Faculty of Economics and Business Administration

Academic year 2000-2001

Modeling Expected Returns and Discount Factors for European Stock Markets

John Crombez

Promotor: Prof. Dr. Rudi Vander Vennet

Co-promotor: Prof. Dr. Jan Annaert

Dissertation submitted to obtain the degree of PhD in Economics

April 2001

Faculty of Economics and Business Administration

Academic year 2000-2001

Modeling Expected Returns and Discount Factors for European Stock Markets

John Crombez

Promotor: Prof. Dr. Rudi Vander Vennet

Co-promotor: Prof. Dr. Jan Annaert

Dissertation submitted to obtain the degree of PhD in Economics

April 2001

Preface

Deze doctoraatsthesis is het resultaat van mijn (nog steeds groeiende) interesse in financiële markten, maar meer nog van de omgeving waarin ik heb gewerkt. In de eerste plaats wil ik mijn promotor, Rudi Vander Vennet bedanken voor de kans die hij mij heeft geboden om te doctoreren. Daarnaast natuurlijk ook voor de wetenschappelijke steun en de nodige sturing. Een betere kans en omgeving om te doctoreren kan een promotor een jonge onderzoeker niet bieden. Ook de aanwezigheid van Jan Annaert heeft dit doctoraat gemaakt tot wat het nu is. Hij is steeds bereid geweest mijn niet ophoudende stroom vragen te beantwoorden.
Vervolgens wil ik de volledige vakgroep danken voor de omkadering, hun functie als praatgroep en veel dingen meer. In het bijzonder wil ik Gert eruit lichten, niet alleen omdat hij vier jaar tolerant met mij een bureau heeft gedeeld, maar ook omdat hij een stimulans is geweest bij het doctoreren zelf.
Graag wil ik ook de overige leden van mijn doctoraatscommissie, William De Vijlder, Freddy Van den Spiegel, Marc Deceuster en Geert Bekaert, bedanken voor hun bereidwilligheid om tijd in mijn doctoraat te investeren. Het is een plezier en een geruststelling om suggesties en opmerkingen te onvangen van mensen met heel wat ervaring in de materie.

Laat mij stellen dat het proces van het schrijven van het doctoraat een speciaal karakter heeft. Ik ben er vrij zeker van dat, in de periodes waarin er voor de doctorandus niets anders meer bestaat dan zijn doctoraat, verdraagzaamheid vereist wordt van zijn directe omgeving. De steun van mijn ouders en broer, Heidi en haar ouders en Olivier zijn voor mij onschatbaar geweest. Ik zal hen dan ook nooit genoeg kunnen bedanken. Verder wil ik alle vrienden en kenissen bedanken voor aanwezigheid en steun in de hele periode. In het bijzonder een aantal mensen die zeer aanwezig waren: Patrick, Patricia, Sien, Benoît, Carmen, Tonie, Anneke, Pieter, de drummer en de gitarist.

John
Gent, juni, 2001.

Contents

Chapter I

Introduction

Introduction

1.Expected returns, factor models and discount factors

This dissertation assesses different aspects in the modeling of expected returns and discount factors for European stock markets. The relevance of this topic is easily motivated for financial markets. First, a lot of the empirical literature in finance is dedicated to the study of which risk factors are priced. Finding what factors are priced and hence, what the appropriate asset-pricing model is, implies a framework for portfolio diversification, risk management and investment fund analysis. The first two chapters of this dissertation analyze the relevance of a well-known set of risk factors for European stock markets. In these chapters, we also evaluate European sector value-growth portfolio strategies and the sources of the difference in return between value and growth portfolios. Second, financial agents are confronted with complex decisions. Examples of such decisions are the determination of the initial price for an initial public offering, a judgment about the cost of equity, etc. But also common decisions fit into the decision-making aspect of the modeling of expected returns. It concerns the decision to buy or sell a stock for a retail investor or a portfolio manager, or to increase or reduce the weight of an asset in his wealth portfolio for a consumer. These financial estimation and decision-making problems are inherent to the modeling of expected returns and discount factors and are studied in the fourth chapter.

We will first introduce a general asset-pricing framework that covers most of the questions handled in this dissertation. The reason why we start with such a general framework is to give some intuition about the relationship between expected returns, factor models and the stochastic discount factor. In a lot of the research in financial markets and also in this thesis, the relationship between these three concepts covers most of the problems concerning the selection of asset-pricing models and decisions a financial agent has to make.
All asset-pricing models can be regarded as derivations from the general standard consumption-based asset-pricing model. We start from this general concept of relating asset prices to the investor's decisions about consumption and saving. Consumers evaluate how much of their wealth they want to invest in risky assets and how much they want to consume immediately. They will continue to buy or sell assets until the marginal cost of not instantaneously consuming equals the marginal benefit of holding additional risky assets. This concept is expressed by the first-order condition from an inter-temporal choice problem of an investor and provides the pricing equation [equation 1: Campbell et al., 1997, Cochrane, 2001]. From [1], we

evaluate the first-order condition analyzing the investor's optimal consumption and portfolio plan (equation 2) [Cochrane, 2001].

[1] $$p_t = E_t\left[d\,\frac{u'(c_{t+1})}{u'(c_t)}x_{t+1}\right],$$

[2] $$p_t u'(c_t) = E_t[d\,u'(c_{t+1})x_{t+1}],$$

where p_t is the asset price, x_{t+1} is the asset payoff, c_t is the consumption at *t*, d is the time discount factor, *E* denotes expectation and u' denotes marginal utility.

The left side of equation 2 stands for the loss in utility in the case where the investor purchases one additional unit of the asset and hence decides not to consume the unit. The right side expresses the increase in expected utility when the investor invests in the asset at *t* and receives an extra payoff at *t+1*. The investor buys and sells the asset until marginal cost equals marginal benefit.

In equation 1, $d\,u'(c_{t+1})/u'(c_t)$ is denoted at the stochastic discount factor or pricing kernel m_{t+1}, so directly from equation 1, the basic pricing formula can than be expressed in the following form:

[3] $$p_t = E_t(m_{t+1}x_{t+1}).$$

Often, this pricing equation is presented in its unconditional form: $p = E(mx)$. What is important is that this specification states that all asset-pricing models are different representations of linking the stochastic discount factor to the data [Cochrane, 2001]. It means that we link the following description of the stochastic discount factor,

[4] $$m_{t+1} = f(data, parameters),$$

to the pricing equation in [3] to model asset prices. Intuitively speaking, a certain choice of a model *f(.)* leads to predictions using equation 4 stated in terms of a expected return-beta representation. A well-known formulation of [3] is the following:

[5] $$E(r) = r^f + b_m l_m\ ,$$

where $E(r)$ denotes the expected return for an asset, *r* is the return on any asset, r^f is the risk-free rate of return, b_m is the factor loading of the asset on the discount factor *m* and l_m denotes the price of risk, common to every asset.

What is required to see that the relation between factor models, expected returns and discount factors can be derived from the general equation 3 to the well-known formulation in [5]? Equation 3 is now stated in terms of a pricing kernel for asset prices. The same general formulation can be derived for asset returns as a special case of [3]:

$$1 = E(mr). \tag{6}$$

Rewriting equation 6 by using the covariance decomposition of the right-hand side term and using $r^f = 1/E(m)$, we can rewrite [6] as:

$$\begin{aligned} &E(r) - r^f = -r^f \operatorname{cov}(m, r) \\ &or \\ &E(r) - r^f = \frac{-\operatorname{cov}(u'(c_{t+1}), r_{t+1})}{E(u'(c_{t+1}))} \end{aligned} \tag{7}$$

This pricing equation for asset returns implies that there is a direct relation between the discount factor and returns. Next, the discount factor can be regarded as a combination of factors or factor portfolios that are ex ante mean-variance efficient (further: MVE). The MVE frontier determines how much expected return is given for a certain level of risk. A portfolioor asset return or a linear combination of portfolio or asset returns[1] that is on the MVE frontier can express all returns on this frontier. A well-known example is given by the CAPM, where the only factor is the return on the market portfolio. In the existence of a risk-free rate, the discount factor for this simple one-factor model is given by a linear combination of the risk-free rate and the return on the market portfolio. If this one-factor model holds, this linear combination of factors will ex ante be MVE. This MVE return is expressed by equation 7. If we denote this MVE return as r^{mve}, this means that

$$r^{mve} = r^f + x(r^m - r^f) \tag{8}$$

defines all returns on the efficient frontier (for some scalar x). In a single-beta representation, expected returns can be expressed using any MVE return:

$$E(r) = r^f + b_{mve}(E(r^{mve}) - r^f), \tag{9}$$

with b_{mve} the beta on the single factor in the model.

[1] In the remainder of this thesis, we will refer to a single explanatory factor or factor portfolio or a linear combination of explanatory factors or factor portfolios as the factor portfolio.

In the case of an expected return-beta model with the vector of factors f and the vector of factor loadings b, we have $m = b' f$, linear in the factors that satisfy the pricing equation $p = E(mx)$ [Cochrane, 2001], and hence this example can be extended to all factor models.

This general formulation of asset pricing accounts for most settings to test expected returns and discount factors. In this dissertation, we will go deeper into three aspects of asset pricing and the modeling of expected returns. First, we analyze asset-pricing models for a large and European dataset, as a direct application of the framework described above. In the second part of the dissertation, we will explore the relation between accounting multiples (as earnings-to-price and book-to-price) and expected return. One of the questions in that part of the dissertation is whether the returns earned by a zero-cost earnings yield strategy can be explained by fundamental risk. In a third and final part, we will discuss some aspects of decision-making in financial markets with respect to expected returns.

2. The cross-section of European stock returns

A large part of past research in financial markets has been devoted to the identification of asset pricing models. In chapter II of the dissertation, we explore the relevance of explanatory factors for asset returns that are found to be useful for U.S. data, for a large new set of European stock returns. The central hypothesis we assess is whether a linear combination of factors is ex ante mean-variance efficient. As we outlined in the previous section, if a factor portfolio or a linear combination of factors is ex ante MVE, it should describe the returns on all assets.

Research about the exact specification of an asset-pricing model is relevant for most problems in investment analysis. The identification of priced risk factors allows us, in the first place, to perform an asset selection that is motivated by risk diversification. If we know what the factor loadings are for different assets on different priced risk factors, asset selection can be performed based on this knowledge within the objective of risk diversification.
Secondly, estimating expected returns for the purpose of the application of the expected return–variance rule is conducted by identifying the right asset-pricing model. Equation 8 of the first section gives a general representation of this idea. When the factor loadings on priced risk factors are estimated and also the price of risk is estimated, the expected return can be calculated and used to determine portfolio weights. However, past research has documented the impact of small deviations in the calculated expected returns on portfolio weights by applying the

expected return –variance rule [as in Jobson and Korkie, 1980]. Hence, information on the exact identification of the asset-pricing model should provide more reliable expected returns.

Although this aspect is not assessed in this thesis, the exact identification of asset-pricing models is important for performance monitoring of investment funds as well. The performance of a fund is often evaluated by its alpha in an asset-pricing model. Given that funds are also assets in the risk-return space and accounting for all priced risks in the market, the alpha expresses the excess performance of the fund given its factor loadings on the priced risk factors. A false identification of the asset-pricing model implies that the estimated alphas are not reliable.

The framework in the second chapter of this dissertation evaluates a large set of questions. First, we use a multivariate test to analyze the pricing relationship for European country portfolios, sector portfolios and size portfolios. Second, we test the ex ante MVE of the factor portfolio for an international CAPM as well as for extensions of this model by additional factors. Third, we evaluate the power of this multivariate test against the alternative model. In this alternative model, we both evaluate the possibility that there is a misspecification of the risk factors (risk-based alternative) and the possibility that there are data-snooping issues or market imperfections (nonrisk-based).

There are several applications for the answers we provide to these questions. First, one asset-pricing model we evaluate is the Fama-French [1993] three-factor model. This model is widely used in practice for portfolio selection purposes and fund performance measurement. Second, the risk-based alternative asset-pricing model we suggest, is the extension of the European market portfolio of stocks by including the return on human capital (as a proxy for the portfolio of labor income). If this model holds for European assets, the CAPM is not dead, its factor portfolio is only misspecified. Third, we identify for three types of European portfolios whether the extensions of the CAPM do price the assets. An answer to this question provides asset-pricing models for European assets. Finally, we evaluate whether the power of these multivariate tests is high enough to conclude that the extensions of the CAPM are reliable.

We extend the international CAPM by using factor-mimicking portfolios as additional explanatory factors. This is one example of the general framework described in equation 4 in the first section of the introduction. The factor portfolios that are tested as an extension of the CAPM are constructed based on size, book-to-market (BTM) and momentum characteristics. The framework is tested under the assumption that a risk-free asset exists and with portfolios as factors [Campbell et al., 1997]. Factor mimicking portfolios are a projection of the factors f on the payoff space the discount factor that consists of these mimicking portfolios carries the same pricing relation on the payoff space as the discount factor defined in equation 4 [Cochrane, 2001].

Table 1.

Empirical evidence on multivariate tests for asset-pricing models

All studies perform multivariate tests for the exact factor pricing relation. All studies use monthly data. The portfolio characteristics on which the dependent portfolios are selected are reported. More details about the portfolio characteristic are provided in the legend. The p-values are reported for the multivariate test that the vector of intercepts is zero. N denotes the number of dependent portfolios and K denotes the number of explanatory factors.

Study	Time period	Portfolio characteristic	*N*	*K*	p-value
CK 88	64/01-83/12	Market value of equity (MV)[1]	10	5	.002
			10	10	.002
FF 93	63/07-91/12	MV and BTM[2],	32	2	.010
		Maturity	32	3	.039
		default risk	32	5	.025
FF 98	1975/1995	Accounting multiples[3]	8	1	.000
			6	2	.194
			6	2	.154
			6	2	.435
			6	2	.036

[1] Connor and Korajczyk [1988] use principal component analysis to determine the explanatory factors.

[2] Fama and French [1993] use 25 stock portfolios based on a two-way sort based on market value (MV) and book-to-market (BTM), 5 government bond portfolios grouped based on maturity and 2 corporate bond portfolios based on default risk..

[3] Fama and French [1998] perform the test for the CAPM on 8 global value and growth portfolios based on book-to-market, earnings yield, cash-flow-to-price and dividend yield. In the additional for cases, they use one of four factor-mimicking portfolios as explanatory portfolios and explain the remaining six.

The ex ante mean variance efficiency of the factor portfolio is evaluated by the multivariate Gibbons-Ross-Shanken test [Gibbons et al., 1989]. The efficiency of portfolios has been extensively evaluated in past literature for different sets of dependent portfolios. The dependent portfolios are the sets of assets that are selected to test ex ante efficiency of the factor portfolio and formed by regrouping the sample of stocks based on a characteristic such as size or sector membership. A sample of the testing of asset-pricing models is presented in Campbell et al. [1997]. We will briefly discuss the most relevant results in the sense that the design of the discussed papers has characteristics we will also evaluate. Table 1 shows the results for a study by Connor and Korajczyk [1988] and Fama and French [1993]. We extend the results reported in Campbell et al. [1997] by an international study by Fama and French [1998]. Based on table 1 and on the results of some other research, we will give an overview of what the findings are for the evaluation of the mean-variance efficiency of factor portfolios for U.S. data.

Table 1 shows that there is no clear support in favor of or against the ex ante mean variance efficiency of the factor portfolios for different types of dependent portfolios. Fama and French [1998] find that they cannot reject that a linear combination of the market portfolio and the factor mimicking portfolio based on valuation multiples are mean-variance efficient. Moreover, Lehmann and Modest [1988] find that the p-values for the multivariate test depend to a large extent on the number of dependent portfolios. They find low p -values using 5 portfolios compared to the tests where they use 20 portfolios, performing the estimations on weekly data.

However, MacKinlay [1987] reports that too large a set of dependent assets lowers the power of the test. Gibbons et al. [1989] find that they cannot reject the ex ante MVE for the market portfolio for the period 1926-1982. Connor and Korajczyk [1988] as well as Lehman and Modest [1988] and Gibbons et al. [1989] find evidence against exact factor pricing for size portfolios after 1960. MacKinlay [1987] finds the opposite for shorter data series (five to ten years).

Finally, Gibbons et al [1989] motivate the importance of a multivariate test compared to univariate tests. These authors state that univariate tests are intuitively appealing, but wrong inferences can be drawn as to which factors should be taken into account to form the efficient portfolio. Otherwise said, we want to determine which factors are needed to model the discount factor. Gibbons et al. [1989] show results where none of the univariate statistics are significant and yet the multivariate test rejects. Hence the characteristics of the residual covariance matrix are important for the determination of an applicable asset-pricing model for a sample of stocks as well.

These findings imply a lot of interesting results for a practitioner using factor models for portfolio selection purposes or performance monitoring. The first and most important finding is that, as table 1 shows, the different studies are inconclusive about the results. There is no determined answer for the question which asset-pricing model is best to use. This is already enough of a motivation to assess the European version of this research question. A second important item we learn from past papers is that univariate analyses are not suitable for the selection of factors. In the one-by-one estimation of the relation between portfolio returns and factors, all constants can be statistically zero and as Gibbons et al. [1989] indicated, the factor portfolio can still not be evaluated as ex ante MVE in a multivariate setting. Therefore, we focus in this analysis on the mutlivariate evaluation of the factor models. Thirdly, as we outlined in the previous section, if a factor portfolio is found to be ex ante MVE, it should price all assets. The results from the Fama-French 1998 study as well as from other papers learn us that this is not the case. It implies a warning for the application of a factor model that is not rejected for one set of portfolios or even for one region (for example U.S stocks) to the specific set a practitioner is using. Observing this discrepancy between theory and empirical observations, we test the different asset-pricing models for three sets of dependent portfolios regrouped based on size, sector membership and country membership. Finally, another indication given in different papers is that the power of the test can vary a lot for different research designs. This enforces us to give a lot of attention to this topic as well.

The factor models that we suggest in this second chapter are evaluated keeping these findings from past research in mind. Eventually, what we look for is to select a set of factors that is large enough to price assets, yet not too large to avoid over-fitting of the data. Most research for the U.S. indicates that the number of factors required to price assets lies between three and five [as in Chen et al., 1986, among others]. This set of factors should capture most of the variation in asset returns, and the selection of factors should be conducted with a powerful test. If practitioners are willing to use a factor model to model the expected return or to evaluate investment funds, they should be aware of the fact that each model is evaluated keeping in mind that the choice of the methodology determines the outcome and that there is a possibility that biases are not entirely avoided in the analysis. This implies that there is no 'true factor model' in the sense that every choice can and will be debated. However, we believe that a carefully designed analysis using a methodology that is econometrically acceptable and that avoids possible biases as much as possible is a better guide to select a factor model for practical purposes than simply using a model that somebody else found to be powerful for a different set of assets and for a different geographical region.

3.Earnings yield forecasts, BTM and expected returns

In the third part of this dissertation, we explore the relation between earnings yield forecasts, book-to-market and expected returns. Past research documents that expected returns are higher for stocks with high earnings yield and lower for stocks with low earnings yield. These empirical findings have led to a successful introduction of what is called contrarian investment strategies. In this thesis, we also denote the contrarian strategies based on accounting multiples as the value-growth strategy.

The basic motivation for these strategies is that investors are susceptible to cognitive biases[2] and make false intuitive judgments. A simple example of this motivation is that investors extrapolate present growth in earnings too far into the future, biding up prices too high, and hence an investment strategy taking the opposite expectation would be profitable. Generally, in practice it implies that buying stocks that are priced low relative to accounting measures of operating performance[3] (often denoted as value stocks) and selling stocks that are priced high relative to accounting measures of operating performance (often denoted as growth stocks) is a profitable strategy.
These strategies have become popular in practice. A large amount of stock funds are managed based on contrarian strategies (i.e. contrarian funds). Hence, a profound analysis of these contrarian strategies for European stock markets seems relevant. Also the fact that we analyze these value-growth strategies on a sector basis is appropriate for European stock markets. Towards the introduction of the single currency, a shift from country-based portfolio selection to sector-based portfolio selection has been observed in European stock markets. The identification of the returns from these strategies as well as the determination of the driving forces behind these returns can improve insights about sector stock selection in European investment management.

In this dissertation, we only handle the case of earnings yield forecasts and BTM and focus on the earnings yield forecasts because of their relation to expected returns. For U.S. data, the same issue has been empirically assessed using price-to-sales and price-to-cash flow ratios. Bakshi and Chan [2000] report evidence of this phenomenon for a large international dataset using one-year forecasts of earnings yield. Table 2 show that for most countries, their studied time-period is the same as the one we study (1987-1998) In most countries they find a positive excess return for the zero-cost strategy of a long position in high forecast earnings yield stocks and a short position in low forecast earnings yield stocks. However, when they look at the proportion of months in which this zero-cost strategy provides a positive return, they find that this proportion is only in half of the cases significant. Moreover, a country as the UK with a large capitalization and a large number of stocks in this international sample shows absolutely no relation between earnings yield forecasts and return. For U.S. data, the finding of a relationship between earnings yield and expected return seems empirically relevant. Different authors find for different time periods and different stock samples that the annualized return premium for the

[2] irrationalities
[3] By accounting measures of operating performance we mean accounting numbers such as earnings, book value, cash-flow, sales and dividends.

zero-cost strategy is 9%. Dechow et al. [1999] report the same return premium for a comparable period (1976-1998).

However, the findings for the international dataset motivate further research on the topic. Not much consensus has been found in past literature for the explanation of this return premium. Some authors explain the premium by the cognitive failure or irrationalities of false intuitive forecasts allowing for contrarian strategies [Lakonishok et al., 1994]. The failure itself would be ascribed to the fact that investors extrapolate past earnings growth too far into the future. Hence, stocks with a high market price relative to the earnings (or low earnings yield stocks) imply high growth prospects. Investors extrapolate these prospects too far into the future leading to a too high valuation of the current stock price and making high earnings yield stocks more attractive. Other researchers do not accept this explanation but explain the return premium by biases in analysts' forecasts [Dechow and Sloan, 1997].

The only subject past researchers seem to agree on with respect to this premium is that it cannot be explained by fundamental risk [Lakonishok et al., 1994 and La Porta, 1996]. However, portfolios are formed based on the valuation multiples as earnings yields and book-to-market, but there is a possibility that other fundamentals and the exposure of the individual stocks in the sub-sample is very different. Therefore, we regroup the available European stocks into sector portfolios, assuming that the exposure of these stocks to market risk is comparable and that also the individual firm characteristics are comparable within one sector for companies that operate in the same businesses. Next, we explore whether we find a return premium for the zero-cost strategy based on earnings yield forecasts for each sector.
Making the assumption that the underlying characteristics for individual stocks are comparable in that the exposure to market risk and the fundamental firm characteristics are comparable in a sector, reduces the possible explanations for the return premium. Especially, the explanation that biases in analysts' forecasts account for the premium weakens when the analysis is conducted on a sector basis. If analysts' forecasts are biased in their level and in their dispersion (the extent to which analysts disagree about the forecast of earnings), this bias is due to the difficulty to predict the earnings. We argue that this difficulty is more determined by the sector the firm belongs to than to the decile based on the accounting multiple. Hence, the bias in analysts' forecasts would have a smaller impact on the return spread within one sector.

Therefore, we explore the explanation that there is a difference in expected returns for portfolios of stocks with high forecast earnings yield and portfolios of stocks with low forecast earnings yield because of fundamental risk. Lakonishok et al. [1994] define fundamental risk by arguing that stocks that are fundamentally riskier because they are expected to perform worse in periods where the marginal utility of wealth is high. This implies that the payoffs of growth stocks should be higher in bad states of the world such as recessions and bearish markets. As mentioned earlier, they find no evidence for this explanation.

Table 2.

International evidence on the relation between earnings yield forecasts and returns

All returns are estimated for the period 1987:01-1998:06 except for Finland (from 1988:01), Portugal (from 1991:04) and the U.S. (from 1976:01). p denotes the proportion of months with a positive return for the value-growth strategy. * indicates whether this proportion is statistically significant different from 0.5.

