choice of the sample period and the set of benchmark portfolios. These findings are consistent with the datasnooping explanation as well as with other non-risk based explanations, such as transaction costs and market liquidity for the small cap stock market segment.

This chapter proceeds as follows. Section 2.2 motivates and explains our methodology. Next, Section 2.3 explains our procedure for selecting stocks and forming benchmark portfolios. Section 2.4, describes and compares the performance of the CAPM and multifactor models relative to the different benchmark sets, the different time periods and different size segments. Finally, Section 2.5 summarizes our findings and gives suggestions for further research.

2.2 Methodology

Since the early time-series tests (see for instance Black, Jensen and Scholes (1972) and cross-sectional tests (see for instance Fama and MacBeth (1973)) the empirical asset pricing methodology has evolved substantially. In this chapter, we will use the pooled time-series-cross-section framework of Generalized Method of Moments (GMM) (see e.g. MacKinlay and Richardson (1991)). Currently, this framework is popular, because it can resolve various methodological problems associated with the early tests, such as heteroskedasticity and correlation for the alphas (see below).

The three models analyzed in this chapter (CAPM, 3FM, 4FM) all link the excess return on a set of risky assets $r_1, \cdots, r_N$ to the returns on a set of orthogonal factor-mimicking portfolios (or factors) $f_1, \cdots, f_K$. In the CAPM, the only relevant factor is the excess market return. The 3FM introduces two additional factors, the "Small Minus Big" (SMB) hedge portfolio return and the "High Minus Low" (HML) hedge portfolio return, and the 4FM further adds the "Winner Minus Loser" (WML) hedge portfolio return as a factor. These additional factors are intended to capture common non-market risk factors that are related to size, BE/ME and momentum.

The equilibrium condition for the general multifactor model is given by the following risk-return relationship:

$$E[r_i] = \sum_{j=1}^{K} \beta_{ij} E[f_j] \quad i = 1, \cdots, N \tag{2.1}$$

with $\beta_{ij} \equiv Cov[r_i, f_j]/Var[f_j]$ for the factor risk loading of the i-th asset for the j-th factor and $E[f_j]$ for the j-th factor risk premium. In words, the expected excess return of every asset $E[r_i]$ should equal the sum product of the asset's factor risk loadings and the factor risk premiums. In this case, the market is said to be multifactor efficient.

In practice, we cannot directly check multifactor efficiency, because the return distribution of the assets and factors is unknown. However, we can estimate the distribution using time-series observations for the assets and factors and employ statistical tests to determine if the equilibrium condition (2.1) is violated to a statistically significant degree. We will represent the return observations for the assets and the factors by $r_{i1},\cdots,r_{iT}$ and $f_{j1},\cdots,f_{jT}$ respectively.

Using the observations, we can compute the following empirical deviations from the equilibrium condition (2.1) or pricing errors:

$$\hat{\alpha}_i \equiv \bar{r}_i - \sum_{j=1}^{K} \hat{\beta}_{ij} \bar{f}_j \quad i = 1,\cdots,N \tag{2.2}$$

with $\bar{r}_i$ for the sample return mean of the *i*th asset, $\hat{\beta}_{ij}$ for the sample risk loading of the *i*th asset for the *j*-th factor and $\bar{f}_j$ for the sample risk premium of the *j*th factor. To test multifactor efficiency, we need to test if the alphas are jointly significantly different from zero.[2] For this purpose, we can aggregate the individual alphas with the following test statistic:

$$JT \equiv \sum_{i=1}^{N} \sum_{s=1}^{N} w_{is} \hat{\alpha}_i \hat{\alpha}_s \tag{2.3}$$

with w_{is}, $i,s = 1,\cdots,N$, for properly chosen weights. Assuming that the observations are serially independently and identically distributed (IID) random draws, the *JT* test statistic (3) obeys an asymptotic chi-squared distribution with *N* degrees of freedom.[3] If we set $w_{is} = 1$ if $i = s$ and $w_{is} = 0$ otherwise, then the *JT* statistic reduces to the sum of squared errors used in OLS regressions. This is a good weighting scheme if the alphas are homoskedastic and independent. However, other weighting schemes are

2 Multifactor asset pricing models should not be judged based on their ability to explain time-series return variations, measured by for example the R-squared of time-series regressions. The CAPM does not predict that the market return is the only common risk factor and allows for non-market risk factors related to for example size, BE/ME and momentum. Rather, the CAPM predicts that the market return is the only common risk factor with a non-zero risk premium. To use the words of Davis, Fama and French (2000, p. 390)): "The acid test of a multifactor model is whether it explains differences in average returns."

3 The choice of *N* degrees of freedom is related to the use of benchmark portfolios rather than individual assets. For individual assets, *N-K* degrees of freedom is more appropriate. The *K* factor portfolios by construction have zero alphas (see also Table II in Section IIA). This has the effect of reducing the alphas of the individual assets that constitute the factor portfolios. In fact, in case *N*=*K*, all assets will have zero alphas and *JT*=0 by construction. In general, *JT* behaves as the sum of squares of (*N-K*) IID distributed random variables. However, in the empirical analysis, we will aggregate the large number of individual assets to a small number of benchmark portfolios and test multifactor efficiency for those benchmark portfolios. While the factor portfolios are linear combinations of the individual assets, they generally cannot be formed by combining the benchmark portfolios. Thus, the loss of degrees of freedom does not occur and we use *N* degrees of freedom (rather than *N-K*).

required in case of heteroskedasticity and correlation. These problems are especially relevant in the context of our study, because small caps generally have highly volatile and highly correlated alphas (see Section 2.3 below). This may yield misleading conclusions when forming benchmark portfolio on size or stock characteristics that are associated with size. In this chapter, we will use the well-known Hansen and Jagannathan (1997) weighting scheme, in which the weights equal the elements of the inverted sample variance-covariance matrix of the asset returns.[4] The associated Hansen and Jagannathan distance measure $HJ \equiv \sqrt{JT/T}$ can be interpreted as the maximal weighted pricing error.

We stress that our methodology estimates the factor risk premiums $E[f_1],\cdots,E[f_K]$ with their sample equivalents $\bar{f}_1,\cdots,\bar{f}_K$, in the spirit of the early time-series tests and the Gibbons, Ross and Shanken (1989) and MacKinlay and Richardson (1991) tests. An alternative approach is to select the estimates that optimize the empirical fit (or minimize the JT statistic), in the spirit of the early cross-section tests. An important drawback of this approach is that it may select "implausible" values for the risk premiums. For example, Wang and Zhang (2004) show that fitted risk premiums often imply arbitrage opportunities and thus violate the basic no-arbitrage principle that underlies all asset pricing models. By contrast, the sample estimates generally take "plausible" values that are arbitrage-free. Related to this, fitting the risk premiums yields less statistical power (probability of detecting multifactor inefficiency) than fixing the risk premiums. After all, we can always select the estimates such that K out of the N alphas are equal to zero. Further, fixing the estimates ensures that a single set of factor risk premiums is used for all benchmark sets. Hence, we cannot for example rationalize the BE/ME effect with one set of risk premiums and rationalize the momentum effect with another set of risk premiums. For these reasons, we *fix* rather than *fit* the risk premiums in the study.

4 A common alternative for the Hansen-Jagannathan weights is to use the "optimal" weights, which correct for the statistical distribution of the risk factors in addition to the asset returns distribution. In our analysis, the "optimal" weights generally yield somewhat lower values for the JT statistic, suggesting a better fit. However, the conclusions regarding the relative goodness of the three models are not affected. Hence, for the sake of brevity, we report the results for the Hansen-Jagannathan weights only.

2.3 Data

2.3.1 Data sources

Our empirical analysis uses individual stock returns, index returns and hedge portfolio returns. The monthly stock returns (including dividends and capital gains) are from the Center for Research in Security Prices (CRSP) at the University of Chicago. The market index is a value-weighted average of all U.S stocks included in this study. The one-month U.S Treasury bill is obtained from Ibbotson Associates. The monthly hedge portfolio returns (SMB, HML, and WML) are taken from the data library of Kenneth French.

Table 2.2 includes descriptive statistics for the four risk factors used in this chapter (market, SMB, HML and WML). The table lists the means and the alphas and associated standard errors for the three competing models (CAPM, 3-FM, and 4-FM). Note that the alphas of the SMB, HML and WML factors differ from the means of these factors, which implies that the risk factors are not orthogonal. This does not affect the computation of the alphas or the statistical inference about the alphas. However, it does affect the interpretation of the factor means as risk premiums. For example, from the high HML mean in the pre-1963 period (0.43% per month), documented also in Fama and French (2000), we may suspect that there is a sizable value premium. However, the CAPM alpha of the HML portfolio is close to zero (0.06% per month) and hence the value premium disappears after correcting for market risk. Note also that the 3-FM alpha of the WML momentum factor is very high in both periods, suggesting a severe momentum problem for the three-factor model.

Table 2.2
Descriptive Statistics for the common risk factors

This table shows the average returns and pricing errors of the common risk factors: market return (market), Small Minus Big (SMB), High Minus Low (HML) and Winner Minus Loser (WML) in the three competing models (CAPM, 3FM and 4FM). Results are shown for the post-1963 period (January 1963-December 2002) and the pre-1963 period (January 1931- December 1962). Standard errors are given in brackets. The alphas of the factors that enter in a model are zero by construction.

	Post-1963				Pre-1963			
	Avg	CAPM $\hat{\alpha}$	3FM $\hat{\alpha}$	4FM $\hat{\alpha}$	Avg	CAPM $\hat{\alpha}$	3FM $\hat{\alpha}$	4FM $\hat{\alpha}$
market	0.46	0.00	0.00	0.00	0.93	0.00	0.00	0.00
	(0.20)	(0.00)	(0.00)	(0.00)	(0.34)	(0.00)	(0.00)	(0.00)
SMB	0.21	0.13	0.00	0.00	0.36	0.18	0.00	0.00
	(0.15)	(0.15)	(0.00)	(0.00)	(0.18)	(0.17)	(0.00)	(0.00)
HML	0.26	0.57	0.00	0.00	0.43	0.06	0.00	0.00
	(0.14)	(0.14)	(0.00)	(0.00)	(0.22)	(0.18)	(0.00)	(0.00)
WML	0.49	0.94	1.07	0.00	0.47	0.89	0.93	0.00
	(0.18)	(0.19)	(0.19)	(0.00)	(0.29)	(0.25)	(0.23)	(0.00)

2.3.2 Data requirements

We select ordinary common U.S stocks listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX) and NASDAQ markets. We exclude ADRs, REITs, closed-end-funds, units of beneficial interest, and foreign stocks. Hence, we only include stocks that have a CRSP share type code of 10 or 11. We require a stock to have 60 months of prior return data available and information about the market capitalization (defined as price times the number of outstanding shares) at formation date. The past returns are needed for calculating beta and long-term performance (reversal) and the market capitalization is required for computing value-weighted returns. Portfolio formation takes place at December of each year (except for momentum portfolios which are formed every month). Thus, to be included at December 1930, a stock must have trading information since January 1926 and a (positive) market capitalization for December 1930. A stock is excluded from the analysis if trading information is no longer available. In case of exclusion, the delisting return or partial monthly return provided by CRSP is used for the last return observation. On average, 1,854 stocks are included in the analysis, starting with 373 (in December 1930) and ending with 3,730 (in December 2002) after reaching a maximum of 3,907 (January 1999).

We use a long time-series that includes the pre-1963 period. For several reasons, the year 1963 is an important date. Prior to 1963, the Compustat database is affected by a survivorship bias. This bias is caused by the back-filling procedure excluding delisted firms, which typically are less successful (Kothari, Shanken and Sloan (1995)). In the context of our study, only the BE/ME portfolios require accounting data (the book value of equity).[5] By contrast, the size, momentum, beta, reversal and industry portfolios do not require Compustat information. Further, from June 1962, AMEX-listed stocks are added to the CRSP database, which includes only NYSE-listed stocks before this month. Since AMEX stocks generally are smaller than NYSE stocks, the relative number of small caps in the analysis increases from June 1962. In 1973, NASDAQ stocks are added to the CRSP database, which further skews the size distribution.

2.3.3 Single-sorted portfolios

Portfolios are formed at the end of each year, starting in December 1930, 60 months after the start of the CRSP reporting for individual stocks (January 1926). Thus, the portfolio returns cover the period from January 1931 to December 2002. We sort stocks into 10 deciles based on a given stock characteristic and compute value-weighted portfolio returns of all stocks in each decile. The stock characteristics used in this chapter are: size, BE/ME, momentum, beta, industry and reversal.

The size, BE/ME and momentum portfolios are more or less "standard". Size and BE/ME portfolio portfolios are taken from Kenneth French homepage.[6] Momentum portfolios are based on the price performance during the period from 12

---

5 We may circumvent this problem by replacing the book-to-market-equity ratio with the HML-beta (or the sensitivity for the HML factor), which does not require accounting information. To analyze if our results for the value portfolios are affected by the Compustat survivorship bias, we form HML-beta portfolios (using our own data requirements and formation procedure). Interestingly, the portfolio characteristics and the test results are not materially affected by the choice between BE/ME and HML-beta. For this reason, we report the results for the familiar BE/ME portfolios only.

6 We use Kenneth French' portfolios because of familiarity and availability to the general readership. The data requirements and portfolio formation procedures used to arrive at these portfolios are somewhat different from the ones we use for the other benchmark sets:

(1) To avoid a look-ahead bias, which arises when using accounting data, Fama and French construct portfolios at the end of June. By contrast, we construct portfolios at the end of December.

(2) For BE/ME portfolios, Fama and French required a firm to have Compustat/Moody's accounting information, which limits the number of included stocks and may introduce survivorship bias. By contrast, we rely solely on the 200212 CRSP database.

(3) For BE/ME portfolios, Fama and French exclude the financials because these firms tend to be highly levered. For all other sorts, the financials are included in the analysis.

(4) Fama and French employ NYSE decile breakpoints to limit the role of small caps in the post-1963 period. We employ NYSE/AMEX/NASDAQ breakpoints and analyze the small cap market segment seperately.

We also form size and value portfolios using our own data requirements and formation procedure. However, the portfolio characteristics and the test results are not materially affected by the choice between the two sets of size and value portfolios. For this reason, we report the results for Kenneth French' portfolios only.

months ago to one month ago (past 12-1 month returns), similar to the procedure of Fama and French (1996). For momentum, portfolio formation takes place on a monthly basis rather than annually.

There are several good reasons for analyzing beta, industry and reversal portfolios in addition to the "standard" portfolios. Beta portfolios can be motivated by theory, as the CAPM stipulates that beta drives average returns. In fact, forming portfolios on other stock characteristics may lead to erroneous rejections of the CAPM due to a lack of variation in the betas and means. Our beta portfolios are based on the historical 60-month betas of the individual stocks.

Contrary to some other benchmark sets, industry portfolios are not suspected of data-snooping problems, because the industry classification does not depend on the outcomes of prior empirical research. Our industry portfolios are based on the 4-digit SIC codes given on Kenneth French' website.

Reversal portfolios are directly related to BE/ME portfolios, because long-term losers (winners) tend to have low (high) prices and hence favorable (unfavorable) multiples. However, contrary to BE/ME portfolios, reversal portfolios do not require accounting data, which makes these portfolios useful for analyzing the pre-1963 period without using Compustat data. Our reversal portfolios are based on the long-term price performance (past 60-month returns), following the definition ofDeBondt and Thaler (1985).

At this point, it is useful to look in closer detail at the problems surrounding small caps. As discussed in Section 2.2, the high volatility and high correlation of small caps introduces the need to weight the alphas. Panel A of Table 2.3 illustrates this point with the sample variance and sample correlation coefficients of the ten size portfolios. For example, the sample variance of the small cap portfolio (ME1) is 88.83, while the sample variance for the large cap portfolio (ME10) is only 21.77, a ratio of four to one. Also, the correlation between ME1 and ME2 is 0.95, while the correlation between ME1 and ME10 is only 0.66. These findings suggest that the alphas of small caps are less reliable than the alphas of large caps and that the alpha of one small cap portfolio contains only little information in addition to alpha of another small cap portfolio. This clearly introduces the need to weight the alphas.

Table 2.3
Size-related heteroskedasticity and correlation and the size distribution

Panel A shows the sample variance and the correlation coefficients for the ten size (ME) decile portfolios for the January 1963 to December 2002 period (T=480). Panel B shows the average firm size within the portfolio (size in billion $) and total portfolio size (% of total market capitalization) of the single-sorted portfolios based on size, BE/ME, momentum, beta, reversal and industry for the January 1963 to December 2002 period (T=480). The size and BE/ME portfolios are taken from Kenneth French's website. The momentum portfolios are constructed at the end of each month based on momentum (cumulative past 12-1 month price performance). The beta portfolios are constructed based on 60-month historical betas. Reversal portfolios are constructed based on long-term return performance (cumulative past 60-month price performance). Industry portfolios are based on the 4-digit SIC codes (1 nondurables, 2 durables, 3 oil, 4 chemicals, 5 manufacturing, 6 telecom, 7 utilities, 8 shops, 9 financials and 10 all other).

Panel A: Size-related heteroskedasticity and correlation

Port.	Var.	Correlation coefficients									
		1	2	3	4	5	6	7	8	9	10
ME1	88.83	1.00	0.95	0.92	0.91	0.88	0.84	0.83	0.78	0.76	0.66
ME2	70.02		1.00	0.97	0.96	0.94	0.92	0.91	0.87	0.84	0.75
ME3	55.23			1.00	0.98	0.97	0.95	0.94	0.91	0.89	0.80
ME4	49.58				1.00	0.98	0.97	0.96	0.93	0.91	0.82
ME5	44.98					1.00	0.98	0.97	0.96	0.93	0.85
ME6	39.67						1.00	0.98	0.97	0.95	0.88
ME7	37.28							1.00	0.98	0.97	0.90
ME8	31.79								1.00	0.98	0.92
ME9	28.38									1.00	0.94
ME10	21.77										1.00

Panel B: Size distribution

	Size (ME)		BE/ME		Mom		Beta		Rev		Ind	
Port.	Size	%	Size	%	Size	%	Size	%	Size	%	Size	%
1	0.02	0%	1.38	25%	0.20	3%	0.44	8%	0.12	3%	0.98	8%
2	0.07	1%	1.10	15%	0.44	6%	0.76	12%	0.24	5%	1.28	8%
3	0.13	1%	0.88	12%	0.72	9%	1.11	15%	0.36	6%	1.65	11%
4	0.20	1%	0.75	9%	0.92	11%	1.09	12%	0.53	8%	1.88	11%
5	0.32	2%	0.70	9%	1.08	12%	1.06	11%	0.75	9%	0.74	24%
6	0.48	3%	0.56	8%	1.16	13%	1.21	12%	0.98	11%	3.57	6%
7	0.76	5%	0.51	8%	1.24	13%	1.33	11%	1.22	13%	1.08	8%
8	1.27	8%	0.42	6%	1.26	13%	1.03	8%	1.46	15%	0.69	10%
9	2.44	15%	0.32	5%	1.22	13%	0.86	7%	1.63	15%	0.86	11%
10	11.83	64%	0.15	3%	0.81	8%	0.49	4%	1.99	15%	0.40	3%

There is a strong association between size and the other stock characteristics (BE/ME, momentum, beta, reversal and industry). Panel B of Table 2.3 illustrates this problem by means of the average market capitalization and the total market capitalization of the single-sorted portfolios. For example, the growth portfolio (BE/ME1) involves an average size of $1.38bn and represents 25% of the total market capitalization, while the value portfolio (BE/ME10) involves an average size of $0.15bn and represents 3% of the total market capitalization. Further, the telecom industry portfolio consists of a small number of relatively large stocks ($3.57bn), while the manufacturing industry portfolio includes a high number of relatively small stocks ($0.74bn). In general, value stocks, short-term loser stocks, low-beta stocks and long-term loser stocks tend to be smaller. Hence, the small cap problems, ranging from the high volatility and high correlation of the alphas to the low market liquidity and high transaction costs, are passed on to the other benchmark portfolios.

2.3.4 Double-sorted portfolios

The weighting of the alphas (see Section 2.2) can correct for the differences of volatility and correlation between the alphas. However, this approach cannot correct for the size-related problems of market liquidity and transaction costs. Instead, to control for size, we construct sets of 25 double-sorted portfolios by first sorting the individual stocks into size quintiles (NYSE breakpoints) and next forming quintiles based on another stock characteristic within each size segment. We will subsequently analyze multifactor efficiency within each size quintile separately.

We use the well-known 25 Fama and French size-BE/ME portfolios and complement these portfolios with 25 size-momentum portfolios, 25 size-beta portfolios, 25 size-reversal portfolios and 38 size-industry portfolios. The size-industry portfolios are constructed by sorting on the 4-digit SIC code within every size quintile, yielding 50 portfolios in total. Due to some size-industry combinations not occurring, 12 out of these 50 portfolios are empty for some period of time. For example, the telecom industry includes no small caps throughout the 1930s to 1970s. Excluding these 12 portfolios yields 38 size-industry portfolios. [7]

The data selection criteria and sorting frequency for the double-sorted portfolios are identical to those used for the single-sorted portfolios. All data used in this dissertation are publicly available.[8]

---

7 The 12 missing size-industry portfolios are: oil (small), telecom (small, 2, 3 and 4), utilities (small, 2 and 3) and financials (small, 2, 3 and 4).

8 Monthly return data of the single-sorted benchmark portfolios (momentum, beta, reversal, industry), double-sorted portfolios (size-momentum, size-beta, size-reversal) and the market index (consisting of all stocks included in this study) are available at our online datacenter:

www.few.eur.nl/few/people/wvanvliet/datacenter.

Hedge factors and size (ME) and value (BE/ME) portfolio return data can be found at Kenneth French datalibrary:

http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

2.4 Results

Our test results are summarized in Table 2.4 (size, BE/ME and momentum portfolios), Table 2.5 (beta, reversal and industry portfolios) and Table 2.6 (size quintiles). Below, we discuss the main conclusions that can be drawn from these results.

2.4.1 Size, BE/ME and momentum portfolios

Panel A of Table 2.4 shows the results for the post-1963 period. The results for this period confirm the existing empirical evidence for the multifactor models. Specifically, the CAPM cannot be rejected for size, but has to be rejected for the BE/ME and momentum portfolios. By contrast, the 3FM and 4FM cannot be rejected for the BE/ME portfolios, but still have to be rejected for momentum portfolios.

Consistent with the findings of Gibbons, Ross and Shanken (1989) and Fama and French (1992), there seems to be no "size effect". Substantial improvements do occur for the individual alphas of the size portfolios. Most notably, the alpha for the small cap portfolio (ME1) goes from 0.23 (CAPM) to -/-0.07 (3FM) and -/-0.09 (4FM). However, the small cap alphas are highly volatile and highly correlated (see Table 2.3 panel A). This is reflected in the *JT* statistic (which corrects for heteroskedasticity and correlation) not being significantly greater than zero for the CAPM (p=0.69) and not improving significantly for the multifactor models. Hence, no stand-alone size effect occurs. Still, there is an important interaction effect between size and the other stock characteristics; the BE/ME and momentum effects are most pronounced in the small cap and midcap market (see Section 2.3.4).

Table 2.4
Size, value and momentum portfolios

This table gives the test results for the portfolios based on size, BE/ME (value), and momentum for the three competing asset pricing models (CAPM, 3FM and 4FM). The size and BE/ME portfolios are taken from Kenneth French's website. The momentum portfolios are constructed at the end of each month based on momentum (cumulative past 12-1 month price performance). Every row shows the alphas (Equation 2.2), the JT test statistic (Equation 2.3) and the associated Hansen Jagannathan (1997) distance and p-value. Test results are shown for the post-1963 sample from January 1963 to December 2002 (T=480) and the pre-1963 sample from January 1931 to December 1962 (T=384). P-values below the significance level of 10% are highlighted.

Panel A: Post 1963

	Alphas	1	2	3	4	5	6	7	8	9	10	JT	HJ	p
Size	CAPM	0.23	0.13	0.15	0.11	0.15	0.05	0.08	0.06	0.02	-0.06	7.4	0.124	0.69
	3FM	-0.07	-0.09	-0.07	-0.08	-0.01	-0.09	-0.03	-0.04	-0.04	0.04	5.8	0.110	0.83
	4FM	-0.09	-0.11	-0.05	-0.08	0.00	-0.08	-0.04	-0.02	-0.05	0.04	6.9	0.120	0.73
Value	CAPM	-0.19	-0.03	-0.02	-0.02	0.05	0.17	0.27	0.32	0.36	0.39	16.3	0.184	**0.09**
	3FM	0.15	0.04	-0.05	-0.14	-0.10	-0.03	-0.04	-0.07	-0.07	-0.17	6.0	0.112	0.81
	4FM	0.13	0.06	-0.01	-0.14	-0.12	-0.03	-0.02	-0.07	-0.03	-0.16	8.1	0.130	0.62
Mom	CAPM	-1.18	-0.48	-0.31	-0.06	-0.19	-0.08	0.07	0.23	0.26	0.65	43.8	0.302	**0.00**
	3FM	-1.25	-0.60	-0.41	-0.16	-0.29	-0.16	0.03	0.20	0.27	0.79	51.2	0.327	**0.00**
	4FM	-0.13	0.24	0.23	0.24	-0.09	-0.10	-0.08	-0.07	-0.18	0.09	30.3	0.251	**0.00**

Panel B: Pre 1963

	Alphas	1	2	3	4	5	6	7	8	9	10	JT	HJ	p
Size	CAPM	0.53	0.25	0.21	0.19	0.08	0.14	0.09	0.05	0.03	-0.01	4.0	0.091	0.95
	3FM	0.14	-0.06	-0.02	0.00	-0.08	0.04	0.00	0.00	0.01	0.01	3.7	0.088	0.96
	4FM	0.36	-0.01	0.01	-0.02	-0.05	0.02	-0.05	0.00	0.03	0.01	6.2	0.114	0.80
Value	CAPM	-0.01	0.04	0.01	-0.16	0.12	-0.09	-0.13	0.13	0.06	-0.07	11.3	0.153	0.34
	3FM	0.01	0.04	0.01	-0.16	0.11	-0.11	-0.19	0.08	-0.05	-0.29	12.2	0.159	0.27
	4FM	0.02	0.05	0.02	-0.10	0.13	-0.08	-0.11	0.08	0.02	-0.17	7.8	0.127	0.65
Mom	CAPM	-0.84	-0.53	-0.42	-0.22	-0.11	-0.10	0.04	0.31	0.28	0.66	23.7	0.222	**0.01**
	3FM	-0.96	-0.56	-0.43	-0.21	-0.10	-0.12	0.05	0.32	0.28	0.65	25.6	0.231	**0.00**
	4FM	-0.18	0.14	0.05	0.06	0.17	-0.02	-0.03	0.12	-0.09	0.12	11.0	0.151	0.36

Several further points should be mentioned. The 3FM actually yields a worse fit than the CAPM for the momentum portfolios (*JT* goes from 43.8 to 51.2), which confirms the findings of Fama and French (1996) and the high 3FM alpha of the WML portfolios (see Table 2.2). Adding momentum as a fourth risk factor does not help to better explain the returns on size and BE/ME portfolios. Further, despite the improvements for the individual pricing errors and the overall fit, the 4FM still has to be convincingly rejected for the momentum portfolios (p=0.00).

Panel B of Table 2.4 shows the results for the pre-1963 period. For the size portfolios, more or less the same pattern emerges as for the post-1963 period: the

CAPM cannot be rejected and the multifactor models do not yield significant improvements.

Interestingly, the BE/ME effect disappears in the pre-1963 period, consistent with the finding of Ang and Chen (2003). Specifically, the CAPM alphas are much smaller and not jointly significantly different from zero (p=0.34). The 3FM actually yields a somewhat worse fit than the CAPM (most notably, the alpha of the BE/ME portfolio #10 deteriorates from -/-0.07 to -/-0.29). The 4FM improves the fit for the book/market portfolios, albeit not to a significant degree. At first sight, the disappearing of the value effect seems inconsistent with the high HML mean in the pre-1963 period, which was used by Fama and French (2000) as evidence for the robustness of the value effect. However, as discussed in Section 2.3.1, the HML portfolio is not market-neutral and no value premium remains after correcting for market risk in the pre-1963 period.

The momentum effect is persistent and continues to challenge the CAPM in the pre-1963 period. Again the 3FM is also unable to explain the returns on momentum portfolios. Adding a momentum factors shows a substantial improvement relative to the CAPM (*JT* goes from 23.7 to 11.0). In fact, due to the alphas generally being smaller than in the post-1963 period, the 4FM now cannot be rejected (p=0.36).

In sum, if we move from the post-1963 period to the pre-1963 period, the rationale for using the 3FM disappears, while the 4FM is useful to capture the momentum effect.

2.4.2 Beta, reversal and industry portfolios

The empirical successes for post-1963 BE/ME portfolios (3FM and 4FM) and momentum portfolios (4FM only) are well documented and are often used to justify the use of the multifactor models. However, the empirical validity of these models for other types of benchmark portfolios is not well documented. Table 2.5 shows our results for the beta, reversal and industry portfolios, again using post-1963 and pre-1963 data.

For the beta portfolios, the CAPM cannot be rejected (with a post-1963 p-value of 0.21 and a pre-1963 p-value of 0.65). Also, the multifactor models do not improve the fit; in fact, they generally lead to a small deterioration. Important improvements of the alphas do occur for the high beta portfolios. However, these alphas are relatively unreliable (high beta portfolios are dominated by small caps; see Table 2.3 panel B) and the improvements are offset by deteriorations of the alphas of the medium beta portfolios (which are dominated by large caps and have more reliable alphas). Thus, multiple factors do not improve the fit for beta portfolios. This finding is remarkable, because beta portfolios are an obvious choice for testing the CAPM.

Table 2.5
Beta, Reversal and Industry Portfolios

This table gives the test results for the decile portfolios based on beta, reversal and industry for the three competing asset pricing models (CAPM, 3FM and 4FM). The beta portfolios are constructed based on 60-month historical betas. Reversal portfolios are constructed based on long-term return performance (cumulative past 60-month price performance). Industry portfolios are sorted on 4-digit SIC codes (1 nondurables, 2 durables, 3 oil, 4 chemicals, 5 manufacturing, 6 telecom, 7 utilities, 8 shops, 9 financials and 10 all other). Every row shows the alphas (2), the *JT* test statistic (3) and the associated Hansen Jagannathan (1997) distance and p-value. Test results are shown for the post-1963 sample from January 1963 to December 2002 (*T*=480) and the pre-1963 sample from January 1931 to December 1962 (*T*=384). P-values below the significance level of 10% are highlighted.

	Alphas	1	2	3	4	5	6	7	8	9	10	JT	HJ	p
Panel A: Post 1963														
Beta	CAPM	0.21	0.19	0.02	-0.02	0.11	-0.09	-0.10	-0.22	-0.20	-0.30	13.2	0.166	0.21
	3FM	-0.02	0.08	-0.09	-0.13	0.02	-0.16	-0.16	-0.25	-0.12	-0.18	13.3	0.166	0.21
	4FM	-0.09	0.01	-0.14	-0.16	0.04	-0.15	-0.08	-0.08	0.06	-0.04	14.6	0.174	0.15
Rev	CAPM	0.41	0.30	0.30	0.22	0.19	0.16	0.08	0.00	-0.10	-0.19	7.0	0.121	0.73
	3FM	-0.10	-0.05	-0.01	-0.02	0.00	0.00	-0.05	-0.05	-0.05	0.08	2.6	0.073	0.99
	4FM	0.41	0.34	0.24	0.15	0.10	0.04	-0.04	-0.09	-0.16	-0.05	6.8	0.119	0.75
Ind	CAPM	0.23	0.00	0.19	0.12	-0.11	0.03	0.02	0.06	0.02	-0.14	8.9	0.136	0.54
	3FM	0.11	-0.12	0.08	0.20	0.05	0.07	-0.26	0.11	-0.20	-0.41	21.0	0.209	**0.02**
	4FM	0.11	-0.05	0.00	0.15	0.14	0.18	-0.28	0.24	-0.13	-0.38	20.2	0.205	**0.03**
Panel B: Pre 1963														
	Alphas	1	2	3	4	5	6	7	8	9	10	JT	HJ	p
Beta	CAPM	0.14	0.01	0.14	0.03	0.13	-0.03	-0.03	-0.11	-0.16	-0.23	7.8	0.143	0.65
	3FM	0.14	0.02	0.14	0.03	0.11	-0.06	-0.07	-0.18	-0.25	-0.32	9.5	0.141	0.48
	4FM	0.20	0.07	0.14	0.04	0.07	-0.07	-0.13	-0.16	-0.14	-0.22	7.3	0.124	0.69
Rev	CAPM	0.23	0.00	0.19	0.12	-0.11	0.03	0.02	0.06	0.02	-0.14	4.4	0.096	0.93
	3FM	0.11	-0.12	0.08	0.20	0.05	0.07	-0.26	0.11	-0.20	-0.41	4.3	0.095	0.93
	4FM	0.11	-0.05	0.00	0.15	0.14	0.18	-0.28	0.24	-0.13	-0.38	4.9	0.101	0.90
Ind	CAPM	0.12	0.16	0.22	0.11	-0.13	0.17	-0.12	0.07	0.15	-0.34	11.1	0.152	0.35
	3FM	0.12	0.15	0.25	0.14	-0.14	0.17	-0.13	0.07	0.13	-0.39	12.6	0.162	0.25
	4FM	0.11	0.18	0.13	0.13	-0.09	0.24	-0.05	0.10	0.18	-0.26	9.4	0.140	0.50

A similar pattern is found for the reversal portfolios; no significant reversal effect occurs. In the post-1963 period, some individual CAPM alphas are large, for example 0.41 for the long-term loser portfolio #1, and the 3FM lowers these alphas (as documented also in Fama and French (1996)). However, as for the high beta portfolios, these alphas are highly unreliable (loser portfolios are dominated by small caps; see Table 2.2 panel B) and do not weight heavily in determining the overall goodness-of-fit. The advantage of using multiple factors for reversal portfolios disappears entirely in the pre-1963 period. This pattern is comparable to that of the BE/ME portfolios and may reflect the close link between reversal and BE/ME. Note that adding the WML factor (4FM) worsens the fit relative to the CAPM and 3FM.

While the three competing models do not yield significantly different results for the beta and reversal portfolios, the results are significantly different for the industry portfolios. Specifically, in the post-1963 period, the CAPM cannot be rejected, with generally small alphas. Surprisingly, the multifactor models lead to severe deteriorations of the alphas and have to be convincingly rejected. For example, the alpha for the utilities portfolio (#7) goes from 0.02 (CAPM) to -/-0.26 (3FM) and -/-0.28 (4FM). Overall, the *JT* statistic increases from 8.9 (0.54) to 21.0 (0.02) and 20.2 (0.03) for the three models respectively. Clearly, multiple factors hurt rather than help for industry portfolios. This finding is remarkable, because the industry classification (contrary to size, BE/ME and momentum) is not motivated by known patterns in historical return series and hence allows for comparing the competing models with less risk of data snooping.

2.4.3 Size quintiles

It is important to control for size because small caps generally involve relatively low market liquidity and high transaction costs. By weighting the alphas, the GMM framework corrects for the high volatility and correlation of small caps, but it does not correct for liquidity and costs. For this reason, we complement the above analysis of single-sorted portfolios with a further analysis of size-BE/ME, size-momentum, size-beta, size-reversal and size-industry portfolios.

Table 2.6 shows the results within the different size quintiles of the double-sorted portfolios. The general conclusion that can be drawn from these results is that the CAPM cannot be rejected for the large caps, representing roughly 80% of the total stock market capitalization. For post-1963 BE/ME portfolios, Loughran (1997) found that a substantial portion of the pricing errors is driven by the low returns of small newly-listed growth stocks. Similarly, Hong, Lim and Stein (2000) showed that the profitability of momentum strategies declines sharply with firm size. Our results suggest that similar size-effects occur also for stock characteristics other than BE/ME and momentum. For example, the CAPM can be rejected relative to beta and industry within the smallest stock quintile. The interaction with size disappears during the early pre-1963 period. This may reflect the smaller fraction of small caps in the analysis due to the focus on NYSE-listed stocks prior to June 1962.

Table 2.6
Size quintiles

This table gives the test results for the double-sorted portfolios based on BE/ME (value), momentum, beta, reversal and industry for the three asset pricing models (CAPM, 3FM and 4FM). In December of each year, stocks are first sorted into size quintiles (NYSE quintiles) and subsequently further divided based on other stock characteristics. The 25 Size-BE/ME portfolios are obtained from Kenneth French. The 25 size-momentum, 25 size-beta, 25 size-reversal and 38 size-industry portfolios are formed using the 200212 CRSP database. Every row shows the *JT* test statistic (3) and p-value. Test results are shown for the post-1963 sample from January 1963 to December 2002 (T=480) and the pre-1963 sample from January 1931 to December 1962 (T=384). p-values below the significance level of 10% are highlighted.

Panel A: Post 1963

		Value		Momentum		Beta		Reversal		Industry	
		JT	p	JT	p	JT	p	JT	p	JT	p
Small	CAPM	59.9	**0.00**	71.9	**0.00**	23.2	**0.00**	14.6	**0.01**	12.5	**0.05**
	3FM	20.0	**0.00**	67.3	**0.00**	5.6	0.35	4.2	0.51	14.6	**0.02**
	4FM	17.0	**0.00**	14.0	**0.02**	3.5	0.63	12.7	**0.03**	11.8	**0.07**
2	CAPM	33.1	**0.00**	42.3	**0.00**	19.0	**0.00**	10.8	**0.06**	7.0	0.43
	3FM	7.1	0.21	35.4	**0.00**	3.8	0.58	4.9	0.43	12.8	**0.08**
	4FM	6.1	0.30	2.3	0.80	1.7	0.89	12.7	**0.03**	10.2	0.18
3	CAPM	20.4	**0.00**	22.2	**0.00**	11.2	**0.05**	5.9	0.32	7.4	0.39
	3FM	2.7	0.75	26.9	**0.00**	1.0	0.96	2.4	0.79	5.8	0.57
	4FM	2.1	0.84	5.2	0.39	2.0	0.85	2.5	0.78	4.6	0.71
4	CAPM	18.4	**0.00**	14.1	**0.01**	12.9	**0.02**	7.5	0.19	3.1	0.93
	3FM	9.6	**0.09**	19.5	**0.00**	7.0	0.22	1.3	0.93	7.7	0.46
	4FM	6.4	0.27	5.7	0.34	5.8	0.33	4.3	0.51	7.6	0.48
Large	CAPM	4.7	0.46	8.7	0.12	5.4	0.37	4.8	0.44	14.3	0.16
	3FM	6.0	0.31	16.7	**0.01**	1.4	0.93	1.2	0.95	25.8	**0.00**
	4FM	5.4	0.37	16.6	**0.01**	0.7	0.98	5.3	0.39	25.1	**0.01**

Panel B: Pre 1963

		Value		Momentum		Beta		Reversal		Industry	
		JT	p	JT	p	JT	p	JT	p	JT	p
Small	CAPM	13.7	**0.02**	4.8	0.44	12.9	**0.02**	3.3	0.97	1.9	0.93
	3FM	12.4	**0.03**	3.7	0.60	11.9	**0.04**	2.5	0.99	1.3	0.97
	4FM	9.5	**0.09**	4.1	0.54	11.5	**0.04**	2.6	0.99	1.1	0.98
2	CAPM	1.2	0.94	30.6	**0.00**	18.3	**0.00**	1.5	1.00	5.0	0.66
	3FM	1.1	0.95	28.0	**0.00**	17.1	**0.00**	1.6	1.00	5.8	0.57
	4FM	2.1	0.84	11.7	**0.04**	18.6	**0.00**	2.3	0.99	5.2	0.64
3	CAPM	2.4	0.79	12.3	**0.03**	15.6	**0.01**	3.2	0.98	9.2	0.24
	3FM	2.5	0.78	11.9	**0.04**	14.7	**0.01**	4.3	0.93	8.5	0.29
	4FM	0.8	0.97	0.7	0.98	11.0	**0.05**	3.1	0.98	4.3	0.74
4	CAPM	3.2	0.71	24.4	**0.00**	14.4	**0.01**	4.6	0.92	18.5	**0.02**
	3FM	3.6	0.71	23.9	**0.00**	14.8	**0.01**	5.1	0.88	19.0	**0.01**
	4FM	3.2	0.95	4.8	0.44	16.1	**0.01**	5.2	0.88	16.8	**0.03**
Large	CAPM	3.0	0.67	12.2	**0.03**	6.5	0.26	3.0	0.98	10.5	0.39
	3FM	2.9	0.61	13.1	**0.02**	7.9	0.16	3.6	0.96	11.7	0.31
	4FM	1.1	0.67	0.7	0.98	8.4	0.14	3.2	0.98	8.7	0.56

The multifactor models yield important improvements for the small caps and medium caps, most notably for BE/ME, beta and reversal in the post-1963 period (3FM and 4FM) and for momentum portfolios in both periods (4FM only). However, multiple factors do not help for the large cap market segment. In fact, for post-1963 large cap momentum portfolios, the CAPM cannot be rejected (with a p-value of 0.12), while the multifactor models perform significantly worse and have to be rejected (with a p-value of 0.01 for both models). Figure 2.1 illustrates this remarkable finding by means of the alphas of the 5 momentum portfolios in the large cap segment. Most notably, for the large winner portfolio (#5) the 3FM alpha is 0.50 (0.31 for CAPM) and for the large loser portfolios (#1) the 4FM alpha is 0.41 (-/-0.20 for CAPM). Similarly, for post-1963 large industry portfolios, addition of multiple factors lead to a significant deterioration of the models fit (p-value goes from 0.16 to 0.00 and 0.01). For example, the pricing error of the large cap utility portfolio deteriorates from -/-0.06 (CAPM) to -/-0.33 (3FFM) and -/-0.36 (4FM).

These results clearly suggest that the multifactor models help for small caps but hurt in the large cap stock market. This is an important finding given the economic significance of large caps; to repeat, these stocks generally involve high market liquidity and low transaction costs.

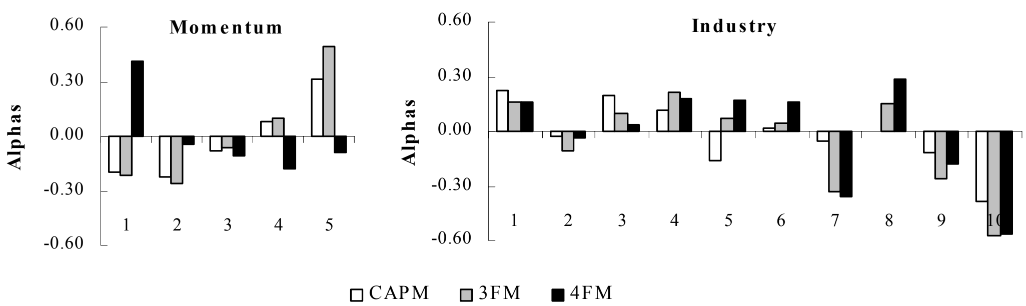


Figure 2.1: Large cap momentum and industry alphas (1963-2002). This figure shows the alphas for the large cap momentum and the large cap industry portfolios for the January 1963 – December 2002 period. Momentum portfolio #1 consists of the large loser stocks and portfolio #5 consists of large winner stocks. Large cap industry portfolios are based on the 4-digit SIC codes (1 nondurables, 2 durables, 3 oil, 4 chemicals, 5 manufacturing, 6 telecom, 7 utilities, 8 shops, 9 financials and 10 all other). The CAPM alphas (clear bars) are not jointly significantly different from zero, while the 3FM alphas (grey bars) and 4FM alphas (black bars) are jointly significantly different from zero.

2.5 Concluding remarks

Multiple factors are known to improve the empirical fit relative to the single-factor CAPM for value portfolios and momentum portfolios (only 4FM) in the post-1963 period. Our study suggests that these results are highly specific to the sample period and the set of benchmark portfolios. Specifically, the evidence in favor of multifactor alternatives largely disappears if we (1) analyze the pre-1963 period rather than to the typical post-1963 period, (2) employ size, beta, reversal and industry portfolios rather than BE/ME and momentum portfolios, and/or (3) focus on the large cap market segment, roughly 80% of the total market capitalization. In fact, in the cases of the post-1963 industry portfolios and large cap momentum portfolios, the multifactor models yield a significantly worse fit than the CAPM. These findings support the hypothesis that the existing evidence in favor of the multifactor models is the result of data snooping or other non-risk based explanations such as transaction costs and market liquidity of small stocks. Combined with the weak theoretical underpinnings, this in our opinion casts serious doubt on the use of multifactor models, especially for large cap stocks.

Chapter 3

SD efficiency of the stock market portfolio[1]

ABSTRACT: This chapter analyzes if the value-weighted stock market portfolio is stochastic dominance (SD) efficient relative to benchmark portfolios formed on size, value, and momentum. In the process, we also develop several methodological improvements to the existing tests for SD efficiency. Interestingly, the market portfolio is SD efficient relative to all benchmark sets. By contrast, the market portfolio is inefficient if we replace the SD criterion with the traditional mean-variance criterion. Combined these results suggests that the mean-variance inefficiency of the market portfolio is caused by the omission of return moments other than variance. Especially downside risk seems to be important for explaining the high average returns of small/value/winner stocks.

3.1 Introduction

EFFICIENCY OF THE STOCK MARKET PORTFOLIO is a much-debated topic in financial economics. Asset pricing models that employ a representative investor, including the mean-variance based CAPM, predict that the market portfolio is efficient.[2] At first glance, market portfolio efficiency is also consistent with the popularity of passive mutual funds and exchange traded funds that track broad value-weighted equity indexes. Nevertheless, several empirical studies suggest that the market portfolio is highly and significantly inefficient. Most notably, the market portfolio seems mean-variance (MV) inefficient relative to stock portfolios formed on variables such as market capitalization (size), book-to-market equity ratio (value) and price momentum (see for instance Basu (1977), Banz (1981), Fama and French (1992) (1993) and Jegadeesh and Titman (1993)).

The empirical results may reflect the limitations of the MV criterion. It is well-known that asset returns cannot be described by mean and variance alone. For example, the monthly returns of many stocks exhibit positive skewness and excess kurtosis. Also, a wealth of psychological research on decision-making under risk suggests that the perception of risk is more complex than variance. Especially the

---

1 This chapter is based on the paper entitled "downside risk and asset pricng" co-authored by Thierry Post(2004b). For the most recent version of this paper go to: http//ssrn.com/abstract=503142

2 These models typically use one of two theoretical motivations: (1) the capital market is complete or (2) investor preferences are sufficiently similar; see for instance Rubinstein (1974).

phenomena of skewness preference and loss aversion have attracted much attention among financial economists. This provides a rationale for replacing the MV criterion with a more general efficiency criterion that accounts for higher-order central moments (such as skewness and kurtosis) and lower partial moments (such as expected loss and semi-variance).

One popular approach to extend the MV criterion is by changing the maintained assumptions on investor preferences. If we do not restrict the shape of the return distribution, then the MV criterion is consistent with expected utility theory only if utility takes a quadratic form.[3] Extensions can be obtained by using alternative classes of utility. For example, Kraus and Litzenberger (1976) and Harvey and Siddique (2000) assume that utility can be approximated using a third-order polynomial (to account for skewness), and Dittmar (2002) uses a fourth-order polynomial (to account for skewness and kurtosis).

Unfortunately, the researcher faces two possible obstacles in implementing this approach. First, economic theory does not forward strong predictions about the shape of investor preferences and the return distribution. The theory specifies general regularity conditions such as nonsatiation (utility of wealth is increasing) and risk aversion (utility is concave), but not a functional form for preferences and distributions. This introduces a serious risk of specification error. For example, a fourth-order polynomial is not sufficiently flexible to account for lower partial moments, which require a non-differentiable utility function.

Second, it is often difficult to impose the regularity conditions in practice. The regularity conditions are needed to ensure that the results are economically meaningful. Also, these regularity conditions can improve the statistical power (=ability to detect inefficient portfolios) of efficiency tests. Further, from a mathematical perspective, a concave increasing utility function is required in order to justify the common approach of checking the first-order condition to test for market portfolio efficiency.[4] Unfortunately, it can be difficult to impose the regularity conditions on a parametric utility function. For example, we cannot restrict a quadratic utility function to be globally increasing and we cannot restrict a cubic utility function to be globally concave (see for example Levy (1969)). Moreover, not imposing the regularity conditions frequently leads to severe violations of the

3 Less restrictive assumptions are obtained if we do restrict the shape of the return distribution; see, for example, Berk (1997).

4 For a non-concave utility function the first-order condition is a necessary but nut sufficient condition for establishing a global maximum for expected utility; the first-order condition also applies for possible local optima and a possible global minimum. This important result is sometimes ignored when using non-concave utility functions.

regularity conditions. For example, Dittmar (2002, Section IIID) shows that the apparent explanatory power of a quartic utility function (which accounts for skewness and kurtosis) disappears if we impose (necessary conditions for) risk aversion. Similarly, Wang and Zhang (2004) show that the results of unrestricted asset pricing studies frequently imply serious arbitrage opportunities.

These empirical and theoretical considerations provide strong arguments for using the criteria of stochastic dominance (Hadar and Russell (1969), Hanoch and Levy (1969), Levy and Hanoch (1970) and Whitmore (1970)). These criteria avoid parameterization of investor preferences and the return distribution, while ensuring that economic regularity conditions are satisfied. Different SD criteria have been developed for different sets of regularity conditions. In this chapter, we focus on the relatively powerful third-order stochastic dominance (TSD) criterion (Whitmore (1970)). This criterion imposes the standard regularity conditions of nonsatiation and risk aversion. Also, TSD assumes a preference for positive skewness (marginal utility is convex). Empirical evidence suggests that investors indeed display this kind of skewness preference (see for instance Arditti (1967), Kraus and Litzenberger (1976), Cooley (1977), Friend and Westerfield (1980), and Harvey and Siddique, 2000).

We analyze if the value weighted CRSP total return index, a popular proxy for the stock market portfolio, is TSD efficient. To implement the TSD criterion, we first extend Post's (2003) empirical test for stochastic efficiency in several respects. First, we change the focus from second-order stochastic dominance (SSD), which imposes nonsatiation and risk aversion, to TSD, which also requires skewness preference. Second, we derive the asymptotic sampling distribution of the TSD test statistic under the true null of efficiency rather than the restrictive null of equal means that was used earlier. This extension is intended to avoid rejection of efficiency in cases where the market portfolio is efficient but the assets have substantially different means. Third, we derive a linear programming test for MV efficiency that can be compared directly with the TSD test. This allows us to attribute differences between the two tests to omitted moments exclusively.

With the resulting tests, we show that the CRSP index is TSD efficient relative to common benchmark portfolios formed on size, value, and momentum. By contrast, we find that the market portfolio is significantly MV inefficient, consistent with the existing evidence on size, value and momentum portfolios. The TSD criterion is especially successful in rationalizing MV inefficiencies that occur in the 1970s and the early 1980s. This suggests that the asset pricing puzzles that exist in the MV framework can be explained by omitted return moments during this period. The difference with the results of Post (2003), who rejects market portfolio efficiency, can be attributed to the use of the restrictive null of equal means and the use of an extended sample period.

The remainder of this chapter is structured as follows. Section 3.2 introduces the notation, assumptions, definitions and tests that will be used throughout the text. Section 3.3 discusses the data used in our analysis. Section 3.4 empirically analyzes the TSD efficiency of the market portfolio. Finally, Section 3.5 summarizes our conclusions and presents directions for further research.

3.2 Methodology

3.2.1 Assumptions

We consider a single-period, portfolio-based model of investment that satisfies the following assumptions:

Assumption 1 Investors choose investment portfolios to maximize the expected utility associated with the return of their portfolios. Investor preferences are characterized by nonsatiation, risk aversion and skewness preference. Throughout the text, we will denote utility functions by $u:\Re \rightarrow P$, $u \in U_{TSD}$, with U_{TSD} for the set of increasing and concave, once continuously differentiable, von Neumann-Morgenstern utility functions with convex marginal utility, and P for a nonempty, closed, and convex subset of $\Re$.[5,6] For the sake of comparison, we will also use the subset of quadratic utility functions that underlies mean-variance efficiency, that is, $U_{MV} \equiv \{u \in U_{TSD} : u(x) = ax + 0.5bx^2\}$. We follow the definition of MV efficiency by Hanoch and Levy (1970): a portfolio is efficient if and only if there exists an increasing and concave, quadratic utility function that rationalizes the portfolio. Further, we will use u' to denote the first-order derivative or marginal utility. If the utility function belongs to U_{TSD}, then marginal utility is a positive and decreasing

5 Throughout the text, we will use $\Re^N$ for an N-dimensional Euclidean space, and **Error! Objects cannot be created from editing field codes.**denotes the positive orthant. Further, to distinguish between vectors and scalars, we use a bold font for vectors and a regular font for scalars. Finally, all vectors are column vectors and we use **Error! Objects cannot be created from editing field codes.** for the transpose of **Error! Objects cannot be created from editing field codes.**.

6 Post (2003) does not assume that the utility function is continuously differentiable, so as to allow for, for instance, piecewise linear utility functions. However, in practice, we typically cannot distinguish between a kinked utility function and a smooth utility function with rapidly changing marginal utility. Nevertheless, using subdifferential calculus, we may obtain exactly the same characterization of the sampling distribution if utility is not continuously differentiable. Further, Post (2003) requires utility to be strictly increasing. To remain consistent with the original definition of TSD, we require a weakly increasing utility function. This is one of our reasons for adopting a novel standardization for the gradient vector; see Section IB.

function. In the special case of quadratic utility, marginal utility is a positive and decreasing, linear function.[7]

Assumption 2 The investment universe consists of N-1 risky assets and a riskless asset. Throughout the text, we will use the index set $\mathrm{I} \equiv \{1, \cdots, N\}$ to denote the different assets, with N for the riskless asset. The returns $\boldsymbol{x} \in \Re^N$ are serially independent and identically distributed (IID) random variables with a continuous joint cumulative distribution function (CDF) $G : \Re^N \to [0,1]$.

Assumption 3 Investors may diversify between the assets, and we will use $\boldsymbol{\lambda} \in \Re^N$ for a vector of portfolio weights. We focus on the case where short sales are not allowed, and the portfolio weights belong to the portfolio possibilities set $\Lambda \equiv \{\boldsymbol{\lambda} \in \Re_+^N : \boldsymbol{e}^{\mathrm{T}} \boldsymbol{\lambda} = 1\}$, with $\boldsymbol{e}$ for a unity vector with dimensions conforming to the rules of matrix algebra. The simplex Λ excludes short sales. Short selling is typically difficult to implement in practice due to margin requirements and explicit or implicit restrictions on short selling for institutional investors (see, for instance, Sharpe (1991) and Wang (1998)).[8]

Under these assumptions, the investors' optimization problem can be summarized as $\max_{\boldsymbol{\lambda} \in \Lambda} \int u(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\lambda}) dG(\boldsymbol{x})$. A given portfolio, say $\boldsymbol{\tau} \in \Lambda$, is optimal for a given utility function $u \in U_{TSD}$ if and only if the first-order condition is satisfied:

$$\int u'(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau})(x_i - \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\tau}) dG(\boldsymbol{x}) \leq 0 \quad \forall i \in \mathrm{I}, \tag{3.1}$$

The inequality should hold with strict equality for all assets that are included in the evaluated portfolio, that is, $i \in \mathrm{I} : \tau_i > 0$. If all assets are included in the evaluated portfolio ($\boldsymbol{\tau} > \boldsymbol{0}$), as is true for the value-weighted market portfolio, then inequality

7 Of course, if marginal utility is decreasing, then it can be positive only over a bounded interval. In our analysis, we require marginal utility to be positive and decreasing over the sample interval of returns on the market portfolio.

8 Nevertheless, we may generalize our analysis to include (bounded) short selling. The TSD test is based on the first-order optimality conditions for optimizing a concave objective function over a convex set. The analysis can be extended to a general polyhedral portfolio possibilities set. We basically have to check whether there exists an increasing hyperplane that supports the extreme points of the portfolio possibilities set. One approach is to enumerate all extreme points and to include all extreme points as virtual assets.

(1) automatically reduces to $\int u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(x_i - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})dG(\boldsymbol{x}) = 0 \quad \forall i \in \mathrm{I}$.[9] Interestingly, this equality represents the least favorable case for our test of efficiency, that is, the probability of a Type I error (wrongly classifying an efficient portfolio as inefficient) achieves its maximum if all assets are included. The proof to Theorem 1 in Section 3.2.3 uses this result.

Following the asset pricing terminology, we refer to violations of the first-order condition as pricing errors. We may measure the maximum pricing error as:[10]

$$\zeta(\boldsymbol{\tau}, G, u) \equiv \max_{i \in \mathrm{I}} \left\{ \int u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(x_i - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})dG(\boldsymbol{x}) \right\} \tag{3.2}$$

Portfolio $\boldsymbol{\tau} \in \Lambda$ is optimal relative to $u \in U_{TSD}$ if and only if $\zeta(\boldsymbol{\tau}, G, u) = 0$. The TSD efficiency criterion basically checks if this condition is satisfied for some $u \in U_{TSD}$. Similarly, the MV efficiency criterion checks if the first-order condition is satisfied for some $u \in U_{MV}$. To test for TSD efficiency or MV efficiency, we introduce the following measure:

$$\xi(\boldsymbol{\tau}, G, U) \equiv \min_{u \in U} \zeta(\boldsymbol{\tau}, G, u)\,. \tag{3.3}$$

with $U \in \{U_{TSD}, U_{MV}\}$. The efficiency criteria can equivalently be formulated in terms of the set of utility functions that rationalize the evaluated portfolio:

$$\Psi(\boldsymbol{\tau}, G, U) \equiv \{u \in U : \zeta(\boldsymbol{\tau}, G, u) = 0\}. \tag{3.4}$$

Note that the evaluated portfolio may be optimal for multiple utility functions, and hence $\Psi(\boldsymbol{\tau}, G, U)$ may contain multiple elements.

---

9 This equality is a variation to the well-known Euler equation **Error! Objects cannot be created from editing field codes.**.

10 Our focus on the maximum error reflects the short sales restriction. A negative 'pricing error' for a given asset does not constitute a violation of the first-order condition if the asset is not included in the evaluated portfolio (τ_i =0), as the investor can then improve the evaluated portfolio only by short selling the asset (which is not allowed) and not by reducing the weight of the asset in the portfolio. By contrast, a positive error is always problematic, because an investor can then improve the evaluated portfolio by increasing the weight of the asset and decreasing the weight of the other assets included in the portfolio. Of course, we generally do not know the number or the identity of the positive pricing errors in advance. However, the maximum pricing error is always positive.

Definition 1 Portfolio $\boldsymbol{\tau} \in \Lambda$ is efficient if and only if it is optimal for at least some $u \in U$, $U \in \{U_{TSD}, U_{MV}\}$, that is, $\xi(\boldsymbol{\tau}, G, U) = 0$, or, equivalently, $\Psi(\boldsymbol{\tau}, G, U)$ is non-empty. Portfolio $\boldsymbol{\tau} \in \Lambda$ is inefficient if and only if it is not optimal for all $u \in U$, that is, $\xi(\boldsymbol{\tau}, G, U) > 0$, or, equivalently, $\Psi(\boldsymbol{\tau}, G, U)$ is empty.

To test the null of efficiency, that is, $H_0 : \xi(\boldsymbol{\tau}, G, U) = 0$, we need full information on the CDF $G(\boldsymbol{x})$. In practical applications, $G(\boldsymbol{x})$ generally is not known and information is limited to a discrete set of T time series observations.

Assumption 4 The observations are serially independently and identically distributed (IID) random draws from the CDF. Throughout the text, we will represent the observations by the matrix $\mathbf{X} \equiv (\boldsymbol{x}_1 \cdots \boldsymbol{x}_T)$, with $\boldsymbol{x}_t \equiv (x_{1t} \cdots x_{Nt})^{\mathrm{T}}$. Since the timing of the draws is inconsequential, we are free to label the observations by their ranking with respect to the evaluated portfolio, that is, $\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} < \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} < \cdots < \boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau}$.

Using the observations, we can construct the following empirical distribution function (EDF):

$$F_{\mathbf{X}}(\boldsymbol{x}) \equiv \frac{1}{T}\sum_{t=1}^{T} 1(\boldsymbol{x}_t \leq \boldsymbol{x}), \qquad (3.5)$$

Since the observations are assumed to be serially IID, $F_{\mathbf{X}}(\boldsymbol{x})$ is a consistent estimator for $G(\boldsymbol{x})$, and we may use $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U)$ as a consistent estimator for $\xi(\boldsymbol{\tau}, G, U)$.

3.2.2 Linear programming test statistics

Following Post (2003), we may derive the following linear programming formulation for $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{TSD})$:[11]

11 This LP problem can be solved with minimal computational burden, even with spreadsheet software run on a desktop computer. Nevertheless, for applications where the number of time-series observations (T) is very large (for example, thousands of observations), it is useful to use a simplified formulation. A simplified version can be obtained by using **Error! Objects cannot be created from editing field codes.**, with **Error! Objects cannot be created from editing field codes.** and **Error! Objects cannot be created from editing field codes.**. Specifically, substituting **Error! Objects cannot be created from editing field codes.** in Equation (3.6) and rearranging terms yields

Error! Objects cannot be created from editing field codes.. (3.6')

$$\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{TSD}) = \min_{\boldsymbol{\beta}\in \mathrm{B}_{SSD},\theta}\left\{\theta : \sum_{t=1}^{T}\beta_t(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} - x_{it})/T + \theta \geq 0 \quad \forall i \in \mathrm{I}\right\} \tag{3.6}$$

with

$$\mathrm{B}_{TSD} \equiv \left\{\boldsymbol{\beta} \in \Re_+^T : \beta_1 \geq \cdots \geq \beta_T; \frac{\beta_1 - \beta_2}{\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}} \geq \cdots \geq \frac{\beta_{T-1} - \beta_T}{\boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau}}; \sum_{t=1}^{T}\beta_t / T = 1\right\} \tag{3.7}$$

In this formulation, $\boldsymbol{\beta}$ represents the gradient vector $(u'(\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau})\cdots u'(\boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau}))^{\mathrm{T}}$ for some well-behaved utility function $u \in U_{TSD}$. B_{TSD} represents the restrictions on the gradient vector that follow from the assumptions of nonsatiation, risk aversion, skewness preference and the standardization $\sum_{t=1}^{T} u'(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})/T = 1$

Note that the original test of Post (2003, Thm 2) uses the standardization $\beta_T = 1$ rather than $\sum_{t=1}^{T}\beta_t / T = 1$. The original standardization has an important drawback. Specifically, the higher the degree of risk aversion of the utility function, the higher the values of all betas. Hence, increasing the level of risk aversion tends to inflate the value of test statistic relative to the case with risk neutrality ($\boldsymbol{\beta} = \boldsymbol{e}$). This lowers Post's (2003, Thm 3) p-value for testing efficiency, possibly leading to erroneous rejections of efficiency, as the p-value is based on the risk neutral case and it does not account for the level of the betas. To circumvent this problem, we use the standardization $\sum_{t=1}^{T}\beta_t / T = 1$ in this study.[12] This standardization allows for risk aversion without inflating the test statistic, because the average level of the betas is fixed. Also, the novel standardization allows utility to be weakly increasing, as some betas may equal zero.