Annualized average monthly returns in % and for the local currency, 1987:01-1998:06				
	low E/P	high E/P	return premium	p
Austria	0.93	11.22	10.29	.68*
Belgium	12.74	13.74	1.00	.56
Denmark	4.30	12.27	7.97	.62*
Finland	7.35	15.57	8.22	.55
France	6.14	8.59	2.45	.54
Germany	3.78	9.81	6.03	.57
Ireland	5.96	14.54	8.58	.54*
Italy	-0.06	6.83	6.89	.57*
Netherlands	11.32	13.01	1.69	.44
Portugal	11.99	12.52	0.53	.53
Spain	7.14	14.30	7.16	.60*
Sweden	16.01	12.58	-3.43	.54
Switzerland	0.28	12.21	11.93	.69*
UK	6.78	6.99	0.16	.46
US	9.46	20.29	10.65	.65*

Using samples grouped on sector membership, we re-assess this question. Starting from the consumption-based CAPM as described in the first section, we model the discount factor as proposed by Campbell and Cochrane [1999, 2000]. The discount factor depends in this specification on the habit of consumption and on aggregate consumption. Using variables for habit and aggregate consumption, we implicitly assume to identify the relation between the returns on the value-growth strategy and the state of the world. We also test whether the discounted returns for value and growth portfolios are the same using the factor models we identified for sector portfolios in chapter II.

In chapter III, we discuss the use of earnings yield forecasts with respect to valuation models and returns. The aim of this part is to describe the theoretical framework in which we analyze the relation between valuation multiples[4] and expected returns with respect to consumer decisions. A following section of chapter III empirically evaluates the return premiums for the value-growth strategies in European sectors. Finally, in this second section, we analyze whether the states of the world explain these returns. In other words, we analyze the evidence that the states of the world account for the return premium paid using this strategy. If returns of value stocks covary negatively with marginal utility, they should earn an additional risk premium. Hence, what we test for is whether the return premiums for value-growth strategies can be explained by risk.

The finance and accounting literature have identified a return premium from investing in value stocks and short-selling growth stocks. However, this literature meets the problem of identifying the source of these returns. For that reason, it is labeled as an anomaly. There are no risk-based explanations for these returns and often, these returns are assigned to the irrational behavior of investors. Therefore, it is important to know if this anomaly exists for a different dataset. We analyze the contrarian strategies for European sector samples. If investment funds are launched in Europe based on contrarian strategies, a lot of these funds are managed on a sector basis. There has been some research questioning the return from a value-growth strategy for European stock samples, however mostly for country samples. Hence it is important to learn the about the existence of these returns in European sector samples, and as important, to identify the source of these returns. The fact that contrarian funds have been successful in the past is not a good reason to simply assume that these strategies work unconditionally. The results from this chapter learn that this statement is not exaggerated. Especially the part about the dynamics of these returns allows a fund manager to be critical about his investment strategy and be careful about the execution of it.

[4] In this dissertation we use the term valuation multiple as a synonym of accounting multiple indicating a ratio of an accounting measure of operating performance to the market price of the firm.

4. Rational decisions and expected returns

Investor decisions are complex. Some authors argue that there are too many examples of observed irrationalities to justify the assumption of universal rationality in financial markets [as in De Bondt and Thaler, 1995]. These observations justify the search for descriptive[5] behavioral models to describe decision-making in finance. However, the basic assumptions for these models are arbitrarily chosen and it is difficult to see whether results of psychological tests on a sample of people apply to investors. Therefore, in chapter four, we go back to evaluating investor decisions in a normative[6] framework. More specifically, we evaluate decisions about the next period's return.

One of the axioms of normative models is that people are unbiased Bayesian forecasters or, in other words, make choices based on rational expectations. In chapter four, we evaluate Bayesian forecasts about the next period's return. In a first part, we analyze this decision and its possible deficiencies to illustrate the importance of the study of normative decisions in financial markets. In a second part, we study the value of Bayesian forecasts about returns for stock selection purposes and in a third part we analyze Bayesian forecasts of the mean vector and its application in asset allocation decisions.

Why is it important to study investor decisions from this angle and what are the novelties presented in chapter four? First, recent literature draws the attention to information diffusion in financial markets [as in Hong et al., 2000 among others]. We focus on this topic in the evaluation of the rational decisions. Second, it is important to evaluate Bayesian forecasts or rational expectations to understand the problems an investor is faced with when making his decisions. A better understanding of these problems can contribute to the discussion about efficiency of financial markets. If rational expectations are not optimal ex post because of other reasons than irrationalities, this questions some of the assumptions used in behavioral finance. Third, we use analyst forecasts to model the prior distribution or the base rates for an investor. Starting from the idea that analysts posses costly public information, these forecasts are valuable for the investor. Bayesian forecast have been used in finance before but the problem a practitioner and a researcher is faced with is to define a proper prior that is applicable to individual assets or asset classes. The analyst forecasts suggested in this chapter are an example of an individual prior and the use of these forecasts imply that the loss of information is reduced. Often, in finance, a common prior is used for all assets meaning that there is a lot of individual information that does not enter the decision-making process.

---

5 descriptive models describe how people *do* behave

6 normative models describe how people *should* behave

An example of the practical interest of this chapter concerns portfolio selection procedures. We argued before that the estimation of the expected return vector is crucial for the ex post optimality of optimized portfolios. A study of the ex ante rational forecasts, and the environment in which these judgments are made, can reveal the possible problems of this decision process. The empirical evidence of the existence of beneficial value-growth strategies, mentioned in the previous part is another example. Finally, the large equity premium has been puzzling for years and is difficult to explain by rational models. Is there really a case for irrational agents? Or are there different causes and can the market's microstructure be improved in order to increase market efficiency?

De Bondt and Thaler [1995] recognize the problem that, in finance, there is little attention to the investor decision process or the quality of the judgment made by investors. There is an increased attention from academics over the past years to study the observed market anomalies starting from cognitive failures that are documented by psychologists. Cognitive failures in general explain the deviations from assumed rationality. A rational investor is modeled making his forecasts in the decision process by applying Bayes' rule. The entire set of normative axioms explaining the investor behavior consists of the axiom that investors behave as risk-averse expected utility maximizers and Bayesian forecasters [De Bondt and Thaler, 1995]. The assumptions are often made without knowledge of their applicability in the behavior of investors. In the fourth part of this dissertation, we argue that the definition of the relevant environment is crucial for their relevance in descriptive models that explain the observed anomalies in financial markets.

In chapter four of this dissertation, we handle three topics related to rational forecasts of returns in financial markets. In the first section, we study a simplified environment that maintains its realism to evaluate the rational decisions that are made. The objective is to get insight in the characteristics of these rational decisions and evaluate these rational decisions compared to cognitive failures that are assumed.

This evaluation both concerns practitioners and academics. It is important for practitioners to be aware of the problems concerned with rational forecasts and caused by the environment the decision is made in. Suppose that a fund manager makes a portfolio selection and assume that this decision is ex-ante rational. After one period, this portfolio manager can be evaluated as having a poor performance. In other words, his decisions are sub-optimal ex-post. It is in this case very well possible that the portfolio manager has a simplified environment where he knows the first two moments of the multivariate distribution of the returns of the assets and he receives additional information from a research team, which he uses as prior evidence. The portfolio manager is aware that his decision is a complex one and he uses the evidence from the expert team as his prior expert evidence or his base rates. A possible bias in his base prior expert evidence can cause his rational decision to be sub-optimal ex-post. De Bondt and Thaler [1995] describe the importance to compare actual decisions with their normative benchmarks for the decision-making behavior in corporate finance. The stakeholders, such as shareholders, management, customers and suppliers, all make decisions that govern the firm. Examples are the shareholder's preference for dividends and managerial behavior in the evaluation of projects.

For academics, it is important to benchmark their assumptions of cognitive failures in descriptive models. There are a large amount of possible assumptions of irrationalities and it is difficult to choose which assumptions are relevant and on top of that, it is hard to quantify these assumptions [Shiller, 1998]. Moreover, it is not always clear whether the assumptions of cognitive failures that are found to be relevant in psychological experiments are also applicable to investor behavior. This motivates the benchmarking of irrational assumptions conditional on the environment.

In the second part of chapter IV, we evaluate rational forecasts of returns with respect to their relevance for stock selection. The environment in which the rational forecasts are made is analyzed as well. The objective of this section is to evaluate whether ex ante rational forecasts of returns are also optimal ex post when they are used for stock selection purposes. Furthermore, we explore what the characteristics of the statistical evidence and the base rates are for the judgments about the return. The most important aspect of the analysis is the evaluation of the market of information. The question raised is whether market imperfections can explain anomalies.

The third part of chapter IV evaluates the rational estimation of the mean vector as an input parameter for the expected return – variance rule. We previously mentioned the importance of a reliable estimator for the mean vector with respect to the stability of the optimized portfolio weights. In this section, we extend the Jorion [1991] paper using the mean vector of rational forecasts as an additional estimator for the mean vector. Again, we stress the importance of the characteristics of the environment in which the judgment is made. For different types of forecasts, we evaluate the ex-post performance of optimized portfolios. These portfolios are composed both with and without short-sales and evaluated before and after transaction costs.

Summarizing, we analyze rational forecasts of returns in the objective to evaluate rational decisions made by Bayesian forecasters compared to different model assumptions. Overall, we analyze the importance of the environment in which the judgment is made and draw inferences from this framework that can both be interesting for practitioners and academics.

Chapter II

The Cross-Section of Expected Returns for European Stock Portfolios

The Cross-Section of Expected Returns for European Stock Portfolios[*]

Abstract

The empirical testing of single-factor and multi-factor asset pricing models attempts to identify the relevant risk factors that should be used by investors to value risky cash flows and tries to assess the ability of various models to produce estimated expected returns without misspecification. In this paper we evaluate different model specifications for European stock market data over the period 1979-1998 and for different portfolio types. We find indications that a one-factor market model accurately describes European country portfolios while more factors are required to describe the returns on sector and size portfolios. We find no evidence against the observation that a linear combination of the market portfolio and the momentum factor portfolio is mean-variance efficient. However, with respect to sector portfolios, we find that it is possible that the market portfolio is misspecified and that the return on human capital should be included. Furthermore, we find no evidence against an exact factor pricing relation using the market portfolio, the size factor portfolio and the momentum factor portfolio for the pricing of size portfolios. An evaluation of the power of the tests reveals that this evidence is robust.

[*] Co-authored by R. Vander Vennet.

1. *Introduction*

Estimating and testing the ex ante mean-variance efficiency (further: MVE) for a factor portfolio of both single- and multi-factor models has been studied extensively over the last three decades. Following the finding of various anomalies associated with the single-factor market model, various types of multi-factor models have been developed to capture non-market risks. Portfolios as factors based on the book-to-market ratio, firm size and momentum effects feature among the most prominent additional variables used in empirical testing. The challenge for any model is to identify those pervasive risk factors that are consistently priced by the representative investor. In this paper we acknowledge the existence of a broad class of competing asset pricing models. Rather than to test the relative explanatory power of the variables using univariate regressions, we test the mean-variance efficiency of the factor portfolios in a multivariate setting. We investigate the ex ante MVE of the factor portfolio in the case of portfolios as factors in the presence of a risk-free asset. This analysis is done for a carefully constructed database of European stocks. Multivariate tests for asset-pricing models have been explored for international databases before [as in Fama and French, 1998]. We extend this international testing framework by using three types of dependent portfolios (i.e. portfolios based on a characteristic used as dependent variables). Next, we also apply a larger European database compared to other studies. This is the main contribution of this paper.

Asset-pricing models that have been extensively tested for the U.S. database are evaluated for their power on a new and large European dataset. The testing of the ex ante efficiency is important because of its use in the estimation of the firm's cost of equity as well as for asset allocation purposes. Using models in the case where there is no ex ante efficiency can lead to a misspecification of the expected returns, with a potentially large impact both on the estimation of the cost of equity and active portfolio strategies based on expected return-risk optimization procedures.

Testing asset pricing models and exact factor pricing

Fama and MacBeth [1973] proposed an important contribution to the development of a test design of asset-pricing models. They suggested a two-pass estimation method in order to reduce the errors-in-variable problem. Gibbons [1982] developed a direct test of the CAPM to avoid this errors-in-variable problem using maximum likelihood estimation. All these tests are designed to validate the mean-variance efficiency of the market portfolio. Following the development of the arbitrage pricing theory [Ross 1976], a broad set of extensions of the one-factor model has been investigated [Basu, 1977, Banz, 1981, Chen et al., 1986]. As a consequence, the testing framework of asset pricing models has been extended to multifactor models [MacKinlay, 1987, Gibbons, Ross and Shanken, 1989, henceforth GRS].

Ex ante MVE of a factor portfolio implies that the elements of the vector of intercepts for an asset-pricing model (APM) are zero. If the factor portfolio is MVE, this is sometimes denoted in the literature as exact factor pricing using a specified model. The test of a vector of intercepts is the principal hypothesis in the testing of APMs [Gibbons et al., 1989, and Campbell et al., 1997]. If all elements of the vector of intercepts are zero, a linear combination of the explanatory factors forms the tangency portfolio, for which the squared Sharpe ratio is equal to the squared Sharpe ratio of the efficient portfolio [Gibbons et al., 1989]. In fact, the tests for efficiency of some portfolio can be regarded as the ex-ante test of the mean-variance efficiency of the market portfolio (in the case of the CAPM) or of a linear combination of a set of (factor) portfolios (in the case of multi-factor models).

Explaining deviations from exact factor pricing

The possible sources of deviations from ex ante efficiency can be summarized by two broad categories. A first popular approach to explain deviations from ex ante efficiency is the reformulation of the suggested APM by adding an extra risk factor or using different risk factors. This approach is called the risk-based alternative. In other words, the risk-based alternative consists of an extension or a reformulation of the set of (factor) portfolios for which a linear combination is mean-variance efficient.

However, this approach implies that different explanations of deviations from exact factor pricing are not considered. Of course, this different choice of risk factors can be explained by improved identification procedures of the right APM [Campbell et al., 1997]. But different possible interpretations are (1) that there is a possibility that the improvement of the APM is based on a good within-sample fit through data-snooping [Lo and MacKinlay, 1990], or (2) based on a relation with sample selection bias [Breen and Korajczyk, 1993], or (3) based on market inefficiencies. These sources of explanations are labeled as nonrisk-based alternatives. A second reason why the risk-based alternatives are not always a straightforward choice is that the risk factors are often selected without theoretical motivation of the choice of the factors.

Empirical results

A survey of results on these issues is provided by Campbell et al. [CLM, 1997]. They present results for APMs that are tested for statistical factors (using principal components and factor analysis) and theoretical factors (using firm characteristics and macro-economic variables). According to their summary, the strongest evidence against exact factor pricing is found for dependent portfolios that are based on market capitalization and book-to-market. Portfolios based on dividend yield and variance of the individual return series provide little evidence against ex ante efficiency of the factor portfolio. More important is the indication that it is difficult to find strong and clear evidence in favor of or against ex ante efficiency of the factor portfolio for one asset-pricing model. Therefore, it is informative to apply the multivariate testing framework to a large unexplored stock database.

Another interesting debate CLM report is the one about the relevant number of factors for the construction of an APM. Fama and French [1993] find that three factors are sufficient, while Chen et al. [1986] report that five factors are optimal. Finally, Lehman and Modest [1988] show that the number of dependent portfolios included in the analysis determines the strength of the tests. They find lower power for the tests when the number of dependent portfolios is low (i.e. equal to 5 in their study). These are important findings with respect to the evaluation of APMs and will be considered in the setup of this paper.

This paper

This paper tests the ex ante efficiency of factor portfolios for a large European dataset and different specifications of the APM. The CAPM is tested as well as extensions of the CAPM using factor portfolios as factors assuming that a risk-free asset exists. By adding extra risk factors, the first and most popular question of deviations from exact factor pricing is assessed : are there missing risk factors? Also, different regroupings of the stock sample into dependent portfolios are tested. The overview provided by CLM [1997] showed that testing an APM for one group of dependent portfolios does not give sufficient information about exact factor pricing for a different set of dependent portfolios. This makes an extension to different types of dependent portfolios worthwhile. We first regroup stocks into portfolios based on the country characteristic. This implies an indirect test of market integration for the European stock markets.

Second, we evaluate the power of the tests for possible deviations from exact factor pricing. We investigate the sensitivity of the hypothesis testing to the choice of the factor, or the linear combination of factors, which is evaluated to be mean-variance efficient [GRS, 1989]. MacKinlay [1987] reports that the multivariate tests are sometimes not powerful. With respect to the risk-based alternative, we test the possibility that the market portfolio is misspecified. The power of the test is analyzed, evaluating the null hypothesis against an alternative with a different specification of this risk factor. The nonrisk-based alternative APM is analyzed using a simulation procedure. The specification of the vector of intercepts which we used in this paper, is conventional and related to the methods reported by MacKinlay [1987] and CLM [1997].

The use of a previously unexplored dataset allows us to alleviate the potentially disturbing impact of data-snooping biases. Lo and MacKinlay [1990] argue that portfolio formation based on a characteristic identified in prior empirical research instead of being based on theoretical arguments could induce this bias. Other studies have used European stock returns [e.g., Fama and French 1998], but the coverage of our European sample is much broader. In order to perform genuine tests, we group the 2427 European stocks in three different types of dependent portfolios: country portfolios, sector portfolios and size portfolios. This particular choice is made because asset allocation, as it is practiced by institutional investors, has been and is still widely being conducted using a country-based or sector-based evaluation of the universe of investable stocks. We also analyze size portfolios because it provides the opportunity to compare our findings with previous studies. Moreover, we try to avoid a survivorship bias by including non-surviving stocks.

This chapter has three additional focuses. First, we use a different definition of the European wealth portfolio as an alternative APM to evaluate the hypothesis of a misspecified risk factor. Roll [1977 and 1978] concluded that the test of the one-factor market model may not be a genuine test of the CAPM because proxies for the stock market portfolio may fail to capture the true wealth portfolio. The misspecification of the total wealth portfolio may cause biases in the description of the cross-section of stock returns. Jagannathan and Wang [1996] show that the inclusion of labor income next to capital income improves the results for stock returns in the U.S. Hence, this paper deals with the definition of the market portfolio and the problem of data-snooping biases. The CAPM test is performed using a market-capitalization-weighted portfolio of all stocks in the sample. We analyze the value added of using an alternative wealth portfolio by including labor income, next to stock market returns.

A second additional issue in testing ex ante efficiency is the power of the performed tests. MacKinlay [1987] acknowledges that little attention has been devoted to this issue. GRS [1989] and Affleck-Graves and McDonald [1990] explore different possibilities for the formulation of the alternative hypothesis. The evaluation of the tests using various alternatives turns out to produce substantial differences [MacKinlay 1987, CLM 1997]. MacKinlay [1987] reports that alternative assumptions about the risk-free return and the existence of a second factor next to the market return are not sufficient to explain the deviations from exact factor pricing. He also reports that the observed deviations are best explained by the nonrisk-based alternative hypotheses.

A third issue we consider in this study is the stationarity assumption [MacKinlay 1987, GRS 1989, Affleck-Graves and McDonald 1990]. The GRS-test requires stationarity of the excess returns in order to use an estimate of the variance-covariance matrix, and thus limits the number of assets with respect to the number of time periods under consideration because the degrees of freedom that depend on the length on the time period minus the number of assets would be too low. We look at the 20-year window of monthly observations and four subperiods of 5 years. We also estimate a part of the models for distinctive periods of monetary policy to evaluate the characteristics of the APMs in these sub-periods of macro-economic regimes.

The rest of the paper is organized as follows. Section 2 describes the European dataset of stock returns. Section 3 presents the asset pricing equations in an international framework, develops the mutivariate tests used to evaluate ex ante MVE of factor portfolios and discusses the main results. Section 4 deals with the power of the testing framework. Section 5 contains a number of conclusions and suggestions for future research.

2. Data, portfolio construction and variables

The dataset consists of 2427 individual European stocks, aggregated into country portfolios, sector portfolios, and size portfolios (appendix 1 explains the construction of the portfolios). We collected a basic sample containing all European stocks representing at least 80% of the market capitalization in each of 17 European countries at the last trading day of December 1998. We augmented this sample with the stocks that were delisted prior to December 1998. Common reasons for delisting are merger, acquisition and failure. The 80% market capitalization threshold is also used for the dead stocks. This selection procedure implies that the smallest stocks in each country are not included. The main justification is that these stocks may suffer from infrequent trading and low volumes, which may lead to inefficient pricing (see Lo and MacKinlay 1990). From the initial list, preferred stocks were deleted for those companies with both listed ordinary and preferred shares, as well as stocks listed on a stock exchange outside their home country[1]. For the remaining 2427 stocks, we retrieved the monthly returns (calculated as procentual changes in the return index) from January 1979 until December 1998 from Datastream. This dataset is composed of 2070 stocks listed in December 1998 and 357 dead stocks. The inclusion of a sub-sample of dead stocks is intended to reduce possible survivorship bias. Return series are calculated as the relative changes in the return index on a monthly basis. The returns for the dead stocks are calculated up to the last complete trading month of the stock. The monthly returns cover the period January 1979 to December 1998, but the number of stocks in the sample is different every month because a number of stocks were listed later than January 1979, while others were delisted prior to December 1998.

All returns are expressed both in Deutschmark (DEM) and in synthetic euro. The synthetic euro is calculated for the period before the start of the EMU as a GDP-weighted average of the constituent currencies. For the analysis in euro, the risk-free rate is calculated in a similar way. The risk-free rate used in the analysis in

[1] E.g. Nokia is listed in Finland and Germany; only the returns on the Finnish Stock Exchange are used.

DEM is the monthly return on three-month German treasury bills. The use of two currencies allows us to investigate whether the currency of denomination influences the results. In this paper, we use the DEM because it can be considered to have played the role of anchor currency for the countries that are now part of the Eurozone, but also, e.g. for Switzerland. It moreover implies that the portfolios used to perform the tests are investable, not only for German investors, but also for the investors in the countries whose currencies were linked to the DEM in the setting of the European exchange rate mechanism (ERM)[2]. The implicit assumption underlying the analysis based on returns in synthetic euro is that the portfolios are investable by investors from the Eurozone, which is probably less realistic because it would have required extensive use of hedging techniques in the early stages of the ERM. This is a fictive assumption, because the investor did not know the weights of the currencies in the synthetic euro basket. The synthetic euro basket is composed using GDP weights and hence, these are the weights this investor uses prior to the introduction of the euro.
A final reason to perform the tests for both currencies is that, if the results are found to be comparable for the DEM and the synthetic euro series, this would imply that the conclusions will probably be relevant for future investment decisions in European stocks for an investor in the Eurozone which, from 1999 onwards, are all expressed in euro. Overall, we find no different conclusions using the two currencies of denomination.

We present a number of descriptive statistics related to the return characteristics of the different regroupings of individual stocks. In this chapter, we report all results for the DEM sample. Most of the results for the synthetic euro sample are shown in the appendix and referred to in the text. Table 1 shows the difference in relative weights (in %) for the three types of stock portfolios (country, sector and size). The full names of all portfolios are listed in appendix 1. The weights are reported at three points in time : 1979, 1988 and 1998. The number of country portfolios is 12 because there were fully available samples for 12 countries, while 5[3] out of the 17 countries only had a very small number of stocks in the early years of the sample period and are therefore excluded for reasons of stability. The number of sectors is 14, based on an industry regrouping using Financial Times indices and STOXX regroupings of industries into sectors. Allocation to a specific sector is constant over the 20-year period (assuming that companies do not frequently shift their main business) and determined by the sector affiliation at the end of December 1998[4].

[2] The data cover the period January 1979 – December 1998, from the start of the ERM to the launch of EMU.

[3] These five countries are Spain, Luxembourg, Greece, Finland and Portugal.

[4] A shift in the main business implies that the company will be affiliated in a totally different sector. A shift of some stocks from banking to financial services is not expected to influence the results. The number of shifts between totally different sectors for stocks with a reasonable market capitalization is not that high.

The actual regrouping used in the paper is explained in appendix 2[5]. The number of size portfolios is chosen to be 13. Lehman and Modest [1988] report that the number of dependent portfolios included in the analysis is critical for the evaluation of the power of the tests. Hence, we choose to keep the number of dependent portfolios comparable across regrouping procedures because we want to avoid that this characteristic has an impact on the evaluation of the tests.