While Equation (3.6) involves T variables and N+2T-2 constraints, Equation (3.6') involves T variables and only N+T-1 constraints, which yields a large reduction in computational burden. We have effectively removed the T-1 restrictions **Error! Objects cannot be created from editing field codes.**, which are now satisfied by construction, as **Error! Objects cannot be created from editing field codes.**and **Error! Objects cannot be created from editing field codes.** imply **Error! Objects cannot be created from editing field codes.**. This simplification is similar to the one used by Post (2003, Proof to Thm 2) to arrive at his simplified dual test statistic.

12 Also, we will derive the p-value under the null of efficiency rather than the null of equal means.

Using linear interpolation, we may recover a full marginal utility function from the optimal solution ($\boldsymbol{\beta}^*$) as

$$p'(x|\boldsymbol{\beta}^*) \equiv \begin{cases} \beta_1^* & x \le \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} \\ \beta_1^* + (\beta_2^* - \beta_1^*)(x - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau})/(\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau}) & \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{\tau} \le x \le \boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\tau} \\ \vdots \\ \beta_{T-1}^* + (\beta_T^* - \beta_{T-1}^*)(x - \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau})/(\boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} - \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau}) & \boldsymbol{x}_{T-1}^{\mathrm{T}}\boldsymbol{\tau} \le x \le \boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} \\ \beta_T^* & x \ge \boldsymbol{x}_T^{\mathrm{T}}\boldsymbol{\tau} \end{cases} . \tag{3.8}$$

A full utility function is then found as $p(x|\boldsymbol{\beta}^*) \equiv \int_{z=-\infty}^{x} p'(z|\boldsymbol{\beta}^*)dz$.

We can derive a linear programming test for mean-variance efficiency in the spirit of (3.6). As discussed in Section 3.2.1, mean-variance analysis is the special case of TSD where utility takes a quadratic form and marginal utility takes a linear form. Put differently, the gradient vector $\boldsymbol{\beta}$ must belong to $\mathrm{B}_{MV} \equiv \mathrm{B}_{TSD} \cap \{\boldsymbol{\beta} \in \Re^T : \beta_t = a + b\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} \quad \forall t \in \Theta\}$. Hence, we obtain a linear programming test for MV efficiency by simply adding the restrictions $\beta_t = a + b\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} \quad \forall t \in \Theta$:[13]

$$\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U_{MV}) = \min_{\boldsymbol{\beta} \in \mathrm{B}_{MV}, \theta} \left\{ \theta : \sum_{t=1}^{T} \beta_t (\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau} - x_{it})/T + \theta \ge 0 \quad \forall i \in \mathrm{I} \right\}. \tag{3.9}$$

This test statistic differs in several respects from the traditional MV efficiency tests, such as the Gibbons, Ross and Shanken (1989) or GRS test. First, the test is consistent with TSD by adhering to the Hanoch and Levy (1969) definition of MV efficiency. For example, the GRS test may classify the market as inefficient if all assets have the same mean. However, in this case, the market is TSD efficient, because the market is optimal for the risk neutral investor. Second, our MV test excludes short selling; see Section 3.2.1. Third, our test focuses on the maximum

13 In fact, the additional restrictions simplify the problem, because there are now only two unknown variables. Substituting **Error! Objects cannot be created from editing field codes.** in equation (3.6) and rearranging terms, we find **Error! Objects cannot be created from editing field codes.Error! Objects cannot be created from editing field codes.**. The full utility function is simply **Error! Objects cannot be created from editing field codes.**, with **Error! Objects cannot be created from editing field codes.**for the optimal solution.

pricing error rather than a weighed average of all squared pricing errors, so as to allow for the case where not all assets are included in the evaluated portfolio.

3.2.3 Asymptotic sampling distribution

Our objective is to test the null hypothesis that $\boldsymbol{\tau} \in \Lambda$ is TSD or MV efficient, that is, $H_0 : \xi(\boldsymbol{\tau}, G, U) = 0$, $U \in \{U_{TSD}, U_{MV}\}$. Again, since the observations are serially IID, $F_{\mathbf{X}}$ is a consistent estimator for $G(\boldsymbol{x})$, and $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U)$ is a consistent estimator for $\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U)$. However, $F_{\mathbf{X}}(\boldsymbol{x})$ generally is very sensitive to sampling variation and the test results are likely to be affected by sampling error in a nontrivial way. The applied researcher must therefore have knowledge of the sampling distribution in order to make inferences about the true efficiency classification. Post (2003) derived the asymptotic sampling distribution of his SSD test statistic under the null hypothesis that all assets have the same mean, that is, $H_1 : E[\boldsymbol{x}] = \mu \boldsymbol{e}$, $\mu \in \Re$. In general, H_1 gives a sufficient condition for the true null of efficiency, that is, H_0. In fact, under the null, all portfolios $\boldsymbol{\lambda} \in \Lambda$ are efficient, because they are optimal for investors with utility function $v(x) \equiv x$, that is, the risk neutral investors. However, H_1 does not give a necessary condition for H_0, and rejection of H_1 generally does not imply rejection of H_0 and there is no guarantee that H_1 is sufficiently close to H_0. Hence, the sampling distribution H_1 may lead to erroneous conclusions about H_0. The purpose of this section is to analyze the asymptotic sampling distribution of $\xi(\boldsymbol{\tau}, G, U) = 0$ under the true null of efficiency rather than the null of equal means.

Using $\Phi(\cdot|\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u))$ for a *N*-dimensional multivariate normal distribution function with mean $\mathbf{0}$ and (singular) variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u) \equiv (\mathbf{I} - \boldsymbol{e}\boldsymbol{\tau}^{\mathrm{T}}) \boldsymbol{\Omega}(G, u)(\mathbf{I} - \boldsymbol{e}\boldsymbol{\tau}^{\mathrm{T}})^{\mathrm{T}}$, with $\boldsymbol{\Omega}(G, u) \equiv \int (u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})^2 \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}) dG(\boldsymbol{x})$, the following theorem characterizes the asymptotic sampling distribution under the true null:

THEOREM 1 *Asymptotically, the p-value* $\Pr[\xi(\boldsymbol{\tau}, F_{\mathbf{X}}, U) > y | H_0]$, $U \in \{U_{TSD}, U_{MV}\}$, $y \geq 0$, *is bounded from above by* $\Gamma(y, \boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u)/T) \equiv (1 - \int_{z \leq ye} d\Phi(z | \mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u)/T))$ *for all* $u \in \Psi(\boldsymbol{\tau}, G, U)$.

The theorem provides an upper bound to the *p*-value $\Pr[\xi(\boldsymbol{\tau},F_{\mathbf{X}},U)>y|H_0]$. It is difficult to derive the exact *p*-value, because $\Psi(\boldsymbol{\tau},G,U)$ generally contains multiple elements under H_0. Also, the theorem considers the least favorable case where all assets are included in the evaluated portfolio, and $\Gamma(y,\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)/T)$ represents an upper bound for the *p*-value in cases where some assets are excluded. While it is possible to identify an element of $\Psi(\boldsymbol{\tau},G,U)$ (see below), we generally do not know the element that minimizes the *p*-value $\Gamma(y,\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)/T)$. Nevertheless, the upper bound can be used in the same way as the true *p*-value. Specifically, we may compare $\Gamma(\xi(\boldsymbol{\tau},F_{\mathbf{X}},U),\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u))$, $u\in\Psi(\boldsymbol{\tau},G,U)$, with a predefined level of significance $a\in[0,1]$, and reject efficiency if $\Gamma(\xi(\boldsymbol{\tau},F_{\mathbf{X}},U),\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u))\leq a$. Equivalently, we may reject efficiency if the observed value of $\xi(\boldsymbol{\tau},F_{\mathbf{X}})$ is greater than or equal to the critical value $\Gamma^{-1}(a,\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)) \equiv \inf_{y\geq 0}\{y:\Gamma(y,\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u))\leq a\}$, $u\in\Psi(\boldsymbol{\tau},G,U)$. The statistical size or the probability of a Type I error (wrongly classifying an efficient portfolio as inefficient) of this approach is almost always smaller than the nominal significance level *a*.

Two results are useful for implementing the above approach in practice. First, computing *p*-values and critical values requires the variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)$ for some $u\in\Psi(\boldsymbol{\tau},G,U)$. Unfortunately, G is not known. Nevertheless, we know that $F_{\mathbf{X}}$ converges to G and that $p(x|\boldsymbol{\beta}^*)$ asymptotically belongs to $\Psi(\boldsymbol{\tau},G,U)$ under H_0. Hence, we may estimate $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)$ in a distribution-free and consistent manner using the sample equivalent $\boldsymbol{\Sigma}(\boldsymbol{\tau},F_{\mathbf{X}},p(x|\boldsymbol{\beta}^*))$ with elements $\omega_{ij}(\boldsymbol{\tau},F_{\mathbf{X}},p(x|\boldsymbol{\beta}^*))=\sum_{t=1}^{T}\beta_t^{*2}(x_{it}-\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})(x_{jt}-\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})/T$, $i,j\in\mathrm{I}$.

Second, we may approximate $\Gamma(\xi(\boldsymbol{\tau},F_{\mathbf{X}},U),\boldsymbol{\Sigma}(\boldsymbol{\tau},F_{\mathbf{X}},p(x|\boldsymbol{\beta}^*)))$, using Monte-Carlo simulation. In this chapter, we will use the following approach. We first generate *S*=10,000 independent standard normal random vectors $\boldsymbol{w}_s\in\Re^{N-1}$, $s\in\{1,\cdots,S\}$, using the RNDN function in Aptech Systems' GAUSS software. Next, each random vector $\boldsymbol{w}_s$ is transformed into a multivariate normal vector $\boldsymbol{z}_s\in\Re^N$ with variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)$ by using $\boldsymbol{z}_s=(\mathbf{I}-\boldsymbol{e}\boldsymbol{\tau}^{\mathrm{T}})\mathbf{D}(G,u)\boldsymbol{w}_s$, where $\mathbf{D}(G,u)\in\Re^{N\times(N-1)}$ is a matrix with the first $(N-1)$ rows equal to the Cholesky factor of the nonsingular $(N-1)\times(N-1)$ variance-covariance matrix of risky assets, and with zeros for the *N*th row. Finally, $\Gamma(\xi(\boldsymbol{\tau},F_{\mathbf{X}},U),\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u))$ is approximated by the relative frequency of the transformed vectors z_s, $s \in \{1,\cdots,S\}$, that fall outside the integration region $\left\{z \in \Re^N : z \leq y\boldsymbol{e}\right\}$.

Theorem 1 subtly differs from Post's characterization of the sampling distribution under H_1. That characterization used the variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,v)$, with $v(x)=x$, in place of $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)$. This replacement is valid only if $v \in \Psi(\boldsymbol{\tau},G,U)$, that is, if the evaluated portfolio is optimal for the risk neutral investor. This reflects the replacement of the null of efficiency (H_0) by the null of equal means (H_1); under H_1, all portfolios are optimal for the risk neutral investor. Obviously, it is relatively simple to reject H_1 and hence the p-values and critical values under this null are likely to underestimate the true values under H_0. Consequently, a test procedure that uses the sampling distribution under H_0 will involve a more favorable statistical size (=relative frequency of Type I error) and a less favorable statistical power (=one minus the relative frequency of Type II error) than a test procedure that uses the sampling distribution under H_1.

3.2.4 Simulation experiment

Using a simulation experiment, Post (2003, Section IIIC) demonstrates that his SSD test procedure based on H_1 involves relatively low power in small samples generated from the return distribution of the well-known 25 Fama and French stock portfolios formed on size and value. Since H_0 is more general than H_1, the power of our test procedure based on H_0 will be not as good. Nevertheless, the procedure may be sufficiently powerful to be of practical use for data sets with a smaller cross-section. The relatively low power in the Post experiment probably reflects the difficulty of estimating a 25-dimensional multivariate return distribution in a nonparametric fashion. It is likely that the power increases (at an increasing rate) as the length of the cross-section is reduced to for example ten benchmark portfolios, which is common in asset pricing tests.

To shed some light on the statistical properties of our test procedure based on H_0 in smaller cross-sections, we extend the original simulation experiment. The simulations involve ten assets with a multivariate normal return distribution. The joint population moments are equal to the sample moments of the monthly excess returns of ten Fama and French stock portfolios formed on BE/ME during Post's sample period from July 1963 to October 2001.[14] We will analyze the statistical properties of our TSD test procedure (reject efficiency if and only if $\Gamma(\xi(\tau, F_{\mathbf{X}}, U_{TSD}), \Sigma(\tau, F_{\mathbf{X}}, p(x|\beta^{*})) \leq a)$ and our MV test procedure (reject efficiency if and only if $\Gamma(\xi(\tau, F_{\mathbf{X}}, U_{MV}), \Sigma(\tau, F_{\mathbf{X}}, p(x|\beta^{*})) \leq a)$ by the rejection rates of these procedures for certain test portfolios in random samples drawn from this multivariate normal distribution.

For the true distribution, the equal weighted portfolio (EP) is known to be MV and TSD inefficient.[15] Hence, we may analyze the statistical power of the TSD and MV test procedures by the ability to correctly classify the EP inefficient. By contrast, the ex ante tangency portfolio (TP) is MV and TSD efficient and we may analyze the statistical size by the relative frequency of random samples in which this portfolio is wrongly classified as inefficient.[16]

We draw 10,000 random samples from the multivariate normal population distribution through Monte-Carlo simulation. To each sample, we also add 'observations' for a riskless asset with a return of zero in every month (recall that we use excess returns). For every random sample, we apply the MV and TSD test procedures to the efficient TP and the inefficient EP. For both procedures, we compute the size as the rejection rate for TP and the power as the rejection rate for EP. This experiment is performed for a sample size (T) of 25 to 1,500 observations and for a significance level (a) of 2.5, five, and ten percent.

Figure 3.1 shows the results. The size is generally substantially smaller than the nominal level of significance a, and it converges to zero. In fact, using a level of significance of ten percent, the size is smaller than five percent for samples as small as 100 observations. This reflects our focus on the least favorable distribution, which minimizes Type I error.

As discussed in Section 3.2.3, $\xi(\tau, F_{\mathbf{X}}, U), U \in \{U_{TSD}, U_{MV}\}$, converges to $\xi(\tau, G, U)$, and we expect minimal Type II error in large samples. Indeed, for both procedures, the power goes to unity as we increase the sample size. However, in small samples, the TSD procedure is substantially less powerful than the MV procedure.

14 The power depends on the degree of inefficiency of the evaluated portfolio. We selected the set of ten BE/ME portfolios, because the degree of inefficiency of the equal weighted portfolio is 'medium'; it is higher than for the ten size portfolios, but lower than for the momentum portfolios.

15 It is possible to achieve a substantially higher mean given the standard deviation of EP and hence EP is MV inefficient. Since we assume a normal distribution in the simulations, the TSD criterion coincides with the MV criterion and EP is also TSD inefficient.

16 The tangency portfolio consists of 18.22%, 2.04% and 79.74% invested in the fifth, sixth and eight BE/ME portfolios respectively.

For example, using a ten percent significance level, the MV procedure achieves a rejection rate of about 50 percent already for samples of about 250 observations. By contrast, The TSD procedure achieves this rejection rate only for samples of about 400 observations. At that sample size, the rejection rate of the TSD procedure increases rapidly. Interestingly, this sample size is not uncommon for this type of application. For example, our empirical tests will use samples of 378 to 840 monthly observations. Hence, the TSD procedure appears sufficiently powerful to be of practical use in this type of application.

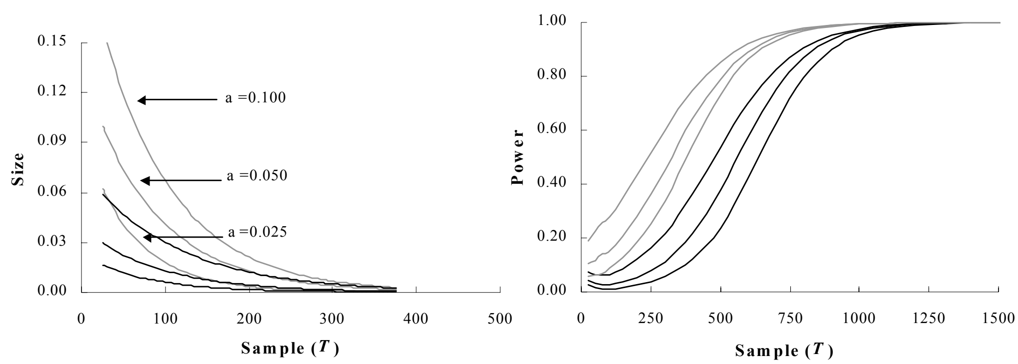


Figure 3.1: Statistical properties of the TSD test procedure. The figure displays the size and power of the test procedures for various numbers of time-series observations (T) and for a significance level (a) of 2.5, five and ten percent. The results are based on 10,000 random samples from a multivariate normal distribution with joint moments equal to the sample moments of the monthly excess returns of the ten B/M portfolios and the U.S. Treasury bill for the period from July 1963 to October 2001. The dark lines show the results for the TSD test, and the gray lines show the results for the MV test. Size is measured as the relative frequency of random samples in which the efficient tangency portfolio (TP) is wrongly classified as inefficient. Power is measured as the relative frequency of random samples in which the inefficient equally weighted portfolio (EP) is correctly classified as inefficient.

3.2.5 Further tests

Our main results rest on the MV and TSD tests derived in Section 3.2.2. Nevertheless, we will also make use of three other tests in order to interpret our test results. First, as discussed in Section 3.2.2, our MV tests uses the Hanoch-Levy definition of MV efficiency that requires nonsatiation. To determine if the MV results are due to the difference between this definition and the traditional MV efficiency definition (TMV, which does not require nonsatiation), we consider two alternative tests that do adhere to the traditional definition. First, we apply TMV, a relaxed version of our MV test that drops nonsatiation.[17] In addition, we also apply the well-known GRS test of market efficiency. This test not only drops nonsatiation, but it also allows for short sales. Finally, we apply Post's (2003) original SSD efficiency test. Recall that this test uses the sampling distribution under the null of equal means rather than the null of efficiency and hence it may erroneously reject efficiency if the assets have different means.

3.3 Data

We analyze if the CRSP all-share index is efficient. This value-weighted index consists of all common stocks listed on NYSE, AMEX, and Nasdaq. To proxy the investment universe of individual assets, we use three different sets of ten portfolios that cover the three well-known puzzles of size, value, and momentum. To calculate monthly excess returns, we subtract the risk-free rate defined as the U.S 30 day T-bill rate maintained by Ibbotson.

First, we use the widely used decile portfolios formed on market capitalization. Second, we use ten benchmark portfolios formed on book-to-market-equity ratio (value). The size and value portfolios are formed using NYSE decile breakpoint data for the total sample period also when (smaller) Nasdaq and AMEX shares are added to the CRSP database. For detailed data description and selection procedures we refer to Fama and French (1992) (1993). In our analysis of size and value, we focus on a long 70-year sample period starting in January 1933 to December 2002 (840 months). Third, we use ten portfolios sorted on price momentum as described in Jegadeesh and Titman (2001). This benchmark set ranges from January 1965 to December 1998 (408 months). These three sets of decile portfolios capture the separate effects of size, value and momentum.

It is perhaps even more interesting to consider the combined effect of the different puzzles. For this purpose, we also analyze the 25 Fama and French (FF25) benchmark portfolios formed on size and value and the 27 Carhart (C27) portfolios formed on value, size, and momentum.[18] For the former benchmark set, data are available also from January 1933 to December 2002 (840 months). For the latter benchmark set, data are limited to the period ranging from July 1963 to December

17 This means that we no longer require the gradient vector to be non-negative. It is straightforward to verify that this yields a necessary and sufficient test for mean-variance efficiency in the traditional definition but with short sales excluded.

18 We thank Kenneth French, Narasimhan Jegadeesh, and Mark Carhart for generously providing us with these data.

1994 (378 months). These two benchmark data sets allow us to determine if the same type of utility function can explain the various puzzles combined. Also, they allow us to analyze if the TSD criterion can capture the interaction between the size, value and momentum effects that may occur if investors can adopt a mixed strategy of investing in for instance small value stocks or small value winners stocks.

One should bear in mind that empirical research carried out before 1999 could give different results due to the 'delisting bias' first noticed by Shumway (1997). Especially Nasdaq listed stocks with low market capitalizations were severely affected by this bias (Shumway and Warther (1999)). Since 1999, CRSP has carried out a series of projects to improve the quality of the delisting returns database eliminating the delisting bias (CRSP (2001)).[19] Because momentum (single and triple sorted) portfolios rely on older versions of the CRSP database, we treat these with more caution than the size and value sorted portfolios.

Table 3.1 gives descriptive statistics for the five sets of benchmark portfolios: ten size portfolios, ten value portfolios, ten momentum portfolios, 25 size-value portfolios (FF25) and 27 value-size-momentum portfolios (C27). Clearly, the returns of these portfolios do not obey a normal distribution. This provides an important rationale for adopting the TSD test, which account for the full return distribution rather than mean and variance alone.

19 For example, consider the Fama and French (1996) sample period ranging from July 1963 to December 1993 (366 months). In FF Table I the average monthly excess return on the BV (SG) portfolio was 0.71 (0.31), but with the new 2002 database (employed in this study) this number is downwards revised to 0.59 (0.25).

Table 3.1

Descriptive Statistics Data Sets

The table shows descriptive statistics for the monthly excess returns of the value-weighted CRSP all-share market portfolio and the ME10 (size), BM10 (value), past six month returns (momentum), size/value (FF25) and value/size/momentum (C27) sorted data sets. The size, value and FF25 data are taken from the homepage of Kenneth French, the momentum data are from Jegadeesh (2001) and the C27 data are from Carhart (1997).The sample period is from January 1933 to December 2002 (T=840) for the size, value and FF25 data sets, from January 1965 to December 1998 (T=408) for the momentum data set and from July 1963 to December 1994 (T=378) for the C27 data set. Excess returns are computed from the raw return observations by subtracting the return on the one-month US T-bill.

Panel A: Single sort: Size, BM (T=840 months)

		Mean	Stdev.	Skewness	Kurtosis	Min	Max
	Market	0.714	4.937	0.156	9.18	-23.67	38.17
10 Size Portfolios	Small	1.328	9.425	2.988	29.49	-34.59	95.97
	2	1.173	8.368	2.399	28.24	-32.93	33.30
	3	1.107	7.432	1.599	21.00	-29.65	95.19
	4	1.047	7.041	1.311	16.88	-30.07	78.59
	5	1.019	6.706	1.086	15.77	-28.89	64.25
	6	0.962	6.298	0.799	13.05	-26.89	57.52
	7	0.928	6.105	0.774	13.77	-29.07	53.07
	8	0.854	5.638	0.379	10.11	-24.90	54.42
	9	0.811	5.327	0.656	12.95	-23.80	45.86
	Big	0.657	4.665	0.093	7.95	-22.99	49.10
10 Value Portfolios	Growth	0.630	5.366	0.193	7.32	-23.30	38.45
	2	0.695	5.104	-0.192	6.68	-25.19	71.70
	3	0.676	4.967	-0.196	6.87	-26.47	28.74
	4	0.723	5.365	1.210	19.30	-24.26	27.24
	5	0.843	4.982	0.752	13.89	-24.58	56.29
	6	0.898	5.276	0.621	13.40	-26.20	46.15
	7	0.912	5.778	1.424	18.55	-25.62	48.79
	8	1.077	5.910	1.074	14.79	-29.08	59.17
	9	1.136	6.943	1.470	18.83	-30.87	52.43
	Value	1.164	8.215	1.496	18.97	-45.76	62.24

Panel B: Single sort: Momentum (T=408 months)

		Mean	Stdev.	Skewness	Kurtosis	Min	Max
	Market	0.513	4.472	-0.505	5.42	-23.09	16.05
10 Mom. Portfolios	Loser	-0.105	6.878	0.001	5.20	-28.59	28.98
	2	0.374	5.674	-0.085	5.50	-22.68	20.05
	3	0.520	5.156	-0.150	5.80	-23.01	26.82
	4	0.586	4.873	-0.334	6.47	-25.18	25.13
	5	0.606	4.709	-0.491	6.87	-25.90	23.85
	6	0.640	4.701	-0.694	6.93	-26.61	22.51
	7	0.666	4.770	-0.923	7.26	-27.80	20.72
	8	0.750	4.995	-1.041	7.30	-29.18	18.98
	9	0.860	5.391	-1.038	6.73	-30.27	17.31
	Winner	1.123	6.561	-0.882	5.52	-32.74	15.84

Panel C: Double sort, Size/Value (T=840 months)

			Mean	Stdev.	Skewness	Kurtosis	Min	Max
	Market		0.714	4.937	0.156	9.18	-23.67	38.17
al ue	S	G	0.575	11.006	1.872	17.81	-42.95	99.95
	S	2	0.992	9.193	2.119	22.22	-34.49	87.77

S	3	1.229	8.513	2.039	20.36	-36.52	77.98
S	4	1.426	8.017	3.125	41.40	-34.00	105.62
S	V	1.523	8.937	3.346	38.38	-34.15	105.30
2	G	0.734	7.960	0.462	8.27	-33.32	54.01
2	2	1.049	7.286	1.760	24.40	-32.37	81.58
2	3	1.170	6.974	2.300	30.12	-28.27	81.55
2	4	1.242	6.930	1.860	23.70	-28.28	71.97
2	V	1.329	7.816	1.204	14.71	-34.41	57.98
3	G	0.802	7.244	0.930	12.36	-30.03	59.99
3	2	0.989	6.240	0.424	10.79	-29.49	44.72
3	3	1.030	6.182	1.059	16.17	-28.34	56.03
3	4	1.136	6.087	1.073	15.04	-24.99	54.46
3	V	1.250	7.410	1.029	13.98	-36.04	62.32
4	G	0.731	5.961	-0.073	5.84	-26.02	32.80
4	2	0.812	5.870	0.966	17.24	-29.45	58.25
4	3	1.010	5.696	0.500	12.02	-27.13	48.35
4	4	1.003	5.997	0.694	11.10	-29.48	47.42
4	V	1.183	7.841	1.595	20.44	-34.75	75.34
B	G	0.662	5.078	0.128	7.43	-22.61	35.27
B	2	0.648	4.799	-0.033	6.94	-23.53	27.95
B	3	0.819	4.915	0.771	13.76	-22.46	46.40
B	4	0.891	5.823	1.470	19.37	-27.14	60.77
B	V	0.943	7.083	0.799	13.92	-34.62	56.24

Panel D: Triple sort, Value/Size/Momentum (T=378 months)

27 Value/Size/Momentum Portfolios

			Mean	Stdev.	Skewness	Kurtosis	Min	Max
Market			0.387	4.399	-0.394	5.60	-23.09	16.05
G	S	L	-0.279	6.644	0.003	5.78	-31.41	30.45
G	S	M	0.349	6.070	-0.461	6.19	-31.33	27.54
G	S	W	0.933	6.755	-0.704	5.49	-33.52	19.45
G	M	L	-0.089	5.952	0.088	5.76	-26.49	27.54
G	M	M	0.143	5.341	-0.461	5.21	-26.81	18.96
G	M	W	0.895	6.033	-0.550	5.26	-29.90	20.73
G	B	L	0.105	5.218	0.178	5.21	-20.40	25.40
G	B	M	0.250	4.507	-0.161	4.88	-20.73	17.86
G	B	W	0.611	5.348	-0.298	5.03	-23.68	21.21
N	S	L	0.306	5.975	0.474	8.36	-24.81	37.51
N	S	M	0.631	4.957	-0.201	8.04	-27.05	27.02
N	S	W	1.157	6.088	-0.730	6.57	-32.60	23.17
N	M	L	0.400	5.396	0.501	6.47	-19.08	31.20
N	M	M	0.513	4.498	-0.344	7.72	-25.68	23.05
N	M	W	0.772	5.325	-0.896	6.68	-30.54	17.52
N	B	L	0.438	4.801	0.477	5.29	-19.66	21.74
N	B	M	0.367	4.211	0.158	5.31	-15.57	21.03
N	B	W	0.595	4.856	-0.347	5.56	-24.00	19.95
V	S	L	0.552	6.440	0.974	10.77	-28.26	45.34
V	S	M	1.062	5.583	0.199	9.22	-29.09	33.99
V	S	W	1.341	6.246	-0.289	7.56	-32.00	30.20
V	M	L	0.582	6.078	0.469	7.68	-25.79	38.07
VV	M	M	0.909	5.303	-0.016	9.04	-28.30	31.03
V	M	W	1.287	5.881	-0.806	8.23	-33.86	25.76
V	B	L	0.600	5.569	0.718	7.41	-16.99	35.34
V	B	M	0.589	4.752	-0.053	5.72	-23.47	20.41
V	B	W	0.923	5.352	-0.254	5.85	-24.78	22.84

Figure 3.2 shows the five benchmark sets in mean-standard deviation space, including the individual assets (the clear dots), the market portfolio (the black square), the MV tangency portfolios (black dots), and the mean-standard deviation frontier (with and without the risk free asset). Clearly, the market portfolio is (ex post) inefficient in terms of mean-variance analysis relative to all benchmark sets. The distance from the mean-standard deviation frontier is smallest for the size portfolios and it is largest for the momentum portfolios. For example, in the momentum data set, it is possible to achieve an average monthly return of 0.33% per month (or 4.0% per annum) in excess of the market average and given the market standard deviation. If size, value and momentum are combined the distance to the MV-frontier increases. Of course, these ex post mean-standard deviation diagrams do not reveal if the MV classification is statistically significant. Also, the diagrams are silent on return moments other than mean-variance (such as higher-order central moments and lower partial moments). The next section gives more details on the mean-variance efficiency and TSD-efficiency of the market portfolio.