Size portfolios are regrouped every month based on the market capitalization at the last trading day of the previous month. As table 1 indicates, by construction, there is a difference in capitalization across the size portfolios, while this is less the case for country and sector portfolios. Within the country portfolios, the UK has the largest capitalization, followed by Germany and France. The largest sector portfolios are banking, insurance and cyclical services. Also by construction, the size portfolios contain an equal number of stocks, while this is not the case for country and sector portfolios. Table 1 reveals that among the country portfolios, the largest changes in relative market capitalization are observed for France, Italy (both upwards) and Germany (downwards). In the sector portfolios, the growing sectors in terms of relative market capitalization are pharmaceuticals, banks, insurance, utilities and especially telecom. A downward trend is observable, e.g., for resources and chemicals, and most pronounced for cyclical consumer goods. The weights of the size portfolios are more stable over time.

Table 2 presents the first two moments of the time series (of 240 monthly returns) for all types of portfolios expressed in DEM. The magnitude of both the expected returns and the standard deviations for the different portfolios is comparable for the two currencies of denomination (see table 2bis in appendix 6 for the euro based statistics). From table 2 it is clear that the dispersion of the returns is more or less similar for the three types of portfolios of European stocks. The dispersion of risk is somewhat larger for country and sector portfolios than for size portfolios. This could indicate that the underlying characteristics of the country and sector portfolios are more diverse than those of the size portfolios.
Of the country portfolios, Italy, Sweden and Norway are the most volatile ones. Also notable is that the country portfolios exhibit a different size-related effect. For some countries (and for the German investor), the market-capitalization weighted return is higher than the equally weighted return (e.g., Ireland, Switzerland) while for most other countries the reverse is true. Return

[5] The sectors are constructed based on a combination of the Financial Times indices and the STOXX indices. The comparable industry groups from both methodologies are treated as a relevant sector. The remaining industries are allocated to their corresponding sector according to the FT typology. For all stocks in the sample, we used the FT industry code to allocate them to one of the 14 resulting sectors. This treatment ensures that the sectors remain comparable over time, even when the index companies periodically alter the index compositions.

Table 1.

Relative weights in the country, sector and size portfolios

The weights for the three types of portfolios are calculated as the percentage of capitalization represented by the portfolio relative to the total sample capitalization. Weights are reported at three different time periods : the first year of the sample (December 1979), the middle of the sample (December 1988) and the end of the sample (December 1998). In panel A country portfolio weights are reported, sector portfolio weights appear in panel B and size portfolio weights in panel C.

Panel A

		AUS	BEL	FRA	GER	IRE	ITA	NET	DEN	NOR	SWE	SWI	UK
Country	12-1979	1.0	2.8	7.7	20.4	0.5	2.0	9.1	0.9	0.7	7.4	9.4	38.2
	12-1988	2.2	3.2	8.3	13.2	0.5	7.3	6.9	1.0	0.6	2.4	6.5	41.6
	12-1998	0.5	3.1	12.5	14.9	0.9	6.6	8.5	1.2	0.7	3.7	10.1	29.3

Panel B

		Reso	Bmat	Chem	Cycc	Ncyc	Phar	Cycs	Bank	Insu	Fina	Indu	Tech	Tele	Util
Sector	12-1979	10.9	4.0	6.4	13.4	9.3	5.2	9.6	11.9	6.1	4.9	9.4	6.0	0.3	2.6
	12-1988	8.8	3.9	5.3	4.2	11.6	5.7	14.3	12.2	9.2	5.0	7.8	5.5	3.7	2.8
	12-1998	6.6	2.6	2.8	3.7	8.7	8.4	11.8	15.8	11.0	3.2	7.7	5.0	8.5	4.4

Panel C

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
Size	12-1979	0.1	0.2	0.4	0.7	1.1	1.6	2.1	2.9	4.2	5.8	8.8	15.3	56.7
	12-1988	0.1	0.4	0.6	0.9	1.3	1.7	2.3	3.2	4.4	6.2	9.2	16.9	52.8
	12-1998	0.1	0.4	0.5	0.8	1.0	1.3	1.8	2.3	3.2	4.6	7.8	15.0	61.1

Table 2.

Average return and risk for European stock portfolios from January 1979 to December 1998.

All average returns and standard deviations are reported in % for the country, sector and size portfolios. The standard deviations are the numbers in italic below the returns. All numbers are calculated for market capitalization (MCAP) and equally weighted portfolios. The values are reported for the Deutschmark as the currency of denomination.

	AUS	*BEL*	*FRA*	*GER*	*IRE*	*ITA*	*NET*	*DEN*	*NOR*	*SWE*	*SWI*	*UK*
MCAP	0.68	1.04	0.97	0.78	1.10	1.03	1.27	0.99	1.01	1.47	1.02	1.14
	4.91	*5.23*	*6.09*	*5.02*	*6.41*	*7.46*	*4.67*	*5.09*	*8.50*	*7.78*	*4.89*	*5.83*
EW	0.87	1.20	1.11	0.73	1.02	1.13	1.19	1.09	1.09	1.45	0.61	1.24
	4.73	*5.01*	*6.07*	*4.56*	*5.85*	*7.61*	*4.64*	*4.92*	*8.04*	*7.71*	*4.23*	*5.96*

	reso	*bmat*	*chem*	*cycc*	*ncyc*	*phar*	*cycs*	*bank*	*insu*	*fina*	*indu*	*tech*	*tele*	*util*
MCAP	1.22	0.53	0.89	1.01	1.19	1.36	0.94	0.88	1.18	0.94	0.80	0.67	1.07	1.00
	6.14	*5.33*	*4.99*	*6.48*	*4.78*	*5.21*	*5.35*	*5.21*	*5.32*	*4.86*	*5.09*	*5.43*	*6.21*	*3.85*
EW	1.00	0.83	1.02	0.84	1.12	1.18	1.16	1.01	1.26	1.16	0.87	0.88	1.20	1.11
	6.17	*5.43*	*5.04*	*4.99*	*4.19*	*4.45*	*4.81*	*4.13*	*4.89*	*5.17*	*4.63*	*5.21*	*5.65*	*3.43*

	S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*	*S11*	*S12*	*S13*
MCAP	1.43	1.30	1.25	1.12	1.04	0.90	0.89	0.82	0.84	0.90	0.92	0.81	1.08
	4.42	*4.37*	*4.50*	*4.35*	*4.09*	*4.59*	*4.63*	*4.69*	*4.47*	*4.56*	*4.70*	*4.74*	*4.73*
EW	1.45	1.31	1.25	1.13	1.03	0.91	0.89	0.83	0.85	0.90	0.90	0.82	1.01
	4.50	*4.33*	*4.49*	*4.35*	*4.68*	*4.59*	*4.63*	*4.69*	*4.47*	*4.56*	*4.73*	*4.74*	*4.72*

volatilities are more comparable across sectors. Again, a mixed size effect can be observed for sector portfolios. Capitalization weighted returns are higher than equally weighted returns for some sectors (resources, cyclical consumer goods, pharmaceuticals) and lower for others. The volatility of returns is comparable across size portfolios.

The incidence of the size effect in European stock returns is ambiguous. If any effect can be detected in the lower panel of table 2, it seems to be restricted to the relatively small stocks. The post-formation returns of the size portfolios decrease from S1 to S8, where the lowest return is recorded for the market-capitalization weighted returns expressed in DEM.

The portfolio of very large stocks (S13, which represents approximately 55% of the total market capitalization), however, exhibits a higher return than the 7 preceding size-ranked portfolios. The evidence in table 2 suggests that the size effect is not strong in this more recent period in European stock markets, which would confirm US studies [see Jegadeesh and Titman, 1999]. We still observe the largest returns for the small stock portfolios. But the relation between returns and size is not monotone, and amongst the large capitalization deciles, the largest decile performs best. Also, remind that the real small stocks are not included because they can be susceptible to infrequent trading among other things (as reported earlier). Moreover, we find that in the last five-year period (1994-1998), the average size effect is a highly insignificant 0.08% (for the DEM sample).

Portfolios as factors

Table 3 reports the statistics for the DEM sample for the portfolios that are used as factors in this analysis. We choose to extend the one-factor model using the high minus low book-to-market portfolio, the small minus big size portfolio, and a fourth factor portfolio based on momentum. Past empirical findings reported that momentum partially explains the cross-section of returns [Chan et al., 1996]. Also, the momentum-effect is the major market anomaly that seems to be persistent through time [Jegadeesh and Titman, 1999] but also moves with the business cycle [Chordia and Shivakumar, 2000].

Table 3 compared to table 3.bis in appendix 6 shows that the market portfolio expressed in DEM has a lower average return (1.133% per month) than the market portfolio in synthetic euro (1.304% per month). The second factor is the BTM factor made popular by Fama and French [1992, 1993, 1996]. The book value of a company is defined as the value of equity capital plus reserves minus total intangibles[6]. The return differential consists of a long position in the 30%

[6] Since we use book-to-market ratios of companies headquartered in different countries, differences in accounting standards could influence the rankings. For that reason, we re-rank all stocks after subtracting the cross-sectional mean BTM ratio of each month and of their home country. We find that the average rank correlation between the original series and the deviations from the country mean is

highest BTM stocks minus a short position in the 30% lowest BTM stocks [HML, as in Fama and French, 1996]. All available stocks in December of year t-1 are ranked according to their BTM value at the last trading day of December t-1. In most studies the subsequent return analysis covers the period July t to June t+1 in order to ensure there is no look-ahead bias. However, we prefer to perform the return analysis for the period January to December of year t because this treatment increases the number of stocks included in the estimations, especially in the beginning of the sample period. A value-weighted monthly HML return is calculated for the 12 months of year t[7]. Starting the first ranking in December 1978 and ending the ranking in December 1997 produces a time series of monthly returns from January 1979 until December 1998, both for the synthetic euro and the DEM sample.

Table 3.

Average return and standard deviation of the factor portfolios and the risk-free rate.

Average returns and standard deviations for the factor portfolios are presented (in %) in panel A. All factor portfolio statistics are reported in Deutschmark (DEM). The global market portfolio is a market-capitalization weighed average of all available stocks. The equity premium is the difference between the average return on the global market portfolio and the risk-free rate (three month German treasury bills). HML is the factor portfolio based on the return differential between the high book-to-market multifactor minimum variance portfolio and the low book-to-market MMV portfolio. SMB is the factor portfolio based on the return differential between the small size MMV portfo lio and the big size MMV portfolio. LMOM is the factor portfolio based on the return differential between the local losers MMV portfolio and the local winners MMV portfolio. Panel B reports the correlations between the factors. * denotes statistical significance of a non-zero return at the 5% level.

Panel A.	**Average return (%)**	**Stdev (%)**
Global market portfolio	1.133*	5.10
Equity Premium	0.637*	4.92
HML	0.197	2.73
SMB	0.716*	2.70
LMOM	-0.642*	2.08
Risk-free	0.496*	0.18

Panel B : correlations	M-F	HML	SMB	LMOM	Risk-free
Equity premium	1				
HML	-0.102	1			
SMB	-0.040	0.148	1		
LMOM	-0.030	0.136	0.238	1	
Risk-free	-0.127	-0.026	-0.201	-0.019	1

0.9001 (with a low of .7779 and a high of .9636), indicating that differences in accounting standards should not have a large impact on the calculation of the BTM portfolios (see also Lewellen, 1999, for sector BTM and appendix 3).

[7] We checked the robustness of the results for this treatment by comparing the estimation results for the common method of calculation (BTM ranking in December t-1 and returns from July t to June t+1) and our approach (BTM ranking in December t-1 and returns from January to December of year t) (see appendix 3). In the 39 DEM regressions for the two-factor model including HML, the alphas were indistinguishable and the average correlation of the residuals of the regressions is 0.9903. (tables are available upon request). This supports the hypothesis that book values are common knowledge among stock analysts. Moreover, it is consistent with the finding by Fama and French (1995) that BTM ratios exhibit a high degree of persistence over time.

The third factor is the small minus big factor based on market capitalization. We applied the exact procedure decribed by Fama and French [1992]. However, we use quintiles instead of deciles. The fourth factor portfolio is based on the individual stock's momentum (Jegadeesh and Titman, 1993 and 1999). The factor LMOM is calculated as the return differential of a long position in the 25% stocks with the lowest six-month trailing return ('losers') minus a short position in the portfolio containing 25% of the stocks with the highest previous six-month performance ('winners') [8]. All stocks with data available from t-6 months to t-1 month are ranked at the end of t-1 [as in Jegadeesh and Titman, 1993, 1999 and Rouwenhorst, 1998]. The parameter used to perform the ranking is the 6-month local momentum, which is the six-month cumulative stock return minus the six month home market return. The momentum portfolio rebalancing is performed on a monthly basis and the return differential is calculated as a difference between two equally weighted portfolios. Table 3 shows the averages and standard deviations for the factors. The return premia associated with the factors is similar for DEM and euro (see also table 3 bis in appendix 6). Since the cross-correlations between the factors are relatively low, they are assumed not to cause any estimation problems (panel B of table 3).

3. The pricing framework

The main focus of this paper is to test the relative efficiency of the following asset pricing kernels in a European setting: an international CAPM, a two-, three- and four-factor ICAPM. To motivate our ICAPM, we assume investors are not concerned about deviations from purchasing power parity across the European Union [as in Fama and French, 1998] although there is some general evidence that exchange rate risk is priced [Dumas and Solnik, 1995]. Moreover, there is empirical evidence that factor loadings on international risk factors may vary through time [Ferson and Harvey, 1993]. In order to restrict the dimensions of the pricing models, we assume the absence of time-variation in both the factor loadings and the risk premiums [as in Fama and French, 1998]. As in Fama and French [1998], the test results suggest that this simple approach provides an acceptable story for the cross-section of returns.

For the three types of portfolios, the performance of the pricing models is evaluated by means of the Gibbons-Ross-Shanken [GRS 1989] multivariate test. A pricing model should be able to model the dynamics of any stock or portfolio return, but the overview in CLM [1997, p.241] indicates that this is not always the case. Hence, we test the accuracy of different pricing models on three kinds of portfolio regroupings.

[8] As in Rouwenhorst (1998) we use quartile instead of decile portfolios in order to ensure that the portfolios contain a sufficient number of stocks, especially in the beginning of the sample period.

3.1 The choice of a multivariate testing framework

In evaluating the efficiency of a factor portfolio, we can apply time-series or cross-sectional regressions, a GMM approach or a maximum likelihood approach. However, the design of these tests and the test statistics are quite similar [Cochrane, 2001]. We will first give an overview of the important determinants that explain the choice of our testing framework.

As was outlined in the introduction, the regression approach involves the estimation of two equations:

[1]
$$\begin{aligned} r_t - r^f &= a + b\, f_t + e_t \\ E(r) - r^f &= b\, l \end{aligned} \quad .$$

In equation 1, r_t is the time-series of returns on an asset, r^f is the risk-free rate of return, a is the regression intercept, b is the fector of factor loadings on the factors f, e_t is the vector of residuals from the time-series regression. $E(r)$ denotes the expected return for the asset, and l denotes the vector of prices of risk for the explanatory factors. If we compare the time-series regression to the expectation of the time-series regression, it is clear that the only implication of the model is that all regression intercepts should be equal to zero. These intercepts are the pricing errors in the test.

In this chapter, we use the regression-based multivariate tests to evaluate the ex ante MVE of the factor portfolio. First, this test has been and is widely used in empirical research tests of this kind. For reasons of comparability this seems a reasonable choice. Second, we both report the results for the asymptotic test distribution and the finite-sample test distribution. Both these results should account for the possible improper characteristics of the data. We explain this statement by giving more detail about both test distributions.

The multivariate test is described below. The test statistic J for the finite-sample Gibbons-Ross-Shanken test (equation [2]) has a central F-distribution with degrees of freedom N and T-N-k under the null hypothesis [see GRS 1989; MacKinlay 1987; Affleck-Graves and Mc Donald 1990 and Campbell et al. 1997]. The test statistic in [2] is a generalization for multifactor models. In this test, T is the number of periods of the time series (here 60 for 5 years of monthly data or 240 for 20 years), N is the number of portfolios and k is the number of independent factors. The test has an F-distribution with a non-centrality parameter (n, equation [3]) that equals zero under the null hypothesis (expression [4]). The asymptotic test using n has a c^2 distribution with N degrees of freedom.

[2]
$$J = \frac{T-N-k}{N}\left[1+\hat{m}_k^{'}\hat{\Omega}_k^{-1}\hat{m}_k\right]^{-1}\hat{a}^{'}\hat{\Sigma}^{-1}\hat{a}$$

[3] $$n = T\left[1+\hat{m}_k^{'}\hat{\Omega}_k^{-1}\hat{m}_k\right]^{-1}\hat{a}^{'}\hat{\Sigma}^{-1}\hat{a}$$

[4] finite-sample test: $J \sim F_{N,\,T-N-k}(n)$

[5] asymptotic test: $n \sim c_N^2$

The components of the test-statistic derived from the pricing equations are : $\hat{a}$ the $(N\times 1)$ vector of asset return intercepts, $\hat{\Sigma}$ the $(N\times N)$ variance-covariance matrix of disturbances, $\hat{m}_k$ the $(k\times 1)$ vector of means of the factor portfolios and $\hat{\Omega}_k$ the variance-covariance matrix of factor portfolio returns.

GRS (1989) provide a geometric interpretation of the test as expressed in equation [6]. The interpretation is that the test statistic J evaluates exact factor pricing by comparing the squared Sharpe ratio of the portfolio of risk factors (sr_m) with that of the tangency portfolio (sr_t). The alternative hypothesis assumes that a linear combination of portfolios is not the tangency portfolio with the highest Sharpe ratio. In equation [6] we use the Sharpe ratio of the global market portfolio. It is important to note that J is an increasing function of the difference between the squared Sharpe ratios of the tangency portfolio and the portfolio of factors.

[6] $$J = \frac{T-N-k}{N}\left[\frac{sr^2 - sr_m^2}{1+sr_m^2}\right]$$

Most important is that the test distributions have the following different characteristics:

the GRS F-test:

- recognizes the sampling variation in $\hat{\Sigma}$
- assumes that the residuals are normal as well as uncorrelated across assets and homoscedastic

The asymptotic c^2 test:

- assumes that Σ and Ω have converged to their probability limits
- the test is asymptotically valid even though the residuals are estimated and it does not require e rrors to be normal, relying on the central limit theorem.

Hence, reporting both tests allows us to draw fairly robust conclusions and at the same time allows us to compare results to results from past research. Gibbons et al. [1989] argue that the normality assumption is a good working approximation for monthly stock returns. However, they do notice that deviations from normality can be observed. With respect to this subject, MacKinlay [1985] finds that the F-test is fairly robust to misspecifications of the distribution of asset returns.

A next issue that is important in these multivariate tests is raised by Roll [1977, 1977]. He mentions that regression tests probably have low power. Grouping the assets may may even further reduce the power. Affleck-Graves and McDonald tested factor models, by allowing the number of assets (*N*) to be larger than the length of the time-series (*T*). They find that an alternative statistic for a large number for *N* does worse than the common portfolio tests. Evaluating the power of the test, Gibbons et al. [1989] suggest that using *N* being one third of *T* provides a proper choice of *N* and *T*. A commonly made choice of *N* between 10 and 20 and *T* equal to 60 seems reasonable. Finally, we are concerned about the estimates of the covariance matrix in the tests. Therefore, we also use a *T* of 240. Summarizing, we perform multivariate tests for the efficiency of the factor portfolio using an asymptotic test and a finite-sample test using *N* between 10 and 20 and *T* equal to 60 or 240.

3.2 The ICAPM

When testing the international CAPM, the stock returns are explained by their exposure to a global market portfolio. The global market portfolio is constructed as the market-capitalization weighted portfolio of all stocks in the sample with a continuous listing during the month in which the return is calculated. We estimate [using OLS for this international dataset as in Fama and French, 1998] the sensitivity of the excess portfolio return ($r - r^f$) to the excess return of the market index ($r^m - r^f$), as expressed by equation [7].

[7] $$r - r^f = a + b(r^m - r^f) + e$$

To test whether this pricing equation accurately describes the cross-section of the returns of the three types of portfolios, we test whether the vector of α's is multivariate zero.

Estimation results for the one-factor model

Table 4 presents the estimation results for the one-factor pricing model [7] expressed in DEM[9]. In order to determine ex ante efficiency, we are interested in the behavior of the vector of estimated constants in the regression. For the one-factor model applied to country portfolios, table 4 shows that the alphas are small and always insignificant. For sector portfolios, the alphas are small, although significant for the basic materials and technology stocks. In the case of size

[9] The results for the synthetic euro are almost identical (available upon request).

portfolios we notice that there is a pattern both in the size and the significance of the estimated constants. The smallest size portfolios (S1, S2 and S3) exhibit a positive and significant alpha. The one but last portfolio (S12) has a significantly negative alpha. The pattern in the alphas coincides with the previously mentioned size effect (see table 2) and with the findings for the U.S. data for a different period as in GRS [1989].

Table 4.

Estimation results for the single-factor pricing equation.

The factor loadings are reported for the single-factor pricing equation in DEM. The t-values of the factor loading are in parentheses. a is the estimated constant, β is the factor loading on the excess return of the global market portfolio.

DEM	$r - r^f = a + b(r^m - r^f) + e$							
	α	β		α	β		α	β
AUS	-.002	0.580	Reso	.001	0.930	S1	.005	0.687
	(-0.74)	(11.62)		(0.53)	(18.71)		(2.75)	(19.75)
BEL	.001	0.656	Bmat	-.005	0.915	S2	.004	0.711
	(0.47)	(12.77)		(-3.25)	(27.96)		(2.18)	(22.68)
FRA	-.001	0.818	Chem	-.001	0.848	S3	.003	0.762
	(-0.16)	(14.51)		(-0.89)	(26.85)		(1.81)	(26.10)
GER	-.001	0.648	Cycc	-.001	0.927	S4	.002	0.739
	(-0.51)	(13.53)		(-0.28)	(16.57)		(1.08)	(26.59)
IRE	.001	0.866	Ncyc	.001	0.877	S5	.000	0.833
	(0.16)	(14.61)		(1.26)	(40.76)		(0.06)	(32.80)
ITA	.000	0.816	Phar	.003	0.826	S6	-.001	0.816
	(0.02)	(10.40)		(1.70)	(21.26)		(-0.88)	(33.14)
NET	.003	0.755	Cycs	-.002	0.982	S7	-.001	0.829
	(1.69)	(22.38)		(-1.53)	(41.42)		(-1.10)	(34.33)
DEN	.001	0.562	Bank	-.002	0.856	S8	-.002	0.842
	(0.51)	(10.51)		(-0.89)	(23.70)		(-1.69)	(34.85)
NOR	-.001	0.985	Insu	.001	0.877	S9	-.002	0.809
	(-0.24)	(11.31)		(0.66)	(23.91)		(-1.50)	(36.66)
SWE	.004	0.982	Fina	-.001	0.901	S10	-.001	0.827
	(0.89)	(13.03)		(-1.27)	(43.93)		(-1.05)	(37.58)
SWI	.001	0.681	Indu	-.003	0.886	S11	-.001	0.874
	(0.39)	(15.54)		(-1.69)	(29.82)		(-1.35)	(45.71)
UK	-.000	1.074	Tech	-.004	0.948	S12	-.002	0.889
	(-0.33)	(43.03)		(-2.63)	(30.13)		(-2.72)	(49.88)
			Tele	.000	0.829	S13	.000	0.902
				(0.14)	(14.35)		(0.15)	(63.02)
			Util	.002	0.553			
				(0.88)	(16.45)			

Since the estimated alphas cannot directly be used to assess the ex ante efficiency abilities of the tested pricing equations in a multivariate setting, we report the GRS J-statistics. All estimations and calculations are done for the entire period (20 years of monthly data) as well as for four five-year sub-periods. Previous empirical work has documented that asset-pricing models behave differently in varying monetary, and hence interest rate, regimes [Jensen et al, 1996]. Consequently, we perform all tests for the entire sample period and the 4 sub-periods of five years. We also compare the estimates of the test statistics for the two last periods of five years, where the interest rate regime shifted a lot. In the period 1989-1993, we observe a period of distinctive rising interest rates and in the period 1994-1998 we observe decreasing interest rates in the run-up to EMU. In each sub-period T=60 months. A graphical justification for the choice of the sub-periods where there is a clear interest rate regime, is given in appendix 5 where both the short-term interest rate for the DEM and the synthetic euro are displayed.

Table 5 reports the results of the ex ante efficiency tests for the country, sector and size portfolios and both currencies of denomination. The GRS-statistic n is the estimated non-centrality parameter for the one-factor market model and is evaluated by the asymptotic test using a c_N^2 critical value (at the 5% level). The F-statistic and its associated p-value are calculated from equations 2 and 4. We also report the average absolute alphas for each set of dependent portfolios in order to compare them for the different APMs and between the sets of dependent portfolios. Factor portfolios that are found to be ex ante mean-variance efficient for one set of assets are indicated by a shaded are for the relevant test.

For country portfolios the ICAPM appears to provide an accurate description of the pricing dynamics. The non-centrality parameter for country portfolio is not statistically different from zero in the full sample and the sub-periods, irrespective of the currency of denomination for both the asymptotic test and the finite-sample test (see also table 5bis in appendix 6). All p-values for the F-test are much larger than 5% confirming the choice of the ICAPM. This implies that the null hypothesis of a multivariate zero alpha vector cannot be rejected and that a one-factor European market model captures the pricing of country stock portfolios. The calculated confidence interval of the estimated alphas for country portfolios ranges from -29 basis points to 39 basis points. The only caveat is that alphas are still found to be high in absolute values for the sub-periods (in the period 1979-1983, the average absolute alpha for country portfolios is 0.61%). Apparently, the values in the inverted residual covariance matrix are low, indicating that diversification across country portfolio is beneficial.
By contrast, for sector portfolios the null hypothesis of a zero α-vector is rejected for the full sample period (p-values for the F-test of 0.032). This would imply a rejection of the ICAPM as a relevant model for the pricing of European sector portfolios. However, when the sub-periods are considered, the p-values are generally higher but do not allow strong inferences about ex ante efficiency because the asymptotic test reveals ex ante efficiency in only one of the sub-periods. For all the sub-periods, the p-values are lowest in the period of falling interest rates (around 10%) and highest in the period of rising interest rates (maximum value of 0.355).

Table 5.

Results for the multivariate GRS test of ex ante efficiency using a one-factor model

For the five different time periods and the DEM, the test values are reported. The first period is the full sample period from January 1979 until December 1998. n is the estimated non-centrality parameter from the one-factor model and is evaluated by a c^2 critical value. For country portfolios, c^2_{12} is 21.03, for sector portfolios c^2_{14} is 23.69 and for size portfolios, c^2_{13} is 22.36. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. N is the number of portfolios and T is the number of observations. The shaded areas indicate ex ante efficiency of the factor portfolio using the asymptotic test for n and using the GRS-test reporting the p-value (both at the 5% level).

$$r - r^f = a + b(r^m - r^f) + e$$

	COUNTRY	SECTOR	SIZE
79:01-98 :12			
average $\|a\|$	.0013	.0021	.0019
n	6.73	27.55	45.73
F-statistic	0.535	1.861	3.341
p-value	0.891	0.032	0.000
79 :01-83 :12			
average $\|a\|$	.0061	.0043	.0024
n	14.76	26.35	25.66
F-statistic	0.997	1.461	1.566
p-value	0.466	0.166	0.131
84:01-88 :12			
average $\|a\|$	.0039	.0024	.0048
n	11.638	24.82	27.57
F-statistic	0.786	1.376	1.682
p-value	0.662	0.205	0.097
89:01-93:12			
average $\|a\|$	.0032	.0033	.0025
n	13.675	20.50	28.02
F-statistic	0.923	1.136	1.709
p-value	0.532	0.355	0.091
94:01-98:12			
average $\|a\|$	.0026	.0034	.0031
n	14.160	29.29	38.462
F-statistic	0.956	1.623	2.347
p-value	0.502	0.110	0.017

The strongest evidence against ex ante efficiency based on the ICAPM is found for the size portfolios. The p-values are below 5% for the full sample period and for the sub-period of falling interest rates. In all other periods, p-values are around 0.10 except for the most recent period. the asymptotic test never indicates ex ante efficiency of the factor portfolio. For both the sector and size portfolio, the confidence interval for the α-vector ranges from -50 to 50 basis points. The findings for the size portfolios are in accordance with the GRS findings on size portfolios for the U.S [as in Gibbons et al. 1989]. It is remarkable however, that in two of the four sub-periods, average absolute alphas are the lowest for the three groups of dependent portfolios. Using the same argument as before, the weights of the alphas from the inverted residual matrix are quite high, indicating that additional risk factors are required.

3.3 Multifactor models

The ICAPM can be augmented by assuming that the fraction of the portfolio returns which is not captured by the global market portfolio is priced by additional global multifactor minimum-variance (MMV) portfolios [Fama and French, 1996]. Similarly, the generalized CAPM initiated by Merton [1973] suggests that investors are concerned about state variable risk next to the mean and variance of their portfolio returns. Following Fama and French [1996 and 1998], the return differentials on two MMV portfolios are added to equation [1] in order to explain the expected portfolio returns. The construction of the long-short strategy of return differences for the three additional factors (on book-to-market, size and local momentum) was described in section 2. As in Fama and French [1996], we assume that the low BTM, the high BTM, the small size, the big size, the local losers and the local winners portfolios are MMV. A combination of the one-factor market model and the three additional mimicking factor portfolios leads to the construction of 7 additional models described in equations 8, 9 and 10. The testable hypothesis is that the vectors of estimated intercepts is zero. The correlations between the three factor portfolios (as shown in table 3) is relatively low, hence we expect that their combination causes no particular estimation problems.

[8] $$r - r^f = a + b(r^m - r^f) + g f_2 + e$$

[9] $$r - r^f = a + b(r^m - r^f) + g f_2 + d f_3 + e$$

[10] $$r - r^f = a + b(r^m - r^f) + g f_2 + d f_3 + h f_4 + e$$

The univariate statistics of the augmented equations are shown in tables 4 a, b and c in appendix 6. For country portfolios (table 4a), the findings for the vector of alphas are very similar to those based on equation [7] (left panel). None of the estimated country alphas is significant at conventional levels. An interesting observation is that the HML factor and the SMB factor are univariate significant for almost all countries, while the momentum factor is generally insignificant. However, none of the individual factors induce an increase of the R^2 in the two-factor models.

For sector portfolios (table 4b), adding the non-market factors does not change the univariate interpretation of the alphas, they are only significant for the basic materials and technology stocks. Again we observe that the HML factor explains some of the variance of the individual portfolio returns (except for cyclical consumer goods, cyclical services and telecom). Only using SMB as an additional factor, alphas are significant for more sectors. Again, there is no increase in the R^2 for the two-factor models.

For size portfolios, the inclusion of the two additional factors has an ambiguous effect on the univariate significance of the alphas. When the HML factor is included, next to the market portfolio, only the α of S1 and some of the large-size portfolios are significant. With the two-factor model, using SMB as a second factor, the structure of alphas changes. The alphas for the small stock portfolios are no longer significant, but alphas are significant for the S5 to S12 portfolios. When the LMOM factor is added, on the other hand, the pattern of significance observed in the results for the market model (left panel) is preserved (α is significant for the smallest size portfolios and S12). The R^2 increase reasonably for the 3 portfolios with the smallest stocks. Interestingly, the local momentum variable seems to explain more of the variance of size portfolio returns than for the other portfolio regroupings. For the Fama-French three-factor model, 14 out of 27 alphas are still significant for size and sector portfolios.

However, one has to bear in mind that the univariate interpretation of the test statistics does not provide information about of the changes in the estimated variance-covariance matrix of residuals when factors are added. This shortcoming is remedied by using the multivariate test statistics.
Finally, contrary to earlier reported findings by Fama and French [1992], and also contrary to the findings of GRS [1989] (both for U.S. data) we find that betas for European size portfolios are low for the portfolios of small stocks and vice versa.

The results of the multivariate ex ante efficiency tests for the extended models are reported in table 6. We highlight four main findings from table 6. First, although it is difficult to compare models based on p -values for this multivariate test, it is interesting to see that the p-value for country portfolios increases the most for the two-factor APM including the local momentum factor. Next, average absolute alphas are the lowest for this factor model compared to the other multi-factor models for country portfolios. Third, augmenting the model to a four-factor model sharply decreases the p-value for country portfolios (from 0.891 for the one-factor model to 0.335). Finally, the average absolute alphas more than double using this four-factor model for country portfolios (from 0.0013 to 0.0029). These conclusions are made in the margin, because we already found that there is no evidence against mean-variance efficiency for the market portfolio for country portfolios.
More important are the findings for sector and size portfolios. For the whole period, we found evidence against the ex ante efficiency using the single-factor model. First, for sector portfolios, we find no evidence against ex ante efficiency for the two-factor model with the market portfolio and the local

Table 6.

Test results for the multivariate GRS test of ex ante efficiency using the DEM as the currency of denomination.

For the entire time period (1979-1998) the test values are reported for 7 augmented models. Av$|a|$ is the average absolute value of the alphas, n is the estimated non-centrality parameter from the one-factor model and is evaluated by a c^2 critical value. For country portfolios, this critical value is 21.03, for sector portfolios 23.69 and for size portfolios 22.36. The F-statistic (F-stat) is the GRS statistic and the p-value is the associated probability of the F-test. N is the number of portfolios and T is the number of observations. The shaded areas indicate ex ante efficiency of the factor portfolio using the asymptotic test for n and using the GRS-test reporting the p-value (both at the 5% level).

$r - r^f = a + b(r^m - r^f) + g f_2 + e$									
	F2=HML			*F2=SMB*			*F3=LMOM*		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av.$\|a\|$	.0015	.0022	.0020	.0020	.0029	.0021	.0014	0.0017	.0018
n	6.717	36.29	49.40	10.34	46.24	33.20	5.64	20.57	40.19
F-stat	0.532	2.440	3.593	0.818	3.109	2.415	0.446	1.383	2.292
p-value	0.893	0.003	0.000	0.632	0.000	0.005	0.942	0.163	0.001
$r - r^f = a + b(r^m - r^f) + g f_2 + d f_3 + e$									
	F2=HML, F3=SMB			*F2=HML, F3=LMOM*			*F2=SMB, F3=LMOM*		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av.$\|a\|$	.0024	.0029	.0024	.0016	.0021	.0018	.0023	.0026	.0017
n	11.14	54.09	38.94	7.58	30.43	41.19	10.28	36.21	23.46
F-stat	0.878	3.620	2.819	0.597	2.037	2.982	0.810	2.423	1.699
p-value	0.570	0.000	0.001	0.843	0.016	0.001	0.640	0.004	0.062
$r - r^f = a + b(r^m - r^f) + g f_2 + d f_3 + h f_4 + e$									
	F2=HML, F3=SMB, F4=LMOM								
	Country	*Sector*	*Size*						
av.$\|a\|$	.0029	.0028	.0021						
n	14.44	46.08	27.57						
F-stat	1.133	3.070	1.987						
p-value	0.335	0.000	0.023						
N	12	14	13	12	14	13	12	14	13

momentum factor (with a p-value of 0.163) for both the asymptotic test and the GRS-test. Average absolute alphas are lowest for sector portfolios using this model (0.0017). No other factor portfolio is found to be ax ante efficient.

For size portfolios, we find no evidence against the ex ante efficiency only in the case of a three-factor model, augmenting the ICAPM by the SMB factor and the local momentum factor (p-value of 0.062) using the GRS-test. For the asymptotic test, the test value is just above the critical value ($n = 23.46$, $c^2_{13} = 22.36$). Again, absolute average alphas for this APM are the lowest for all APMs for size portfolios. Also notice that we find evidence against ex ante efficiency for the Fama and French three factor model both for sector and size portfolios.

The explanation of the cross-sectional returns of the sector and size portfolios varies across the sub-periods. Table 6 a,b,c and d in appendix 6 show that the p-values for the multifactor models applied to the sector portfolios over the four sub-periods are on average higher than the corresponding p-values from table 5. The most pronounced effect for sector portfolios is found for the two-factor model with momentum as the additional factor. In three of the four periods, the p-value for this two-factor model is far above the ones for the other APMs. Especially in the period of increasing interest rates, this model performs well (p-value=0.784). For the sector portfolios, the Fama-French three-factor model adds little or no value to the single-factor model and has on average low p-values. This finding confirms that the two-factor model including the momentum factor performs well for sector portfolios in European stock markets for the whole period and in the sub-periods. Adding the SMB factor or the HML factor does not seem to improve the APM.

For the size portfolios, the performance of the multifactor models is reasonable in the period 1984-1988 (where interest rates were less volatile), with p-values in table 6b being consistently higher than in table 5 for the single-factor model. However, the augmented models perform poorly in the sub-period of declining interest rates (1994-1998). The one exception is the three-factor model including the momentum factor and the SMB factor. This finding confirms the conclusion for size portfolios from the analysis for the entire period. Also in the other sub-periods, the inclusion of the momentum factor or the size factor or both factors yields the best results.

Hence, our empirical exercise reveals that looking at the univariate test statistics does not always give sufficient information to assess the degree of mis-specification of the model in a cross-sectional framework, which was already reported by GRS [1989]. For example, for sector portfolios, the HML factor seems to provide additional explanatory power for the return series (based on univariate t-statistics and R^2 in table 4b), but it turns out to be the local momentum factor that reduces the level of cross-sectional misspecification.

Based on the reported findings, we conclude that the returns on European country portfolios are fairly accurately described by their sensitivity to a broad market portfolio. Additional factors add little value. This indicates that European stock markets were not segmented over the period studied. A somewhat surprising indication is that the widely used BTM factor does not seem to improve the ex ante efficiency of the return structure of European portfolios in general.

The inclusion of additional MMV factors causes an upward shift in the p -values away from rejection only in some cases for sector or size portfolios. Moreover, we find that the momentum factor seems to be more important for ex ante efficiency in general than the book-to-market factor, which is usually found to add explanatory power based on the univariate tests. The results do not allow us to conclude that the pricing dynamics of European stock portfolios systematically differ across the different monetary policy regimes. However, the finding that in the full period the null hypothesis of an overall zero alpha-vector can be strongly rejected for some specifications, while this is not the case in various sub-periods, could indicate that the factor loadings shift over time. Under this interpretation, factor loadings or risk premiums could be time varying. The best example is that for size portfolios, ex ante efficiency is not rejected in the period of relatively stable interest rates and less so in the other periods.

For the sector portfolios, we find that with an asset-pricing model that includes the market portfolio and the momentum factor, we cannot reject the ex ante efficiency. For European size portfolios, there is no evidence against ex ante efficiency for a three-factor model including the size-factor and the momentum factor.

4. A risk-based alternative and the power of the tests

Two more extensions to this research have to be made to evaluate the robustness of the findings. First, we evaluate the possibility that the European market portfolio is misspecified. We test an alternative model including the labor portfolio, suggested by Jagannathan and Wang [1996] as the alternative static APM. Second, we evaluate the most important conclusions we have drawn from the previous analysis by evaluating the power of the tests against specifications of alternative APMs. MacKinlay [1987] reports that the multivariate tests sometimes lack power. MacKinlay also reports that the power increases with the length of the time period used to perform the test. The reason is that, at a given significance level, the accuracy of the estimation of the residual covariance matrix increases. The alternative market model suggested above is used to test the power of the previous tests against a risk-based alternative. Also, we evaluate our conclusions against a nonrisk-based alternative based on the assumption that there is a possible data-snooping bias or that market irregularities exist.

4.1 A risk-based alternative

One possible alternative hypothesis we consider is that the market portfolio is an incomplete proxy for the wealth portfolio [see Roll 1977]. In most of the asset pricing literature, an equally weighted or market-capitalization weighted sum of returns of all shares is used as the market proxy. However, Jagannathan and Wang [1996] argue that stocks are only a minor (although growing) part of the national wealth and, hence, stock index returns would only proxy for total wealth if the correlation with the total wealth portfolio were perfect. It is indeed an assumption of the CAPM that all assets including human capital are marketable. Consequently, in a first test of the power of ex ante efficiency, we use a different specification for the market proxy as the alternative hypothesis. We refer to this test as a risk-based test.

We augment the ICAPM by a second factor that we denote as labor income. The wealth portfolio assumed here is a linear combination of labor income and capital income (equation 9). As Jagannathan and Wang [1996] describe for the U.S., the part of total income provided by capital income is increasing but still dominated by labor income. The same is observed in Europe. The part of total income generated by capital income increased from 32% in 1979 to 38% in 1998. This shows that unless there is an increase in capital income, labor income is still dominating.

[11] $$r^m = c + w_{VW} r^{VW} + w_{LI} r^{LI} \quad ,$$

where r^m is the return on the wealth portfolio, c, w_{VW}, w_{LI} are assumed to be constants, r^{VW} is the return on the market capitalization weighted stock portfolio, r^{LI} is a proxy for the return on human capital (in this case approximated by the growth rate in per capita labor income).

We calculate the return on human capital as suggested by Jagannathan and Wang [1996]. We use the monthly change in the European Union manufacturing hourly wage earnings index to proxy for the return on human capital. The estimated APM takes the following form:

[12] $$r - r^f = a + b\,(r^{VW} - r^f) + g\,r^{LI} \ .$$

The correlation between the value-weighted stock portfolio and the time-series of returns on human capital is –0.030.

Table 7 shows the multivariate tests for this alternative APM for the three types of portfolios and for the whole period. For country and size portfolios, there is no change in the evidence. But for sector portfolios, we find that including the return on human capital means that we can no longer reject the static CAPM (with a p-value of 0.302). This is an interesting finding and makes the evaluation of the previously made conclusions worthwhile in terms of the power of the multivariate test when we use this risk-based alternative.

Table 7.

Results for the multivariate GRS test of ex ante efficiency for a risk-based alternative APM.

Av $|a|$ is the average absolute value of the alphas, n is the estimated non-centrality parameter from the multifactor models and is evaluated by a c^2 critical value. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. The number of observations is T (=240). The shaded area indicates ex ante efficiency of the factor portfolio using the asymptotic test for n and using the GRS-test reporting the p-value.

	Country	Sector	Size
av. $\|a\|$	.0024	.0024	.0026
n	12.02	17.44	54.71
F-stat	0.947	1.167	3.961
p-value	0.501	0.302	0.000

4.2 A Nonrisk-based alternative

The distribution of the nonrisk-based alternatives is specified by the elements of expression [2], this is $\hat{a}$, $\hat{\Sigma}$ and $\hat{m}_k^{'}\hat{\Omega}^{-1}\hat{m}_k$. We use the sample variance-covariance matrix $(\hat{\Sigma})$ from equations [7], [8], [9] or [10] as well as the squared Sharpe ratios, $\hat{m}_k^{'}\hat{\Omega}^{-1}\hat{m}_k$. Several other studies (mentioned in CLM, 1997) assume $m_k^{'}\Omega^{-1}m_k$ to be zero. In the case of multifactor models we find it more appropriate to use the estimated value of $m_k^{'}\Omega^{-1}m_k$. In the nonrisk-based test we specify values for the intercepts. The assumption that the vector α is normally distributed remains. As in CLM [1997], we take the value for the cross-sectional standard deviation to be 0.002 seeming – given the findings from the univariate regressions – to be a reasonable number for European data. A value of 20 basis points is consistent with possible spreads that can arise from data-snooping [Campbell et al., 1997]. Moreover, 95% of the deviations will be situated between -.004 and +.004 which is close to the estimated values reported in table 4. For each assessment of the power under the assumption of a nonrisk-based alternative, we randomly draw 100 vectors of N alphas (as in CLM, 1997, MacKinlay, 1987, uses 200 drawings) from the specified distribution and use the average power of the simulated alternatives under the non-central F-distributions as a measure of the power of this test. So each vector of alphas is drawn from a normal distribution with mean 0 and standard deviation 0.002 and are hence drawn from a normal distribution with parameters $(\underline{0}, s^2\underline{I})$. For each drawing of alphas, a non-centrality parameter is calculated using the above-mentioned specifications for Σ and $m_k^{'}\Omega^{-1}m_k$. The mean of the 100 calculated non-centrality parameters is used as the specification of the non-centrality parameter of the nonrisk-based alternative.

4.3 A power-based evaluation of the conclusions

In table 8, we present the results for the power tests with respect to the most important evidence we found in the previous sections. It is impossible to report all the power evaluations for a nonrisk-based and a risk-based alternative because the set of tested APMs in this paper is large. The nonrisk-based alternative is the one as described in the previous section and the risk-based alternative is estimated using the different specification of the wealth portfolio. From table 8, we learn that in line with past literature, the power is higher for a nonrisk-based alternative formulation of the APM than for a risk-based alternative.

Table 8.

Power statistics

In table 8, we report the power statistics for the multivariate tests. n is the non-centrality parameter, the power is both presented for the 5% and the 1% level. The left column gives the factors included in each APM. FF denotes the three-factor APM from Fama and French. The second column indicates whether the evaluation is risk-based (RB) or nonrisk-based (NR).

		n	0.05	0.01
		Country portfolios		
M-F	R	12.02	0.577	0.332
	NR	14.31	0.673	0.429
M-F, LMOM	R	10.67	0.515	0.276
	NR	12.92	0.617	0.370
FF	R	19.41	0.832	0.632
		Sector portfolios		
M-F	R	17.44	0.747	0.514
	NR	34.03	0.980	0.924
M-F, LMOM	R	13.74	0.632	0.384
	NR	35.27	0.984	0.936
FF	R	37.68	0.990	0.955
		Size portfolios		
M-F	R	54.71	0.999	0.998
	NR	128.00	1.000	1.000
M-F, LMOM,SMB	R	20.88	0.851	0.661
	NR	143.00	1.000	1.000
FF	R	42.59	0.997	0.983

Overall, the power is high (compared to the MacKinlay [1987] findings for 240 observations), indicating that the evidence we find in the multivariate testing is powerful. Power is however lowest for the country portfolios. Notice that the risk-based power evaluations for the Fama-French model indicate a high power for the finding that there is evidence against a linear combination of the Fama-French model that is mean-variance efficient for European stock markets, for all sets of independent portfolios.

Summarizing, there is reasonable power for the multivariate test for ex ante efficiency of the factor portfolio using the CAPM for European country portfolios. The power of rejecting the null hypothesis given the alternative is even higher for the evidence that the two-factor model using the market portfolio and the momentum factor for the pricing of European sector portfolios. Finally, the evidence in favor of a three-factor model for European size portfolios is also powerful. The conclusions that are drawn from the multivariate tests are hence robust for a possible misspecification of the APMs.

Table 9.
Power statistics for the risk-based alternatives based on missing risk factors.

For the four time periods and results for the DEM sample, the power statistics (at the 5% level) are given for the three-factor models as the alternative APMs to the ICAPMs. For the risk-based alternative, the non-centrality parameter n is reported as well as the power of the test that the specified null hypothesis is not accepted given the alternative. df_1 and df_2 are the degrees of freedom of the F-test under the null hypothesis.

		n	power	df_2
79:01-98:12	Countries			
df_1=12	M-F,HML,SMB	11.14	0.537	227
	M-F,HML,LMOM	7.58	0.362	
	M-F,SMB,LMOM	10.28	0.496	
79:01-98:12	Sectors			
df_1=14	M-F,HML,SMB	54.09	1.000	225
	M-F,HML,LMOM	30.43	0.963	
	M-F,SMB,LMOM	36.21	0.987	
79:01-98:12	Size			
df_1=13	M-F,HML,SMB	38.94	0.993	226
	M-F,HML,LMOM	41.19	0.996	
	M-F,SMB,LMOM	23.46	0.899	

We finally test the power of the extensions of the ICAPM. We evaluate the power for extensions of the ICAPM with respect to three-factor models. In this case, the null hypothesis that a one-factor model is the right APM is evaluated against the risk-based alternative that there are more factors required to find ex ante efficiency of the factor portfolio. These three-factor models are used as a risk-based alternative because, as we previously mentioned, three-factor models are widely used in finance. Moreover, past literature has documented that three factors are sufficient to efficiently price the payoff matrix.

As in the previous power evaluations, we find in table 9 that he power is high, but lowest for country portfolios. The major inference from this analysis is that multivariate tests of suggested extensions of the ICAPM are powerful. Thelevel of these power statistics is high compared to past literature.