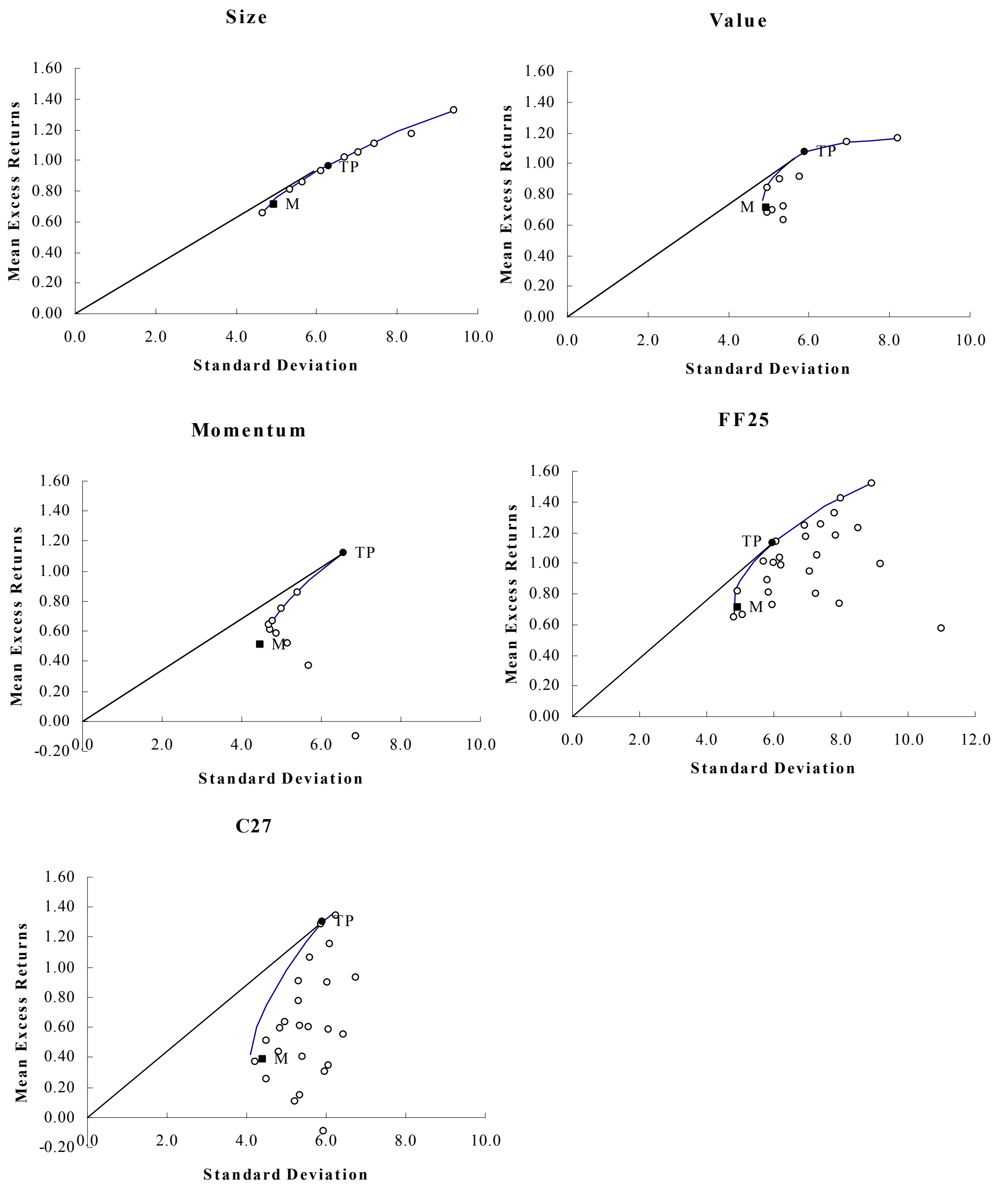


Figure 3.2: Mean-standard deviation diagrams. This figure shows the mean-standard deviation diagram for the size, value, momentum, FF25 and C27 data sets. We show the mean excess return and the standard deviation of the individual benchmark portfolios (clear dots) and the efficient frontier with and without the risk-free asset. The market portfolio is labelled "M" and the mean-variance tangency portfolio "TP".

3.4 Results

3.4.1 Full sample results

Table 3.2 summarizes the results for the full sample from January 1933 to December 2002. The table includes a separate cell for every combination of the five data sets (size, value, momentum, FF25 and C27) and the five efficiency tests (TSD, MV, Post, TMV and GRS). Each cell includes the value of the test statistic and the associated p-value. Also, each cell includes the identity of the optimal portfolio (OP), or the benchmark portfolio with the largest pricing error.

Table 3.2
Efficiency of the stock market portfolio

Formatted: Figure text

We test if the CRSP all-share index is TSD and MV efficient in the size, value, momentum, FF25 and C27 data sets. Results are shown for the full samples. Each cell contains the maximum pricing error (theta), the associated p-value, and the identity of the optimal portfolio (OP) or the portfolio with the largest pricing error. Cells are coloured grey if the p-value falls below 10% and efficiency is rejected with at least 90% confidence. Panel A includes our main results and compares the results of our TSD and MV tests. Panel B shows the outcomes of three alternative tests: the original Post (2003) test, the traditional mean variance (TMV) test (without the restriction of nonsatiation), and finally the Gibbons, Ross, Shanken (1989) test. The GRS p-value corresponds to the optimally weighted minimized sum of squared errors.

				Panel A: TSD versus MV		Panel B: Alternative tests		
	N	*T*		TSD	MV	Post	TMV	GRS
			Theta	0.214	0.341	0.184	0.222	0.314
Size	10	840	p-value	0.251	**0.049**	0.231	0.144	0.391
			OP	Small	Small	Small	Small	Small
			Theta	0.205	0.317	0.211	0.291	0.312
Value	10	840	p-value	0.409	**0.032**	0.135	0.041	0.022
			OP	8	8	8	8	8
			Theta	0.244	0.349	0.215	0.349	0.455
Mom.	10	408	p-value	0.256	**0.051**	0.175	**0.051**	**0.000**
			OP	Winner	Winner	Winner	Winner	Winner
			Theta	0.370	0.579	0.363	0.411	0.556
FF25	25	840	p-value	0.293	**0.014**	0.140	**0.075**	**0.000**
			OP	14	SV	SV	SV	SV
			Theta	0.429	0.795	0.469	0.795	0.902
C27	27	378	p-value	0.295	**0.000**	**0.048**	**0.000**	**0.000**
			OP	VSW	VSW	VSW	VSW	VSW

Panel A compares the results for the MV and TSD criteria. Using a significance level of ten percent, the market portfolio is highly and significantly MV inefficient relative to all benchmark sets. For example, we find a test statistic of 0.317 in the value data set. This figure can be interpreted as a maximal pricing error of 0.317 percent per month (or 3.8% per annum). The associated p-value is 0.032, suggesting that the market is significantly MV inefficient relative to the value portfolios. Similar results are found for the other four data sets. In other words, the three asset pricing puzzles (size, value, and momentum) are clearly present in our data sets. The strongest evidence against MV efficiency is found in the triple-sorted C27 data set that combines the three puzzles, with a maximum pricing error of 0.795 (or 9.5% per annum) and a p-value of 0.000.

In contrast to the MV results, we find that the market portfolio is TSD-efficient relative to all five benchmark sets; in all cases, the test statistic becomes insignificant. For example, for the value portfolios, the test statistic falls from 0.317 to 0.205 (or 2.4% per annum) and the p-value increases from 0.032 to 0.409. Similar results are found for the other benchmark sets. For the triple-sorted C27 data set, the maximum pricing error falls from 0.795 to 0.429 (or 5.1% per annum) and the p-value increases from 0.000 to 0.295. Combined, the MV and TSD results suggest that omitted moments (such as higher-order central moments and lower partial moments) may explain MV inefficiency of the market portfolio.

Panel B of Table 3.2 includes the three alternative tests for market portfolio efficiency (see section 3.2.5). First, the panel shows the results of the Post (2003) test, which differs from our TSD test, because it focuses on SSD efficiency and it uses the restrictive null of equal means rather than the true null of efficiency. Clearly, using the wrong null lowers the p-values in all data sets (even though the SSD criterion has lower rejection rates than the TSD criterion). In fact, for the triple-sorted C27 data set, market portfolio efficiency is rejected with more than 90 percent confidence. However, efficiency cannot be rejected for the other four data sets, even if we use the Post test. These results help to understand the difference between our findings and the finding in Post (2003, Section IV) that the market portfolio is SSD inefficient relative to the FF25 portfolios (p-value was 0.031). First, including the pre-1963 data increases the p-value to 0.140 (was 0.031). Second, replacing the null of equal means with the null of efficiency further increases the p-value further to 0.293.[20]

Second, Panel B includes the results for the MV test after dropping the regularity condition of nonsatiation, which yields a test that is in line with the traditional definition of MV efficiency (TMV). The condition of nonsatiation is binding (the MV and TMV test results are different) for the size, value and FF25 data sets, but not for the momentum and C27 data sets. Presumably, this can be explained by the broad market return interval for the 1933-2002 period. During this period, the market return ranged from -/-23.67 to 38.17 percent. A quadratic utility function can

20 Rather than the TSD criterion, Post (2003) uses the SSD criterion, which does not require skewness preference. Using this criterion, the p-value in our application becomes even higher than reported here.

exhibit only little curvature if it is to be monotone increasing over such a broad interval. During the more recent periods for the momentum and C27 portfolios, the market return interval is narrower (-/-23.09% to 16.05%) and the two tests coincide. The effect of dropping nonsatiation is most substantial for the size data set. In fact, the size effect is no longer significant if we drop nonsatiation; the maximum pricing error falls from 0.341 (or 4.1% per annum) to 0.222 (or 2.7% per annum) and the p-value increases from 0.049 to 0.144. By contrast, the value and momentum effects remain strong and significant.

Third, we give the results of the GRS test. Although this test differs in various ways from our TMV test, the results are fairly similar. We find slightly higher maximal pricing errors and lower levels of significance. However, the market portfolio is efficient relative to size portfolios and inefficient relative to all other benchmark sets. It is encouraging to see that inferences about mean-variance efficiency in our study are not heavily affected by the exact test procedure.

3.4.2 Downside risk

How can we explain the differences between the MV and the TSD results? In principle, any omitted higher-order central moment (other than mean and variance) or lower partial moment could explain why the market seems MV inefficient but TSD efficient. To gain further insight in what explains TSD efficiency, it is useful to look at Figure 3.3. This figure shows the optimal marginal utility function $p(x|\boldsymbol{\beta}^*)$ for the MV, TMV and TSD tests. As discussed in Section 3.2.2, this marginal utility function is the optimal solution for the utility gradient vector ($\boldsymbol{\beta}^*$). Two results are noteworthy. First, the figure shows that under the more restrictive MV criterion, the optimal marginal utility function for the size, value and FF25 data sets are not as steep as under the less restrictive TMV criterion, so as to ensure nonsatiation for the entire sample range of market returns. Second and most important, the TSD utility functions exhibit "crash-o-phobia" or a strong aversion for downside risk. For all five data sets, the optimal marginal utility function assigns a high weight to large losses. In case of the size, momentum, FF25 and C27 data set, the implicit utility function is very sensitive to losses that exceed about -/- 12 percent. In the value data set, the critical return level is even lower, at about -/- 20 percent. The linear marginal utility function is not sufficiently flexible to allow for these patterns of downside risk aversion. Hence, it seems that downside risk may explain why the market portfolio is MV inefficient.

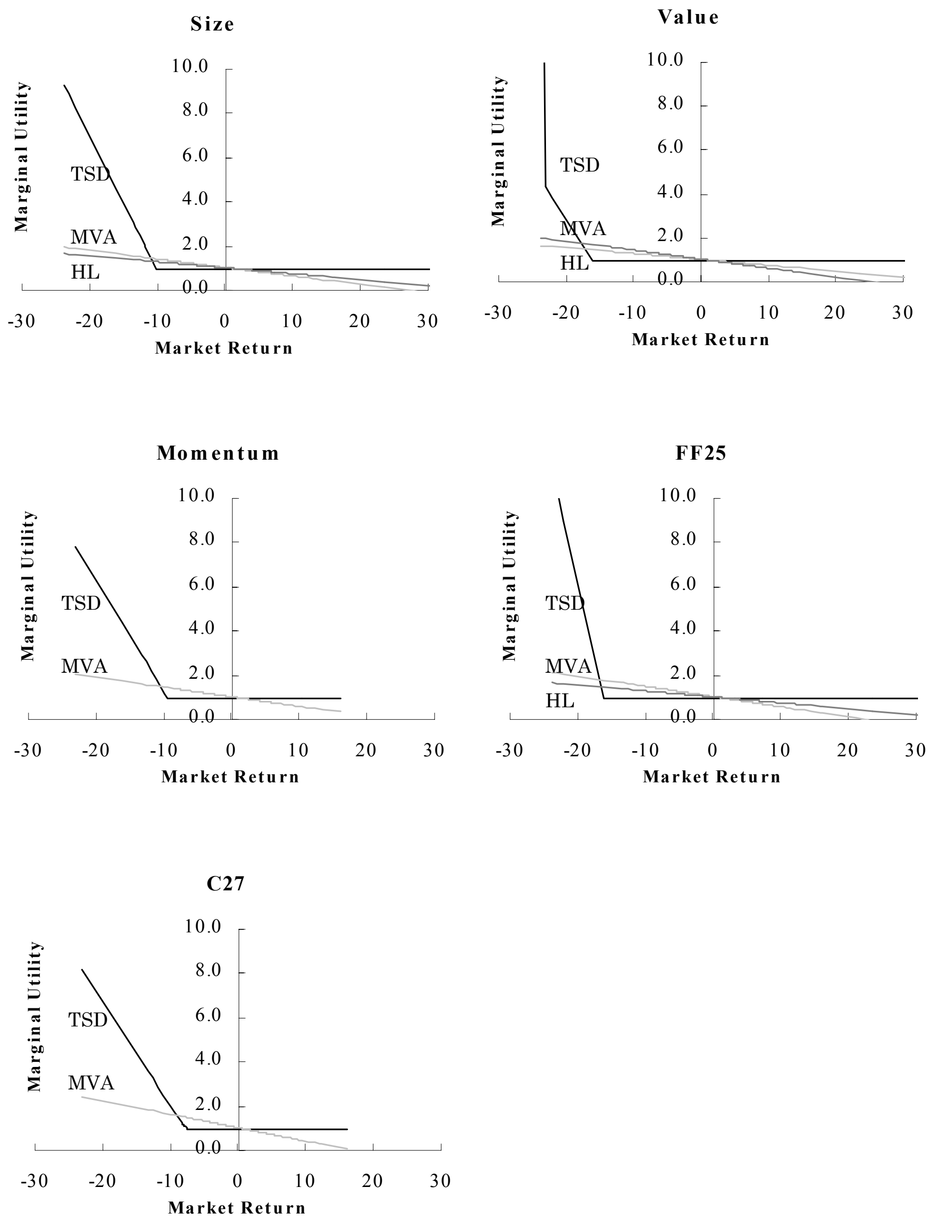


Figure 3.3: Optimal marginal utility functions. This figure shows the optimal marginal utility functions $p'(x|\boldsymbol{\beta}^*)$ for the MV, TMV and TSD tests and the size, value, momentum, FF25 and C27 data sets. The black line shows the marginal utility function under the TSD criterion, the dark grey line shows the linear marginal utility function under the TMV criterion, while the grey line shows the linear marginal utility function under the MV criterion (with nonsatiation over the sample range of market return). The marginal utility functions are standardized such that $\sum_{t=1}^{T} p'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}|\boldsymbol{\beta}^*)/T = 1$.

We do not intend to reduce investment risk to a single measure, because we believe that risk is a multi-dimensional concept; in fact, this is one of our reasons for using the TSD criterion in the first place. Still, it is useful to quantify downside risk so as to illustrate its explanatory power for TSD efficiency. It is especially useful to analyze the contribution of the tangency portfolio (TP) to the downside risk of the market portfolio. After all, the market portfolio is classified as MV inefficient because TP's contribution to the variance risk of the market portfolio cannot explain the high average return of TP. Following Price, Price and Nantell (1982) and Harlow and Rao (1989), we measure the downside risk of the market portfolio by the second lower partial moment (LPM):

$$LPM_2(m,\boldsymbol{\tau}|G) \equiv \int_{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}\leq m} (m-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})^2\, dG(\boldsymbol{x}) \tag{3.10}$$

This statistic measures the average squared deviation below the target rate of return m. We may measure the contribution of the tangency portfolio to the downside risk of the market portfolio by the lower partial moment market beta (LPM beta):

$$\beta_i(m,\boldsymbol{\tau}|G) \equiv \frac{\int_{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}\leq m} (m-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(x_i)\, dG(\boldsymbol{x})}{\int_{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau}\leq m} (m-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})\, dG(\boldsymbol{x})} \tag{3.11}$$

Since the true return distribution G is not known, we use the following sample equivalent:

$$\beta_i(m,\boldsymbol{\tau}|F_{\mathrm{X}}) = \frac{\frac{1}{T}\sum_{t=1}^{T} \max(m-\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau},0)(x_{i,t})}{\frac{1}{T}\sum_{t=1}^{T} \max(m-\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau},0)(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})} \tag{3.12}$$

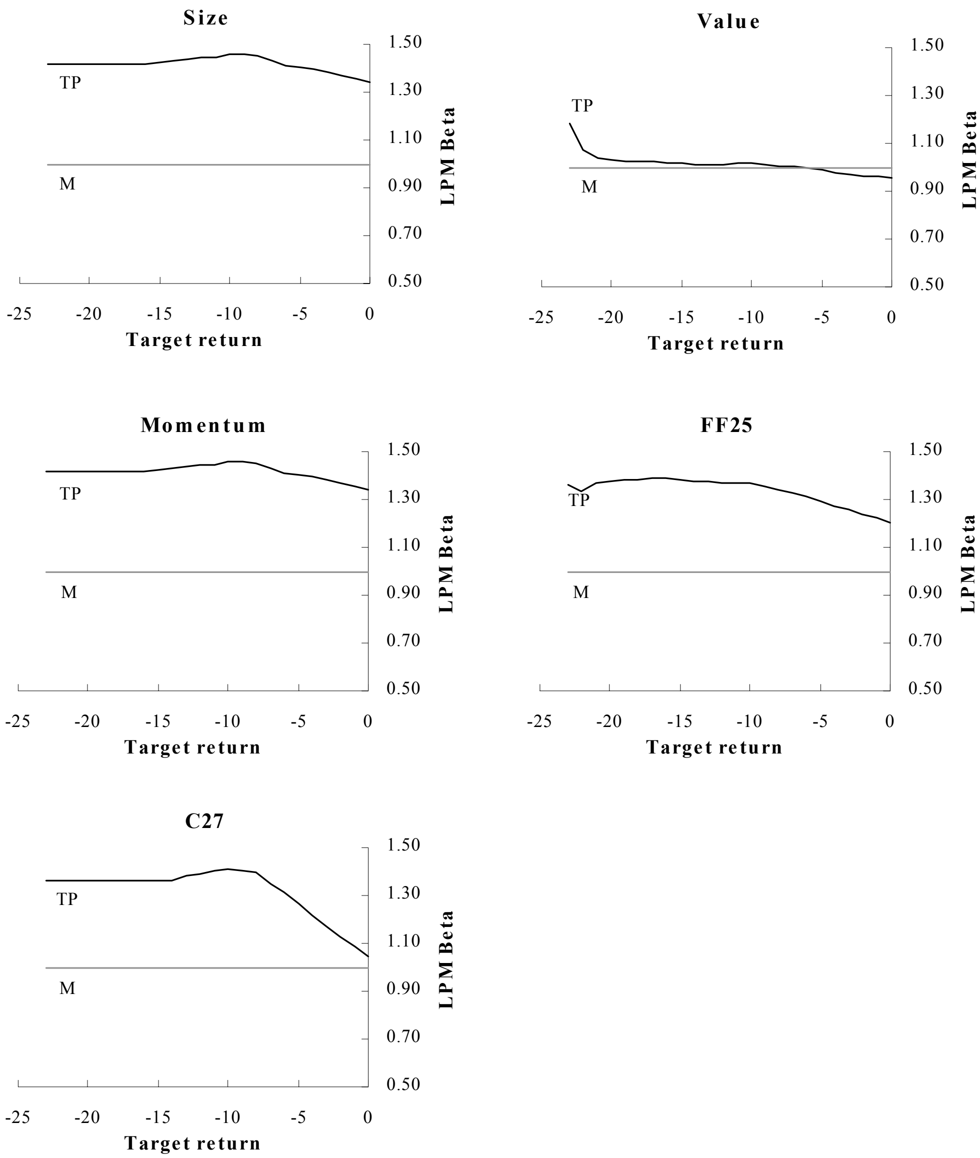


Figure 3.4: Downside betas. This figure shows the lower partial moment beta of the tangency portfolio (the portfolio with the maximal MV pricing error) for different threshold levels (m) and for each of the five data sets. The downside risk of TP increases as the threshold return is lowered. Hence, the systematic risk of the seemingly attractive portfolios (small/value/winner) increases during market crashes.

Figure 3.4 shows the tangency portfolio's LPM beta for threshold levels (m) ranging from 0 to -/-25 percent for the five benchmark portfolio sets. Clearly, for all data sets, the systematic downside risk of the tangency portfolio increases as the target return is lowered. In other words, the portfolio that seems superior to the market in mean-variance terms becomes riskier during market downturns. The pattern of increasing betas during heavy stock market losses is most noticeable for the triple sorted C27 dataset.

The pattern of the LPM betas in Figure 3.4 is similar to the pattern of the optimal utility functions in Figure 3.3. The optimal utility functions tend to kink at return levels where the LPM beta of the optimal portfolio is high. For example, in the value data set, the largest increase in downside risk occurs for target returns below -/-20 percent.

The increase in LPM beta as we lower the target return is in the range of 0.08 to 0.32. At first sight, these increases may seem relatively small. However, combined with a beta premium of about 0.387 to 0.714 percent per month (or 4.6% to 8.6% per annum, see Table 3.1), these increases yield a substantial decrease in the pricing error of the tangency portfolio.

3.4.3 Rolling window analysis

How robust are the findings for the sample period under consideration? It is possible that equity return distributions are conditionally normal, but unconditionally non-normal. After all, the risk profile of stocks and the risk preferences of investors change through time.[21] Tests for market portfolio efficiency could be affected by this time variation in risk and risk premia.

Therefore the market portfolio may be MV inefficient in the total sample, but MV efficient in the subsamples. Perhaps the TSD efficiency classification picks up this conditional pattern in the risk return relation. In addition, the degree of efficiency may change over time. For example, Gibbons, Ross and Shanken (1989) find no full sample size effect, but reject MV efficiency during the early 1960s, late 1970s and early 1980s. To control for structural variation in risk and risk premia, we employ a rolling window analysis.

With 12-month steps, we consider all 120-month samples from January 1933 to December 2002 (721 samples in total). For every subsample, we compute the p-values of the MV and TSD tests. The TMV results coincide with the MV results in the large majority of the subsamples and are not reported separately.

21 Fama and French (2002) argue that structural economic changes have lowered the expected equity premium during the last four decades. Further, the risk of stocks has changed along: for example the betas of value stocks show a structurally declining trend (for example see: Petkova and Zhang (2003)).

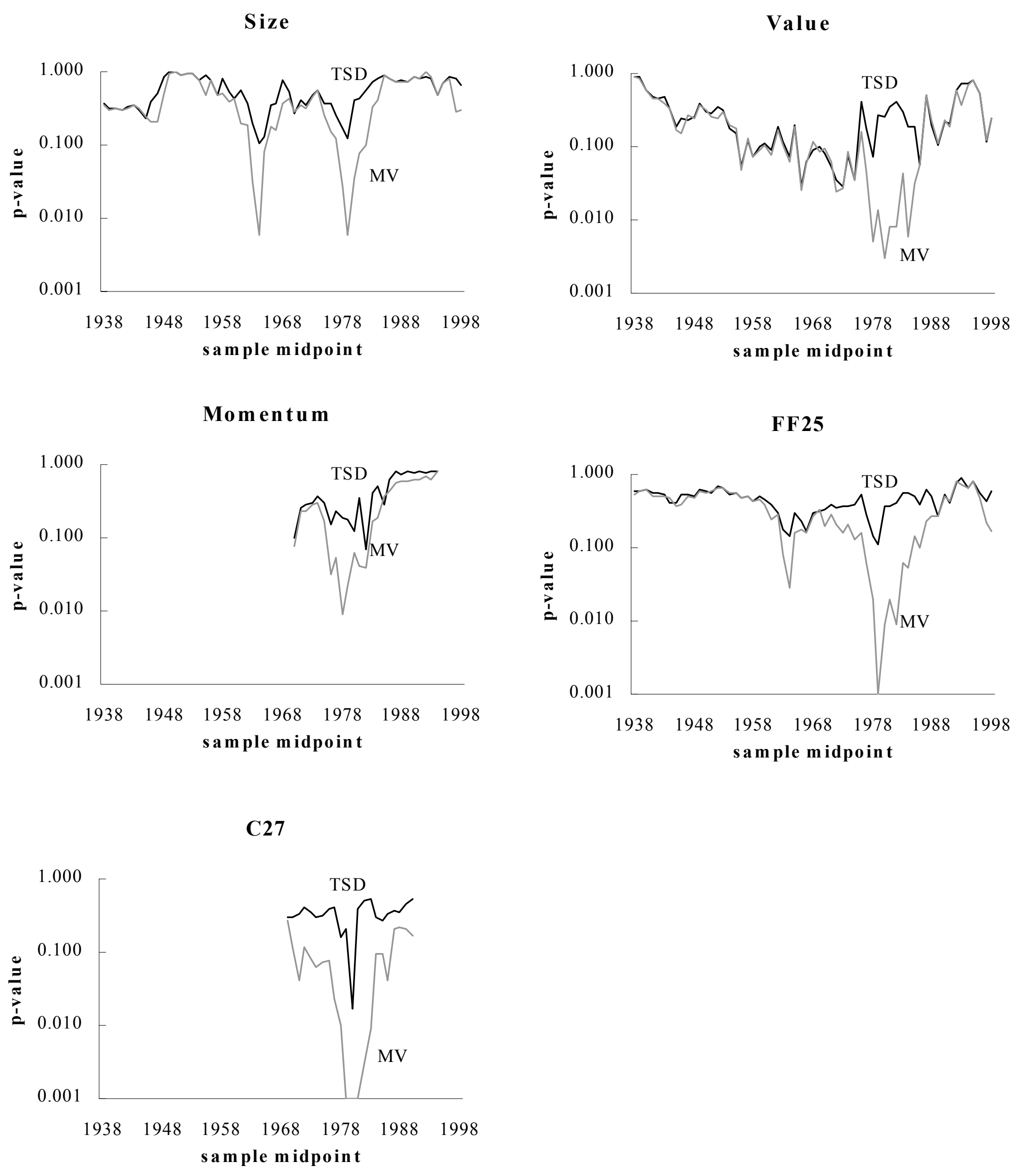


Figure 3.5: Rolling window analysis of efficiency tests. This figure shows the p-values for the MV and TSD efficiency tests in each of the five data sets using a rolling 120-months period (12-month steps). The grey line represents the p-value of the MV test and the dark line shows the TSD p-value. The figure reads as follows: consider the 1980 observation in the value data set, which represents the 120-month period starting in January 1975 to December 1984. For this period, the stock market is MV inefficient ($p=0.004$) but TSD efficient ($p=0.263$). During the 1970s and early 1980s, MV efficiency can be rejected relative to all data sets ($p<0.05$), while TSD efficiency cannot be rejected ($p>0.1$).

Interestingly, we observe in Figure 3.5 that the market portfolio is MV efficient from the early 1930s to the late 1950s. The first empirical tests of the CAPM depended heavily on this sample period. For example, the Black, Jensen and Scholes (1972) sample period ranged from 1931 to 1967. However, starting in 1960s to the early 1990s we find serious violations of MV efficiency for all sorts based on size, value and momentum. The influential Fama and French (1993, (1996) studies focus on this particularly anomalous sample period.

However, during the 1970s and the early 1980s, the MV and TSD results diverge strongly. Specifically, MV efficiency is consistently rejected, with p-values reaching levels far below ten percent. The MV-results are very similar across the various benchmarks sets with significant inefficiencies concentrated in one particular anomalous period, the period during the 1970s and the early 1980s. By contrast, TSD efficiency is not rejected relative to size, value, or momentum for mostly all subsamples. Only in a few subsamples during the 1960s for value, does the TSD p-value fall slightly below the ten percent level. In all other subsamples and for all benchmark sets we cannot reject TSD efficiency. As is true for the full sample, the implicit TSD utility functions during the 1970s and the early 1980s assign a high weight to large losses. Again, this suggests that downside risk can explain the high average returns of small/value/winner stock portfolios.

Figure 3.6 further illustrates the explanatory power of downside risk. The figure shows a rolling window analysis of the tangency portfolio's standard market beta and the LPM market beta. In every subsample, the threshold return (m) is set at the 2.5th percentile of the market return distribution. Hence, the LPM market beta measures the contribution to the left tail of the market return distribution. During most subperiods, the downside risk of the tangency portfolio is smaller than or equal to the standard market beta. Interestingly, during these periods, the market portfolio generally is efficient in terms of both the MV criterion and the TSD criterion. By contrast, during the 1970s and early 1980s, when the MV and TSD efficiency classifications diverge, TP's downside risk also sharply increases; the tangency portfolio involves substantially more downside risk than measured by the standard market beta. This pattern of downside risk is again found for all data sets. The optimal portfolio is riskier than it seems in the mean-variance framework, because its contribution to the left tail of the market return distribution is much larger than its contribution to market variance. The TSD criterion picks up this pattern of downside risk and uses it to rationalize the MV inefficiency of the market portfolio.

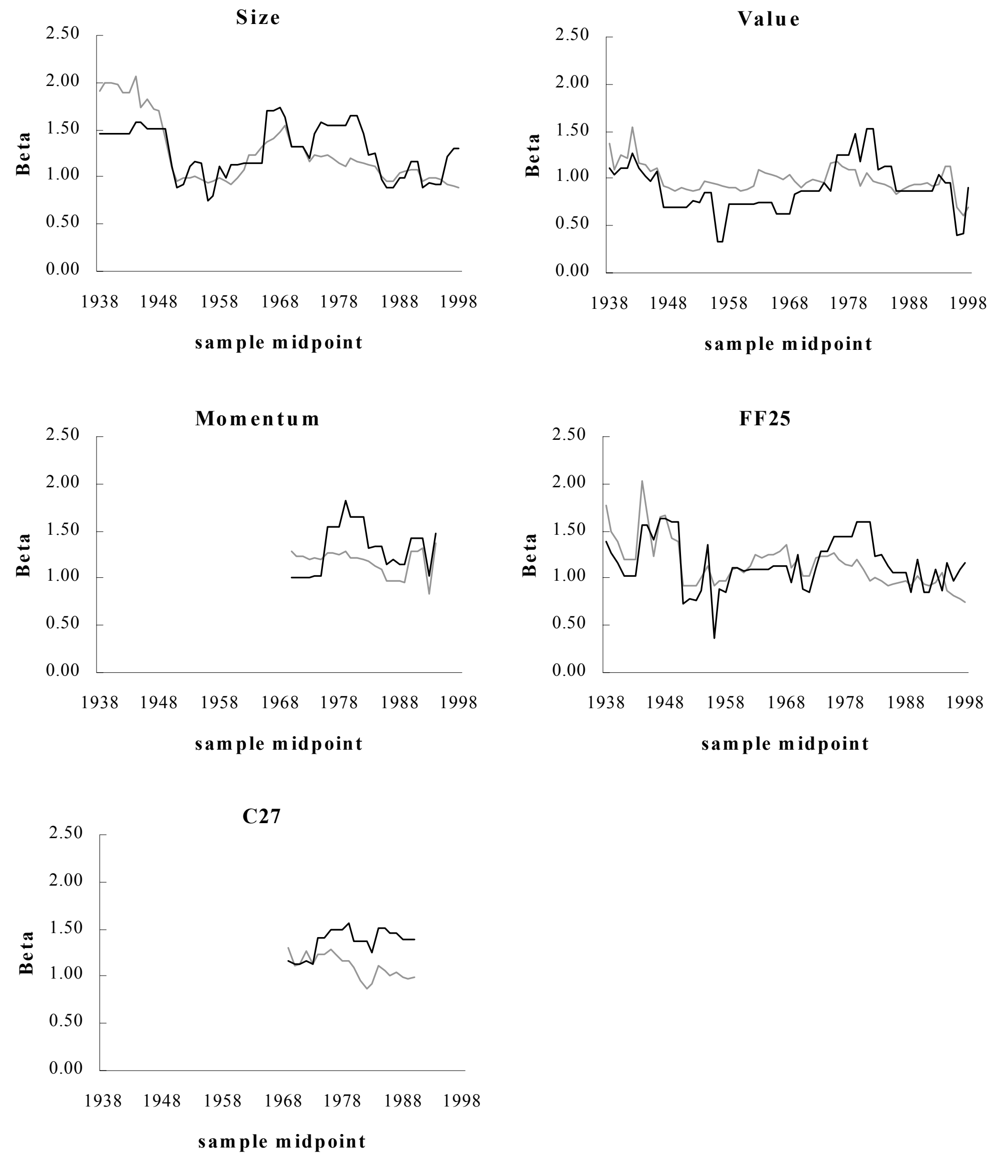


Figure 3.6: Rolling window analysis of downside betas. This figure shows the lower partial moment (LPM) beta $\beta_i(m,\tau|F_X)$ of the tangency portfolio (the portfolio with the maximal MV pricing error) compared to the traditional MV beta, using a rolling 120-months period (12-month steps). In every subsample, the threshold return (m) is set at the 2.5th percentile of the market return distribution. The rolling LPM betas of the tangency portfolio coincide or are lower than the standard MV betas during most subsamples. Interestingly, during these periods, the market portfolio is MV-efficient. However, during the anomalous 1970s and early 1980s downside risk of the tangency portfolio is higher than co-variance risk.