5. Conclusions

This paper evaluates the hypothesis of ex ante mean-variance efficiency of the factor portfolio for single-factor and multifactor asset pricing models for a large sample of European stocks in a multivariate setting. The analysis is explicitly designed to avoid to a maximum degree the potentially disturbing influences of data-snooping, survival bias and model misspecification. The coverage of the European dataset is much broader than in previous studies. Particular attention is devoted to the testing framework, the specification of the alternative hypothesis and the power of the tests.

A first test is conducted with respect to the deviation from ex ante efficiency of the factor portfolio using a simple ICAPM. For country portfolios, the results indicate that the cross-section of returns is accurately described by the one-factor market-model. Sector and size portfolios seem to be priced by more than one factor. The results of this paper indicate that the factor portfolio based on the momentum effect performs better as an additional risk factor for European stock data than the factor portfolio based on the book-to-market effect. Although the univariate statistics are in favor of the book-to-market mimicking portfolio, evidence against ex ante efficiency is lower when the momentum portfolio is included to form a linear combination of portfolios that is mean-variance efficient. This implies a caveat of the use of risk factors that seem to be priced in a univariate analysis. Since the results are similar for the specifications in DEM and in synthetic euro, European investors can rely on these insights to guide their investment decisions in the Eurozone. As an important result from this analysis, we find no evidence against the ex ante efficiency using the market portfolio and the momentum factor portfolio to describe the cross-section of sector returns. Also, we find no evidence against efficient pricing by a three-factor model for size portfolios. These three factors are the market portfolio, the size factor portfolio and the momentum factor portfolio.

These findings are reenforced by conducting the same tests on four sub-periods. We observe different interest rate regimes for these four periods. For sector and size portfolios, however, these results indicate that the assumption of non time-varying factor loadings and risk premia is weak. Several tests revealed that some set of dependent portfolios are better priced in periods of different interest rate regimes than in other. This implies an important caveat for the use of these factor models. However, the finding of a two-factor model for sector portfolios and a three-factor model for size portfolios seems to be robust across time.

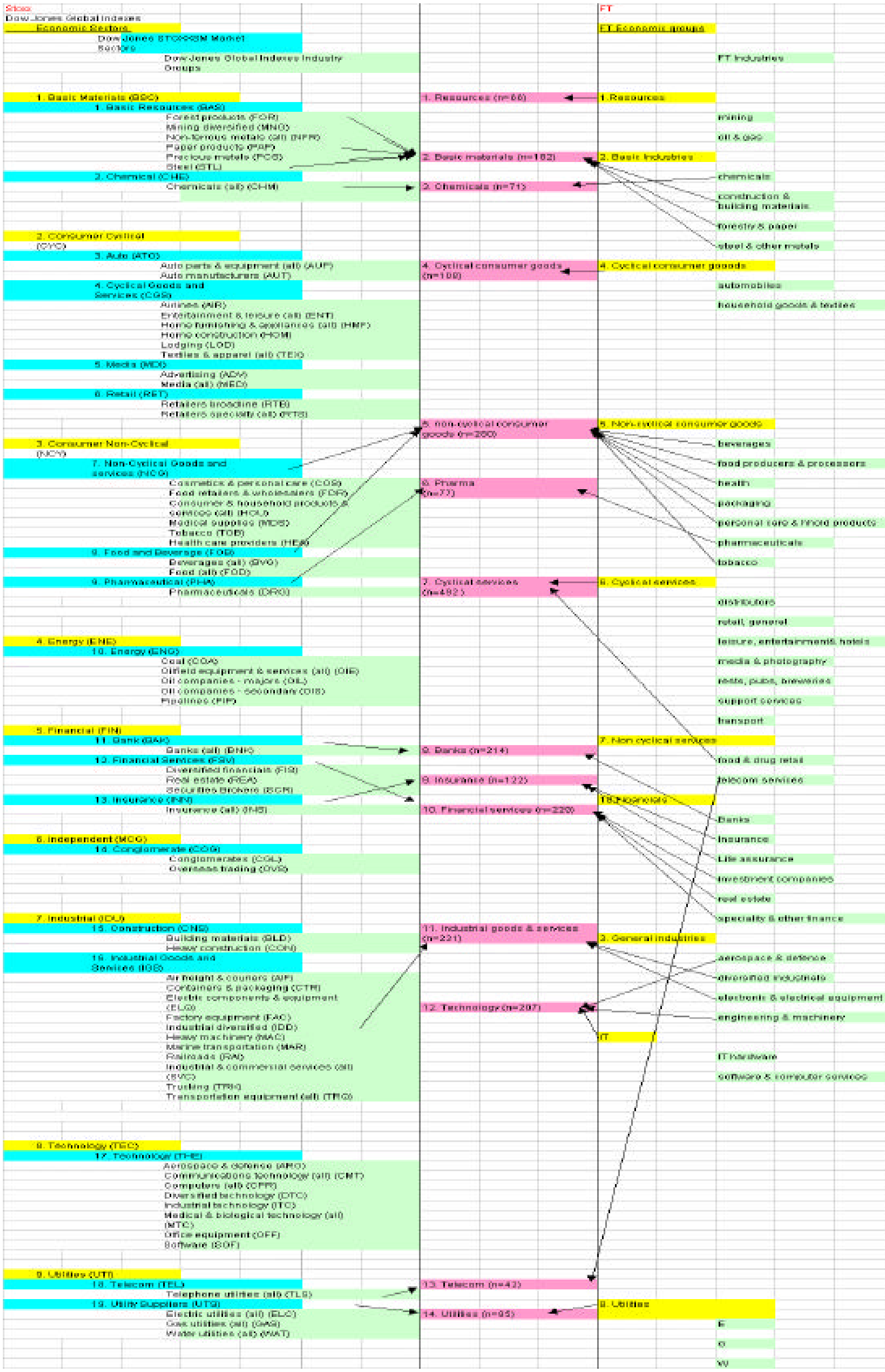
In a test of the existence of a risk-based alternative, we evaluated whether the ICAPM is rejected for size and sector portfolios because of a misspecification of the market portfolio of returns and does not fully correlate with the wealth portfolio. Therefore, we included the return on human capital as an additional explanatory factor. We find that for sector portfolios, this indeed is the case. We fail to reject the hypothesis that the static ICAPM holds for sector portfolios when human capital is accounted for. For size portolios this is not the case, meaning that the APM using a broader specification of the wealth portfolio is still misspecified.

Finally, power evaluations are important in the assessment of ex ante efficiency. In fact, the formulation of an alternative hypothesis is a test to review which portfolios should be combined to be mean-variance efficient. Testing the power using this alternative implies an evaluation of the strength of the conclusions based on the multivariate statistic relative to the misspecification of the set of risk factors. Previous research has generally shown weaker power for the risk-based alternative. In this paper, the different specification of the market portfolio, including both labor income and capital income, yields high power under the risk-based alternative compared to past literature. The nonrisk-based alternative hypothesis also produces significant power, and in line with past research, higher power than the risk –based alternative. The set of chosen risk factors based on the multivariate tests seems to be relevant with respect to the study of mean-variance efficient portfolios for European data. If this is not the case, it is hard to determine for European stock data, whether possible misspecifications of the set of risk factors is due to the risk-based explanation of deviations or the nonrisk-based explanation.
We find that tests based on the extension of the ICAPM by additional factors also yield high power under the risk-based alternative. The most remarkable finding is that we find robust evidence that a linear combination of the three factors from the Fama-French model is not mean-variance efficient for European stock market data. This is an important finding because this model is widely used for portfolio selection and fund performance measurement for European assets as well.

APPENDIX 1 : Identification of the portfolio types.

country portfolios	sector portfolios	size portfolios
AUS = Austria	Reso = resources	S1 = small size
BEL = Belgium	Bmat = basic materials	S2
SPA = Spain	Chem = chemicals	S3
FRA = France	Cycc = cyclical consumer goods	S4
LUX = Luxemburg	Ncyc = non-cyclical consumer goods	S5
GER = Germany	Phar = pharma	S6
GRE = Greece	Cycs = cyclical services	S7
IRE = Ireland	Bank = banks	S8
FIN = Finland	Insu = insurances	S9
ITA = Italy	Fina = financial services	S10
NET = Netherlands	Indu = industrials	S11
POR = Portugal	Tech = technology	S12
DEN = Denmark	Tele =telecom	S13 = big size
NOR = Norway	Util = utilities	
SWE = Sweden		
SWI = Switzerland		
UK = United Kingdom		

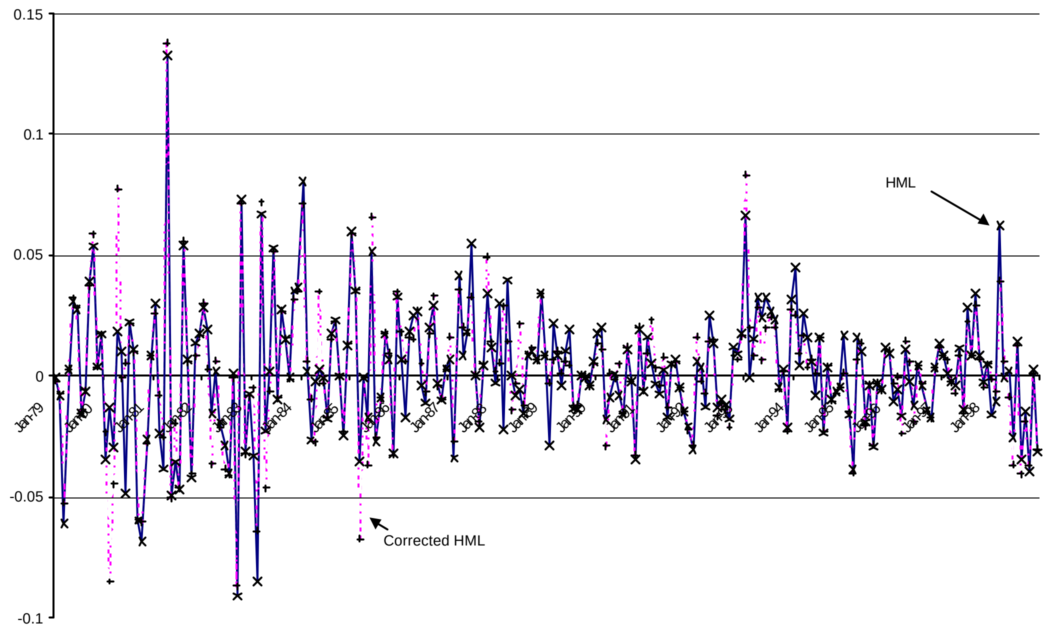
APPENDIX 2. Construction of the sector portfolios based on the Financial Times and STOXX industries.



Appendix 3: HML-factor corrected for the home country level of BTM.

Stocks are re-ranked based on the demeaned BTM. For each stock, the home country average of BTM is subtracted from the stock's BTM in an attempt to avoid country specific accounting methods. The ranking of all stocks based on this new BTM is performed for the whole period. The figure and the numbers below make a comparison of the HML factor used in this paper and the one corrected for possible differences in accounting structures across countries. From the figure we learn that the two series almost superpose each other, meaning that not correcting for the home country is not problematic.

Furthermore, the average difference between the two series is –0.02%, with a 95% interval of [-2.3%, +2.3%]. The correlation between the two factor portfolios based on BTM is 0.908.



Appendix 4 : The GRS test

From a general factor model:
With i = the number of factors in the model.

$$r_t - r_t^f = a + \sum_i b \ f_t + e_t$$

If a portfolio is mean-variance efficient, the following first-order condition must hold:

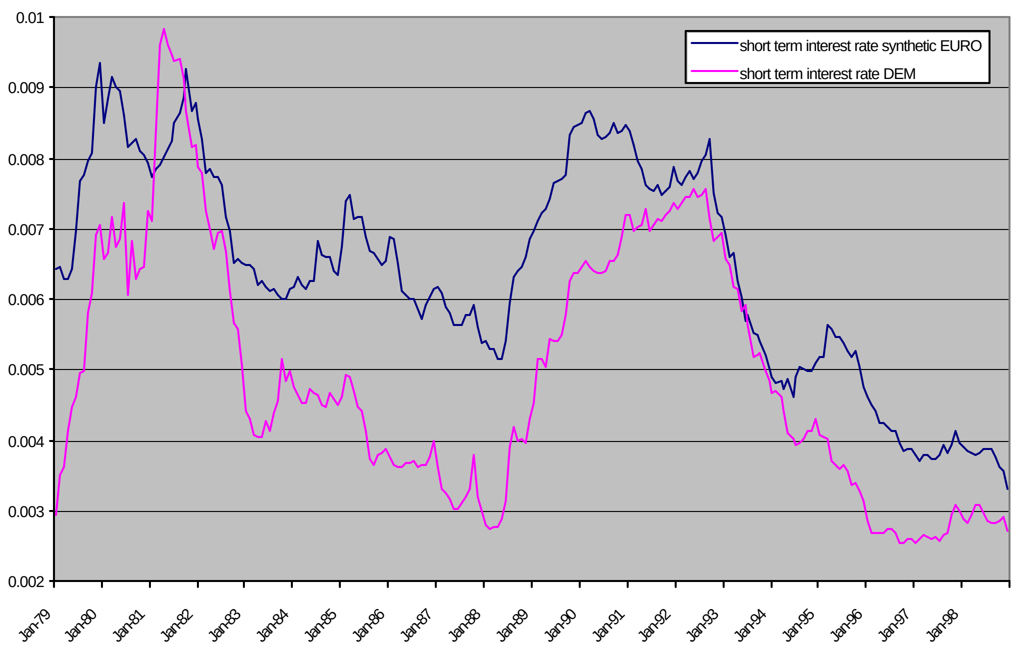
$$E(r) - r^f = \sum b \ E[f]$$

Comparing the unconditional expectation of both expressions forms the basics for the principal hypothesis. From this follows a null hypothesis that contains the parameter restriction:

$H_0 : a = 0$

Where α is the cross-sectional vector of intercepts for N portfolios or assets.

Appendix 5: Short term interest rates in DEM and synthetic euro for the period 01-1979 : 12-1998.



Appendix 6 : Appendix of tables

Table 2.bis

Average return and risk for European stock portfolios from January 1979 to December 1998.

All average returns and standard deviations are reported in % for the country, sector and size portfolios. The standard deviations are the numbers in italic below the returns. All numbers are calculated for market capitalization (MCAP) and equally weighted portfolios. The values are reported for the synthetic euro as the currency of denomination.

	AUS	*BEL*	*FRA*	*GER*	*IRE*	*ITA*	*NET*	*DEN*	*NOR*	*SWE*	*SWI*	*UK*
MCAP	0.78	1.21	1.14	0.96	1.27	1.16	1.42	1.18	1.19	1.64	1.19	1.30
	4.82	*5.17*	*6.01*	*5.01*	*6.27*	*7.17*	*4.63*	*4.93*	*8.41*	*7.60*	*4.86*	*5.63*
EW	1.05	1.39	1.28	0.89	1.19	1.31	1.37	1.30	1.31	1.63	0.78	1.41
	4.75	*4.96*	*5.98*	*4.52*	*5.71*	*7.41*	*4.62*	*4.89*	*8.01*	*7.54*	*4.22*	*5.77*

	reso	*bmat*	*chem*	*cycc*	*ncyc*	*phar*	*Cycs*	*bank*	*insu*	*fina*	*indu*	*tech*	*tele*	*Util*
MCAP	1.39	0.70	1.07	1.18	1.35	1.53	1.10	1.04	1.36	1.11	0.99	0.84	1.22	1.17
	6.04	*5.15*	*4.83*	*6.29*	*4.58*	*5.10*	*5.18*	*5.08*	*5.21*	*4.70*	*4.96*	*5.27*	*6.04*	*3.73*
EW	1.19	1.02	1.20	1.01	1.30	1.38	1.33	1.19	1.44	1.33	1.06	1.04	1.38	1.27
	6.06	*5.27*	*4.86*	*4.81*	*4.03*	*4.37*	*4.65*	*4.03*	*4.71*	*4.98*	*4.48*	*5.06*	*5.55*	*3.25*

	S1	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*	*S11*	*S12*	*S13*
MCAP	1.62	1.44	1.43	1.29	1.22	1.08	0.60	1.01	1.01	1.07	1.07	0.97	1.25
	4.23	*4.22*	*4.29*	*4.17*	*4.54*	*4.41*	*4.39*	*4.52*	*4.33*	*4.41*	*4.45*	*4.56*	*4.57*
EW	1.65	1.45	1.43	1.30	1.22	1.08	1.05	1.01	1.01	1.07	1.06	0.98	1.18
	4.31	*4.18*	*4.28*	*4.17*	*4.53*	*4.42*	*4.47*	*4.52*	*4.33*	*4.42*	*4.53*	*4.57*	*4.58*

Table 3.bis

Average return and standard deviation of the factor portfolios

Average returns and standard deviations of the factor portfolios are presented in % in panel A. All factor portfolio statistics are reported in synthetic euro (EURO). The global market portfolio is a market-capitalization weighted average of all available stocks. The equity premium is the difference between the average return on the global market portfolio and the risk-free rate. HML is the factor portfolio based on the return differential between the high book-to-market multifactor minimum variance portfolio and the low book-to-market MMV portfolio. LMOM is the factor portfolio based on the return differential between the local losers MMV portfolio and the local winners MMV portfolio. Panel B reports the correlations between the factors for the synthetic euro. * denotes significance of a non-zero return at the 5% level.

Panel A.	*Average return (%)*	*Stdev (%)*
Global market portfolio	1.304*	4.92
Equity Premium	0.662*	5.12
HML	0.193	2.73
LMOM	-0.617*	2.12

Panel B : correlations	M-F	HML	LMOM
M-F	1		
HML	-0.106	1	
LMOM	-0.040	0.136	1

Table 4a

Estimation results for the single-factor and multi-factor pricing equations : country portfolios.

The R² and the factor loadings are reported for the 8 estimated pricing equations in DEM. The t-values of the factor loading are in parentheses. The left panel of panel A shows the results for the one-factor model, the second panel reports the estimation results for the two-factor model including the high minus low BTM factor, the third, including the small minus big size factor, the fourth including the local momentum factor *a* is the estimated constant, β is the factor loading on the excess return of the global market portfolio, γ is the factor loading on second factor.

Panel A	$r-r^f = a+b(r^m-r^f)+e$			$r-r^f = a+b(r^m-r^f)+g\,HML+e$				$r-r^f = a+b(r^m-r^f)+g\,SMB+e$				$r-r^f = a+b(r^m-r^f)+g\,LMOM+e$			
	R²	α	β	R²	α	β	γ	R²	α	β	γ	R²	α	β	γ
AUS	0.36	-.002	0.580	0.37	-.002	0.587	0.145	0.37	-.003	0.584	0.199	0.36	-.001	0.58	0.082
		(-0.74)	(11.62)		(-0.87)	(11.76)	(1.55)		(-1.27)	(11.78)	(2.12)		(-0.51)	(11.63)	(0.67)
BEL	0.41	.001	0.656	0.45	.000	0.677	0.383	0.43	-.001	0.662	0.314	0.41	.003	0.66	0.203
		(0.47)	(12.77)		(0.14)	(13.53)	(4.09)		(-0.39)	((13.15)	(3.28)		(0.92)	(12.85)	(1.61)
FRA	0.47	-.001	0.818	0.48	-.001	0.83	0.224	0.47	-.001	0.821	0.129	0.47	.000	0.82	0.060
		(-0.16)	(14.51)		(-0.34)	(14.76)	(2.12)		(-0.47)	(14.56)	(1.21)		(-0.03)	(14.49)	(0.43)
GER	0.44	-.001	0.648	0.46	-.002	0.663	0.274	0.44	-.002	0.651	0.140	0.44	-.001	0.65	-.009
		(-0.51)	(13.53)		(-0.78)	(14.02)	(3.09)		(-0.90)	(13.62)	(1.55)		(-0.51)	(13.50)	(-0.07)
IRE	0.47	.001	0.866	0.49	-.000	0.881	0.278	0.48	-.001	0.871	0.237	0.47	.001	0.866	0.025
		(0.16)	(14.61)		(-.05)	(14.95)	(2.52)		(-0.40)	(14.79)	(2.13)		(0.20)	(14.58)	(0.17)
ITA	0.31	.000	0.816	0.37	-.002	0.853	0.674	0.40	-.006	0.833	0.803	0.31	-.001	0.815	-0.104
		(0.02)	(10.40)		(-0.38)	(11.29)	(4.76)		(-1.47)	(11.29)	(5.74)		(-0.13)	(10.36)	(-0.54)
NET	0.68	.003	0.755	0.72	.002	0.775	0.371	0.68	.002	0.758	0.115	0.68	.002	0.753	-0.130
		(1.69)	(22.38)		(1.29)	(24.65)	(6.30)		(1.17)	(22.54)	(1.81)		(1.17)	(22.39)	(-1.56)
DEN	0.32	.001	0.562	0.36	.001	0.582	0.371	0.33	.000	0.566	0.211	0.32	.001	0.561	-0.061
		(0.51)	(10.51)		(0.20)	(11.13)	(3.79)		(-0.05)	(10.65)	(2.09)		(0.35)	(10.47)	(-0.46)
NOR	0.35	-.001	0.985	0.41	-.003	1.027	0.774	0.37	-.004	0.993	0.405	0.35	.001	0.988	0.284
		(-0.24)	(11.31)		(-0.67)	(12.30)	(4.95)		(-0.88)	(11.52)	(2.48)		(0.15)	(11.36)	(1.33)
SWE	0.42	.004	0.982	0.44	.004	0.960	-0.408	0.43	.002	0.988	0.266	0.42	.005	0.985	0.225
		(0.89)	(13.03)		(1.15)	(12.86)	(-2.92)		(0.38)	(13.16)	(1.87)		(1.20)	(13.07)	(1.22)
SWI	0.50	.001	0.681	0.51	.001	0.690	0.163	0.51	.000	0.685	0.188	0.50	.001	0.680	-0.037
		(0.39)	(15.54)		(0.23)	(15.76)	(1.99)		(-0.21)	(15.75)	(2.28)		(0.28)	(15.49)	(-0.34)
UK	0.89	-.000	1.074	0.89	-.001	1.075	.012	0.90	.001	1.070	-0.213	0.89	-.001	1.074	-0.045
		(-0.33)	(43.03)		(-0.35)	(42.75)	(0.25)		(0.89)	(44.66)	(-4.68)		(-0.53)	(42.95)	(-0.73)

Table 4a

Estimation results for the single-factor and multi-factor pricing equations : country portfolios.

The R² and the factor loadings are reported for the 8 estimated pricing equations in DEM. The significance of the non-zero estimated coefficients at the 5% level is indicated by *. The left panels of panel B shows the results for the three-factor model, the right panel reports the estimation results for the four-factor model. *a* is the estimated constant, β is the factor loading on the excess return of the global market portfolio, γ is the factor loading on the secon factor and δ is the factor loading on the third factor and η is the factor loading on the fouth factor.

Panel B	$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*HML+\delta SMB$				$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*HML+\delta LMOM$				$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*SMB+\delta LMOM$				$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*HML+\delta SMB+\eta LMOM$				
	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ	η
AUS	-.003	0.590 *	0.119	0.181	-.002	0.588 *	0.140	0.058	-.003	0.584 *	0.195 *	0.022	-.003	0.590 *	0.119	0.180	0.006
BEL	-.001	0.680 *	0.345 *	0.263 *	.001	0.677 *	0.369 *	0.139	.000	0.663 *	0.293 *	0.113	-.001	0.680 *	0.340 *	0.252 *	0.067
FRA	-.002	0.832 *	0.209 *	0.099	-.001	0.831 *	0.221 *	0.022	-.001	0.821 *	0.125	0.022	-.002	0.832 *	0.210 *	0.100	-0.007
GER	-.003	0.665 *	0.259 *	0.103	-.002	0.663 *	0.280 *	-0.057	-.003	0.651 *	0.150	-0.055	-.003	0.664 *	0.266 *	0.118	-0.091
IRE	-.002	0.883 *	0.249 *	0.201	.000	0.881 *	0.280 *	-0.024	-.002	0.870 *	0.247 *	-0.051	-.002	0.883 *	0.256 *	0.216	-0.086
ITA	-.006	0.863 *	0.571 *	0.720 *	-.003	0.851 *	0.697 *	-0.226	-.009 *	0.830 *	0.871 *	-0.371 *	-.010 *	0.861 *	0.606 *	0.798 *	-0.454
NET	.002	0.776 *	0.362 *	0.063	.001	0.774 *	0.391 *	-0.198 *	.001	0.756 *	0.147 *	-0.175 *	.000	0.775 *	0.380 *	0.101	-0.227
DEN	-.001	0.584 *	0.348 *	0.160	.000	0.581 *	0.384 *	-0.128	-.001	0.565 *	0.235 *	-0.133	-.002	0.584 *	0.362 *	0.191	-0.183
NOR	-.005	1.031 *	0.731 *	0.299	-.002	1.028 *	0.758 *	0.151	-.003	0.995 *	0.375 *	0.169	-.004	1.031 *	0.726 *	0.287	0.069
SWE	.002	0.964 *	-0.456 *	0.332 *	.006	0.962 *	-0.439 *	0.302	.003	0.989 *	0.238	0.152	.004	0.965 *	-0.473	0.295 *	0.217
SWI	-.001	0.692 *	0.139	0.168 *	.000	0.689 *	0.170 *	-0.067	-.001	0.684 *	0.207 *	-0.101	-.002	0.691 *	0.148	0.189 *	-0.121
UK	.001	1.072 *	0.043	-0.219 *	-.001	1.075 *	0.017	-0.048	.001	1.070 *	-0.217 *	0.022	.001	1.072 *	0.042	-0222 *	0.016

Table 4b

Estimation results for the single-factor and multi-factor pricing equations : sector portfolios.

The R^2 and the factor loadings are reported for the 8 estimated pricing equations in DEM. The t-values of the factor loading are in parentheses. The left panel of panel A shows the results for the one-factor model, the second panel reports the estimation results for the two-factor model including the high minus low BTM factor, the third, including the small minus big size factor, the fourth including the local momentum factor *a* is the estimated constant, β is the factor loading on the excess return of the global market portfolio, γ is the factor loading on second factor.

Panel	$r-r^f=a+b(r^m-r^f)+e$			$r-r^f=a+b(r^m-r^f)+gHML+e$				$r-r^f=a+b(r^m-r^f)+gSMB+e$				$r-r^f=a+b(r^m-r^f)+gLMOM+e$			
	R^2	α	β	R^2	α	β	γ	R^2	α	β	γ	R^2	α	β	γ
Reso	0.60	.001	0.930	0.62	.001	0.947	0.326	0.60	.002	0.928	-0.053	0.60	.000	0.928	-0.134
		(0.53)	(18.71)		(0.24)	(19.43)	(3.57)		(0.65)	(18.65)	(-0.56)		(0.19)	(18.68)	(-1.10)
Bmat	0.77	-.005	0.915	0.80	-.006	0.934	0.357	0.79	-.008	0.921	0.297	0.78	-.004	0.918	0.273
		(-3.25)	(27.96)		(-4.02)	(30.59)	(6.24)		(-4.60)	(29.52)	(5.01)		(-2.18)	(28.69)	(3.46)
Chem	0.75	-.001	0.848	0.77	-.002	0.862	0.257	0.76	-.003	0.851	0.146	0.75	-.001	0.849	0.017
		(-0.89)	(26.85)		(-1.30)	28.22)	(4.48)		(-1.51)	(27.21)	(2.46)		(-0.79)	(26.79)	(0.22)
Cycc	0.54	-.001	0.927	0.54	.000	0.918	-0.168	0.57	-.004	0.937	0.448	0.54	.001	0.930	0.269
		(-0.28)	(16.57)		(-0.14)	(16.37)	(-1.60)		(-1.42)	(17.35)	(4.37)		(0.30)	(16.72)	(1.96)
Ncyc	0.88	.001	0.877	0.88	.001	0.881	0.082	0.88	.002	0.875	-0.070	0.88	.001	0.877	0.004
		(1.26)	(40.76)		(1.10)	(41.03)	(2.05)		(1.67)	(40.82)	(-1.71)		(1.23)	(40.66)	(0.08)
Phar	0.66	.003	0.826	0.66	.004	0.816	-0.181	0.66	.004	0.823	-0.124	0.66	.002	0.823	-0.222
		(1.70)	(21.26)		(1.92)	(21.13)	(-2.51)		(2.08)	(21.26)	(-1.68)		(0.96)	(21.38)	(-2.34)
Cycs	0.88	-.002	0.982	0.88	-.002	0.984	0.032	0.88	-.001	0.981	-0.057	0.88	-.002	0.982	0.021
		(-1.53)	(41.42)		(-1.58)	(41.24)	(0.71)		(-1.15)	(41.39)	(-1.26)		(-1.36)	(41.34)	(0.35)
Bank	0.70	-.002	0.856	0.74	-.003	0.876	0.369	0.71	-.003	0.860	0.197	0.70	-.002	0.856	-0.004
		(-0.89)	(23.70)		(-1.43)	(25.72)	(5.78)		(-1.63)	(24.17)	(2.92)		(-0.86)	(23.64)	(-0.04)
Insu	0.71	.001	0.877	0.73	.001	0.895	0.325	0.71	.000	0.880	0.108	0.71	.001	0.878	0.028
		(0.66)	(23.91)		(0.28)	(25.43)	(4.93)		(0.24)	(24.02)	(1.55)		(0.72)	(23.86)	(0.30)
Fina	0.89	-.001	0.901	0.90	-.002	0.912	0.207	0.89	-.002	0.902	0.084	0.89	-.001	0.901	0.019
		(-1.27)	(43.93)		(-1.83)	(47.09)	(5.71)		(-1.80)	(44.32)	(2.17)		(-1.10)	(43.84)	(0.38)
Indu	0.79	-.003	0.886	0.79	-.003	0.892	0.108	0.79	-.003	0.889	0.124	0.79	-.003	0.886	-0.036
		(-1.69)	(29.82)		(-1.86)	(30.03)	(1.94)		(-2.22)	(30.12)	(2.21)		(-1.76)	(29.74)	(-0.49)
Tech	0.79	-.004	0.948	0.83	-.005	0.968	0.369	0.80	-.005	0.950	0.125	0.79	-.004	0.949	0.072
		(-2.63)(	(30.13)		(-3.43)	(33.39)	(6.80)		(-3.10)	(30.41)	(2.11)		(-2.24)	(30.14)	(0.93)
Tele	0.46	.000	0.829	0.46	.000	0.832	0.057	0.47	-.001	0.832	0.128	0.47	.000	0.828	-0.098
		(0.14)	(14.35)		(0.09)	(14.30)	(0.53)		(-0.17)	(14.39)	(1.16)		(-0.07)	(14.30)	(-0.69)
Util	0.53	.002	0.553	0.56	.001	0.565	0.235	0.53	.002	0.552	-0.022	0.53	.001	0.552	-0.034
		(0.88)	(16.45)		(0.58)	(17.21)	(3.82)		(0.94)	(16.39)	(-0.34)		(0.73)	(16.40)	(-0.41)

Table 4b

Estimation results for the single-factor and multifactor pricing equations :sector portfolios.

The R² and the factor loadings are reported for the 8 estimated pricing equations in DEM. The significance of the non-zero estimated coefficients at the 5% level is indicated by *. α is the estimated constant, β is the factor loading on the excess return of the global market portfolio, γ is the factor loading on the secon factor and δ is the factor loading on the third factor and η is the factor loading on the fouth factor.

Panel B	$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*HML+\delta SMB$				$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*HML+\delta LMOM$				$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*SMB+\delta LMOM$				$r-r^f=\alpha+\beta*(r^m-r^f)+\gamma*HML+\delta SMB+\eta LMOM$				
	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ	η
Reso	.001	0.946*	0.341*	-0.103	-.001	0.946*	0.346*	-0.194	.001	0.927*	-0.030	-0.125	.000	0.945*	0.354*	-0.073	-0.173
Bmat	-.008*	0.937*	0.321*	0.250*	-.005*	0.936*	0.335*	0.214*	-.006*	0.923*	0.261*	0.192*	-.007*	0.938*	0.310*	0.224*	0.150*
Chem	-.003	0.864*	0.241*	0.111	-.002	0.862*	0.259*	-0.028	-.003	0.851*	0.151*	-0.029	-.003	0.864*	0.246*	0.122*	-0.063
Cycc	-.004	0.925*	-0.237*	0.482*	.002	0.920*	-0.199	0.303*	-.003	0.938*	0.422*	0.139	-.002	0.925*	-0.250*	0.453*	0.174
Ncyc	.002	0.880*	0.094*	-0.083*	.001	0.881*	0.083*	-0.010	.002	0.875*	-.074	0.027	.002	0.880*	0.093*	-0.086*	0.014
Phar	.005*	0.815*	-0.167*	-0.099	.003	0.815*	-0.162*	-0.194*	.003	0.822*	-0.088	-0.195*	.003	0.814*	-0.154*	-0.069	-0.174
Cycs	-.002	0.983*	0.041	-0.063	-.002	0.984*	0.030	0.015	-.001	0.981*	-0.064	0.040	-.001	0.983*	0.038	-0.069	0.035
Bank	-.004*	0.878*	0.348*	0.147*	-.003	0.876*	0.376*	-0.069	-.004	0.860*	0.210*	-0.068	-.004*	0.878*	0.357*	0.167*	-0.117
Insu	.000	0.896*	0.316*	0.061	.000	0.985*	0.328*	-0.030	.000	0.880*	0.109	-0.006	.000	0.896*	0.320*	0.070	-0.050
Fina	-.002*	0.913*	0.199*	0.055	-.002	0.912*	0.209*	-0.017	-.002	0.902*	0.085*	-0.007	-.002*	0.912*	0.202*	0.061	-0.034
Indu	-.004*	0.893*	0.092	0.110*	-.003*	0.892*	0.114*	-0.056	-.004*	0.888*	0.138*	-0.078	-.004	0.893*	0.099	0.126*	-0.092
Tech	-.006*	0.969*	0.359*	0.073	-.005*	0.968*	0.368*	0.008	-.005*	0.951*	0.119	0.036	-.006*	0.969*	0.360*	0.075	-0.014
Tele	-.001	0.834*	0.040	0.122	.000	0.832*	0.069	-0.110	-.002	0.831*	0.154	-0.145	-.002	0.833*	0.052	0.148	-0.152
Util	.001	0.565*	0.243*	-0.058	.000	0.565*	0.243*	-0.076	.001	0.552*	-0.017	-0.029	.001	0.564*	0.248*	-0.047	-0.063

Table 4c

Estimation results for the single-factor and multi-factor pricing equations : size portfolios.

The R² and the factor loadings are reported for the 8 estimated pricing equations in DEM. The t-values of the factor loading are in parentheses. The left panel of panel A shows the results for the one-factor model, the second panel reports the estimation results for the two-factor model including the high minus low BTM factor, the third, including the small minus big size factor, the fourth including the local momentum factor *a* is the estimated constant, β is the factor loading on the excess return of the global market portfolio, γ is the factor loading on second factor.

Panel	$r-r^f = a+b(r^m-r^f)+e$			$r-r^f = a+b(r^m-r^f)+g\,HML+e$				$r-r^f = a+b(r^m-r^f)+g\,SMB+e$				$r-r^f = a+b(r^m-r^f)+g\,LMOM+e$			
	R²	α	β	R²	α	β	γ	R²	α	β	γ	R²	α	β	γ
S1	0.62	.005	0.687	0.67	.004	0.705	0.343	0.76	.000	0.700	0.618	0.63	.006	0.689	0.207
		(2.75)	(19.75)		(2.43)	(21.42)	(5.56)		(0.27)	(25.25)	(11.75)		(3.37)	(20.02)	(2.45)
S2	0.68	.004	0.711	0.72	.003	0.728	0.307	0.79	.000	0.722	0.526	0.69	.004	0.713	0..112
		(2.18)	(22.68)		(1.84)	(24.47)	(5.50)		(-0.23)	(28.00)	(10.75)		(2.51)	(22.76)	(1.45)
S3	0.74	.003	0.762	0.79	.002	0.782	0.370	0.81	-.001	0.771	0.452	0.75	.004	0.764	0.174
		(1.81)	(26.10)		(1.37)	(29.57)	(7.48)		(-0.43)	(31.04)	(9.57)		(2.46)	(26.44)	(2.45)
S4	0.75	.002	0.739	0.80	.001	0.758	0.356	0.80	-.001	0.747	0.370	0.76	.003	0.742	0.237
		(1.08)	(26.59)		(0.56)	(30.18)	(7.57)		(-0.87)	(30.07)	(7.84)		(2.09)	(27.32)	(3.54)
S5	0.82	.000	0.833	0.85	-.001	0.849	0.297	0.86	-.003	0.841	0.358	0.82	.001	0.835	0.168
		(0.06)	(32.80)		(-0.50)	(36.26)	(6.76)		(-2.13)	(37.66)	(8.46)		(0.85)	(33.31)	(2.72)
S6	0.82	-.001	0.816	0.85	-.002	0.832	0.281	0.85	-.003	0.823	0.305	0.83	.000	0.818	0.174
		(-0.88)	(33.14)		(-1.51)	(36.43)	(6.57)		(-2.80)	(36.74)	(7.17)		(-0.01)	(33.72)	(2.91)
S7	0.83	-.001	0.829	0.87	-.002	0.847	0.318	0.87	-.004	0.836	0.314	0.84	.000	0.831	0.179
		(-1.10)	(34.33)		(-1.89)	(39.03)	(7.83)		(-3.16)	(38.50)	(7.62)		(-0.18)	(35.01)	(3.06)
S8	0.84	-.002	0.842	0.88	-.003	0.861	0.346	0.87	-.004	0.848	0.296	0.85	-.001	0.845	0.231
		(-1.69)	(34.85)		(-2.67)	(40.64)	(8.72)		(-3.63)	(38.52)	(7.07)		(-0.50)	(36.04)	(4.01)
S9	0.85	-.002	0.809	0.88	-.002	0.826	0.305	0.87	-.003	0.814	0.237	0.85	-.001	0.811	0.143
		(-1.50)	(36.66)		(-2.40)	(42.25)	(8.34)		(-3.13)	(39.53)	(6.06)		(-0.69)	(37.19)	(2.66)
S10	0.86	-.001	0.827	0.88	-.002	0.842	0.262	0.87	-.003	0.832	0.206	0.86	-.001	0.829	0.107
		(-1.05)	(37.58)		(-1.73)	(41.59)	(6.90)		(-2.42)	(39.75)	(5.18)		(-0.43)	(37.86)	(1.98)
S11	0.90	-.001	0.874	0.92	-.002	0.889	0.267	0.90	-.002	0.877	0.111	0.90	-.001	0.876	0.123
		(-1.35)	(45.71)		(-2.24)	(52.63)	(8.45)		(-2.13)	(46.62)	(3.11)		(-0.54)	(46.34)	(2.65)
S12	0.91	-.002	0.889	0.93	-.003	0.902	0.232	0.91	-.003	0.890	0.040	0.91	-.002	0.890	0.084
		(-2.72)	(49.88)		(-3.68)	(56.17)	(7.72)		(-2.94)	(49.92)	(1.18)		(-2.06)	(50.19)	(1.92)
S13	0.94	.000	0.902	0.94	.000	0.902	0.005	0.94	.000	0.901	-0.004	0.94	.000	0.901	-0.020
		(0.15)	(63.02)		(0.13)	(62.58)	(0.18)		(0.18)	(62.83)	(-0.15)		(-0.02)	(62.88)	(-0.57)

Table 4c

Estimation results for the single-factor and multifactor pricing equations :size portfolios.

The R^2 and the factor loadings are reported for the 8 estimated pricing equations in DEM. The significance of the non-zero estimated coefficients at the 5% level is indicated by *. a is the estimated constant, β is the factor loading on the excess return of the global market portfolio, γ is the factor loading on the secon factor and δ is the factor loading on the third factor and η is the factor loading on the fouth factor.

Panel B	$r\text{-}r^f=\alpha+\beta*(r^m\text{-}r^f)+\gamma*HML+\delta SMB$				$r\text{-}r^f=\alpha+\beta*(r^m\text{-}r^f)+\gamma*HML+\delta LMOM$				$r\text{-}r^f=\alpha+\beta*(r^m\text{-}r^f)+\gamma*SMB+\delta LMOM$				$r\text{-}r^f=\alpha+\beta*(r^m\text{-}r^f)+\gamma*HML+\delta SMB+\eta LMOM$				
	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ	α	β	γ	δ	η
S1	.000	0.713*	0.260*	0.580*	.005*	0.706*	0.328 (*	0.150	.001	0.700*	0.615*	0.019	.000	0.713*	0.261*	0.583*	-0.017
S2	-.001	0.735*	0.236*	0.492*	.003*	0.728*	0.301*	0.059	-.001	0.722*	0.536*	-0.052	-.001	0.734*	0.243*	0.507*	-0.086
S3	-.001	0.787*	0.312*	0.406*	.003	0.783*	0.359*	0.111	.000	0.771*	0.445*	0.038	-.001	0.787*	0.313*	0.407*	-0.005
S4	-.002	0.763*	0.309*	0.324*	.002	0.759*	0.338*	0.178*	.000	0.748*	0.346*	0.131*	-.001	0.763*	0.303*	0.309*	0.089
S5	-.003*	0.854*	0.250*	0.322*	.000	0.850*	0.285*	0.118*	-.002	0.841*	0.347*	0.061	-.003*	0.854*	0.248*	0.317*	0.027
S6	-.004*	0.835*	0.242*	0.270*	-.001*	0.832*	0.268*	0.127*	-.003*	0.823*	0.289*	0.085	-.003*	0.835*	0.238*	0.261*	0.052
S7	-.004*	0.850*	0.279*	0.273*	-.001	0.848*	0.305*	0.126*	-.003*	0.837*	0.298*	0.088	-.004*	0.851*	0.275*	0.265*	0.050
S8	-.005*	0.864*	0.310*	0.250*	-.002*	0.862*	0.328*	0.174*	-.003*	0.850*	0.268*	0.149*	-.004*	0.865*	0.302*	0.232*	0.108*
S9	-.004*	0.828*	0.277*	0.196*	-.002*	0.826*	0.296*	0.091	-.003*	0.814*	0.223*	0.074	-.003*	0.828*	0.274*	0.190*	0.037
S10	-.003*	0.844*	0.237*	0.171*	-.001	0.842*	0.255*	0.062	-.002*	0.832*	0.198*	0.046	-.003*	0.844*	0.236*	0.169*	0.014
S11	-.002*	0.890*	0.257*	0.073*	-.001	0.889*	0.259*	0.078	-.001	0.877*	0.094*	0.094*	-.002*	0.890*	0.252*	0.063	0.060
S12	-.003*	0.902*	0.231*	0.006	-.003*	0.902*	0.228*	0.044	-.002*	0.891*	0.026	0.076	-.003*	0.902*	0.228*	-0.001	0.045
S13	.000	0.902*	0.006	-0.005	.000	0.902*	0.007	-0.021	.000	0.901*	0.000	-0.020	.000	0.902*	0.007	-0.001	-0.021

Table 5.bis

Results for the multivariate GRS test of exact factor pricing using a one-factor model

For the four different time periods and the synthetic euro, the test values are reported. The first period is the full sample period from January 1979 until December 1998. The second period is a period of stable interest rates, from January 1983 until December 1987. The third period is a period of rising interest rates, from January 1988 until December 1992. The last period is a period of declining interest rates, from January 1994 until December 1998. n is the estimated non-centrality parameter from the one-factor model. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. N is the number of portfolios and T is the number of observations.

$R\text{-}F=\alpha+\beta(M\text{-}F)$			
	COUNTRY	SECTOR	SIZE
79:01-98 :12			
n	12.351	36.553	56.561
F-statistic	0.982	2.468	4.132
p-value	0.467	0.003	0.000
83 :01-87 :12			
n	3.740	21.495	25.542
F-statistic	0.253	1.191	1.558
p-value	0.994	0.315	0.133
88:01-92 :12			
n	17.152	31.151	32.938
F-statistic	1.158	1.726	2.010
p-value	0.340	0.084	0.042
94:01-98:12			
n	14.160	29.292	38.462
F-statistic	0.956	1.623	2.347
p-value	0.502	0.110	0.017

Table 6a.

Test results for the multivariate GRS test of exact factor pricing using the synthetic as the currency of denomination,1979-1983.

For the entire time period (1979-1998) the test values are reported for 7 augmented models. Av$|a|$ is the average absolute value of the alphas, n is the estimated non-centrality parameter from the multifactor models. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. N is the number of portfolios and T (=60) is the number of observations.

	$r-r^f = a + b(r^m - r^f) + e$			$r-r^f = a + b(r^m - r^f) + g\, f_2 + e$					
				F2=HML			*F2=SMB*		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av. $\|a\|$	.0061	.0043	.0024	.0063	.0046	.0024	.0057	.0041	.0017
n	14.76	26.35	25.66	16.67	33.82	33.04	11.63	24.64	23.24
F-stat	0.997	1.461	1.566	1.102	1.833	1.972	0.769	1.335	1.387
p-value	0.466	0.166	0.131	0.382	0.064	0.046	0.678	0.227	0.203
				$r-r^f = a + b(r^m - r^f) + g\, f_2 + d\, f_3 + e$					
	F3=LMOM			F2=HML, F3=SMB			F2=HML, F3=LMOM		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av. $\|a\|$	.0070	.0054	.0027	.0060	.0044	.0021	.0074	.0059	.0028
n	16.30	37.94	29.93	13.53	30.47	30.63	22.09	61.40	39.24
F-stat	1.077	2.056	1.786	0.875	1.614	1.788	1.428	3.251	2.290
p-value	0.401	0.035	0.076	0.577	0.114	0.076	0.189	0.002	0.021
				$r-r^f = a + b(r^m - r^f) + g\, f_2 + d\, f_3 + h\, f_4 + e$					
	F2=SMB, F3=LMOM			*F2=HML, F3=SMB, F4=LMOM*					
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*			
av. $\|a\|$	.0067	.0052	.0022	.0072	.0058	.0027			
n	13.32	40.85	26.31	19.39	62.16	35.38			
F-stat	0.861	2.163	1.535	1.226	3.215	2.018			
p-value	0.590	0.027	0.143	0.297	0.002	0.043			
N	12	14	13	12	14	13	12	14	13

Table 6b.

Test results for the multivariate GRS test of exact factor pricing using the DEM as the currency of denomination,1984-1988.

For the entire time period (1979-1998) the test values are reported for 7 augmented models. Av$|a|$ is the average absolute value of the alphas, n is the estimated non-centrality parameter from the multifactor models. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. N is the number of portfolios and T (=60) is the number of observations.

$r - r^f = a + b(r^m - r^f) + e$				$r - r^f = a + b(r^m - r^f) + g\, f_2 + e$					
				F2=HML			*F2=SMB*		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av.$\|a\|$	.0039	.0024	.0048	.0044	.0032	.0040	.0039	.0038	.0012
n	11.64	24.82	27.57	13.83	27.58	23.00	11.33	36.32	10.26
F-stat	0.786	1.376	1.682	0.914	1.495	1.373	0.749	1.968	0.612
p-value	0.662	0.205	0.098	0.541	0.153	0.210	0.697	0.044	0.831
				$r - r^f = a + b(r^m - r^f) + g\, f_2 + d\, f_3 + e$					
	F3=LMOM			F2=HML, F3=SMB			F2=HML, F3=LMOM		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av.$\|a\|$	.0039	.0022	.0045	.0043	.0046	.0011	.0044	.0031	.0037
n	13.76	22.55	25.14	15.69	45.12	11.80	13.40	25.16	21.12
F-stat	0.909	1.222	1.501	1.014	2.389	0.689	0.867	1.332	1.232
p-value	0.545	0.294	0.155	0.452	0.015	0.763	0.585	0.229	0.290
				$r - r^f = a + b(r^m - r^f) + g\, f_2 + d\, f_3 + h\, f_4 + e$					
	F2=SMB, F3=LMOM			*F2=HML, F3=SMB, F4=LMOM*					
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*			
av.$\|a\|$	.0045	.0039	.0010	.0050	.0046	.0010			
n	12.73	34.51	9.72	15.44	43.51	11.25			
F-stat	0.823	1.827	0.567	0.976	2.251	0.641			
p-value	0.626	0.066	0.867	0.486	0.022	0.806			
N	12	14	13	12	14	13	12	14	13

Table 6c.