3.5 Conclusions

The value-weighted stock market portfolio is TSD efficient, but MV inefficient relative to benchmark portfolios formed on size, value and momentum. The TSD criterion is especially successful in rationalizing the persistent MV inefficiencies that occur in the 1970s and the early 1980s. During this period, the mean-variance tangency portfolio has relatively high downside risk, and no other portfolio yields a significantly better trade-off between mean and downside risk than the market portfolio.

We may ask if TSD efficiency of the market portfolio simply reflects a lack of power due to the use of minimal assumptions. Personally, we see the use of minimal assumptions as the strength rather than the weakness of our approach; TSD reduces the specification error that follows from ad hoc parameterization, while increasing power by imposing economically meaningful regularity conditions (nonsatiation, risk aversion and skewness preference). Also our main results rely on a relatively long time-series (as high as 840) and a narrow cross-section (as low as 10). Our simulation study shows that the TSD test is sufficiently powerful in this case.

Presumably, the difference with the results of Post (2003), who rejects market portfolio efficiency, can be attributed to Type I error caused by the use in the earlier study of the restrictive null of equal means rather than the true null of TSD efficiency and the focus on the post-1963 period. By contrast, our analysis uses the sampling distribution under the true null and also includes the pre-1963 period.

Our analysis assumes that return observations are serially identical and independent distributed (IID). However, there exists a wealth of evidence that the risk/return characteristics of securities show structural and cyclical variation. Notwithstanding the arguments in favour of time-variation, our main finding is that an unconditional model can explain the size, value and momentum puzzles. Also, conditional models entail a large risk of specification error, as a conditional model has to specify how each aspect of investor preferences and the return distribution depends on the state-of-the-world. Unfortunately, economic theory gives minimal guidance about the evolution of investor preferences and the return distribution. In addition, the problem of imposing the regularity conditions is very severe; with a conditional model, we have to make sure that the utility function is well-behaved for all possible states-of-the-world. In fact, the results of many conditional asset pricing studies can be shown to reflect severe violations of the basic conditions; see for instance Wang and Zhang (2004).

We use a simple single-period, portfolio-oriented, rational model of a competitive and frictionless capital market that is very similar to the traditional CAPM. Our explanation rests solely on a generalization of the way risk is measured in the CAPM. Of course, alternative explanations may exist for the MV inefficiencies, for example based on a multi-period, consumption-oriented model, a model with imperfect competition or market frictions, or a model where investor behave according to non-expected utility theory. It may be impossible to empirically distinguish between some of these explanations and our explanation based on downside risk. For example, liquidity effects and downside risk may be indistinguishable, because liquidity typically dries up when the largest losses occur and, in turn, liquidity dry-up may cause or amplify the losses. Similarly, the subjective overweighing of the probability of large losses will result in similar predictions as a high marginal utility for large losses. Our point is simply that a simple risk-based generalization of the CAPM suffices to explain the high average returns of small, value and winner stocks.

Appendix 3.1

Proof to Theorem 1: By construction, $\xi(\boldsymbol{\tau},F_{\mathbf{X}})$ is bounded from above by $\zeta(\boldsymbol{\tau},F_{\mathbf{X}},u)=\min_{i\in\mathrm{I}}\left\{\sum_{t=1}^{T}u'(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})(x_{it}-\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})/T\right\}$, and hence

$$\Pr[\xi(\boldsymbol{\tau},F_{\mathbf{X}})>y|H_0]\leq\Pr[\zeta(\boldsymbol{\tau},F_{\mathbf{X}},u)>y|H_0], \tag{A3.1}$$

for all $u\in U$. Interestingly, we can derive the asymptotic sampling distribution of $\zeta(\boldsymbol{\tau},F_{\mathbf{X}},u)$, $u\in\Psi(\boldsymbol{\tau},G,U)$, under H_0 from known results. Since the observations $\boldsymbol{x}_t$, $t=1,\cdots,T$, are serially IID, the vectors $u'(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}_t-\boldsymbol{e}\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})$, $t=1,\cdots,T$, are also serially IID. In general, these vectors have mean $\boldsymbol{\mu}(\boldsymbol{\tau},G,u)\equiv\int u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}-\boldsymbol{e}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})dG(\boldsymbol{x})$ and variance-covariance matrix $\int(u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}-\boldsymbol{e}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})-\boldsymbol{\mu}(\boldsymbol{\tau},G,u))(u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}-\boldsymbol{e}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})-\boldsymbol{\mu}(\boldsymbol{\tau},G,u))^{\mathrm{T}}dG(\boldsymbol{x})$. We adhere to the statistical convention of using the least favorable distribution that maximizes the *p*-value under the null. Under H_0, $\boldsymbol{\mu}(\boldsymbol{\tau},G,u)\leq 0$, and $\Pr[\zeta(\boldsymbol{\tau},F_{\mathbf{X}},u)>y|H_0]$ is maximal if $\boldsymbol{\mu}(\boldsymbol{\tau},G,u)=0$. (Note that this represents the case where all assets are included in the evaluated portfolio, that is, $\boldsymbol{\tau}>\mathbf{0}$; in this case, the first-order condition (3.1) must hold with strict equality.) Hence, for the least favorable distribution, the vectors $u'(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}_t-\boldsymbol{e}\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})$, $t=1,\cdots,T$, are serially IID draws with mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)\equiv\int u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})^2(\boldsymbol{x}-\boldsymbol{e}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}-\boldsymbol{e}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\tau})^{\mathrm{T}}dG(\boldsymbol{x})=(\mathbf{I}-\boldsymbol{e}\boldsymbol{\tau}^{\mathrm{T}})\boldsymbol{\Omega}(G,u)(\mathbf{I}-\boldsymbol{e}\boldsymbol{\tau}^{\mathrm{T}})^{\mathrm{T}}$. Therefore, the Lindeberg-Levy central limit theorem implies that the vector $\sum_{t=1}^{T}u'(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})(\boldsymbol{x}_t-\boldsymbol{e}\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\tau})/T$ obeys an asymptotically normal distribution with mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)/T$. Hence, asymptotically, $\zeta(\boldsymbol{\tau},F_{\mathbf{X}},u)$ is the largest order statistic of *N* random variables with a joint normal distribution, and we find

$$\Pr[\zeta(\boldsymbol{\tau},F_{\mathbf{X}},u)>y|H_0]=(1-\int_{z\leq y\boldsymbol{e}}d\Phi(z|\mathbf{0},\boldsymbol{\Sigma}(\boldsymbol{\tau},G,u)/T)). \tag{A3.2}$$

Combining (A3.1) and (A3.2), we asymptotically find

$$\Pr[\xi(\boldsymbol{\tau}, F_{\mathbf{X}}) > y \big| H_0] \leq (1 - \int_{z \leq ye} d\Phi(z \big| \mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\tau}, G, u) / T)), \tag{A3.3}$$

for all $u \in \Psi(\boldsymbol{\tau}, G, U)$.

Q.E.D.

Chapter 4

Conditional downside risk[1]

ABSTRACT: The mean-semivariance CAPM better explains the cross-section of US stock returns than the traditional mean-variance CAPM. If regular beta is replaced by downside beta, the cross-sectional risk-return relationship is restored. Especially during bad-states of the world, when the equity premium is high, we find a near-perfect relation between risk and return. Further, conditional downside risk (1) explains returns within the size deciles, (2) is not related to distress risk, and (3) partly explains the momentum effect.

4.1 Introduction

A WELL-KNOWN LIMITATION OF THE MEAN-VARIANCE (MV) CAPM is that variance is a questionable measure of investment risk. While investors typically assign greater importance to downside volatility than to upside volatility (due to decreasing absolute risk aversion), this measure treats downside volatility and upside volatility in the same manner. This is a powerful argument for replacing variance with measures of downside risk, as already advocated by Markowitz (1959). The mean-semivariance (MS) CAPM (Hogan and Warren (1974) and Bawa and Lindenberg (1977)) replaces variance with semivariance and replaces the regular beta with a downside beta that measures the co-movements with the market portfolio in a falling market.

The MS CAPM preserves all key characteristics of the MV CAPM, including the two-fund separation principle, efficiency of the market portfolio and the linear risk-return relationship. The only difference is the use of the relevant risk measures – variance and regular beta vs. semivariance and downside beta. The importance of this difference depends on the shape of the return distribution. For symmetrical return distributions, regular beta and downside beta are identical. However, for asymmetrical distributions, such as the lognormal, the two models diverge.

Price, Price and Nantell (1982) show that the historical downside betas of U.S stocks systematically differ from the regular betas. Specifically, the regular beta underestimates the risk for low-beta stocks and overestimates the risk for high-beta stocks. This finding may help to explain why low-beta stocks appear systematically

1 This chapter is a slightly revised version of the paper Post and van Vliet (2004c).

underpriced and high-beta stocks appear systematically overpriced in empirical tests of the MV CAPM (see for example Black, Jensen and Scholes (1972), Fama and MacBeth (1973) and Reinganum (1981)).

The MS CAPM is distinctly different from models that add higher-order central moments such as skewness and kurtosis to the mean-variance framework (see for example Kraus and Litzenberger (1976), Friend and Westerfield (1980), Harvey and Siddique (2000) and Dittmar (2002)). Specifically, the MS CAPM involves fewer parameters as it replaces variance with another risk measure and does not introduce additional risk measures. Also, the multi-moment models are not sufficiently flexible to model downside risk and it is generally difficult to restrict these models to obey the standard regularity conditions of nonsatiation (no-arbitrage) and risk aversion (see for example Levy (1969)). In fact, multi-moment models seem best suited for modeling risk seeking rather than downside risk aversion.

Surprisingly, despite the intuitive appeal of semivariance, the empirical problems of the MV CAPM and the known differences between regular betas and downside betas, the MS CAPM thus far has not been subjected to rigorous empirical testing. The few studies devoted to testing the model suffer from problems related to the data and the methodology. Jahankhani (1976) focuses on the relatively short sample period 1951-1969 that does not include the important bear markets of the 1930s, 1970s and 2000s. This may critically affect his conclusion that the MS model does not fare any better than the MV CAPM.[2] By contrast, Harlow and Rao (1989) report strong evidence in favor of the general mean-lower partial moment (MLPM) CAPM, which replaces the regular beta with a general LPM beta.[3] Unfortunately, they do not actually estimate the LPM beta due to a flaw in their empirical methodology (see Appendix 3.1 at the end of this chapter).[4] Thus, we may conclude that the MS CAPM thus far has not been subjected to unambiguous testing.

The purpose of this chapter is to fill the void by providing an empirical comparison of the MV and MS models. The study has three distinctive features. First, we pay special attention to obtaining economically meaningful results. Specifically, we require the pricing kernel to be economically well behaved in the sense that they obey the basic regularity conditions of nonsatiation and risk aversion. One approach

2 Price, Price and Nantell (1982, Footnote 25) also question Jahankani's use of very short (two-year) subsamples.

3 The downside beta is the special case of the second-order LPM beta with the riskless rate as the target rate of return.

4 The problem is that their asymmetric response model (ARM) generally does not estimate their LPM beta; see the Appendix 4.1.

to achieve this is by using non-parametric stochastic dominance tests (see Post (2003)); these tests start from the regularity conditions rather than a parameterized model. Interestingly, using these tests, Post and van Vliet (2004b) show that downside risk helps to explain the high returns earned by small caps, value stocks and recent winner stocks. This study takes an alternative, parametric approach; we fix the model parameters in the spirit of the time-series methodology of Gibbons, Ross and Shanken (1989). In this respect, our study differs from the recent downside risk paper by Ang, Chen and Xing (2004) Ang, Chen and Xing (2004), where the model parameters are fitted (rather than fixed) to optimize the statistical fit using a cross-sectional regression methodology.

Second, we employ data sets that are specially designed for the analysis of downside risk. When analyzing risk and risk preferences, it is particularly important to include periods during which investment risks are high and investors are sensitive to risk. For this reason, we use an extended sample (1926-2002) that includes the prolonged bear markets of the 1930s, 1970s and early 2000s. Further, we will use benchmark portfolios that are formed on downside beta. After all, if downside beta drives asset prices, then sorting on other stock characteristics may lead to a lack of variation in means and downside betas and thus erroneous rejections of the MS CAPM.

Third, we carry out unconditional tests as well as conditional tests that account for the economic state-of-the-world. The conditional models are particularly relevant given the mounting evidence in favor of time-varying risk and time-varying risk aversion (e.g. Jagannathan and Wang (1996) and Lettau and Ludvigson (2001)). Guaranteeing a well-behaved kernel is especially important for conditional tests. Such tests frequently calibrate the model parameters to optimize the statistical fit of the model. Unfortunately, this approach may yield economically questionable results. Specifically, the results of unrestricted conditional tests frequently violate the basic regularity conditions of nonsatiation and risk aversion. For example, Dittmar Dittmar (2002, Section IIID), shows that the apparent explanatory power of skewness and kurtosis in addition to variance can be attributed almost entirely to violations of risk aversion. Further, Wang and Zhang (2004) show that conditional models frequently imply serious arbitrage opportunities.

In this chapter, we find strong indications that conditional downside risk drives asset prices. The MS CAPM outperforms the traditional MV CAPM, both in unconditional and conditional tests. The low (high) beta stocks involve more (less) systematic downside risk than expected based on their regular betas. This pattern is especially pronounced during bad states-of-the-world, when the market risk premium is high. Further, conditional downside risk (1) explains average returns within the size deciles, (2) is not related to distress risk and (3) can partly explain the momentum effect.

Figure 4.1 illustrates our main findings with the empirical risk-return relationship of ten beta decile portfolios. Later on in this chapter we will discuss the data formation and the test methodology in greater detail. Panel A shows the typical, flat relationship between regular betas and mean returns. Low (high) beta stocks are materially under (over) priced relative to the MV CAPM. As shown in Panel B, the results improve if the regular beta is replaced by the downside beta of the MS CAPM. Panel C shows that during months that are preceded by high dividend yields (above median value), the relation between regular betas and mean return also improves. Finally, Panel D shows a near perfect fit between means and downside betas during bad states-of-the world. The mean spread is consistent with the beta spread and the equity premium during bad states.

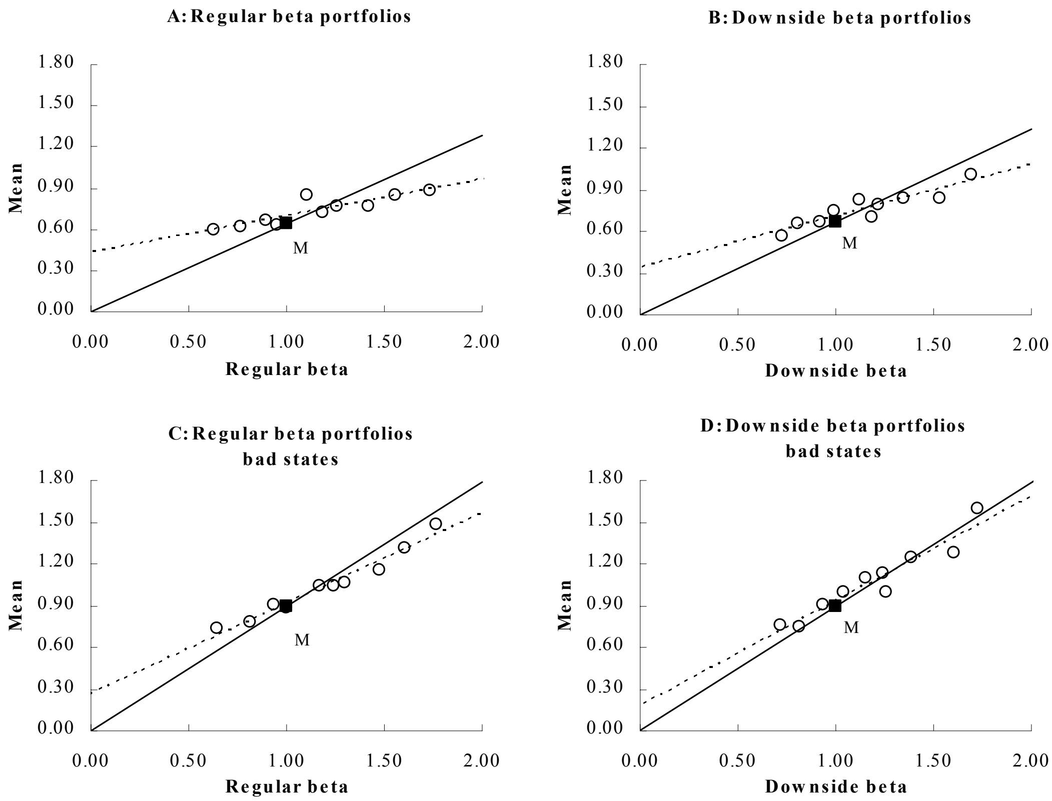


Figure 4.1: Risk-return relationship of beta portfolios 1931-2002. This figure shows the mean-beta relation for the ten beta portfolios (clear dots) and the value-weighted stock market portfolio (filled square). The straight line through the origin and the market portfolio represents the equilibrium situation. The dotted line shows the line-of-best-fit (OLS). Panel A shows the well-known relationship between regular beta and mean returns. Panel B shows the mean-beta relationship when regular beta is replaced with downside beta. Panel C shows the relation between regular beta and mean returns in bad states-of-the-world (defined as months following a bad state when the dividend yield is above its median value). Finally, Panel D shows the relation when regular beta is replaced by downside beta. For a detailed description of the data and the portfolio formation procedure, see Section 4.3.

The remainder of this chapter is structured as follows. Section 4.2 first formulates the competing capital market models in terms of pricing kernels and explains how we will select the unknown kernel parameters and determine the empirical support for the competing models. Section 4.3 discusses the data used to test the competing models. Next, Section 4.4 presents our results. Subsequently, Section 4.5 provides a discussion of the results. Finally, Section 4.6 summarizes our findings and gives some suggestions for further research. Appendix 4.1 discusses the flaw in the methodology of Harlow and Rao (1989).

4.2 Competing asset pricing models

4.2.1 Kernels

MV CAPM and MS CAPM are relatively simple single-period, portfolio-oriented, representative-investor models of a perfect capital market. Both models predict that the value-weighted market portfolio of risky assets (*M*) is efficient and that the expected return of individual assets is determined solely by their contribution to the risk of the market portfolio. In our analysis, it is useful to formulate both capital market models in terms of a pricing kernel.

The investment universe consists of N risky assets with excess returns $\boldsymbol{r} \equiv r(r_1 \ldots r_N)^{\mathrm{T}}$ and a riskless asset with a zero excess return.[5] The return on the market portfolio is given by $r_M \equiv \boldsymbol{r}^{\mathrm{T}} \boldsymbol{\tau}$, where $\boldsymbol{\tau} \equiv (\tau_1 \ldots \tau_N)^{\mathrm{T}}$ denotes the weights of the market portfolio or the relative market capitalization of the risky assets. Capital market equilibrium can be characterized using a pricing kernel $m(r_M)$ that assigns weights to the return of the market portfolio. Specifically, in equilibrium, the following equality must hold:

$$E[m(r_M)\boldsymbol{r}] = \mathbf{0}_N \tag{4.1}$$

In words, the average risk-adjusted excess return of all assets must equal zero. For different specifications of the pricing kernel, this equality is our null hypothesis throughout this chapter.

The pricing kernel can be seen as the marginal utility function of the representative investor and Equality (4.1) as the first-order condition for the portfolio optimization problem of the representative investor. In this chapter, we use this preference-based perspective. The shape of the pricing kernel and the restrictions placed on its parameters are governed by the properties of a well-behaved utility function (most notably, nonsatiation and risk aversion).

It is useful to reformulate (4.1) as the following risk-return relationship:

$$\boldsymbol{\mu} = \mu_M \boldsymbol{\beta} \tag{4.1'}$$

5 To distinguish between vectors and scalars, we use a bold font for vectors and a regular font for scalars. All vectors are column vectors and we use **Error! Objects cannot be created from editing field codes.** for the transpose of **Error! Objects cannot be created from editing field codes.**.

In equilibrium, the mean returns $\boldsymbol{\mu} \equiv E[\boldsymbol{r}]$ equal the market risk premium $\mu_M = \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\tau}$ times the market betas $\boldsymbol{\beta} \equiv \dfrac{(E[m(r_M)\boldsymbol{r}] - E[m(r_M)]E[\boldsymbol{r}])}{(E[m(r_M)r_M] - E[m(r_M)]\mu_M)}$. The betas are generalizations of the traditional market betas (see also Cochrane (2001, Section 8.4)). Specifically, the betas measure the covariance of the assets with the pricing kernel, standardized with the covariance of the market portfolio with the pricing kernel. For the unconditional mean-variance model, the generalized betas reduce to the traditional variance-based betas.

Different capital market models impose different assumptions about the pricing kernel. Using $\delta_- \equiv 1(r_M \leq 0)$, we will analyze the following four models in our analysis:

Model	Kernel ($m(r_M)$)
Unconditional mean-variance (UMV)	$b_0 + b_1 r_M$
Unconditional mean-semivariance (UMS)	$b_0 + b_1 r_M \delta_-$
Conditional mean-variance (CMV)	$(b_0 + b_1 z) + (b_2 + b_3 z) r_M$
Conditional mean-semivariance (CMS)	$(b_0 + b_1 z) + (b_2 + b_3 z) r_M \delta_-$

In the unconditional mean-variance (UMV) CAPM, the kernel is a linear function of market return. The unconditional mean-semivariance model (UMS) deviates from the MV model by using a kernel that is a linear function of market return in case of losses ($r_M < 0$) only; for gains, the kernel is flat. In the conditional versions of these models (CMV and CMS), the two parameters are linear functions of a single conditioning variable *z.* In the empirical analysis, we will condition on the dividend yield (a popular proxy for the state-of-the world) and show that similar results are obtained for other conditioning variables, such as the credit spread and the earnings yield.

In practice, we cannot directly check the equilibrium condition (4.1), because the return distribution of the assets is unknown. However, we can estimate the return distribution using time-series return observations and employ statistical tests to determine if the equilibrium condition is violated to a significant degree. Throughout the text, we will represent the observations by the matrix $\mathbf{R} \equiv (\boldsymbol{r}_1 \cdots \boldsymbol{r}_T)$, with $\boldsymbol{r}_t \equiv (r_{1t} \cdots r_{Nt})^{\mathrm{T}}$. The values of the kernel will be denoted by the vector $\boldsymbol{m} \equiv (m(r_{M1}) \cdots m(r_{MT}))^{\mathrm{T}}$. Finally, we will use $\boldsymbol{r}_M \equiv \mathbf{R}^{\mathrm{T}}\boldsymbol{\tau}$, for the market return observations.

The empirical deviations from the equilibrium equation (4.1), also known as pricing errors or alphas, are defined as

$$\hat{\boldsymbol{\alpha}} \equiv T^{-1}\mathbf{R}\boldsymbol{m} \tag{4.2}$$

The alphas can equivalently be formulated as

$$\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\mu}} - \hat{\mu}_M \hat{\boldsymbol{\beta}} \tag{4.2'}$$

with $\hat{\boldsymbol{\mu}} \equiv T^{-1}\mathbf{R}\boldsymbol{e}$, $\hat{\mu}_M = \hat{\boldsymbol{\mu}}^{\mathrm{T}}\boldsymbol{\tau}$ and $\hat{\boldsymbol{\beta}} \equiv (\mathbf{R}\boldsymbol{m} - \hat{\boldsymbol{\mu}}(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{m}))(\boldsymbol{r}_M^{\mathrm{T}}\boldsymbol{m} - \hat{\mu}_M(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{m}))^{-1}$ for the sample means and sample betas respectively.

In practice, the empirical researcher faces two issues: selecting the kernel parameters and statistical inference about the equilibrium condition based on the alphas.

4.2.2 Selecting the model parameters

Some empirical asset pricing studies select the parameters of the pricing kernel so as to minimize the alphas. Unfortunately, this approach can yield economically questionable parameter values. Most notably, the parameter values may imply arbitrage possibilities (a kernel that takes negative values) and/or risk seeking (a kernel that increases with market return).

Arbitrage opportunities are inconsistent with the basic economic concept of increasing utility of wealth, or nonsatiation. Risk seeking entails two economic problems. First, risk seeking is inconsistent with the basic economic concept of diminishing marginal utility of wealth. Second, the interpretation of the test results in terms of utility maximizing investors breaks down if we allow for risk seeking. Recall that the equilibrium condition (4.1) can be seen as the first-order condition for portfolio optimization. The first-order condition in general is not a necessary condition for a global maximum, because minima and local maxima may arise in case of risk seeking. For these reasons, a good statistical fit may come at the cost of poor economic realism. Section 4.4.1 will give some striking examples of this problem; when selected to minimize the alphas, the CMV and CMS kernels take negative values and are increasing for favorable states-of-the-world.

The UMV efficiency test of Gibbons, Ross and Shanken (GRS, 1989) circumvents this problem by fixing the kernel independently of the alphas. Specifically, this test requires a zero alpha for the market portfolio:

$$T^{-1}\boldsymbol{r}_M^{\mathrm{T}}\boldsymbol{m} = 0 \tag{4.3}$$

Also, the test standardizes the kernel by setting its sample average equal to unity:

$$T^{-1}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{m} = 1 \tag{4.4}$$

Combined, the restrictions (4.3) and (4.4) completely fix the two parameters of the UMV kernel (b_0 and b_1). The resulting kernel typically is well-behaved, that is, it obeys nonsatiation and risk aversion provided the historical market risk premium takes a moderate and positive value.[6] We will use these two restrictions for all models evaluated in this chapter. This means that our UMV alphas are identical to the GRS alphas. As for the UMV kernel, imposing (4.3) and (4.4) completely fixes the UMS kernel. Provided the historical market risk premium is positive, the resulting kernel will be well behaved.[7] For the CMV and CMS models, with four unknown parameters, the two restrictions are not sufficient to guarantee well-behavedness for every value of the conditioning variable *z*, and further restrictions are required. For this purpose, we introduce an "utopia state", characterized by an extremely favorable value for the conditioning variable, say z^*. For example, our analysis below will condition on the dividend yield (D/P) and will use a zero dividend yield for the utopia state. We assume that the representative investor is satiated (the kernel equals zero) and risk neutral (the kernel is flat) in the utopia state. This boils down to imposing the following two restrictions:

---

6 If we solve (3) and (4) for b_0 and b_1 of the UMV kernel, we find **Error! Objects cannot be created from editing field codes.** and **Error! Objects cannot be created from editing field codes.**, with **Error! Objects cannot be created from editing field codes.**. The kernel **Error! Objects cannot be created from editing field codes.Error! Objects cannot be created from editing field codes.** is decreasing (risk aversion) provided the historical market risk premium **Error! Objects cannot be created from editing field codes.** is positive. Further, the kernel is non-negative (nonsatiation) provided **Error! Objects cannot be created from editing field codes.**. This condition generally holds in samples where the historical market risk premium takes a moderate and positive value and the market return distribution is not extremely positively skewed.

7 If we solve (4.3) and (4.4) for b_0 and b_1 of the UMS kernel, we find **Error! Objects cannot be created from editing field codes.** and **Error! Objects cannot be created from editing field codes.**, with **Error! Objects cannot be created from editing field codes.** and **Error! Objects cannot be created from editing field codes.**. The kernel **Error! Objects cannot be created from editing field codes. Error! Objects cannot be created from editing field codes.** is decreasing (risk aversion) provided the historical market risk premium **Error! Objects cannot be created from editing field codes.** is positive. Further, the kernel is always non-negative (nonsatiation), because **Error! Objects cannot be created from editing field codes.**.

$$(b_0 + b_1 z^*) = 0 \tag{4.5}$$

$$(b_2 + b_3 z^*) = 0 \tag{4.6}$$

The four parameters of the conditional models are completely fixed by the four equalities (4.3)-(4.6). By imposing satiation and risk neutrality for the utopia state, we effectively avoid the possibility of a negative and/or increasing kernel for less favorable states-of-the-world. Later on, in the discussion section, we will analyze how the fixing of the model parameters influences the test results.

We stress that our approach of fixing the kernel necessarily leads to a worse statistical fit than optimizing the kernel. However, our approach ensures that the kernel is economically well behaved, in the sense that arbitrage possibilities and risk seeking are excluded. Related to this, our approach involves a higher statistical power (probability of detecting inefficiencies) than optimizing the kernel. An additional advantage of our approach is that a single kernel can be used for different benchmark sets. Thus, we do not explain different benchmark sets with different kernels.

4.2.3 Statistical inference

We now turn to the issue of statistical inference about the equilibrium condition (4.1) based on the estimated alphas. Under the null, the alphas have means $E[\hat{\boldsymbol{\alpha}}] = \mathbf{0}_N$. The covariance matrix $\boldsymbol{\Omega} \equiv E[\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\alpha}}^{\mathrm{T}}]$ of the alphas can be estimated in a consistent manner by

$$\hat{\boldsymbol{\Omega}} \equiv T^{-1}(\boldsymbol{m}^{\mathrm{T}} \otimes \mathbf{R})(\boldsymbol{m} \otimes \mathbf{R}^{\mathrm{T}}) \tag{4.7}$$

In the spirit of the Generalized Method of Moments, we can use the following test statistic to aggregate the individual alphas:

$$JT \equiv T\hat{\boldsymbol{\alpha}}^{\mathrm{T}}\hat{\boldsymbol{\Omega}}^{-1}\hat{\boldsymbol{\alpha}} \tag{4.8}$$

Assuming that the observations are serially independently and identically distributed (IID) random draws, the test statistic obeys an asymptotic chi-squared distribution with N-1 degrees of freedom. The "loss" of one degree of freedom occurs due to the restriction that the alpha of the market portfolio should equal zero (3). Thus, in case of a single risky asset (N=1), the market portfolio is fully efficient and

$JT = 0$ by construction. More generally, for N assets, the test statistic behaves as the sum of squares of N-1 contemporaneously IID random variables.