Test results for the multivariate GRS test of exact factor pricing using the DEM as the currency of denomination,1989-1993.

For the entire time period (1979-1998) the test values are reported for 7 augmented models. Av$\lvert a \rvert$ is the average absolute value of the alphas, n is the estimated non-centrality parameter from the multifactor models. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. N is the number of portfolios and T (=60) is the number of observations.

$r - r^f = a + b(r^m - r^f) + e$				$r - r^f = a + b(r^m - r^f) + g\, f_2 + e$					
				F2=HML			*F2=SMB*		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av. $\lvert a \rvert$	.0032	.0033	.0025	.0038	.0036	.0035	.0032	.0032	.0023
n	13.68	20.50	28.02	16.31	28.92	29.96	14.16	23.41	28.13
F-stat	0.923	1.136	1.709	1.078	1.567	1.788	0.936	1.268	1.679
p-value	0.532	0.355	0.091	0.400	0.128	0.075	0.521	0.265	0.099
				$r - r^f = a + b(r^m - r^f) + g\, f_2 + d\, f_3 + e$					
	F3=LMOM			F2=HML, F3=SMB			F2=HML, F3=LMOM		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av. $\lvert a \rvert$	.0027	.0023	.0010	.0034	.0033	.0025	.0038	.0028	.0017
n	9.08	12.49	19.97	15.64	29.65	29.41	12.80	21.15	19.79
F-stat	0.600	0.677	1.912	1.011	1.570	1.716	0.828	1.120	1.155
p-value	0.831	0.784	0.316	0.455	0.128	0.091	0.622	0.369	0.343
				$r - r^f = a + b(r^m - r^f) + g\, f_2 + d\, f_3 + h\, f_4 + e$					
	F2=SMB, F3=LMOM			*F2=HML, F3=SMB, F4=LMOM*					
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*			
av. $\lvert a \rvert$	.0033	.0028	.0018	.0038	.0029	.0020			
n	10.38	16.36	19.26	13.35	25.70	19.76			
F-stat	0.671	0.866	1.124	0.844	1.330	1.127			
p-value	0.770	0.598	0.366	0.607	0.232	0.364			
N	12	14	13	12	14	13	12	14	13

Table 6d.

Test results for the multivariate GRS test of exact factor pricing using the DEM as the currency of denomination,1994-1998.

For the entire time period (1979-1998) the test values are reported for 7 augmented models. Av$|a|$ is the average absolute value of the alphas, n is the estimated non-centrality parameter from the multifactor models. The F-statistic (F-stat) is the GRS statistic (equation [2]) and the p-value is the associated probability of the F-test. N is the number of portfolios and T (=60) is the number of observations.

$r-r^f = a+b(r^m-r^f)+e$				$r-r^f = a+b(r^m-r^f)+g\,f_2+e$					
				F2=HML			*F2=SMB*		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av. $\|a\|$	.0026	.0034	.0031	.0030	.0029	.0026	.0035	.0046	.0037
n	14.16	29.29	38.46	14.07	27.15	38.60	18.50	69.72	37.44
F-stat	0.956	1.623	2.347	0.930	1.471	2.304	1.223	3.778	2.235
p-value	0.502	0.110	0.017	0.526	0.163	0.019	0.297	0.000	0.023
				$r-r^f = a+b(r^m-r^f)+g\,f_2+d\,f_3+e$					
	F3=LMOM			F2=HML, F3=SMB			F2=HML, F3=LMOM		
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*	*Country*	*Sector*	*Size*
av. $\|a\|$	.0023	.0022	.0023	.0033	.0041	.0032	.0023	.0018	.0023
n	9.22	21.08	29.73	16.94	66.16	35.21	9.11	19.58	30.02
F-stat	0.609	1.143	1.774	1.095	3.503	2.054	0.589	1.037	1.752
p-value	0.823	0.351	0.078	0.387	0.001	0.038	0.839	0.438	0.083
				$r-r^f = a+b(r^m-r^f)+g\,f_2+d\,f_3+h\,f_4+e$					
	F2=SMB, F3=LMOM			*F2=HML, F3=SMB, F4=LMOM*					
	Country	*Sector*	*Size*	*Country*	*Sector*	*Size*			
av. $\|a\|$	.0043	.0033	.0029	.0038	.0031	.0024			
n	13.09	46.08	24.09	11.86	44.33	22.68			
F-stat	0.846	2.440	1.406	0.750	2.293	1.293			
p-value	0.604	0.013	0.195	0.670	0.019	0.254			
N	12	14	13	12	14	13	12	14	13

Chapter III

Accounting multiples
and
Expected Returns:

A Fundamental Risk Analysis
for
European Sectors

Accounting multiples and Expected Returns: A Fundamental Risk Analysis for European Sectors*

Abstract

The relationship between fundamental accounting multiples and stock returns has been an important research topic for the past 20 years. In this paper we analyze the relationship between earnings yield forecasts, the book-to-market ratio and returns on a European sector basis. We notice that there is a persistent difference in performance between high earnings yield stock portfolios and low earnings yield stock portfolios, but in most sectors, the difference is time-varying. A second part of this paper looks at the economic forces behind this time-varying behavior. Using the standard consumption-based asset pricing framework, we find evidence of a relationship between consumption data and return spreads between high and low earnings yield stocks.

* Co-authored by R. Vander Vennet

1. Introduction

There is a lot of evidence in the past literature that buying value stocks (i.e. stocks that are priced low relative to accounting measures of operating performance) and selling growth stocks earns positive returns. Most evidence is provided for U.S. cross-sections, but this return premium seems to exist outside the U.S. as well. In this paper, we analyze the return premium of this value-growth strategy for European sector samples. The contribution of this paper is twofold. First, we explore the return premium on this zero-cost strategy in European samples where stocks are assumed to have the same characteristics because they are in the same business. They do not only belong to the same sector, but they are also assumed to have the same level of systematic risk[1]. Hence, it is interesting to see whether the value premium[2] is unconditionally found in these samples as well. Second, we evaluate the possible differences in returns for fundamental risk. In the past literature, not much evidence has been found for the risk story for country cross-sections. Hence, we reassess this question for a new dataset, but more importantly, for different sub-samples.

In the asset pricing literature, several studies have documented that valuation multiples or accounting information, such as earnings-to-price or book-to-price, serve as useful information variables for a stock investment strategy. However, it is a difficult task to identify the sources of the observed excess returns. One of the most investigated examples is the difference in performance between value and growth stocks [see Lakonishok, Shleifer and Vishny, 1994]. Buying stocks with low prices relative to their accounting fundamentals is found to yield superior returns. A common empirical finding for U.S. data is that value stocks earn a yearly excess return of about 9% [Lakonishok et al., 1994; La Porta, 1996; Dechow and Sloan, 1997]. International evidence on this subject reveals that this return on a zero-cost value-growth strategy exists in most countries, but not in all [Bakshi and Chan, 2000].
Some researchers have explored the returns on contrarian strategies within industries. Already in 1968, Breen found that value stocks perform slightly better than growth stocks in industry samples. Dreman and Lufkin [1997] find that value stocks generally outperform growth stocks in 44 industry samples. In their investigation, they used a total of 4210 companies over the period 1970-1995. However, it is noteworthy that when they look at all the five-year sub-periods from 1970 on, contrarian strategies based on earnings-to-price or book-to-price do not yield a statistically significant positive return in a lot of the case.

[1] First, stocks within an industry are more correlated than across industries. Second, in order to estimate the cost-of-equity, betas are often estimated using the pure-play analysis. This implies that estimating the betas for stocks that have comparable businesses is less noisy than individual estimates.
[2] The value premium is the return on a value-growth strategy, often referred to in the literature as contrarian strategy.

The overall results seem to be driven by an extreme outperformance of value stocks over growth stocks in the period 1975-1979.

Several possible explanations for this excess return on contrarian strategies have been suggested. However, the empirical evidence supporting the various hypotheses for the explanation of this return spread is somewhat contradictory. One plausible explanation is that an investment strategy based on earnings yield is contrarian to naive strategies in which extrapolation of past fundamentals causes overreaction [Lakonishok et al., 1994]. La Porta [1996] and Dechow and Sloan [1997] find analogous excess returns but they argue that stock prices reflect analysts' biased forecasts of future earnings growth. An alternative hypothesis is that value stocks may be fundamentally riskier [Lakonishok et al., 1994; La Porta, 1996]. It can be expected that the relatively lowly priced stocks will underperform the relatively highly priced stocks in states of the world where marginal utility of wealth is high. La Porta [1996] argues that naive investors are not able to make a difference between market risk and company-specific risk, which could lead to overestimation of the risk of value stocks. However, Lakonishok et al. [1994] and La Porta [1996] acknowledge that the empirical support for the risk arguments has been weak in the past. Finally, the empirical finding of a difference in performance associated with different investment styles could be due to biases [Lo and MacKinlay, 1990, Kothari et al., 1995].

Valuation multiples

This paper focuses on two valuation multiples, i.e. the one-year price-earnings IBES consensus forecast and the book-to-market ratio. There are several reasons for this choice. First of all, past research has reported a possible gain from trading strategies based on analysts' forecasts [Dimson and Marsh, 1984, Elton et al., 1986]. Even in very recent analyses, researchers still find valuation multiples based on analysts' forecasts to be informative [Bakshi and Chan, 2000]. On the other hand, we use the (IBES) consensus forecasts for reasons that have been developed in the extant literature. One argument is that, in order to avoid the selection bias, it is necessary to use forecasts from more than one brokerage firm [Elton et al., 1986]. Moreover, an additional motivation to use a consensus forecast is given by Dimson and Marsh [1984] who find that the pooling of forecasts from different sources improves their predictive capacity.

Bakshi and Chan [2000] argument that there are several advantages in using the earnings forecasts rather than observed earnings. First, there is a frequent revision of earnings forecasts implying that active portfolio strategies are accessible using this measure. Second, using this earnings measure implies that there is no possibility for a look-ahead bias. Third, and essential in this paper, is that earnings forecasts are less prone to differences in accounting standards and the country-specific application of accounting rules which allows us to study the European cross-section of this earnings yield valuation multiple. A final advantage is given by the research by Dechow et al. [1999] who find that one-year analyst' forecasts of earnings hold a lot of information about contemporaneous stock prices. Using this variable, they find little additional information in book values. The authors call this finding surprising because the short-term forecasts do not capture the long-term dynamics.

Although we focus on the relation between the earnings yield multiple and returns, we also study the book-to-market ratio. Next to the fact that this valuation multiple seems successful in explaining the cross-section of returns, there are two main reasons why it is worthwhile studying it. First, against the general intuition, Penman states that the P/E only provides information about future profitability when P/E is conditioned on current return on equity. The price-to-book ratio, on the other hand, may be a proxy for earnings growth since it refers to future expected return on equity [Penman, 1996]. Second, Fama and French [1992] and Lakonishok et al. [1994] find that the explanatory power of of earningsyield is contained within book-to-market [further: BTM].

Finally, we exploit the inherent relationship between the price-earnings ratio and the book-to-market ratio (or the inverse price-to-book, P/B) in order to construct sector-dependent benchmarks for the subsequent portfolio formation. Penman [1996] argues that a combination of P/E and P/B multiples contains information about future firm performance. Penman [1998] shows how the price of a stock can be calculated as a multiple of book value and earnin gs. The relationship between P/E and P/B is constructed by referring to the residual income valuation as an alternative to the more classical Gordon model [see Ohlson, 1995; Penman and Sougiannis, 1998]. We use these accounting insights in the portfolio construction procedure based on P/E and P/B. An important observation made by Penman [1996] is that the four combinations of (high and low) price-earnings and (high and low) price-to-book all occur in the market. In this paper, we use the one-year analysts' forecasts of P/E, but the difference in interpretation with the Penman results, which are based on realized P/E ratios, are minor.

This chapter

In this chapter, we explore the return characteristics of portfolios based on the earnings yield multiple and the BTM multiple for a large European sample of stocks. The objective is to see whether the value-growth strategy that is found to generate positive returns in - and outside the U.S. can also be found in sector samples where stocks are assumed to have the same characteristics because they have the same businesses. Consistent with institutional investor practice, we present our results in a European sector-based stock selection framework. In the period before the introduction of the euro currency, European portfolio selection has shifted from country-based to sector-based allocation strategies.

We first review some of the empirical findings from the large bulk of past research about the relationship between earnings yield, BTM and stock returns. In the next part of the chapter, we explore the theoretical relationship between earnings yield and expected return. We use the residual income valuation model [as in Ohlson, 1995] to show that, based on a number of assumptions, there is a direct relationship between the one year price-earnings forecast and the firm's cost-of-capital or expected return.

This paper investigates two hypotheses. The first hypothesis is that earnings yield and BTM are fundamental determinants of stock return performance. We find that this is the case in some European sectors. While the return difference between high earnings yield portfolios and low earnings yield portfolios is often not statistically significant, it is economically important in some sectors over the time period studied. More important is that this investment strategy is not unconditionally applicable to European sector samples. We observe time-varying behavior in the return differences in most sectors. To explain these differences, we focus on the fundamental risk explanation for this observation, based on asset-pricing models that were identified in part II and the standard consumption-based asset pricing model developed by Campbell and Cochrane [1999]. Hence, the second hypothesis is that any observed return difference can be explained by risk factors or by different states of the world (recessions, crashes, …) in a consumption-based setting.

Section 2 describes the theoretical framework to use earnings yield as an information variable in the modeling of expected returns and gives an overview of a sample of past empirical findings on the relation between the valuation multiples and returns. Section 3 presents the database and the regrouping of individual stocks into portfolios. Section 4 reports the main empirical results of the analysis of high versus low earnings yield stocks. Section 5 explores a fundamental risk story for differences in expected returns and the impact of economic state variables on the sector-specific performance differentials. Section 6 concludes.

2. Earnings yield, BTM and expected return: theory and empirical findings

2.1 Theoretical considerations

In this part, we give an overview of the important theoretical underpinnings of the relation between earnings yield and returns. Next, we will extend it with the theoretical relationship, explaining returns in a consumption-based setting. In this framework, the difference in return between high and low earnings yield portfolios is related to the state of the world. This setting allows us to investigate whether the difference in return can be explained by fundamental risk.

Dechow and Sloan [1997] argue that accounting multiples should reflect variation in required returns and growth rates. From the dividend discount model, and under the assumption of a dividend payout ratio of one, the earnings yield ratio equals the required return minus the growth rate (equation 1). In equation 1, E/P is the expected earnings yield ratio, r is the required rate of return and g is earnings growth.

[1] $$E / P = r - g$$

Practitioners use this simple 'assumed' relationship in stock selection or allocation procedures based on earnings-to-price [see, e.g., Strobaek, 1997]. The assumption made is that, since an asset-pricing model provides an estimate of the expected return, and since the E/P is a proxy for expected return (under some assumptions, see Part III a), the two can be used as an identity. Equation 2 shows this specification assuming that the cross-section of stock returns is described by one priced factor. In equation 2, a, b are the estimates of the regression[3] of the single-index market model, r_m is the market return.

[2] $$E/P = a + b\,[r_m]$$

We next motivate the identification of the E/P ratio as a proxy for expected returns without accounting for the growth rate g, and the link with different economic states of the world (the fundamental risk argument). The first step is a review of classic asset pricing models that describe expected returns in their most general form. This relationship describes the link between the risk perception in a state of the world and the performance of groups of stocks in these states in a consumption-based setting. In intertemporal equilibrium models, consumption and savings decisions of investors are related to asset prices. Consumption-based models describe investor behavior by a utility function (comparing current additional consumption with current additional investments). This leads to a general formulation of expected returns (equation 3). The interpretation is formalized by Campbell [1998] who finds that, in order to capture asset market behavior, the model used should take into account a market price of risk that is correlated with the state of the economy. He furthermore finds that the market price of risk, i.e. the degree of risk aversion, should be high and time-varying. This basic model is described in equation 3 where u indicates utility, c denotes consumption, r_f is the risk-free rate of return and r_i is the rate of return on asset i and e denotes expectation [Cochrane, 2001].

[3] We use OLS estimates.

[3] $$[r_i]^e = r_f - \frac{\operatorname{cov}\left[u'(c_{t+1}), r_{i,t+1}\right]}{\left[u'(c_{t+1})\right]^e}$$

The intuition behind this specification is that riskier securities must offer higher expected returns. These high expected returns are considered to be the high earnings yield stocks (see previous chapter). In a consumption-based framework, assets that exhibit a positive covariation with consumption make consumption more volatile. Hence, in a fundamental risk framework, high-risk stocks must offer a lower payoff in bad states of the world. Examples of these bad states are recessions, political crises, declining asset markets, etc. In all different states of the world, investors will try to smooth their consumption and invest in stocks that have a relatively low correlation with consumption. However, while this story coincides with the theoretical concept of a consumption-based asset-pricing model, very little supportive empirical evidence has been reported thus far. A different representation of equation 3 provides us with a formulation based on the discount factor m (equation 4) [Cochrane, 2001].

[4] $$\left[r_{i,t}\right]^e = \frac{1}{[m]^e} - \frac{1}{[m]^e}\operatorname{cov}(m, r_{i,t})$$

$$\text{where } 1/[m]^e = r_f \quad , \quad m = \frac{d\,u'(c_{t+1})}{u'(c_t)}$$

In equation 4, d is the subjective time discount factor. In equation 2, the covariance between the discount factor and the returns is modeled by the market model. A more general representation is given by the general concept that the discount factor is described by macroeconomic variables or variables that forecast macroeconomic events [Campbell and Cochrane, 2000]. Starting with Chen, Roll and Ross [1986], a wide range of models have been developed in which factors reflecting different states of the world are assumed to explain the structure of asset prices. In that case, m describes the marginal utility of an investor and is directly modeled by economic variables. In other words, $m = b'f$ with f a vector of variables containing information to describe the state of the world.

We have established that, relying on the consumption-based asset-pricing model, we expect different performance of different types of stocks in different states of the world because of their capacity to smooth consumption. The second step is to identify a theoretical relationship between earnings yield and expected return (as in equation 1). Equation 1 has the problem that the level of the earnings yield is subject to a set of stringent assumptions in order to serve as a proxy for returns.

The relevant task is to identify which assumptions are necessary for this identity to hold, which is done in the previous chapter. The eventual relation used is based on the information dynamics of abnormal earnings in the residual income valuation model as was reported in Dechow et al. [1999]:

$$\frac{E_{t+1}}{P_t} = r \qquad [5]$$

In equation [5], we do not have to assume that the dividend payout ratio is 1. A second improvement in this formulation is that the earnings yield multiple is no longer related to expected return and growth rates but to expected return only. Dechow et al. [1999] describe two conditions for this relationship. One is that future abnormal earnings are entirely based on the information in current abnormal earnings, when the other information is not all persistent. The second possibility is that other information is persistent, when abnormal earnings are not. In those two cases, there is a direct relation between earnings yield and returns. Hence, in this section we showed that, under some assumptions, there is a direct theoretical relationship between expected return and earnings yield. Moreover, in an intertemporal equilibrium framework, time-varying expected returns are driven by risk factors which describe the state of the world. These theoretical concepts allow us to test two hypotheses. First, we investigate whether or not there is a difference in performance between portfolios of stocks with low E/P or BTM characteristics and portfolios of stocks with high E/P or BTM characteristics. Second, we test whether or not the return spread between high and low yield portfolios is caused by a set of pervasive risk factors.

2.2 Past empirical findings

Cross-sectional regression tests

Given the large body of prior research on the subject, it is instructive to make a short evaluation of past findings about the relationship between earnings yield and stock return. This survey does not pretend to be exhaustive, we review some relevant findings from well documented papers. As in the entire chapter, we will focus on the earnings yield multiple rather than on BTM.

A first set of results can be drawn from studies which use OLS estimates to evaluate the relationship between returns and earnings yield. In most of these analyses, negative earnings are deleted from the cross-sectional earnings yield variable, because they do not proxy for expected return [Fama and French, 1992]. Instead, the regressions are performed adding a dummy variable with a value of one when a negative earnings yield is observed. Some of these papers also report evidence on the relationship between return and earnings yield after correcting for the book-to-market and size factor, two factors that are often found to partially explain the cross-section of returns.

Within this first group of papers, Fama and French [1992, henceforth FF] and Lakonishok et al. [1994, henceforth : LSV] are comparable in terms of time coverage, since they present US evidence for the period 1965 to 1990. Both papers study a cross-sectional relationship between lagged realized earnings yield and returns on an annual basis. An important difference between the two is that FF also include NASDAQ stocks in their analysis, next to the equities listed on the NYSE and AMEX. The reason why LSV do not include this sample is to avoid possible look-ahead bias. Both studies find a positive and significant univariate relationship between earnings yield and stock returns. The size of the estimated effect is, however, quite different. The regression coefficient in the FF study is 4.72 (t-statistic of 4.57), while the point estimate in LSV is 0.53 (t-statistic of 2.54). In addition, both studies correct their estimations for a possible missing variable problem. Next to the earnings yield variable, the negative earnings dummy, a size factor and a book-to-market factor. LSV also add a sales-growth factor. Correcting for these factors, FF no longer find a significant relationship between earnings yield and returns (t-statistic of 1.23), whereas the LSV paper still reports a significant effect (t-statistic of 2.01). The t-statistic for the BTM factor in these regressions is 4.46 for the FF study and 1.04 for the LSV analysis.

Two other relevant studies are La Porta [1996, henceforth LP] and Dechow and Sloan [1997, henceforth DS]. Both papers present results based on a dataset running from 1981 to 1992. LP limits his regression results to a sample which contains IBES, CRSP and Compustat data, while DS also include NASDAQ stocks. LP fails to find a significant relationship between returns and earnings yield (t-statistic of 0.57), whereas DS report a significant relationship between returns and earnings yield at the 1% level, be it that the estimated effect is small (the regression coefficient is 0.09). Furthermore, in both papers the univariate regressions are done with other explanatory variables related to earnings yield. Both papers report a comparable significant negative relationship between returns and long-term analysts' forecasts of earnings growth. Including this variable in the earnings yield regression in the DS paper does not change anything to the conclusions. LP also reports a univariate regression including analysts' forecasted earnings yield as an explanatory variable, without finding a significant relationship.

A comparison of these results leads to a rather inconclusive picture of the value of earnings yield as a factor explaining asset returns. Including the NASDAQ stocks in the analysis provides a significant univariate relationship, but solving the missing variable problem in the FF paper wipes this result out.

Return spread between high and low earnings yield portfolios

Next to regression analyses, there is a lot of empirical evidence given about the difference in return between high earnings yield portfolios and low earnings yield portfolios. FF, LSV, DS and Bakshi and Chan [2000] all report an annual return spread for high versus low earnings yield stocks ranging from 8.50% to 10% for the U.S. Notice that Bakshi and Chan use the contemporary relation between the one-year analyst forecast of earnings yield and returns, finding the same results for the U.S. as for the portfolios formed based on realized earnings yield. However, when they try to explain this spread, they come up with conflicting evidence for a naive investor hypothesis and little or no evidence for the fundamental risk hypothesis.

Bakshi and Chan [2000] expand the geographic coverage of their sample by including other countries, next to the U.S. Their conclusion is that the U.S. findings cannot be generalized to all other regions. For some countries they find an annualized return spread comparable to the one found for U.S. stocks (e.g. Austria, Finland, Ireland and Switzerland). In 26 out of 29 countries they find that a high earnings yield strategy produces a positive excess return, but often the spread is much smaller than in the U.S. Moreover, the observed positive excess returns are not significant in a lot of the cases. The lowest return spread was found in Sweden where the high earnings yield portfolios even performed worse than the low earnings yield portfolios (the difference is –3.43%).

Industry strategies

Dreman and Lufkin study 44 U.S. industries over the period 1970-1995. They find that, in general, there is a profitable contrarian strategy in U.S. industries. For earnings yield and BTM industry strategies, they find an annual excess return of about 4%. However, when they split the sample into five five-year periods, the earnings yield based contrarian strategy yields a significant positive return only in 2 out of the 5 sub-periods. The BTM based contrarian strategy yields a statistically significant positive return only in one of the five sub-periods.