4.3 Data

4.3.1 Data sources

In the empirical analysis we use individual stock returns, index returns, hedge portfolio returns and conditional variables. The monthly stock returns (including dividends and capital gains) are from the Center for Research in Security Prices (CRSP) at the University of Chicago. The CRSP total return index is a value-weighted average of all U.S stocks included in this study. The one-month U.S Treasury bill is obtained from Ibbotson. We subtract the risk-free rate from nominal returns to obtain excess returns. The dividend and earnings yield are obtained from Robert Schiller's homepage. The credit spread is the difference between the Aaa and Baa corporate bond yields and are from the St. Louis Fed.[8] The monthly hedge portfolio returns (SMB and HML) are taken from the data library of Kenneth French.

4.3.2 Stock selection

We select ordinary common U.S stocks listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX) and NASDAQ markets. We exclude ADRs, REITs, closed-end-funds, units of beneficial interest, and foreign stocks. Hence, we only include stocks that have a CRSP share type code of 10 or 11. We require a stock to have (1) 60 months of data available (for beta estimation) and (2) information about market capitalization (defined as price times the number of outstanding shares) at formation date. Portfolio formation takes place at December of each year (except for momentum). For example, to be included at December 1930 a stock must have trading information since January 1926 and a positive market capitalization for December 1930. A stock is excluded from the analysis if there is no more price information available. In that case, the delisting return or partial monthly return provided by CRSP is used as the last return observation.

8 Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
Robert Schiller: http://www.econ.yale.edu/~shiller/data.htm
St. Louis Fed: http://research.stlouisfed.org/fred2.

4.3.3 Sample period

When analyzing risk, it is particularly important to include periods during which investment risks are high and investors are sensitive to risk. In this respect, the failure of the MS model to improve upon the MV model in the analysis of Jahankhani (1976) is presumably caused by the focus on a sample period (1951-1969) that excludes the important bear markets of the 1930s, 1970s and 2000s. Nowadays, empirical researchers often confine themselves to the post-1963 period to avoid biases associated with the Compustat database. Nevertheless, since the revision of 1999, the CRSP database is free of delisting bias and survivorship bias for the total 1926-2002 period. Therefore, when only CRSP data are used (without Compustat requirements) there is no reason to exclude the pre-1963 period. In fact, the early period seems particularly useful because it includes the bear market of the 1930s. This study will use the entire sample period of the 2002 CRSP database: January 1926 – December 2002. Furthermore, we analyze the role of downside risk in different subsamples.

4.3.4 Benchmark portfolios

Rather than analyzing all individual stocks, empirical studies generally evaluate a small set of benchmark portfolios formed from the individual stocks. This reduces the computational burden of having to analyze thousands of individual stocks and also allows the researcher to control for particular stock characteristics and for changes of those characteristics (by periodically rebalancing the portfolios). The main part of our analysis focuses on benchmark portfolios that are based on regular market beta and downside market beta. If the MV CAPM applies and regular beta drives asset prices, then sorting on other stock characteristics may lead to a lack of variation in means and erroneous rejections of the MV CAPM. Similarly, if the MS CAPM applies, then sorting stocks on downside beta maximizes the mean spread and minimizes the probability of erroneous rejections of the MS CAPM.

At the end of December of each year, all stocks that fulfill our data requirements are sorted based on beta and grouped into ten decile portfolios.[9] The sorting starts in December 1930, 60 months after the beginning of the CRSP files, because 60 months of prior data are needed for estimating the betas of the individual stocks when sorting on beta. Next, for each portfolio value-weighted returns are

9 The results are not affected by the sorting frequency. When sorting takes place on a monthly basis (instead of in December of each year) we find similar portfolio characteristics and test results.

calculated for the following next 12 months.[10] When a stock is delisted or removed from the database after formation date, the portfolio return is calculated as the average for the remaining stocks in the portfolio during the holding period.

To disentangle the effect of regular beta and downside beta, we also use double-sorted portfolio. Stocks are sorted first into regular (downside) beta quintiles and next every quintile is further divided into downside (regular) beta quintiles, giving 25 portfolios in total. Later on, in the discussion section, we will also control for size. We then first place stocks in NYSE size deciles and subsequently sort on regular beta and downside beta, yielding 100 portfolios. Also, we will use momentum portfolios to investigate if (conditional) downside risk can help to explain the returns of momentum strategies. In sum, we employ (1) regular-beta portfolios (2) downside-beta portfolios, (3) double sorted beta portfolios, (4) double sorted size/beta portfolios and (5) momentum portfolios. All data are publicly available.[11]

4.4 Results

4.4.1 Pricing kernels

Figure 4.2 shows the conditional and unconditional pricing kernels for the mean-variance and mean-semivariance models. Recall that the pricing kernels assign a weight to each scenario and are standardized such that they take an average value of unity. The shape of the unconditional kernels is determined by the historical market risk premium. Since the UMS model explains the market risk premium only with the distribution of losses, the degree of risk aversion (for losses) in this model exceeds the degree of risk aversion in the UMV model. Specifically, the slope of the UMV kernel is -/-0.022, while the slope of the UMS kernel in the loss segment is -/-0.049.

The shape of the conditional kernels is determined by the unconditional historical market risk premium and the requirement of satiation and risk neutrality in the utopia state. Since the kernels are flat in the utopia state, the slope during the worst states is much higher than the slope of the unconditional models. Specifically, for a dividend yield of 10 percent, the slope of the CMV kernel is -/-0.052 and the slope of the CMS kernel in the loss segment is -/-0.130. The conditional kernels increase with the dividend yield, reflecting that marginal utility during bad states is

10 We prefer value-weighted returns to equal-weighted returns because equal-weighted returns require continuous portfolio updating, which in practice involves high transaction costs. By contrast, value-weighted returns closely resemble a buy-and-hold strategy with relatively low transaction costs.

11 All data can be downloaded at our datacenter: www.few.eur.nl/few/people/wvanvliet/datacenter.

higher than during good states; investors fear negative stock returns during bad states-of-the-world most.

A loss experienced during good states may be assigned a lower weight than a gain experienced during bad states. For example, in October 1974, the excess return on the market portfolio was 16.1 percent. The stock market strongly recovered from a prolonged bear market during which the dividend yield had increased to 5.27. The CMV kernel takes a value of 0.87 for this month. By contrast, the October 1987 crash, with an excess return of -/-23.1 on the market, followed a prolonged bull market during which the dividend yield had fallen to 2.72. In this month, the CMV kernel takes the value 0.99, only marginally higher than the value of 0.87 for the October 1974 gain of 16.1 percent. In an unconditional model, such small differences in the weights are possible only in case of near risk neutrality. However, conditional models recognize that marginal utility is higher during bad states than during good states.

The fixed kernels are well-behaved, that is, they obey nonsatiation and risk aversion over the sample range of the market return and the dividend yield. Imposing these regularity conditions is the key motivation for fixing the kernels. By contrast, selecting the kernel to optimize the statistical fit can result in very ill-behaved kernels; see Section 4.4.1.

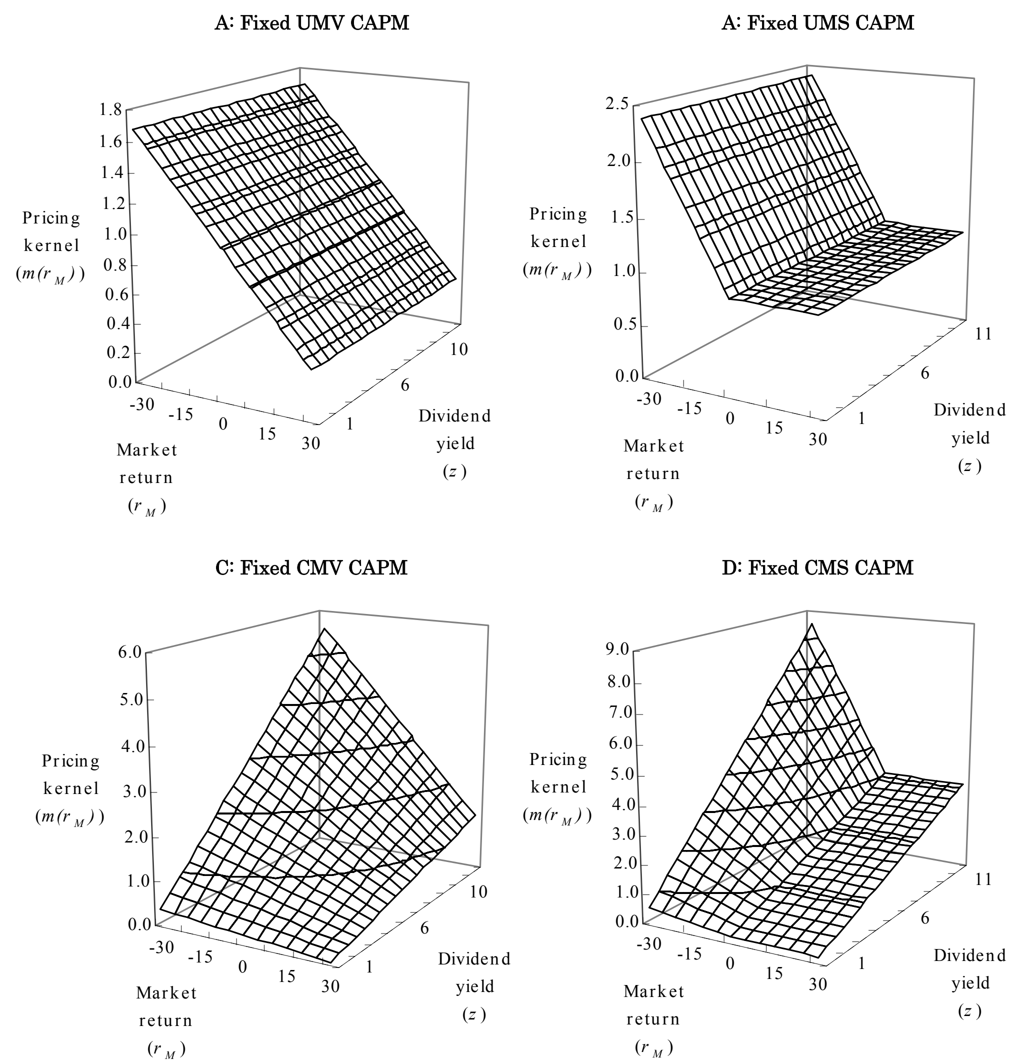


Figure 4.2: Fixed pricing kernels. The figure shows the unconditional and conditional pricing kernels for the MV CAPM and MS CAPM in the full sample (January 1931 - December 2002). The unconditional kernels are found by solving the Equalities (4.3) and (4.4) for the unknown parameters. The conditional kernels are obtained by using the one-month lagged dividend yield as the conditional variable and solving the Equalities (4.3)-(4.6) for the unknown parameters. The resulting kernels are given by $m(r_M)=1.015-0.022r_M$, (UMV) $m(r_M)=0.921-0.049r_M\delta_-$ (UMS), $m(r_M)=0.145z-0.003zr_M$ (CMV), and $m(r_M)=0.219z-0.013zr_M\delta_-$ (CMS).

4.4.2 Regular-beta-sorted portfolios

Panel A of Table 4.1 shows the descriptive statistics and test results for the regular-beta portfolios. Low (high) beta portfolios have low (high) returns, low (high) variance, negative (positive) skewness and low (high) kurtosis. From the values for skewness and kurtosis we can see that the return distribution is not normal and hence that the MV CAPM and MS CAPM can be expected to give different results. Consistent with other studies employing beta-sorted portfolios, low-beta stocks are underpriced and high-beta stocks are overpriced in the UMV CAPM. The lowest-beta portfolio and the highest-beta portfolio have regular betas of 0.63 and 1.74 respectively, a beta spread of 1.11. Given the market risk premium of 0.64 percent, this beta spread is too large compared to the mean spread of 0.29 (0.89-/-0.60).

Consistent with Price, Price and Nantell (1982), the downside betas are higher than the regular betas for the low-beta portfolios, while the downside betas are smaller than the regular betas for the high-beta portfolios. For example, the downside beta of the lowest-beta portfolio is 0.66, while the highest-beta portfolio has a downside beta of 1.68. The beta spread thus decreases from 1.11 to 1.02. As a result, the UMS alphas are smaller than the UMS alphas and the overall *p*-value rises from 0.14 to 0.25.

Apart from downside risk, time-variation also helps to explain the returns of the beta portfolios. Specifically, the betas of the low (high)-beta stocks increase (decrease) during bad times, when the market risk premium is high. For example, the conditional market beta of the lowest-beta portfolio is 0.72, an increase of 0.09 relative to the UMV model, while the conditional market beta of the highest-beta portfolio is 1.55, a decrease of 0.19, and the beta spread falls from 1.11 to 0.83. Overall, the conditional model gives a substantially better fit than the unconditional model; the *p*-value increases from 0.14 to 0.83.

The best fit is obtained with the CMS model, which combines the explanations of downside risk and time-variation. Low beta (high beta) stocks are substantially riskier (less risky) than the unconditional regular beta suggests. For example, the conditional downside beta of the lowest-beta portfolio is 0.78, an increase of 0.15 relative to the UMV model, and the conditional downside beta of the highest-beta portfolio is 1.41, a decrease of 0.33, reducing the beta spread to 0.63. Compared with the UMV model, the alphas show substantial reductions. The largest positive pricing error drops from 0.18 to 0.08 and the largest negative pricing error goes from -/-0.27 to -/-0.06. Overall, the CMS model gives a near-perfect fit, with a *p*-value of 0.98.

In brief, while the UMV model performs poorly for beta-sorted portfolios, accounting for downside risk and for time-variation substantially improves the fit. In fact, the combined effect is strikingly good.

Table 4.1
Descriptives, alphas and betas for beta portfolios

This table shows descriptive statistics for the monthly excess returns of the ten regular-beta portfolios (N=10) and the ten downside-beta portfolios. The sample period is from January 1926 to December 2002 of which the first five years are used for beta estimation. Portfolio returns cover the January 1931 to December 2002 period (T=864 months). In December of each year stocks are sorted in ten decile portfolios based on historical betas. The portfolios are constructed such that each portfolio contains an equal number of stocks. In Panel A and B, the results for the regular-beta and downside-beta portfolios are showed respectively. The alphas and betas for each of the following four models are shown: (1) unconditional mean-variance CAPM, (2) conditional mean-variance CAPM, (3) unconditional mean-semivariance CAPM, and finally the (4) conditional mean-semivariance CAPM. The last two columns show the test results for the joint hypothesis that the alphas equal zero. The optimally weighted alphas (JT) are chi-squared distributed with 9 (N-1) degrees of freedom.

Panel A: Regular-beta portfolios

		Low	2	3	4	5	6	7	8	9	High	*JT*	*p*
Stats.	Mean	0.60	0.62	0.67	0.63	0.85	0.72	0.77	0.77	0.85	0.89		
	Stdev	4.26	4.64	5.23	5.49	6.34	6.77	7.19	8.22	9.15	10.39		
	Skewness	-0.15	0.40	0.55	0.56	0.81	0.77	0.77	1.45	1.48	1.19		
	Kurtosis	6.1	8.0	9.4	8.9	10.2	10.4	10.5	14.4	14.7	11.0		
Alphas	UMV	0.18	0.11	0.07	0.00	0.12	-0.07	-0.07	-0.18	-0.19	-0.27	13.5	0.14
	UMS	0.16	0.11	0.08	0.01	0.14	-0.05	-0.06	-0.13	-0.13	-0.23	11.3	0.25
	CMV	0.12	0.05	0.08	0.03	0.06	-0.04	-0.06	-0.13	-0.11	-0.15	5.0	0.83
	CMS	0.08	0.05	0.08	0.04	0.07	-0.01	-0.04	-0.03	0.02	-0.06	2.4	0.98
Betas	UMV	0.63	0.77	0.89	0.95	1.10	1.18	1.26	1.42	1.56	1.74		
	UMS	0.66	0.77	0.88	0.93	1.07	1.15	1.23	1.35	1.48	1.68		
	CMV	0.72	0.85	0.88	0.91	1.19	1.14	1.24	1.35	1.44	1.55		
	CMS	0.78	0.85	0.88	0.89	1.17	1.09	1.21	1.19	1.25	1.41		

Panel B: Downside-beta portfolios

		Low	2	3	4	5	6	7	8	9	High	*JT*	*p*
Stats.	Mean	0.56	0.66	0.67	0.75	0.83	0.70	0.79	0.83	0.84	1.01		
	Stdev	4.30	4.70	5.20	5.81	6.55	7.23	7.26	8.26	9.80	10.86		
	Skewness	-0.35	0.29	0.05	0.93	0.80	1.55	0.89	1.53	1.96	1.42		
	Kurtosis	5.1	6.8	6.4	12.0	10.2	16.8	10.3	15.4	19.5	12.0		
Alphas	UMV	0.11	0.13	0.07	0.06	0.06	-0.14	-0.05	-0.12	-0.27	-0.17	12.3	0.12
	UMS	0.08	0.12	0.06	0.09	0.08	-0.09	-0.02	-0.07	-0.18	-0.12	8.2	0.52
	CMV	0.11	0.07	0.08	0.05	0.07	-0.12	0.00	-0.07	-0.23	-0.04	10.7	0.30
	CMS	0.05	0.06	0.05	0.08	0.10	-0.05	0.05	0.04	-0.04	0.06	3.6	0.93
Betas	UMV	0.67	0.80	0.90	1.02	1.15	1.26	1.26	1.43	1.66	1.76		
	UMS	0.72	0.81	0.92	0.99	1.12	1.19	1.22	1.35	1.54	1.69		
	CMV	0.67	0.88	0.88	1.04	1.14	1.24	1.18	1.35	1.61	1.57		
	CMS	0.77	0.90	0.92	1.00	1.09	1.12	1.11	1.18	1.33	1.42		

Panel A of Figure 4.3 further illustrates the role of conditional downside risk. The figure shows the regular beta and the downside beta for the lowest-regular-beta portfolio and the highest-regular-beta portfolio as a function of the dividend yield (our proxy for the state-of-the-world). In the figure, we see a narrowing of the beta spread during the bad states (high dividend yield), which helps to explain the success of the conditional models. This narrowing is most pronounced for the downside betas.

Specifically, while the regular beta and the downside beta of the low-beta portfolio show a similar increase in bad states, the downside beta of the high-beta portfolio falls significantly below the regular beta (which does not exhibit a clear trend) in bad states. This reflects the improvement of the CMS model compared to the CMV model.

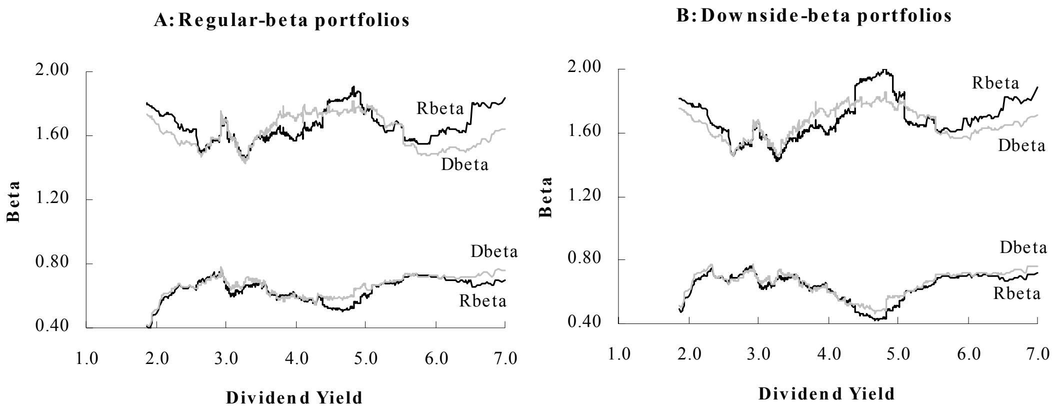


Figure 4.3: Conditional MV and MS betas. This figure shows the regular beta (dark line) and the downside beta (pale line) of the lowest-beta portfolio and the highest-beta portfolio. We use a rolling 120-months window (1-month steps) after sorting the data based on the one-month-lagged dividend yield. Panel A shows the results for the regular-beta portfolios and Panel B shows the results for the downside-beta portfolios.

4.4.3 Downside-beta-sorted portfolios

Panel B of Table 4.1 shows the results for the downside-beta portfolios. As in Ang, Chen and Xing (2004), the variation in mean returns of the downside-beta portfolios increases relative to the regular-beta portfolios. Specifically, the mean spread increases from 0.29 to 0.45 percent per month, while the downside beta spread slightly decreases (1.02 to 0.97). This already gives a first indication that downside beta is more relevant than regular beta. The results of the chi-squared test confirm this impression.

As for the regular-beta portfolios, time-variation and downside risk lead to substantial reductions of the alphas. However, time-variation becomes less important, while downside risk becomes more important. Panel B of Figure 4.3 illustrates this finding. Due to the betas of low-downside-beta stocks being higher than those for the low-regular-beta stocks in good states (low dividend yield), there no longer is a clear narrowing of the beta spread; the regular-beta spread increases slightly, while the downside-beta spread decreases slightly. This illustrates the limited role of conditioning information for the downside-beta portfolios. By contrast, the difference between regular beta and downside beta becomes more important.

4.4.4 Double-sorted portfolios

The above results for regular-beta portfolios and downside-beta portfolios provide evidence that downside beta, rather than regular beta, drives expected returns. Still, regular beta and downside beta are highly correlated. To disentangle the effect of the two risk measures, we apply a double-sorting routine. We sort stocks first into quintile portfolios based on regular-beta and then subdivide each regular-beta quintile into five portfolios based on downside beta. In addition, we sort first on downside beta and then on regular beta. The two resulting datasets of 25 portfolios isolate the separate effects of downside beta and regular beta.

The results in Table 4.2 further strongly support the conclusion that downside beta is more relevant than regular beta. In Panel A we see that average return is positively related with downside beta within each regular-beta quintile. Overall, the average return of low downside-beta portfolios is 0.71 percent compared to 0.89 for the high downside-beta portfolios. Thus controlled for regular-beta, the positive relation between mean and downside beta remains intact. By contrast, Panel B shows that the positive relation between average return and regular beta disappears (becomes flat/negative) within the downside-beta quintiles. Controlled for downside beta, the average return of low regular-beta portfolios is 0.83 percent compared to 0.71 for the high regular-beta portfolios. Apparently, the positive relation between regular beta and mean returns in Table 4.1 and Panel A of Table 4.2 is due to the fact that regular beta and downside beta are so highly correlated. Separating the effects of the two betas shows that downside beta drives average returns.

Table 4.2
Double-sorted beta portfolios

This table shows the average monthly excess returns of 25 regular-beta/downside-beta portfolios and 25 downside-beta/regular-beta portfolios. The sample period (T=864 months), data requirements, and sorting frequency are identical to those for the beta portfolios. The portfolios are constructed such that each portfolio contains an equal number of stocks. In Panel A, the stocks are first sorted into five regular-beta quintile portfolios and then into five downside-beta quintile portfolios. In Panel B the stocks are first sorted into five downside-beta quintile portfolios and then into five regular-beta portfolios. The last rows and columns give the (equal weighted) average returns across the portfolios.

Panel A: Regular beta / Downside beta

		Downside beta					
		Low	2	3	4	High	Avg.
Reg. beta	Low	0.57	0.64	0.57	0.78	0.74	0.64
	2	0.54	0.72	0.75	0.77	0.78	0.67
	3	0.70	0.76	0.85	0.96	0.96	0.80
	4	0.85	0.76	0.73	0.87	0.97	0.81
	High	0.77	0.89	0.77	1.04	1.04	0.84
	Avg.	0.71	0.76	0.75	0.88	0.89	0.76

Panel B: Downside beta / Regular beta

		Regular beta					
		Low	2	3	4	High	Avg.
Down. beta	Low	0.56	0.66	0.62	0.60	0.65	0.62
	2	0.72	0.74	0.78	0.73	0.65	0.71
	3	0.94	0.81	0.87	0.83	0.71	0.80
	4	0.98	0.93	0.78	0.82	0.83	0.84
	High	1.14	0.74	0.97	0.90	1.00	0.93
	Avg.	0.83	0.79	0.80	0.76	0.71	0.76

4.4.5 Further analysis

Table 4.3 shows how robust our results are for (1) the specific sample period and (2) the conditioning variable. The conditional models are not included for the subsamples, because splitting the sample greatly reduces the variation in the conditioning variables, hence reducing the added value of these models.

Panel A shows the split sample results. The sample is divided into subsamples of equal length based on historical time period and state-of-the-world (dividend yield). Clearly, the role of downside risk is most pronounced in the first subsample (1931-1966) and the bad-state subsample. Both subsamples include the bear market of the 1930s. This illustrates the importance of including this specific period when analyzing downside risk. In the more recent subsample (1967-2002) the UMS and

UMV models show similar performance for both datasets. However, sorting on downside beta still yields a better fit than sorting on regular beta.

Panel B further investigates how the results are affected by the choice for the specific conditioning variable (z). The dividend price ratio is possibly influenced by a structural change in dividend policy. Nowadays, firms use share repurchases as a way of returning earnings to stockholders, which structurally lowers the dividend yield. Therefore, we also employ the earnings yield (EY) and credit spread (CS) as conditioning variables. For the conditional models, the utopia state (z^*) is again characterized by a zero value for these variables. In brief, we find that using the earnings yield leads to a worse fit and using the credit spread leads to a better fit, especially for the CMV model. While the CMV model depends heavily on the choice for the specific conditional variable, the CMS model gives a good fit (p>0.68) for all conditioning variables.

Table 4.3
Robustness analysis

This table shows the split sample results for the regular-beta (RB) and downside-beta (DB) portfolios, as well as results for different conditional variables. The total sample is divided into subsamples of equal length (T=432 months) based on historical time period and the state-of-the world (dividend yield). In Panel A results are shown for the unconditional mean-variance (UMV) and mean-semivariance (UMS) models. Panel B shows the results for the conditional mean-variance (CMV) and mean-semivariance (CMS) if dividend yield is replaced with the one-month lagged earnings yield (EY) and credit spread (CS). The test statistic (JT) and levels of significance (p) are reported.

Panel A: Split samples

Sample		1931-2002		1931-1966		1967-2002		Bad state		Good state	
		JT	*p*	*JT*	*p*	*JT*	*p*	*JT*	*p*	*JT*	*p*
RB	UMV	13.5	0.14	7.7	0.56	13.3	0.15	3.2	0.95	18.5	0.03
	UMS	11.3	0.25	4.4	0.88	13.5	0.14	1.5	1.00	19.3	0.02
DB	UMV	12.3	0.12	12.3	0.20	8.7	0.47	15.3	0.08	9.3	0.41
	UMS	8.2	0.52	7.1	0.63	8.3	0.50	7.2	0.61	9.6	0.38

Panel B: Other conditional variables

Variable		DY		EY		CS	
		JT	*p*	*JT*	*p*	*JT*	*p*
RB	CMV	5.0	0.83	11.6	0.24	4.3	0.89
	CMS	2.4	0.98	6.6	0.68	3.8	0.92
DB	CMV	10.7	0.30	17.1	0.05	2.3	0.99
	CMS	3.6	0.93	5.2	0.81	1.0	1.00

4.5 Discussion

When discussing our findings with colleagues, we have often encountered various questions. Below, we briefly summarize the most common questions (seven in total) and attempt to answer these questions.

4.5.1 Kernel fitting

In your study, you have fixed the kernel to impose the regularity conditions of nonsatiation and risk aversion. Do fitted kernels really exhibit strong violations of the regularity conditions?

Yes. To illustrate the need to impose economic structure, Figure 4.4 shows the kernels that are obtained if the parameters are selected to optimize the statistical fit (*JT*). The kernels take negative values (violating nonsatiation) and are positively sloped (violating risk aversion) for a large fraction of the observations. Such kernels are economically irrelevant because they violate the no-arbitrage rule and also the Euler equation is no longer a sufficient optimization condition if concavity is violated. A statistically good fit for such kernels does not mean that we have found an economically meaningful explanation.

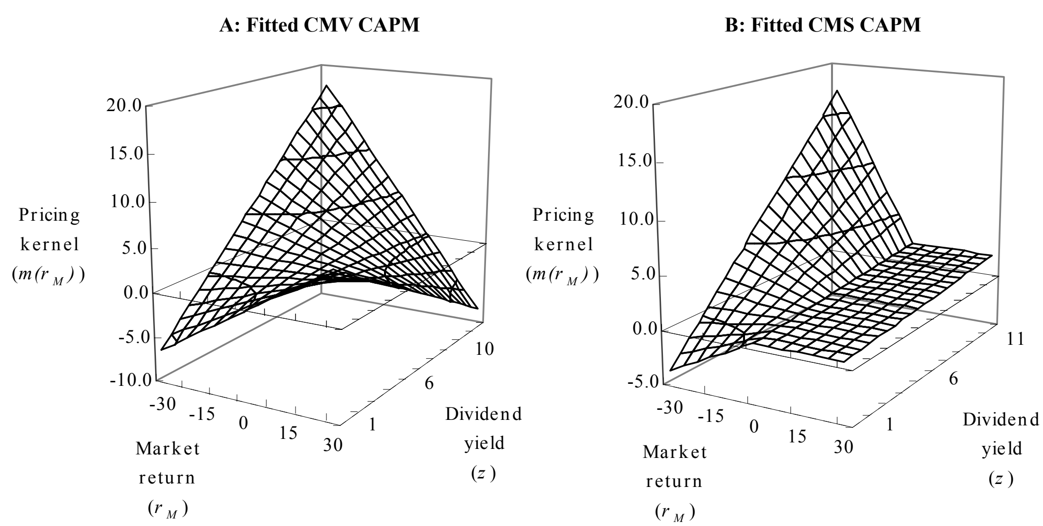


Figure 4.4: Fitted conditional pricing kernels. The figure shows the fitted CMV and CMS pricing kernels for the full sample (January 1931 - December 2002) and with the one-month lagged dividend yield as the conditioning variable. The fitted kernels are determined by maximizing the empirical fit (*JT*) relative to the ten regular beta-sorted portfolios, while maintaining conditions (3) and (4). The resulting kernels are given by $m(r_M) = -0.493 + 0.363z + (0.252 - 0.049z)r_M$ (CMV), and $m(r_M) = 0.454 + 0.120z + (0.187 - 0.049z)r_M\delta_-$ (CMS).

4.5.2 Skewness

The difference between variance and semi-variance seems especially important for skewed return distributions. Does the mean-variance-skewness model of Kraus and Litzenberger (1976) give the same results as the mean-semivariance model?

No. The three-moment (3M) CAPM of Kraus and Litzenberger replaces the traditional linear pricing kernel with a quadratic pricing kernel. Unfortunately, the explanatory power of skewness is very limited if we require the kernel to obey risk aversion (see for example Dittmar (2002), Section IIID). Specifically, it follows from the theoretical analysis of Tsiang (1972) that a linear kernel gives a good approximation for any continuously differentiable and decreasing kernel over the typical sample range of asset returns, and that a quadratic kernel is unlikely to improve the fit. Interestingly, this argument does not apply to semi-variance, because this risk measure is associated with a two-piece linear kernel, for which a linear function generally cannot give a good approximation.

Figure 4.5 illustrates this point using our data set of regular-beta portfolios. Panel A shows the cubic kernel $m(r_M) = b_0 + b_1 r_M + b_2 r_M^2$ with the parameters selected to optimize the fit (*JT*) under conditions (4.3) and (44.). The resulting kernel clearly is ill-behaved, as it severely violates risk aversion. Panel B shows the results that are found if we require the kernel to obey nonsatiation and risk aversion over the sample range of market return. The resulting kernel comes very close to UMV CAPM kernel. Clearly, a cubic kernel is not sufficiently flexible to capture downside risk aversion if risk seeking for gains is excluded. Indeed, the restricted 3M CAPM gives a worse fit than the UMS CAPM (*JT*=12.5 vs. *JT*=11.3), even though the model has one additional parameter that is calibrated to optimize the fit.

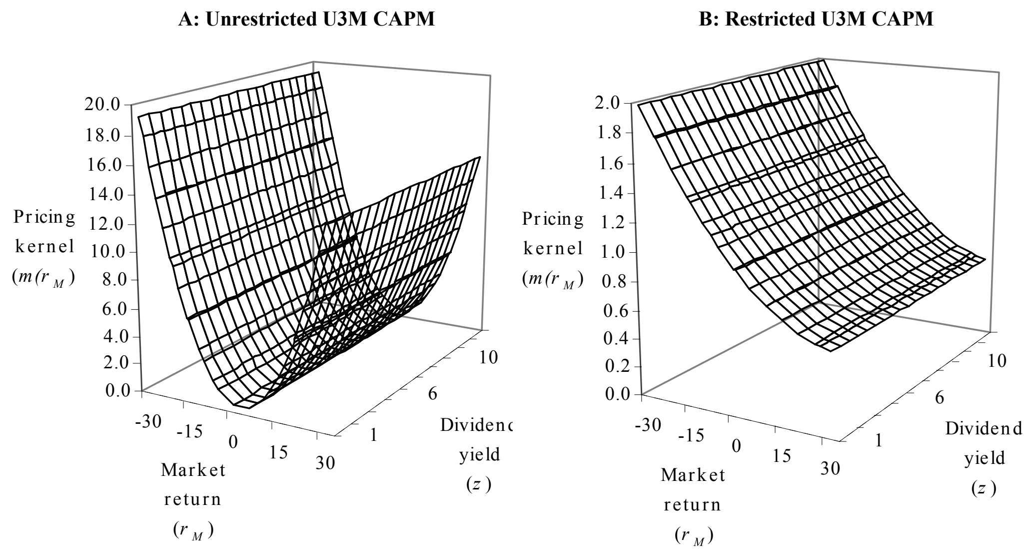


Figure 4.5: Quadratic pricing kernels. The figure shows the unconditional cubic pricing kernels for the three-moment (3M) CAPM using the full sample (January 1931 - December 2002). The kernel in Panel A is determined by maximizing the empirical fit (*JT*) relative to the ten beta-sorted portfolios, while maintaining conditions (4.3)-(4.4). Panel B shows the results obtained if we add the restrictions of nonsatiation and risk aversion for the sample range of the market return. The resulting kernels are given by $m(r_M) = 0.520 - 0.091 r_M + 0.018 r_M^2$ (unrestricted), and $m(r_M) = 1.006 - 0.024 r_M + 0.003 r_M^2$ (restricted).

4.5.3 Other LPMs

The MS CAPM uses semi-variance, which is the second-order lower partial moment (LPM) with the riskless rate for the target rate of return. Since there exist few prior arguments for selecting the order or the target rate, it would be interesting to see the results for other LPMs.

The general mean-LPM (MLPM) CAPM can be represented by the pricing kernel $m(r_M) = (b_0 + b_1 z) + (b_2 + b_3 z)\min(r_M - c)^{k-1}$ with c for the target rate of return and k for the relevant order of the LPM norm. Again we fix the model parameters by the four equalaties (3)-(6), but only alter the LPMs. The MS CAPM model is the special case with c=0 and k=2. Panel A of Table 4.4, unconditional LPM, shows the *JT* statistic for various combinations of c and k. For the regular-beta portfolios, the best fit is obtained for c=-/-10 and k=2, that is, the variance below a return level of -/-10 percent. This suggests that tail beta rather than downside beta even better captures the returns of the regular-beta portfolios. The best fit is obtained for c=-/-15 and k=1 for the downside beta portfolios, while the second best fit is for the MS model.

Panel B of Table 4.4 shows that results for conditional LPM models confirm the results for the uncondional LPM. Again, for the regular-beta portfolios, the best fit is obtained for the variance below a return level of -/-10 percent. Similarly, for the downside beta portfolios semi-variance is once more the optimal LPM. In brief, the beta portfolios are best described by a MLPM CAPM with a low target rate, but the MS CAPM gives the best fit for the downside-beta portfolios.

Table 4.4
Sensitivity LPM model

This table shows the test results for different LPM norms. The general MLPM CAPM is represented by the pricing kernel **Error! Objects cannot be created from editing field codes.**. The threshold (c) varies from -15 percent to +15 percent with 5 percent steps and the LPM order (k) varies from 1 (=expected loss below the target) to 4 (=kurtosis below the target). Each cell contains the test statistic (*JT*). The table shows the test results for regular beta portfolios (RB) and downside beta portfolios (DB). Panel A and B show the results for unconditional LPM and conditional LPM respectively.

Panel A: Unconditonal LPM

		RB				DB			
Order (*k*)		1	2	3	4	1	2	3	4
Threshold (*c*)	-15%	9.0	9.2	10.4	12.6	**7.8**	8.5	10.0	12.0
	-10%	8.8	**7.7**	9.1	10.3	11.1	8.6	8.9	9.9
	-5%	11.1	9.0	8.5	9.1	8.2	8.5	8.8	9.2
	0%	13.9	11.3	9.5	9.0	8.2	8.2	8.5	8.8
	5%	13.5	12.1	10.6	9.6	14.3	8.7	8.5	8.6
	10%	13.4	12.4	11.3	10.3	26.8	9.9	8.8	8.6
	15%	16.7	12.5	11.6	10.8	27.8	11.1	9.3	8.8

Panel B: Conditonal LPM

		RB				DB			
Order (*k*)		1	2	3	4	1	2	3	4
Threshold (*c*)	-15%	2.2	1.9	2.1	3.0	4.3	4.9	6.1	7.8
	-10%	3.4	**1.8**	1.9	2.2	5.2	4.5	5.1	6.0
	-5%	2.2	2.0	1.8	1.9	3.5	4.1	4.5	5.1
	0%	4.7	2.4	1.8	1.8	3.9	**3.6**	4.1	4.6
	5%	5.6	2.9	2.1	1.8	9.4	4.0	3.9	4.3
	10%	9.4	3.2	2.4	2.0	26.3	5.1	4.1	4.1
	15%	10.0	3.5	2.6	2.2	28.3	6.2	4.5	4.2

4.5.4 Relation to size

The risk of stocks seems related to market capitalization (ME). Fama and French (1992) convincingly show that the MV CAPM fails within different size deciles. Does the MS CAPM do any better in this respect?

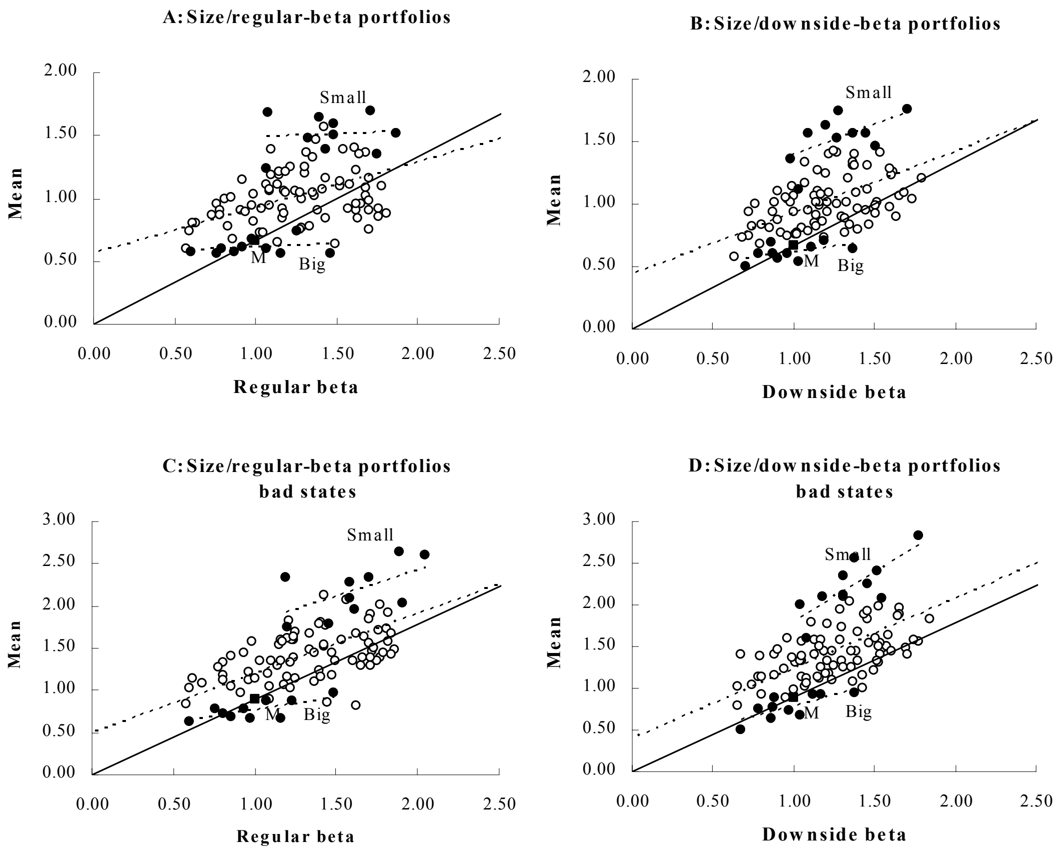


Figure 4.6: Risk-return relationship of size/beta portfolios 1931-2002. This figure shows the mean-beta relation of 100 double size/beta portfolios (clear dots). The smallest and largest decile portfolios (filled dots) and the value-weighted stock market portfolio (filled square) are shown separately. The individual are sorted first on NYSE size decile breakpoints and then on beta. The straight line through the origin and the market portfolio represents the equilibrium situation. The dotted lines give the line-of-best-fit (OLS) for (1) all portfolios and (2) the portfolios within the largest and smallest size deciles. Panel A shows that mean and beta are not related within the different size deciles. Panel B shows the mean-beta relationship when regular beta is replaced with downside beta. Finally, Panel C and D show the corresponding relationship during bad states-of-the-world (defined here as states with a dividend yield above its median value).

Yes. Figure 4.6 shows that the positive risk-return relation is restored within the different size deciles if we replace beta with downside beta. In Panel A of the figure, the risk-return relation is flat within the smallest and largest size deciles. Although Fama and French (1992) employ a shorter sample (1941-1990) that excludes the

1930s and 1990s, they find similar results in their sample. Panel B shows how the beta spread decreases and the mean spread increases when regular beta is replaced by downside beta. This pattern is most pronounced within the smallest size deciles. Panel C and D show further improvements in the risk-return relationship during bad-states of the world. As can be seen from the figures, a residual size effect remains. We emphasize that the MV and MS models assume a perfect capital market and ignore transaction costs and market liquidity, which seem especially relevant for the small market segment. Still, the CMS model cannot be rejected within the smallest size decile (p=0.57) nor within the largest size decile (p=0.96). Thus, downside risk seems to drive asset prices, both within the small and the large market segment.

4.5.5 Three-factor model

The most successful competitor of the MV CAPM is the three-factor model (TFM) of Fama and French (1993). How does this model perform relative to the MS CAPM in explaining beta portfolio returns?

Not very good. To answer this question, the TFM can be represented by the kernel $b_0 + b_1 r_M + b_2 smb + b_3 hml$, where SMB and HML stand for "small (cap) minus big" and "high (book/market) minus low". We fix the model parameters following the multifactor generalization of the GRS methodology by Fama (1996). Table 4.5 compares the fit of the TFM with that of the conditional downside risk model. The high-beta portfolios have higher TFM betas than CMV betas, thus leading to larger pricing errors. In fact, the TFM model exhibits a rather weak performance relative to the beta portfolios (p=0.07 and p=0.01). Thus, the TFM does not capture the underpricing/overpricing of low/high beta stocks and conditional downside risk seems unrelated to "distress risk". Of course, this does not disqualify the TFM. Our analysis uses regular-beta portfolios and downside-beta portfolios in order to test if semi-variance better captures investor preferences than variance. By contrast, the TFM is known to be successful for other benchmark portfolios, most notably double sorted size-value portfolios.

Table 4.5
Three-factor model and conditional downside risk

This table shows the alphas ($\hat{\alpha}$) and betas ($\hat{\beta}$) for the conditional mean-semivariance (CMS) model and the Fama and French three-factor model (TFM). The last two columns show the test results for the joint hypothesis that the alphas equal zero. The test statistic (JT) is chi-squared distributed with 9 (N-1) degrees of freedom. Panel A shows the results for the regular-beta portfolios and Panel B shows the results for the downside-beta portfolios.

		Low	2	3	4	5	6	7	8	9	High	JT	p
Panel A: Regular-beta portfolios													
$\hat{\alpha}$	CMS	0.08	0.05	0.08	0.04	0.07	-0.01	-0.04	-0.03	0.02	-0.06	2.4	0.98
	TFM	0.15	0.10	0.04	-0.03	0.06	-0.14	-0.14	-0.29	-0.32	-0.41	16.0	0.07
$\hat{\beta}$	CMS	0.78	0.85	0.88	0.89	1.17	1.09	1.21	1.19	1.25	1.41		
	TFM	0.68	0.78	0.94	0.99	1.19	1.29	1.35	1.59	1.76	1.94		
Panel B: Downside-beta portfolios													
		Low	2	3	4	5	6	7	8	9	High	JT	p
$\hat{\alpha}$	CMS	0.05	0.06	0.05	0.08	0.10	-0.05	0.05	0.04	-0.04	0.06	3.6	0.93
	TFM	0.12	0.15	0.07	0.05	0.04	-0.23	-0.13	-0.24	-0.47	-0.40	23.3	0.00
$\hat{\beta}$	CMS	0.77	0.90	0.92	1.00	1.09	1.12	1.11	1.18	1.33	1.42		
	TFM	0.66	0.76	0.89	1.04	1.19	1.40	1.38	1.61	1.96	2.11		

4.5.6 Momentum

Your study focuses on explaining the beta-effect. However, the 'anomalie-du-jour' is the momentum effect (see for example Jegadeesh and Titman (2001)). Does the MS model fare any better in explaining the returns of momentum strategies than the MV model?

Yes. The MS and MV models assume a perfect capital market and we do not expect these models to completely explain the returns of investment strategies that involve a high turnover and correspondingly high transactions costs, such as momentum strategies (see for instance Lesmond, Schill and Zhou (2004)). Still, momentum portfolios are an interesting test case for comparing the MV and MS models, because the returns to momentum strategies generally are characterized by asymmetry and hence the regular betas and downside betas can be expected to differ substantially. For this reason, we applied all five tests (UMV, UMS, CMV, CMS, and TFM) to ten momentum decile portfolios. Table 4.6 reports the results. The momentum-effect is strongly present; the portfolio of past losers has the lowest mean (0.01% per month) and the highest UMV beta (1.57), while the portfolio of past winners has the highest mean (1.33% per month) and one of the lowest UMV betas (1.00). Interestingly, downside risk and conditioning lead to substantial improvements in the fit. Most notably, in the CMS model, which combines the two explanations, the beta of the loser portfolio falls from 1.57 to 1.15, while the beta of the winner portfolio rises from 1.00 to 1.20. Apparently, past losers involve less downside risk in bad states than suggested by their unconditional regular betas and for past winners the opposite is true. While the improvements are not large enough to rationalize the entire momentum effect and both models have to be rejected, the sizeable reductions of the alphas again confirm our conclusion that the MS CAPM strongly outperforms the MV CAPM - especially during bad states.

Table 4.6
Momentum and conditional downside risk

This table shows descriptive statistics for the monthly excess returns of the ten momentum portfolios (N=10). The sample period and data requirements are identical to those for the beta portfolios. Each month, stocks are sorted based on 12-month price momentum (cumulative past 12-1 month returns). The portfolios are constructed such that each portfolio contains an equal number of stocks. The alphas ($\hat{\alpha}$) and betas ($\hat{\beta}$) for each of the following five models are shown: (1) unconditional mean-variance CAPM, (2) conditional mean-variance CAPM, (3) unconditional mean-semivariance CAPM, (4) conditional mean-semivariance CAPM and finally (5) Fama and French three-factor model. The last two columns show the test results for the joint hypothesis that the alphas equal zero. The test statistic (JT) is chi-squared distributed with 9 (N-1) degrees of freedom.

		Momentum Portfolios											
		Loser	2	3	4	5	6	7	8	9	Winn	*JT*	*p*
Stats.	Mean	0.01	0.41	0.44	0.62	0.57	0.60	0.72	0.92	0.94	1.33		
	Stdev	10.2	8.61	7.25	6.62	6.32	5.98	5.69	5.68	5.96	6.70		
	Skewness	1.92	2.03	1.66	1.41	1.34	0.66	0.01	0.40	-0.17	-0.28		
	Kurtosis	16.2	21.1	18.4	14.9	16.7	11.1	6.7	6.3	4.4	2.2		
Alphas	UMV	-1.03	-0.51	-0.36	-0.13	-0.15	-0.09	0.05	0.27	0.27	0.66	61.5	0.00
	UMS	-0.95	-0.43	-0.30	-0.09	-0.12	-0.09	0.04	0.27	0.23	0.59	49.2	0.00
	CMV	-0.94	-0.52	-0.30	-0.15	-0.17	-0.15	0.04	0.31	0.34	0.65	48.8	0.00
	CMS	-0.75	-0.36	-0.17	-0.07	-0.10	-0.15	0.01	0.29	0.25	0.53	29.4	0.00
	FF	-1.23	-0.65	-0.44	-0.20	-0.21	-0.13	0.06	0.28	0.29	0.71	73.8	0.00
Betas	UMV	1.57	1.38	1.20	1.12	1.09	1.04	0.99	0.98	1.00	1.00		
	UMS	1.44	1.27	1.12	1.06	1.04	1.04	1.02	0.98	1.05	1.11		
	CMV	1.43	1.40	1.12	1.15	1.11	1.13	1.01	0.92	0.90	1.01		
	CMS	1.15	1.16	0.93	1.04	1.00	1.13	1.06	0.94	1.03	1.20		
	FF	1.86	1.59	1.32	1.22	1.17	1.09	0.99	0.97	0.97	0.93		

4.5.7 Other results

How do your results compare with those of Ang, Chen and Xing (ACX, 2004)?

As in our study, ACX conclude that downside risk is important to explain the cross-section of stock returns. However, the data and methodology of ACX differ from ours. The conclusion of ACX relies on them using (1) equal-weighted portfolio returns and (2) the Fama and MacBeth (1973) cross-sectional methodology that selects the model parameters that give the best empirical fit. In fact, using value-weighted portfolio returns and our methodology, we find relatively weak evidence favoring the UMS model over the UMV model in the ACX sample (1963-2001). This confirms our finding that the role of downside risk in the second half of the 20th century is limited (see Panel A of Table 4.3). Using equal-weighted portfolio returns rather than value-weighted returns has the effect of placing greater weight on the small cap segment. As illustrated in Figure 4.6 above, downside risk is relatively more important for the small caps than for the large caps. The cross-sectional methodology further inflates the explanatory power of downside risk by allowing a high intercept (far above the historical riskless rate) and a low slope (far below the historical equity premium). These two factors explain why the evidence of ACX disappears in our approach that uses value-weighted returns and fixes the model parameters in the spirit of the GRS time-series methodology. In contrast to ACX, our case for the MS CAPM rests of the pattern of downside risk in the earlier years and the bad states-of-the-world. This pattern occurs also for the large caps and if the intercept and slope are fixed.

4.6 Conclusions

Surprisingly, despite the theoretical limitations of variance, the differences between regular beta and downside beta, and the empirical problems of the mean-variance (MV) CAPM, the mean-semivariance (MS) CAPM has not been subjected to rigorous empirical testing thus far. In an extended sample (1931-2002) we employ unconditional MV and MS tests as well as conditional tests that account for the economic state-of-the-world.

We find that the MS CAPM strongly outperforms the traditional MV CAPM in terms of its ability to explain the cross-section of U.S stock returns. Especially during bad-states of the world we find a near-perfect relation between (downside) risk and return. Further, conditional downside risk (1) explains average returns within the size deciles, (2) is not related to distress risk and (3) can partly explain the momentum effect. In sum, our results provided evidence in favor of market portfolio efficiency, provided we account for conditional downside risk.

Our results reflect the asymmetry of the return distribution of common stocks, especially during bad states-of-the-world. Further research could investigate the sources of asymmetry, such as financial and operational leverage. Also, it would be interesting to extend our analysis to other securities with asymmetric distributions, such as bonds, derivatives and securities with embedded options. We expect the case in favor of the MS CAPM to be even stronger for such securities than it is for common stocks.

Appendix 4.1

This appendix discusses a pitfall in modeling downside risk: the asymmetric response model (ARM) of Harlow and Rao (1989, Section III) generally does not estimate the lower partial moment (LPM) beta.

Using our notation, Harlow and Rao (1989, Eq. 10) employ the following bivariate regression model to estimate the downside beta:

$$r_i = \alpha_{ARM,i} + \beta_{ARM,i} x + \gamma_{ARM,i} z + \varepsilon_i \tag{A.4.1}$$

where $x \equiv (r_M \delta_- + E[r_M \delta_+]\delta_+ E[\delta_+]^{-1})$, $z \equiv (r_M \delta_+ - E[r_M \delta_+]\delta_+ E[\delta_+]^{-1})$, $\delta_- \equiv 1(r_M \le 0)$ and $\delta_+ \equiv 1(r_M > 0)$. By taking the expectation of both sides of (A.4.1) and assuming $\alpha_{ARM,i} = 0$, we arrive at the following risk-return relationship:

$$E[r_i] = \beta_{ARM,i} E[r_M] \tag{A.4.2}$$

Since x and z are independent by construction, the ARM beta equals the univariate regression coefficient for x:

$$\beta_{ARM,i} = \frac{E[xr_i] - E[x]E[r_i]}{E[x^2] - E[x]^2} = \frac{E[r_M \delta_- r_i] + E[r_M \delta_+]E[r_i \delta_+]E[\delta_+]^{-1} - E[r_M]E[r_i]}{E[r_M^2 \delta_-] + E[r_M \delta_+]^2 E[\delta_+]^{-1} - E[r_M]^2} \tag{A.4.3}$$

This regression coefficient generally is *not* the second-order co-lower partial moment or the LPM beta (Harlow and Rao (1989, Eq. 9)):

$$\beta_{LPM,i} \equiv \frac{E[r_M \delta_- r_i]}{E[r_M^2 \delta_-]} \tag{A.4.4}$$

which enters in the MS CAPM equilibrium equation:

$$E[r_i] = \beta_{LPM,i} E[r_M] \tag{A.4.5}$$

To show that the ARM beta generally differs from the LPM beta, it is useful to consider the case where the betas are identical. In this case, we can substitute

$\beta_{ARM,i} = \beta_{LPM,i}$, $E[r_M \delta_- r_i] = \beta_{LPM,i} E[r_M^2 \delta_-]$ (from (A.4)) and $E[r_i] = \beta_{LPM,i} E[r_M]$ (from (A.4.2) and $\beta_{ARM,i} = \beta_{LPM,i}$) in the numerator of the right-hand side of (A.4.3) to find

$$\beta_{LPM,i} = \frac{\beta_{LPM,i} E[r_M^2 \delta_-] + E[r_M \delta_+] E[r_i \delta_+] E[\delta_+]^{-1} - \beta_{LPM,i} E[r_M]^2}{E[r_M^2 \delta_-] + E[r_M \delta_+]^2 E[\delta_+]^{-1} - E[r_M]^2}$$

$$= \beta_{LPM,i} + \frac{E[r_M \delta_+] E[\delta_+]^{-1} (E[r_i \, \delta_+] - \beta_{LPM,i} E[r_M \delta_+])}{E[r_M^2 \delta_-] + E[r_M \delta_+]^2 E[\delta_+]^{-1} - E[r_M]^2}$$

$$\Leftrightarrow E[r_i \, \delta_+] - \beta_{LPM,i} E[r_M \delta_+] = 0$$

$$\Leftrightarrow \beta_{LPM,i} = \frac{E[r_i \, \delta_+]}{E[r_M \delta_+]} \qquad \text{(A.4.6)}$$

Equality (A.4.6) generally does not hold and hence the ARM beta $\beta_{ARM,i}$ generally differs from the LPM beta $\beta_{LPM,i}$. Therefore, the ARM alpha $\alpha_{ARM,i}$ generally also differs from the LPM alpha $\alpha_{LPM,i} \equiv E[r_i] - \beta_{LPM,i} E[r_M]$ (which is zero in MS CAPM equilibrium).

Numerical example

The differences between the two betas ($\beta_{ARM,i}$ and $\beta_{LPM,i}$) can be demonstrated by means of the following simple example with three assets and three states-of-the-world of equal probability:

State	Prob.	Market (*M*)	Stock 1	Stock 2	Stock 3
1	1/3	-2.0%	-1.6%	-1.0%	-3.0%
2	1/3	-1.0%	-0.8%	-2.0%	-1.0%
3	1/3	+6.0%	+4.8%	+5.4%	+8.2%
$\beta_{LPM,i}$		+1.00	+0.80	+0.80	+1.40
$\beta_{ARM,i}$		+1.00	+0.80	+0.89	+1.37
$\alpha_{LPM,i}$		0%	0%	0%	0%
$\alpha_{ARM,i}$		0%	0%	-0.09%	+0.03%

Stock 1 is constructed to satisfy (A.4.6) and hence the two betas are identical in this case. Stock 2 and 3 are constructed to show the differences between the ARM results and the LPM results. The ARM model overestimates the LPM beta and underestimates the LPM alpha for Stock 2 and underestimates the LPM beta and overestimates the LPM alpha for Stock 3.

Chapter 5

Conclusions

In this thesis, we document the empirical performance of the CAPM and consider risk measures other than variance to better explain stock prices. The research contributes to the empirical asset pricing literature and the literature on downside risk. Few empirical asset pricing studies are devoted to downside risk, arguably due to empirical and methodological hurdles. This thesis offers a downside risk explanation for the cross-section of stock returns.

The mean-variance (MV) criterion of the CAPM has some important shortcomings. For example, a stock is classified as more risky if it goes up faster than the market and goes down in tandem with the market. Downside risk measures are intuitively appealing because they put more emphasis on downside stock movements. These alternative risk measures may better describe investor's preferences. In general, the MV-criterion violates one or more elementary regularity conditions such as nonsatiation and decreasing absolute risk aversion. In our analysis, we try to adhere to the regularity conditions, and therefore replace variance by other measures of risk.

In order to impose these regularity conditions we phrase our research in the pricing kernel framework. This framework is currently popular because it embeds a broad range of risk-based asset pricing models. The parametric tests are cast in a GMM framework to control for the serial correlation and heteroskedasticity effects that typically plague empirical tests of asset pricing models. Due to methodological advances it is now possible to study a wide class of downside risk measures without having to put parametric structure on the model in advance. We place restrictions on the kernel parameters in order to find well-behaved pricing kernels only.

We pay special attention to several important empirical issues that are typical for asset pricing tests: the data selection criteria, portfolio formation procedures, the sample period and type of benchmark portfolios. Each of these issues may materially affect the test results. For example, the exclusion of small caps and inclusion of the pre-1963 period in the analysis tends to diminish the value effect. Further, several macro-economic variables can be used to capture the time-varying aspect of risk and risk preferences. Due to differences in the empirical research design, many findings reported in the literature are not readily comparable. Therefore, we impose identical data requirements, consider a broad range of cross-sectional effects, control for size and consider a long sample period. Furthermore, we employ several conditional variables and perform subsample and rolling window analyses, so as to asses the robustness of the results.

The conclusions laid down in this thesis are threefold. First, most of the empirical evidence against the CAPM is concentrated in the illiquid small-cap market segment in the post-1963 period. Well-known multi-factor models often do not help to explain the returns on large caps and 'neutral' benchmark sets such as beta or industry portfolios. Second, without short-selling, the market is third-order stochastic dominance efficient relative to portfolios sorted on size, value and momentum. This non-parametric finding suggests that systematic downside risk can help to explain MV-inefficiency of the stock market portfolio. Especially during the 1970s and early 1980s, the tail betas of small, value, winner stocks are much higher than their regular betas. Third, the mean-semivariance CAPM, which assigns greater importance to downside volatility, better explains stock returns than the traditional MV-CAPM. If stocks are sorted into portfolios based on downside beta instead of regular beta, then the empirical risk-return relation improves. Especially during economic recession periods, the risk-return relation is near-perfect. Combined these results suggest that downside risk helps to better understand stock prices.

Some issues are left for further research. First, due to an increase in stock return correlations during bear markets it is much more difficult to diversify away downside risk than upside potential. It would be interesting to find out the economic forces behind this empirical pattern. Second, the development of new stochastic dominance (SD) efficiency tests, such as first-order SD and conditional SD tests would be very useful. These tests could shed new light on the empirical performance of a broad class of representative investor models. Third, we expect that inclusion of other financial assets with embedded options (e.g. corporate bonds, or emerging market shares) in the analysis would further emphasize the need to move away from the mean-variance framework. In general, the field of empirical asset pricing is served by the development of large financial databases including high-quality price information of non-U.S and non-equity assets. Finally, it would be interesting to form benchmark portfolios based on a wide range of downside risk measures and in this way derive and 'optimal' risk measure.

Nederlandse samenvatting
(Summary in Dutch)

Introductie

De laatste decennia heeft het onderzoek naar de prijsvorming van effecten een aantal interessante ontwikkelingen doorgemaakt. Een belangrijke mijlpaal was de ontwikkeling van het eerste financiële prijsvormingsmodel in 1964. Dit model is bekend geworden onder de naam "Capital Asset Pricing Model" (CAPM) en de bedenkers zijn in 1990 bekroond met de Nobelprijs. Sinds die tijd heeft men veel vooruitgang geboekt zowel op theoretisch, methodologisch alsook empirisch gebied. Allereerst zijn er allerlei nieuwe theoretische inzichten opgedaan, die hebben geleid tot diverse generalisaties van het klassieke CAPM. Ten tweede heeft econometrisch onderzoek veel nieuwe methoden en technieken voortgebracht, die dankzij de beschikbaarheid van krachtige computers ook direct toepasbaar zijn. Ten derde zijn er zeer betrouwbare gegevensverzamelingen beschikbaar gekomen, die qua precisie en rijkdom ongeëvenaard zijn in vergelijking met de gegevens in andere economische disciplines.

Het CAPM wordt nog steeds veelvuldig toegepast, ondermeer vanwege haar eenvoud en elegantie. Echter, onderzoek heeft ook aangetoond dat het CAPM enige empirische tekortkomingen kent, aangezien dit model de hoge rendementen op verschillende beleggingsstrategieën niet kan rationaliseren. Een mogelijke verklaring voor deze tekortkomingen is een foutieve definitie van het begrip risico. Het CAPM veronderstelt namelijk dat variantie de relevante risicomaatstaf is. Echter, onderzoek heeft aangetoond dat rendementen niet normaal verdeeld zijn en beleggers gevoeliger zijn voor neerwaartse dan voor opwaartse koersbewegingen. Daarom kunnen alternatieve neerwaartse risicomaatstaven, zoals semi-variantie, een betere beschrijving van de risicohouding van beleggers geven.

Het doel van dit proefschrift is het grondig toetsen van het CAPM en te bepalen, of het loslaten van variantie als relevante risicomaatstaf kan helpen aandelenprijzen beter te begrijpen. In het onderzoek bouwen we voort op de nieuwste inzichten uit de literatuur. Het proefschrift heeft een aantal kenmerkende eigenschappen. Wij hechten er bijzonder belang aan, dat de risicohouding van beleggers voldoet aan economische basisveronderstellingen zoals onverzadigbaarheid en afkeer van risico. Door het gebruik van non-parametrische efficiëntie criteria kunnen we een bredere klasse van economisch zinvolle risicomaatstaven onderzoeken. Daarnaast toetsen we voor het eerst een conditioneel neerwaarts-risico model. In dit proefschrift wordt veel aandacht besteed aan zowel de empirische als methodologische onderdelen van financieel onderzoek. De empirische analyses maken gebruik van de gegevensverzameling van het "Center for Research in Securities Prices" (CRSP); deze bevat nagenoeg geen fouten en bestrijkt een lange tijdsperiode (1926-2002). Het onderzoek bevat een groot aantal robuustheidsanalyses. Om risico nog beter te meten gebruiken wij naast verschillende bekende portefeuilles ook portefeuilles die zijn geformeerd op basis van systematisch neerwaarts risico. Alle portefeuilles gebruikt in dit proefschrift zijn publiekelijk beschikbaar via het Internet. Het onderzoek is opgedeeld in drie delen, waarin het CAPM en verschillende uitbreidingen worden getoetst.

CAPM en multi-factor modellen

Multi-factor modellen zijn bedoeld als antwoord op de empirische tekortkomingen van het CAPM. Aan het één-factor CAPM worden extra factoren toegevoegd die het rendement van bepaalde hedge portefeuilles weergeven. Op dit moment is het driefactor model (3FM) van Fama en French erg populair. Dit model identificeert twee hedgefactoren: kleine versus grote aandelen en waarde- versus groeiaandelen. Carhart voegt later ook nog een vierde factor toe, die prijsmomentum representeert. De keuze voor de factoren is vooral empirisch gemotiveerd en de theoretische interpretatie van deze modellen is niet geheel duidelijk.

In Hoofdstuk 2 toetsen we het CAPM en de multi-factor extensies op grondige wijze. In de empirische literatuur is de verklaringskracht van de verschillende prijsvormingsmodellen nog niet volledig en systematisch in kaart gebracht. Wij sorteren aandelen op basis van zes bekende karakteristieken en onderzoeken naast de meer recente periode 1963-2002 ook de vroege periode 1931-1962. Daarnaast splitsen we de aandelenmarkt op in verschillende marktsegmenten op basis van marktkapitalisatie.

Onze conclusie is dat het empirische bewijs tegen het CAPM is geconcentreerd in de kleinere marktsegmenten gedurende de meer recente periode. Toevoegen van additionele factoren helpt om de rendementen van deze aandelen in deze periode te rationaliseren. Er bestaan echter ook een aantal situaties waarin de extra factoren niet helpen of de resultaten zelfs verslechteren. Dit is het geval (1) in de vroegere periode, (2) voor grote aandelen en (3) voor investeringsstrategieën gebaseerd op marktkapitalisatie, bèta, reversal en industrie. Hoewel multi-factor modellen erg bruikbaar kunnen zijn voor andere doeleinden, zoals stijlanalyse van beleggingsfondsen, betwijfelen we of deze tot een beter begrip van aandelenprijzen zullen leiden.

SD efficiëntie van de marktportefeuille

Het ‘mean-variance’ (MV) criterium negeert het feit dat beleggers een voorkeur hebben voor verdelingen met een positieve scheefheid en impliceert in sommige gevallen een verzadigbare belegger. Daarom is het zinvol om naar alternatieve criteria te kijken die meer overeenstemmen met beleggerspreferenties. Het voordeel van het gebruik van (non-parametrische) stochastische dominantie (SD) criteria is dat een bredere klasse van risicomaatstaven beschouwd kan worden zonder vooraf ad hoc specificaties te maken. Het 3e orde SD (third-order SD, TSD) criterium komt overeen met het veronderstellen dat beleggers (1) meer boven minder prefereren (onverzadigbaar) (2) een afkeer van risico hebben en (3) een voorkeur voor positieve scheefheid hebben. Als aandelenrendementen normaal verdeeld zijn, dan valt het TSD-criterium samen met het MV-criterium.

In Hoofdstuk 3 vervangen wij het MV criterium door het 3e orde SD criterium en toetsen we de marktportefeuille op MV en TSD efficiëntie. Anders dan standaardtoetsen veronderstellen wij dat beleggers niet ‘short’ kunnen gaan; de portefeuillegewichten moeten positief zijn. Wij gebruiken de bekende aandelenportefeuilles gesorteerd op marktkapitalisatie, boekwaarde/marktwaarde en prijsmomentum van French, Titman en Carhart. Het is algemeen bekend dat de marktportefeuille MV inefficiënt is ten opzichte van deze portefeuilles.

De centrale bevinding is dat de marktportefeuille wel TSD efficiënt is ten opzichte van al deze portefeuilles. Vooral gedurende de jaren ‘70 en in het begin van de jaren ‘80 blijkt er een groot verschil tussen beide efficiëntiecriteria te bestaan. Dit betekent dat de hoge rendementen op deze bekende beleggingsstrategieën verklaard kunnen worden door een hoger (neerwaarts) risico. Bij bestudering van de impliciete risicohouding blijkt dat er sprake is van een sterke afkeer van beurskrachs.

Conditioneel neerwaarts risico

Een belangrijke tekortkoming van variantie is dat zij een aandeel als risicovol classificeert, wanneer dit aandeel harder omhoog gaat dan de markt. Het is daarom zinvol om variantie door semi-variantie te vervangen. Het ‘mean-semivariance’ (MS) CAPM verwisselt de standaard bèta door neerwaartse bèta als enige risicofactor. Ondanks haar theoretische aantrekkelijkheid is het MS model nog niet grondig en op juiste wijze onderzocht.

Hoofdstuk 4 toetst het MS-CAPM en daarnaast ook het standaard CAPM en het driefactor model. In tegenstelling tot eerdere studies wordt gebruik gemaakt van (1) een verbeterde onderzoeksmethodologie en (2) een lange tijdsperiode, die het grootste deel van de 20e eeuw bestrijkt. Omdat de risicohouding van beleggers kan variëren gedurende de conjunctuurscyclus, voeren we ook conditionele toetsen uit die rekening houden met dit tijdsvariërende patroon. Naast de bekende portefeuilles gebruiken wij nieuwe portefeuilles, gesorteerd op neerwaartse bèta.

Dit onderzoek toont aan dat de verschillen in aandelenprijzen kunnen worden verklaard met verschillen in conditioneel neerwaarts risico. Vanwege de asymmetrie in de rendementsverdeling is het neerwaartse risico van aandelen met een lage (hoge) bèta hoger (lager). Wanneer standaard bèta door neerwaartse bèta wordt vervangen, verbetert hierdoor de empirisch zwakke relatie tussen risico en rendement. Vooral gedurende economische recessies bestaat een bijna perfect verband tussen risico en rendement.

Conclusies

In het kort presenteert dit onderzoek drie conclusies. Ten eerste, het bewijs tegen het CAPM blijkt geconcentreerd te zijn in de kleinere marktsegmenten in de meer recente periode. Daarnaast betwijfelen wij of multi-factor modellen een beter inzicht bieden in de prijsvorming van effecten dan het standaard CAPM. Ten tweede, variantie als relevante risicomaatstaf kent een aantal belangrijke theoretische en empirische tekortkomingen. Alternatieve risicomaatstaven, die meer overeenstemmen met de sterke afkeer van beleggers voor beurskrachs, kunnen de verschillen in aandelenrendementen beter verklaren dan variantie. Ten derde, beleggers blijken een grotere afkeer van neerwaarts risico te hebben tijdens economische recessies dan tijdens economische expansies.

References

Ang, Andrew, and Joseph Chen, 2002, Asymmetric correlations of equity portfolios, *Journal of Financial Economics* 63, 443-494.

Ang, Andrew, and Joseph Chen, 2003, CAPM over the long run, Working paper Columbia University, SSRN 346600.

Ang, Andrew, Joseph Chen, and Yuhang Xing, 2004, Downside Risk, Working paper Columbia University.

Arditti, Fred D., 1967, Risk and the required return on equity, *Journal of Finance* 22, 19-36.

Bansal, Ravi, David A. Hsieh, and S. Viswanathan, 1993, A new approach to international arbitrage pricing, *Journal of Finance* 48, 1719-1747.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Basu, Sanjoy, 1977, Investment performance of common stocks in relationship to their price-earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663-682.

Bawa, Vijay, and Eric B. Lindenberg, 1977, Capital market equilibrium in a mean-lower partial moment framework, *Journal of Financial Economics* 5, 189-200.

Benartzi, Shlomo, and Richard H. Thaler, 1995, Myopic loss aversion and the equity premium puzzle, *Quarterly Journal of Economics* 110, 73-92.

Berk, Jonathan B., 1997, Necessary conditions for the CAPM, *Journal of Economic Theory* 73, 245-257.

Berkelaar, Arjan B., Roy P. Kouwenberg, and Thierry Post, 2005, Optimal portfolio choice under loss aversion, *Review of Economics and Statistics* forthcoming.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: some empirical tests, *Studies in the theory of capital markets* (Praeger).

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 73, 265-296.

Breeden, Douglas T., Michael R. Gibbons, and Robert H. Litzenberger, 1989, Empirical tests of the consumption-oriented CAPM, *Journal of Finance* 44, 231-262.

Burton, Jonathan, 1998, Revisiting the capital asset pricing model, *Dow Jones Asset Manager* May/June, 20-28.

Campbell, John Y., and J. H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997. *The econometrics of financial markets* (Princeton University Press, Princeton, New Jersey).

Campbell, Rachel, Kees Koedijk, and Paul Kofman, 2002, Increased correlation in bear markets: a downside risk perspective, *Financial Analyst Journal* 58, 87-94.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Carhart, Mark M., Robert J. Krail, Robert I. Stevens, and Kelly E. Welch, 1996, Testing the conditional CAPM, Unpublished manuscript, University of Chicago.

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572-621.

Cochrane, John H., 2001. *Asset Pricing* (Princeton University Press, Princeton).

Cooley, Philip L., 1977, A multidimensional analysis of institutional investor perception of risk, *Journal of Finance* 32, 67-78.

Coval, Joshua D., and Tyler Shumway, 2001, Expected option returns, *Journal of Finance* 56, 983-1009.

CRSP, Center for Research in Security Prices, 2001, CRPS Delisting returns, (University of Chicago, Chicago).

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1-33.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389-406.

DeBondt, Werner F.M., and Richard H. Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793-805.

DeBondt, Werner F.M., and Richard H. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557-581.

Dimson, Elroy, and Paul Marsh, 2001, U.K. financial market returns, 1955-2000, *Journal of Business* 74, 1-31.

Dittmar, Robert F., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns, *Journal of Finance* 57, 369-403.

Dybvig, Philip H., and Jonathan E. Ingersoll, 1982, Mean-variance theory in complete markets, *Journal of Business* 55, 233-251.

Fama, Eugene F., 1965, The behavior of stock-market prices, *Journal of Business* 38, 34-105.

Fama, Eugene F., 1991, Efficient capital-markets II, *Journal of Finance* 46, 1575-1617.

Fama, Eugene F., 1996, Multifactor portfolio efficiency and multifactor asset pricing, *Journal of Financial and Quantitative Analysis* 31, 441-465.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk-factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131-155.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene F., and Kenneth R. French, 2002, The equity premium, *Journal of Finance* 57, 637-659.

Fama, Eugene F., and Kenneth R. French, 2003, The CAPM: Theory and evidence, CRSP Working Paper No. 550.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 71, 43-66.

Friend, Irwin, and Randolph Westerfield, 1980, Co-skewness and capital asset pricing, *Journal of Finance* 35, 897-913.

Ghysels, Eric, 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt?, *Journal of Finance* 53, 549-573.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Hadar, Josef, and William R. Russell, 1969, Rules for ordering uncertain prospects, *American Economic Review* 59, 25-34.

Hanoch, Giora, and Haim Levy, 1969, The efficiency analysis of choices involving risk, *Review of Economic Studies* 36, 335-346.

Hanoch, Giora, and Haim Levy, 1970, Efficient portfolio selection with quadratic and cubic utility, *Journal of Business* 43, 181-189.

Hansen, Lars P., and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557-590.

Harlow, W. Van, and Ramesh K.S. Rao, 1989, Asset pricing in a generalized mean-lower partial moment framework: theory and evidence, *Journal of Financial and Quantitative Analysis* 24, 285-311.

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.

Hogan, William W., and James M. Warren, 1974, Toward the development of an equilibrium capital-market model based on semivariance, *Journal of Financial and Quantitative Analysis* 9, 1-11.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad new travels slowly: size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265-296.

Hong, Harrison, and Jeremy C. Stein, 2003, Differences of opinion, short-sales constraints and market crashes, *Review of Financial Studies* 16, 487-525.

Huang, Chi-Fu, and Robert H. Litzenberger, 1988. *Foundations for Financial Economics* (North Holland).

Ibbotson, Associates, 2002. *Stocks, Bonds, Bills and Inflation: 2002* (Ibbotson Associates, Chicago).

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.

Jahankhani, Ali, 1976, E-V and E-S capital asset pricing models: some empirical tests, *Journal of Financial and Quantitative Analysis* 11, 513-528.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Kane, Alex, 1982, Skewness preference and portfolio choice, *The Journal of Financial and Quantitative Analysis* 17, 15-25.

Knez, P. J., and M. J. Ready, 1997, On the robustness of size and book-to-market in cross-sectional regressions, *Journal of Finance* 52, 1355-1382.

Kothari, S. P., Jay Shanken, and Richard G. Sloan, 1995, Another look at the cross-section of expected stock returns, *Journal of Finance* 50, 185-224.

Kraus, Alan, and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085-1100.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

Lesmond, David A., Joseph P. Ogden, and Charles A. Trzcinka, 1999, A new estimate of transaction costs, *Review of Financial Studies* 12, 1113-1141.

Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2004, The illusory nature of momentum profits, *Journal of Financial Economics* 71, 349-380.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238-1287.

Levy, Haim, 1969, Comment: A utility function depending on the first three moments, *Journal of Finance* 24, 715-719.

Levy, Haim, 1992, Stochastic dominance and expected utility: sorvey and analysis, *Management Science* 38, 555-593.

Levy, Haim, and Giora Hanoch, 1970, Relative effectiveness of efficiency criteria for portfolio selection, *Journal of Financial and Quantitative Analysis* 5, 63-76.

Levy, Haim, Thierry Post, and Pim van Vliet, 2003, Skewness preference and risk aversion, ERIM research paper, SSRN 411660.

Lewellen, Jonathan, and Stefan Nagel, 2004, The conditional CAPM does not explain asset pricing anomalies, NBER Working paper.

Lintner, John, 1965, The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431-468.

Lorie, James H., 1966, Some comments on recent quantitative and formal reserach on the stock market, *Journal of Business* 1966, 107-110.

Loughran, Tim, 1997, Book-to-market across firm size, exchange, and seasonality, *Journal of Financial and Quantitative Analysis* 32, 249-268.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3-28.

MacKinlay, A. Craig, and Matthew P. Richardson, 1991, Using generalized method of moments to test mean-variance efficiency, *Journal of Finance* 46, 511-527.

Markowitz, Harry, 1952, The utility of wealth, *Journal of Political Economy* 60, 151-158.

Markowitz, Harry, 1959. *Portfolio selection: Efficient diversification of investments* (Wiley, New York).

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147-175.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Petkova, Ralitsa, and Lu Zhang, 2003, Is value riskier than growth?, (University of Rochester, NY).

Post, Thierry, 2003, Empirical test for stochastic dominance efficiency, *Journal of Finance* 58, 1905-1931.

Post, Thierry, and Pim van Vliet, 2004a, Market portfolio efficiency and value stocks, *Journal of Economics and Finance* 28, 16-21.

Post, Thierry, and Pim van Vliet, 2004b, Downside risk and asset pricing, ERIM Research Paper, SSRN 503142.

Post, Thierry, and Pim van Vliet, 2004c, Conditional downside risk and the CAPM, ERIM Research Paper, SSRN 557220.

Post, Thierry, and Pim van Vliet, 2004d, Do multiple factors help or hurt?, ERIM Research Paper, SSRN 582101.

Post, Thierry, and Philippe Versijp, 2004, GMM tests for the SD efficiency of a given portfolio, ERIM Research Paper, SSRN 558725.

Price, Kelly, Barbara Price, and Timothy J. Nantell, 1982, Variance and lower partial moment measures of systematic risk: some analytical and empirical results, *Journal of Finance* 37, 843-855.

Reinganum, Marc R., 1981, A new empirical perspective on the CAPM, *Journal of Financial and Quantitative Analysis* 16, 439-462.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of economic theory* 13, 341-360.

Rubinstein, Mark, 1974, An aggregation theorem for securities markets, *Journal of Financial Economics* 1, 225-244.

Scott, Robert C., and Philip A. Horvath, 1980, On the direction of preference for moments of higher order than the variance, *Journal of Finance* 35, 915-919.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Sharpe, William, 1991, Capital asset prices with and without negative holdings, *Journal of Finance* 64, 489-509.

Shumway, Tyler, 1997, The delisting bias in CRPS data, *Journal of Finance* 54, 327-340.

Shumway, Tyler, and Vincent A. Warther, 1999, The delisting in CRSP's Nasdaq data and its implications for the size effect, *Journal of Finance* 54, 2361-2379.

Simkowitz, Michael A., and William L. Beedles, 1978, Diversification in a three-moment world, *Journal of Financial and Quantitative Analysis* 13, 927-941.

Smit, Han, and Pim van Vliet, 2002, Real options and stock market anomalies, in Lenos Trigeorgis, ed.: 6th Annual Real Options Conference (Paphos, Cyprus).

Tobin, James, 1958, Liquidity preferences as a behavior toward risk, *Review of Economic Studies* 25, 65-85.

Tsiang, Sho-Chieh, 1972, The rationale of the mean-standard deviation analysis, skewness preference and the demand for money, *American Economic Review* 62, 354-371.

Wang, Zhenyu, 1998, Efficiency loss and constraints on portfolio holdings, *Journal of Financial Economics* 48, 359-375.

Wang, Zhenyu, and Xiaoyan Zhang, 2004, The role of arbitrage in the empirical evaluation of asset pricing models, Columbia University and Cornell University.

Whitmore, G.A., 1970, Third-degree stochastic dominance, *American Economic Review* 60, 457-459.

Erasmus Research Institute of Management

ERIM Ph.D. Series Research in Management

Appelman, J.H., *Governance of Global Interorganizational Tourism Networks; Changing Forms of Co-ordination between the Travel Agency and Aviation Sector*, Promotors: Prof. dr. F.M. Go & Prof. dr. B. Nooteboom, EPS-2004-036-MKT, ISBN 90-5892-060-7, http://hdl.handle.net/1765/1199

Berens, G., *Corporate Branding: The Development of Corporate Associations and their Influence on Stakeholder Reactions*, Promotor: Prof. dr. C. B. M. van Riel, EPS-2004-039-ORG, ISBN 90 –5892–065–8, http://hdl.handle.net/1765/1273

Berghe, D.A.F., *Working Across Borders: Multinational Enterprises and the Internationalization of Employment*, Promotors: Prof. dr. R.J.M. van Tulder & Prof. dr. E.J.J. Schenk, EPS-2003-029-ORG, ISBN 90-5892-05-34, http://hdl.handle.net/1765/1041

Bijman, W.J.J., *Essays on Agricultural Co-operatives; Governance Structure in Fruit and Vegetable Chains*, Promotor: Prof. dr. G.W.J. Hendrikse, EPS-2002-015-ORG, ISBN: 90-5892-024-0, http://hdl.handle.net/1765/867

Brito, M.P. de, *Managing Reverse Logistics or Reversing Logistics Management?* Promotors: Prof. dr. ir. R. Dekker & Prof. dr. M. B. M. de Koster, EPS-2004-035-LIS, ISBN: 90-5892-058-5, http://hdl.handle.net/1765/1132

Campbell, R.A.J., *Rethinking Risk in International Financial Markets*, Promotor: Prof. dr. C.G. Koedijk, EPS-2001-005-F&A, ISBN: 90-5892-008-9, http://hdl.handle.net/1765/306

Chen, Y., *Labour Flexibility in China's Companies: An Empirical Study*, Promotors: Prof. dr. A. Buitendam & Prof. dr. B. Krug, EPS-2001-006-ORG, ISBN: 90-5892-012-7, http://hdl.handle.net/1765/307

Daniševská, P., *Empirical Studies on Financial Intermediation and Corporate Policies*, Promotor: Prof. dr. C.G. Koedijk, EPS-2004-044-F&A, ISBN 90–5892–070–4, http://hdl.handle.net/1765/1518

Delporte-Vermeiren, D.J.E., *Improving the Flexibility and Profitability of ICT-enabled Business Networks: An Assessment Method and Tool*, Promotors: Prof.mr. dr. P.H.M. Vervest & Prof. dr. ir. H.W.G.M. van Heck, EPS-2003-020-LIS, ISBN: 90-5892-040-2, http://hdl.handle.net/1765/359

Dijksterhuis, M., *Organizational Dynamics of Cognition and Action in the Changing Dutch and US Banking Industries*, Promotors: Prof. dr. ing. F.A.J. van den Bosch & Prof. dr. H.W. Volberda, EPS-2003-026-STR, ISBN: 90-5892-048-8, http://hdl.handle.net/1765/1037

Fenema, P.C. van, *Coordination and Control of Globally Distributed Software Projects*, Promotor: Prof. dr. K. Kumar, EPS-2002-019-LIS, ISBN: 90-5892-030-5, http://hdl.handle.net/1765/360

Fleischmann, M., *Quantitative Models for Reverse Logistics*, Promoters: Prof. dr. ir. J.A.E.E. van Nunen & Prof. dr. ir. R. Dekker, EPS-2000-002-LIS, ISBN: 3540 417 117, http://hdl.handle.net/1765/1044

Flier, B., *Strategic Renewal of European Financial Incumbents: Coevolution of Environmental Selection, Institutional Effects, and Managerial Intentionality*, Promotors: Prof. dr. ing. F.A.J. van den Bosch & Prof. dr. H.W. Volberda, EPS-2003-033-STR, ISBN: 90–5892–055–0, http://hdl.handle.net/1765/1071

Fok, D., *Advanced Econometric Marketing Models*, Promotor: Prof. dr. P.H.B.F. Franses, EPS-2003-027-MKT, ISBN: 90-5892-049-6, http://hdl.handle.net/1765/1035

Ganzaroli , A., *Creating Trust between Local and Global Systems*, Promotors: Prof. dr. K. Kumar & Prof. dr. R.M. Lee, EPS-2002-018-LIS, ISBN: 90-5892-031-3, http://hdl.handle.net/1765/361

Gilsing, V.A., *Exploration, Exploitation and Co-evolution in Innovation Networks*, Promotors: Prof. dr. B. Nooteboom & Prof. dr. J.P.M. Groenewegen, EPS-2003-032-ORG, ISBN 90-5892-05-42, http://hdl.handle.net/1765/1040

Graaf, G. de, *Tractable Morality: Customer Discourses of Bankers, Veterinarians and Charity Workers*, Promotors: Prof. dr. F. Leijnse & Prof. dr. T. van Willigenburg, EPS-2003-031-ORG, ISBN: 90–5892–051–8, http://hdl.handle.net/1765/1038

Hermans. J.M., *ICT in Information Services, Use and deployment of the Dutch securities trade, 1860-1970.* Promotor: Prof. dr. drs. F.H.A. Janszen, EPS-2004-046-ORG, ISBN 90-5892-072-0, http://hdl.handle.net/1765/1793

Heugens, P.M.A.R., *Strategic Issues Management: Implications for Corporate Performance*, Promotors: Prof. dr. ing. F.A.J. van den Bosch & Prof. dr. C.B.M. van Riel, EPS-2001-007-STR, ISBN: 90-5892-009-7, http://hdl.handle.net/1765/358

Hooghiemstra, R., *The Construction of Reality*, Promotors: Prof. dr. L.G. van der Tas RA & Prof. dr. A.Th.H. Pruyn, EPS-2003-025-F&A, ISBN: 90-5892-047-X, http://hdl.handle.net/1765/871

Jong, C. de, *Dealing with Derivatives: Studies on the Role, Informational Content and Pricing of Financial Derivatives*, Promotor: Prof. dr. C.G. Koedijk, EPS-2003-023-F&A, ISBN: 90-5892-043-7, http://hdl.handle.net/1765/1043

Kippers, J., *Empirical Studies on Cash Payments*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2004-043-F&A. ISBN 90-5892-069-0, http://hdl.handle.net/1765/1520

Koppius, O.R., *Information Architecture and Electronic Market Performance*, Promotors: Prof. dr. P.H.M. Vervest & Prof. dr. ir. H.W.G.M. van Heck, EPS-2002-013-LIS, ISBN: 90-5892-023-2, http://hdl.handle.net/1765/921

Langen, P.W. de, *The Performance of Seaport Clusters; A Framework to Analyze Cluster Performance and an Application to the Seaport Clusters of Durban, Rotterdam and the Lower Mississippi*, Promotors: Prof. dr. B. Nooteboom & Prof. drs. H.W.H. Welters, EPS-2004-034-LIS, ISBN: 90-5892-056-9, http://hdl.handle.net/1765/1133

Liang, G., *New Competition; Foreign Direct Investment And Industrial Development In China,* Promotor: Prof.dr. R.J.M. van Tulder, EPS-2004-047-ORG, ISBN 90–5892–073–9, http://hdl.handle.net/1765/1795.

Loef, J., *Incongruity between Ads and Consumer Expectations of Advertising*, Promotors: Prof. dr. W.F. van Raaij & Prof. dr. G. Antonides, EPS-2002-017-MKT, ISBN: 90-5892-028-3, http://hdl.handle.net/1765/869

Mandele, L.M., van der, *Leadership and the Inflection Point: A Longitudinal Perspective*, Promotors: Prof. dr. H.W. Volberda, Prof. dr. H.R. Commandeur, EPS-2004-042-STR, ISBN 90–5892–067–4, http://hdl.handle.net/1765/1302

Meer, J.R. van der, *Operational Control of Internal Transport*, Promotors: Prof. dr. M.B.M. de Koster & Prof. dr. ir. R. Dekker, EPS-2000-001-LIS, ISBN:90-5892-004-6, http://hdl.handle.net/1765/859

Miltenburg, P.R., *Effects of Modular Sourcing on Manufacturing Flexibility in the Automotive Industry: A Study among German OEMs*, Promotors: Prof. dr. J. Paauwe & Prof. dr. H.R. Commandeur, EPS-2003-030-ORG, ISBN 90-5892-052-6, http://hdl.handle.net/1765/1039

Mol, M.M., *Outsourcing, Supplier-relations and Internationalisation: Global Source Strategy as a Chinese Puzzle*, Promotor: Prof. dr. R.J.M. van Tulder, EPS-2001-010-ORG, ISBN: 90-5892- 014-3, http://hdl.handle.net/1765/355

Mulder, A., *Government Dilemmas in the Private Provision of Public Goods*, Promotor: Prof. dr. R.J.M. van Tulder, EPS-2004-045-ORG, ISBN: 90-5892- 071-2, http://hdl.handle.net/1765

Muller, A.R., *The Rise of Regionalism: Core Company Strategies Under The Second Wave of Integration*, Promotor: Prof. dr. R.J.M. van Tulder, EPS-2004-038-ORG, ISBN 90–5892–062–3, http://hdl.handle.net/1765/1272

Oosterhout, J. van, *The Quest for Legitimacy: On Authority and Responsibility in Governance*, Promotors: Prof. dr. T. van Willigenburg & Prof.mr. H.R. van Gunsteren, EPS-2002-012-ORG, ISBN: 90-5892-022-4, http://hdl.handle.net/1765/362

Peeters, L.W.P., *Cyclic Railway Timetable Optimization*, Promotors: Prof. dr. L.G. Kroon & Prof. dr. ir. J.A.E.E. van Nunen, EPS-2003-022-LIS, ISBN: 90-5892-042-9, http://hdl.handle.net/1765/429

Popova. V., *Knowledge Discovery and Monotonicity*, Promotor: Prof. dr. A. de Bruin, EPS-2004-037-LIS, ISBN 90-5892-061-5, http://hdl.handle.net/1765/1201

Puvanasvari Ratnasingam, P., *Interorganizational Trust in Business to Business E-Commerce*, Promotors: Prof. dr. K. Kumar & Prof. dr. H.G. van Dissel, EPS-2001-009-LIS, ISBN: 90-5892-017-8, http://hdl.handle.net/1765/356

Romero Morales, D., *Optimization Problems in Supply Chain Management*, Promotors: Prof. dr. ir. J.A.E.E. van Nunen & Dr. H.E. Romeijn, EPS-2000-003-LIS, ISBN: 90-9014078-6, http://hdl.handle.net/1765/865

Roodbergen , K.J., *Layout and Routing Methods for Warehouses*, Promotors: Prof. dr. M.B.M. de Koster & Prof. dr. ir. J.A.E.E. van Nunen, EPS-2001-004-LIS, ISBN: 90-5892-005-4, http://hdl.handle.net/1765/861

Six, F.E., *Trust and Trouble: Building Interpersonal Trust Within Organizations*, Promotors: Prof. dr. B. Nooteboom & Prof. dr. A.M. Sorge, EPS-2004-040-ORG, ISBN 90–5892–064–X, http://hdl.handle.net/1765/1271

Slager, A.M.H., *Banking across Borders*, Promotors: Prof. dr. D.M.N. van Wensveen & Prof. dr. R.J.M. van Tulder, EPS-2004-041-ORG, ISBN 90-5892-066–6, http://hdl.handle.net/1765/1301

Speklé, R.F., *Beyond Generics: A closer look at Hybrid and Hierarchical Governance*, Promotor: Prof. dr. M.A. van Hoepen RA, EPS-2001-008-F&A, ISBN: 90-5892-011-9, http://hdl.handle.net/1765/357

Teunter, L.H., *Analysis of Sales Promotion Effects on Household Purchase Behavior*, Promotors: Prof. dr. ir. B. Wierenga & Prof. dr. T. Kloek, EPS-2002-016-ORG, ISBN: 90-5892-029-1, http://hdl.handle.net/1765/868

Vis, I.F.A., *Planning and Control Concepts for Material Handling Systems*, Promotors: Prof. dr. M.B.M. de Koster & Prof. dr. ir. R. Dekker, EPS-2002-014-LIS, ISBN: 90-5892-021-6, http://hdl.handle.net/1765/866

Waal, T. de, *Processing of Erroneous and Unsafe Data*, Promotor: Prof. dr. ir. R. Dekker, EPS-2003-024-LIS, ISBN: 90-5892-045-3, http://hdl.handle.net/1765/870

Wielemaker, M.W., *Managing Initiatives: A Synthesis of the Conditioning and Knowledge-Creating View*, Promotors: Prof. dr. H.W. Volberda & Prof. dr. C.W.F. Baden-Fuller, EPS-2003-28-STR, ISBN 90-5892-050-X, http://hdl.handle.net/1765/1036

Wijk, R.A.J.L. van, *Organizing Knowledge in Internal Networks: A Multilevel Study*, Promotor: Prof. dr. ing. F.A.J. van den Bosch, EPS-2003-021-STR, ISBN: 90-5892-039-9, http://hdl.handle.net/1765/347

Wolters, M.J.J., *The Business of Modularity and the Modularity of Business*, Promotors: Prof. mr. dr. P.H.M. Vervest & Prof. dr. ir. H.W.G.M. van Heck, EPS-2002-011-LIS, ISBN: 90-5892-020-8, http://hdl.handle.net/1765/920