The conclusion of this limited literature review is that the empirical evidence of the relationship between earnings yield, BTM and stock returns has not yet attained the status of a stylized fact. Not only is there mixed evidence concerning the universal existence of a return spread produced by an investment strategy based on earnings yield. But there is also no consensus about the underlying causes of any observed yield spread. Hence, the hypothesis that firm's earnings yields and BTM serve as a useful linear factor to explain the cross-section of their stock returns receives weak support at best. This does not, however, imply that the earnings yield factor should simply be ignored in quantitative investment strategies, but it does suggest that the dynamics behind this relationship may be complicated. It is e.g., conceivable that the assumed relationship behaves differently across sectors or varies with the business cycle or consumer sentiment.

3. Data and portfolio construction

Our data set consists of an intersection of two databases. For 17 European countries[4], all stocks from the Datastream local market index in December 1998 are collected. The sample consists of frequently traded stocks accounting for about 80% of the total market capitalization of each country. This universe of stocks is the focus of the coverage by the typical European institutional investor. Also, the inclusion of really small European stocks could induce false identification of results because these stocks are susceptible to infrequent trading. In order to reduce survivorship bias, we add a sample of dead stocks for each country. Stocks that merged, defaulted or were delisted prior to December 1998 are selected up to a total market capitalization of 80% (calculated at December 1998) of all dead stocks for each country. We finally obtain a sample of 2453 European stocks. From this list, preferred stocks are deleted as well as stocks listed on a stock exchange outside their home country[5]. The resulting sample contains 2427 European stock return series over the period January 1987 to December 1998 (retrieved from Datastream). For these 2427 stocks, we collect all price-earnings (P/E) one-year consensus analyst forecasts (FY1) from IBES. The intersection of the two databases for the period January 1987 to December 1998 contains a minimum of 821 stocks (June 1987) and a maximum of 1667 stocks (March 1998).

Two comments have to be made here. The first one is that the sample construction procedure leads to a similar sample as the one used by La Porta [1996]. An inherent drawback is that the coverage of firms by analysts is biased towards larger stocks. A second remark is that in our European sample, the coverage of firms by analysts was much smaller in the eighties than it was in the U.S. Consequently, the number of stocks with full information almost doubles over the time period under consideration.

All stocks are classified into a limited number of homogeneous sectors. In order to obtain a stable aggregation of industries into sectors over time, we construct the relevant sectors based on two industry classification methodologies, the FT[6]-indices and the Stoxx indices. Both index providers regroup industries into sectors and combining the two produces 14 relevant sectors (see appendix 1, Part II). All similar industry groups from both methodologies are kept as one sector. The remaining industries are allocated to one of those sectors according to the FT-typology. We use the FT industry code of all stocks to classify them into one of the 14 sectors. In this way, the sectors remain stable through time, even if the

---

[4] The 17 countries are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK.

[5] E.g. Unilever is listed in the Netherlands and in Belgium, the Belgium data are deleted.

[6] FT = Financial Times indices at the end of 1998.

index providers alter their regrouping. This is the same sector classification as presented in part II.

Returns are expressed in synthetic euro, calculated as a GDP-weighted average of the currencies of the countries that entered the third stage of EMU in 1999. The currencies of the non-EMU countries in the sample were converted to the synthetic euro. The book-to-price (B/P) ratios are taken from Datastream. Note that the calculation of the book value in the Datastream database excludes intangibles, which sometimes causes the book value to be negative. Therefore the rankings of the stocks are based on book-to-market (B/P), and not on market-to-book, because the negative book values may lead to a distorted ranking. If BTM proxies for earnings growth, excluding the intangibles could enforce this relation. If two stocks have a book value of 500 and a market value of 1000, they both have a BTM ratio of 0.5. If the first company has intangibles worth 200 and the second company 100, the BTM of the first company becomes 0.3 and the BTM ratio of the second becomes 0.4. The first company will have a higher chance to be classified as a growth company given its lower BTM ratio. If a high level of intangibles such as a lot of goodwill and R&D implies higher growth opportunities, the exclusion of intangibles is beneficial to this analysis.

B/P ratios and P/E ratios are matched by date in Datastream and IBES. In Chapter II, we tested the impact of using B/P values that are possibly backfilled against lagged B/P values. The impact turns out to be negligibly small, indicating that past B/P values are informative for future B/P values. This is consistent with the Fama and French [1995] finding that B/P ratios are very persistent [see also Lewellen, 1999].

Table 1 gives a summary of the data used in this paper. Panel A of table 1 shows the difference in relative weights (in %) for the sector portfolios. The full names of all sector portfolios are listed in appendix 1. Sector market capitalization is reported at two points in time, December 1988 and December 1998. In some studies, financial firms are excluded from the analysis because of their high leverage property. In this study, the financial sectors (banks, insurance and financial services) are tested separately. According to panel A, these three sectors account for over one fourth of the European market capitalization. Panel B presents the first two moments, both equally weighted and market-capitalization weighted, of the time series for all regrouped portfolios. There is a difference (although not statistically significant) both in terms of average returns and volatility of European sector returns.

Although the issue is not explicitly tested in this study, the comparison of market-capitalization weighted and equally weighted sector returns suggest the existence of a size effect in European sector returns, be it that the effect differs across sectors. E.g., the market-capitalization weighted monthly return for the resource portfolio averages 1.15% while the equally weighted return is 0.75%. In

Table 1.

Characteristics of the sector samples

Panel A shows the relative market capitalization – in % - of all sectors[7] at two points in time, December 1988 and December 1998. Panel B shows the average returns and standard deviations of the return series (in %) for each sector over the entire sample. These characteristics are calculated both market capitalization weighted and equally weighted. The sample period is January 1987 to December 1998. Panel C shows the average level of equally weighted analysts' price-earnings forecast for 14 sectors in two sub-periods of 60 months. Panel D shows the average ratio of book-to-price for each sector. Panel E shows the number of observations at one point in time per sector as a result of the intersection of the European Datastream database and the IBES database as well as the total number of stocks in a sector.

	BANK	BMAT	CHEM	CYCC	CYCS	FINS	INDU	INSU	NCYC	PHAR	RESO	TECH	TELE	UTIL
Panel A. Sector market capitalization (%)														
12-1988	12.2	3.9	5.3	4.2	14.3	5.0	7.8	9.2	11.6	5.7	8.8	5.5	3.7	2.8
12-1998	15.7	2.6	2.8	3.7	11.8	3.2	7.7	11.0	8.7	8.4	6.6	5.0	8.4	4.4
Panel B. Sector average returns and standard deviations; 1987-01 / 1998-12 (%)														
Cap. Weighted	0.84	0.44	0.76	0.37	0.77	0.80	0.78	0.78	1.07	1.52	1.15	0.73	1.31	1.33
	(5.59)	(5.91)	(5.25)	(5.90)	(5.13)	(4.84)	(5.37)	(5.54)	(4.73)	(5.34)	(5.45)	(5.74)	(5.51)	(4.04)
Equally weighted	0.93	0.72	0.73	0.51	0.92	0.92	0.67	0.84	0.85	1.17	0.75	0.84	1.29	1.26
	(4.42)	(6.01)	(5.34)	(5.30)	(5.09)	(5.00)	(4.99)	(5.22)	(4.35)	(4.66)	(5.76)	(5.61)	(5.45)	(3.42)
Panel C. Average analysts' sector one-year price-earnings forecasts in the first and the last five-year period of the sample (P/E, FY1)														
1987-1991	11.51	6.78	13.46	7.61	17.41	15.22	16.36	21.28	15.57	15.52	15.13	15.47	13.02	16.98
1994-1998	16.32	16.33	19.37	18.81	21.03	24.15	18.01	27.54	13.37	14.78	17.20	24.47	15.52	16.38
Panel D. Sector book-to-price levels (B/P)														
1987-1991	0.94	0.80	0.60	0.78	0.58	1.11	0.65	0.70	0.57	0.55	0.59	0.67	0.87	0.83
1994-1998	0.81	0.67	0.52	0.52	0.43	0.98	0.51	0.58	0.48	0.27	0.56	0.51	0.52	0.49
Panel E: Number of observations per sector														
Min. Number	49	74	37	33	182	53	75	47	104	21	24	92	6	18
Max. number	145	138	53	76	385	102	155	92	196	50	40	166	23	65
Total number	214	182	71	108	492	229	231	122	280	77	66	207	42	85

[7] The full names of the sectors are: resources, basic materials, chemicals, cyclical consumer goods, non-cyclical consumer goods, pharmaceuticals, cyclical services, banks, insurances, financial services, industrials, technology, telecom and utilities.

the basic materials sector, on the other hand, the market-capitalization weighted average monthly performance is 0.44% versus 0.72% when the portfolio is equally weighted. Panels C and D contain the sector values for the P/E and B/P variables averaged over two sub-periods of 60 months, the first and the last period of the sample. These numbers indicate that European sectors are characterized by different average values for the accounting fundamentals. For example, the average P/E (FY1) ranges from a minimum of 6.78 (basic materials) to a maximum of 21.28 (insurance) in the first sub-period, and from 13.37 (non-cyclical consumer goods) to 27.54 (insurance) in the second. Moreover, the changes over time are not uniform. For example, the average price-earnings forecast for the cyclical consumer goods sector increases from 7.61 in the first period to 18.81 in the second, while the P/E forecast for the non-cyclical consumer goods sector decreases from 15.57 to 13.37. Similar observations can be made for the B/P-variable. Most important is that, especially for the P/E ratio, the sector characteristics evolve different in time. In some sectors there is an increase in average earnings yield forecasts, while in others they stay the same and yet in others they decrease. These different dynamics motivate the sector approach of this problem based on forecasted earnings yield.

Finally, panel E lists the number of stocks available in each sector as the result of combining the stock and P/E databases. The first row of panel E gives the minimum number of stocks in each sector with full data coverage over the entire sample period. While most sectors have a reasonable coverage to perform the analyses, the minimum number of stocks in the telecom, utilities and pharmaceutical sectors is relatively low. Nevertheless, these sectors are treated as separate entities because of their importance in terms of market capitalization and their use in sector-based portfolio allocation strategies by institutional investors. The data for the maximum number of stocks indicate that only the telecom sector may be somewhat problematic in terms of coverage, but panel B suggests that the number of individual stocks is unrelated to the return volatility of this portfolio. A general observation is that analyst coverage steadily increases over the sample period, with the clearest upward trend in the banking sector and the utilities sector.

A final remark is that given the number of stocks that are in the dataset, the three financial sectors exhibit the most extreme values for their sector fundamentals. Insurance is the sector with the highest price-earnings forecasts, while banks and financial services produce the highest B/Ps (high leverage stocks). The total number of stocks in each sector sample is always higher than the maximum number of stocks that is comprised in the analyses. This is because only fully available observations are used. It implies that stocks with no earnings yield forecast or no BTM or no return are excluded as well as stocks with negative valuation multiples.

Studying the cross-section of European accounting information

A last topic we want to address before showing the results is the use of the cross-sectional European accounting data. Some readers may question the impact of different accounting principles across the European Union leading to different accounting multiples for the same intrinsic accounting numbers and the same valuation method for two stocks in two different European countries. Garcia-Ayuso et al. [1998] study this topic. They analyze the value relevance[8] of accounting information in the capital markets of the European Union. We will shortly summarize some of their findings here.

First, these authors report different multiples across European countries. Important for this study is that they find no explanatory power for the hypothesis that these differences are caused by a different degree of accounting conservatism across European countries. They suggest that their results imply that there are differences in the positive investors' market expectations about the companies' future performance across countries. Since this statement means that there is a difference in the degree of extrapolation of present growth into the future across countries, this falls within the research question of this chapter and rankings based on accounting multiples across countries with this underlying factor explaining the differences between the multiples does not change the results of the analysis. The finding that the differences are not due to a different degree of accounting conservatism across European countries makes that we can worry less about the impact of these kind of differences for our analyses. Moreover, Ashiq and Hwang [2000] find that the different country-specific variables (such as conservatism or market-oriented versus bank-oriented financial systems) determining the value relevance of accounting information are very much interrelated.

Second, Garcia-Ayuso et al. [1998] give an example of the impact of cross-country differences in accounting principles. They mention that goodwill is immediately written off in the UK and not in other countries. This implies that earnings and book values would respectively be higher and lower for the UK. Moreover, Adams et al. [1998] report that this is the main factor determining differences in earnings and book values between UK and US GAAP. In this chapter, intangibles (such as goodwill) are excluded from the book values. Therefore, we are comfortable with the cross-sectional ranking based on the book-to-market multiple.

Finally, the cross-section of accounting multiples in this analysis has the same relation with returns before and after country-demeaning. This implies that the average level of the multiple for each country does not have a large impact on the results.

[8] explanatory power of accounting variables for security returns

4. Empirical Findings

Portfolio selection procedure, Earnings yield, BTM

For the empirical analysis we need to allocate all the stocks in the sample to a high versus low earnings yield or BTM portfolio. In order to study the returns on contrarian strategies in European sectors, we form a portfolio of high valuation multiple stocks and a portfolio of low valuation multiple stocks. Because we focus on earnings yield forecasts in this chapter, we monthly rebalance the portfolios mimicking an active portfolio strategy based on the valuation multiples. Each month, all stocks within one sector are ranked according to their valuation multiple.

We also ranked stocks within one sector according to their valuation multiple demeaned by the country average of the valuation multiple. We did so in order to reduce the possible differences in accounting rules across European countries. We observed no significant shifts in the results allowing us to study the European cross-section of absolute values of the multiples[9].

Stocks are ranked based on earnings yield and BTM implying that the portfolio of low valuation multiples is the growth portfolio and vice versa. Each month, the cross-section of stocks within each sector is split into three sub-samples: the bottom and top 40% of all stocks and a mid-section containing 20% of the cross-section. Each month, the equally weighted return of the bottom section (growth stocks) is calculated as well as the equally weighted return of the top section (value stocks). We also calculated the return of contrarian strategy based on earnings yield and BTM for the complete cross-section of European stocks. The annualized return on the strategy based on the earnings yield multiple is 1.71% with a p-value of 0.270. Over the period 1987-1998, a cross-sectional earnings yield strategy was not profitable. The annualized return on the strategy based on the BTM multiple is 5.13% with a p-value of 0.001, whixh means that this strategy is profitable.

Table 2 reports the characteristics of the European contrarian sector strategies based on earnings yield forecasts and BTM. European contrarian sector strategies based on the earnings yield multiple are in general not profitable. Only in four of the fourteen cases (banks, insurances, financial services and utilities) this strategy earns a significantly positive excess return (at the 5% level). It is somewhat surprising that we find a significant excess return especially in the financial sectors. The highest annualized return found for this strategy is 12.91% (insurance) and the lowest return is even negative (-6.42% for basic material stocks).

The characteristics of the returns on a contrarian strategy based on the BTM multiple show a different picture. In nine out of fourteen sectors, this strategy is

[9] Results are available upon request

profitable at the 5% level. The highest annualized return is found in the sector of resources stocks and is 12.96%. There is only one (but highly insignificant) negative return, found in the sector of pharmaceutical stocks (-5.34%).
We also observe that the volatility of the time-series of contrarian returns is sometimes high. This implies that in some sectors, there is a reasonable downside potential for these value-growth strategies.
With respect to the first hypothesis, we find that the value-growth strategy is not unconditionally applicable to European sector strategies. Especially for strategies based on the earnings yield multiple, there is no statistical evidence for the existence of a value premium. Strategies based on BTM generally do yield significant returns. However this is not the case in all sectors.

Earnings yield and BTM

Next, we also study the returns on a contrarian strategy using the relation between earnings yield and BTM for reasons described in the introduction. In order to judge whether a stock's earnings yield is high or low at a specific point in time, given its affiliation to a sector, we need to define a cut-off value. One possibility would be to use the average P/E forecast. However, this portfolio construction procedure would be at odds with the Penman [1996, 1998] arguments that P/E and P/B are inherently related and that all combinations of high/low P/E and P/B are observed in the market. Hence, we exploit the theoretical relationship between P/E and P/B to construct a conditional benchmark value of P/E based on the pricing dynamics in each sector over the sample period. To that extent, we estimate the relationship between the P/E ratio and the P/B variable in each sector in order to discriminate between expected changes in future abnormal earnings and the P/E relative to the current P/B level. We estimate the pooled-cross section relationship between P/E and P/B over consecutive periods of 6 months (from t-6 to t-1)[10]. This relationship is estimated monthly for each sector, yielding a time-varying coefficient of sensitivity of the P/E yield to the stocks' B/P for the period t-6 to t-1. Inserting the observed value of P/B of stock *i* at t-1 provides a conditional (on the P/B level) benchmark value for the earnings yield for each sector in each month. If the actual forecasted P/E level is lower (higher) than this benchmark value, we classify stock *i* in the portfolio of high (low) earnings yield stocks S_1 (S_2) at the end of t-1. This procedure is repeated monthly for each stock in each of the 14 sectors. Expression [6] summarizes the allocation procedure:

$$\frac{P_t^i}{E_{t+1}^i} < E\left[\frac{P_t^s}{E_{t+1}^s}\middle|\frac{B_t^i}{P_t^i}\right] \text{ then } \mathrm{i} \in S_1 \qquad [6]$$

[10] As a consequence, the empirical results are reported from July 1987, and not January, onwards.

Table 2

Characteristics of the return on European contrarian sector strategies

For the fourteen sectors, the time-series characteristics of a zero-cost investment buying value growth and selling growth stocks are reported. The first row of each panel displays the monthly excess returns on this zero-cost strategy in basis points. The second row shows the annualized return on the contrarian sector strategies (in %). The third row gives the standard deviation of the time-series of monthly returns on the value-growth strategy. The final row of each panel shows the p-value for a two-sided test of equal returns on the value portfolio and the growth portfolio.

	BANK	BMAT	CHEM	CYCC	CYCS	FINS	INDU	INSU	NCYC	PHAR	RESO	TECH	TELE	UTIL
	Returns on contrarian strategies based on the earnings yield multiple													
Spread BP.	65	-55	21	10	9	64	1	102	4	-9	44	-16	58	99
Spread An.	8.08%	-6.42%	2.56%	1.17%	1.05%	7.95%	0.12%	12.91%	0.52%	-1.10%	5.46%	-1.93%	7.18%	12.48%
Stdev.	2.67	3.79	3.04	3.04	2.07	3.33	2.21	3.69	2.22	3.19	3.69	2.89	5.97	4.67
p-value 5%	.005	.090	.416	.709	.624	.025	.956	.002	.820	.735	.160	.511	.256	.020
	Returns on contrarian strategies based on the BTM multiple													
Spread BP.	50	79	91	61	34	53	36	77	37	-46	102	36	59	91
Spread An.	6.16%	9.85%	11.47%	7.52%	4.09%	6.57%	4.43%	9.69%	4.54%	-5.34%	12.96%	4.35%	7.24%	11.49%
Stdev.	3.22	3.79	3.41	3.43	1.92	1.95	2.64	2.89	2.18	9.92	4.14	2.94	5.11	3.34
p-value 5%	.071	.007	.002	.040	.043	.002	.110	.002	.047	.590	.004	.159	.181	.002

Table 3

Characteristics of the return on European contrarian sector strategies based on earnings yield and BTM

For the fourteen sectors, the time-series characteristics of a zero-cost investment buying value growth and selling growth stocks are reported. The first row of each panel displays the monthly excess returns on this zero-cost strategy in basis points. The second row shows the annualized return on the contrarian sector strategies (in %). Th ethird row gives the standard deviation of the time-series of monthly returns on the value-growth stra tegy. The final row of each panel shows the p-value for a two-sided test of equal returns on the value portfolio and the growth portfolio.

	BANK	BMAT	CHEM	CYCC	CYCS	FINS	INDU	INSU	NCYC	PHAR	RESO	TECH	TELE	UTIL
	Returns on contrarian strategies based on the earnings yield and BTM multiples													
Spread BP.	47	-33	17	-40	-2	76	-9	63	-10	-15	63	-11	21	43
Spread An.	5.82%	-3.93%	2.04%	-4.74%	-0.20%	9.51%	-1.09%	7.83%	-1.14%	-1.79%	7.84%	-1.27%	2.53%	5.25%
Stdev.	2.04	3.02	3.15	2.83	1.85	2.87	2.36	3.22	1.96	3.56	4.66	3.24	7.98	3.88
p-value 5%	.007	.198	.531	.096	.916	.002	.652	.023	.568	.621	.114	.700	.760	.198

else $i \in S_2$

In expression [6], P^i/E^i is the FY1 analysts' consensus forecast for stock i at the end of the previous month and P^i/B^i is the book-to-price value for that stock at the end of the same month. P^s/E^s is the expected earnings yield level for a sector, conditional on the observed P/B level of stock i. In the estimations, we trim the data series in order to exclude extreme values for the two variables. All observations outside the confidence interval are discarded from the pooled cross-section. The use of this procedure implies that negative values for B/P or P/E may be included in the estimation of the regression. This is intended to maximize the robustness of the regression because it should provide a reliable sector-specific conditional benchmark level of the earnings yield, used to classify stocks into high and low earnings yield portfolios. However, in the composition of the portfolios S_1 and S_2, negative earnings forecasts are excluded because they have been shown to be unrelated to expected returns [see Fama and French, 1995; Bakshi and Chan, 2000].

In this section, we again test the hypothesis that portfolios for sectors in which the dynamics in forecasted earnings yield differs, exhibit different ex-post realized returns. We study the return characteristics of the portfolio with conditionally high earnings yield properties S_1 relative to the portfolio with conditionally low earnings yield properties S_2. Table 3 presents the mean monthly difference in return (in basis points) between an investment in the S_1 portfolio versus one in the S_2 portfolio per sector. In general, the results in table 3 confirm the previously reported finding that the outcome of a forecasted earnings-based investment strategy is sector-dependent. For 7 of the 14 sectors, we find a positive excess return, but most are not significant at conventional levels. The only statistically significant out-performance (at the 5% level) is again found in the three financial sectors (banks, insurance and financial services), with p-values for a test of zero spreads of .007, .023 and .002 respectively. The largest difference is recorded in the financial services sector, with an annualized excess return of 10%. Also in the sector of resources, an economically significant return is observed of 8.08% per year. However, seven of the 14 sectors display negative excess returns, although none is significant at the 5% level. These results confirm previous findings by Bakshi and Chan [2000] and they suggest that a trading strategy based on P/E multiples yields ambiguous results across European sectors.

5. A risk story for the profitability of the value-growth strategy

5.1 Are value-growth strategies always profitable?

If contrarian strategies are profitable, they should be profitable in any period. We perform a very simple analysis to see whether the returns on contrarian strategies are the same through time. Because we selected the value and growth portfolios mimicking an active portfolio strategy with monthly rebalancing of the portfolio, we simply test whether the return on this strategy is the same in three different time periods. We select the first and the last period of 45 months (July 1987 until March 1991, and April 1995 until December 1998). We also select the five-year period in the middle of the studied time period (May 1991 until January 1995). Deleting the intermediate months assures the independence of the sub-samples.

Next we test whether the mean monthly return is the same for an investor applying these contrarian sector strategies in the three sub-periods using a Wald-test[11]. First, we tested whether the variance of the returns in the sub-periods are equal using a Bartlett test. This is the case in only 24 out of 42 tested strategies. It is however interesting to see that in only one case, variances are not equal using BTM as the one ranking variable.

Table 4 displays the results of the Wald test for equality of the means as well as the annualized return for a contrarian strategy in the three sub-periods within a sector. Let us first turn to the cases where the value-growth strategy is profitable. For the fourteen sectors and three ranking strategies that we analyze, we find that in 17 cases there is no indication that the contrarian strategies are not overall profitable. One example is the sector of financial services where we find an annualized excess return of more than 6.5% for all three cases. The return in the sub-periods is in all cases positive for this contrarian sector strategy. For four sectors, the return on any contrarian strategy is positive. These sectors are: financial services, insurances, resources and utilities.

However, we find that in other sectors there are substantial differences in the profitability of the value-growth strategy. A good example is sector of cyclical services, which is the sector with the highest number of stocks. The difference in return between the sub-periods is important in economical terms. In table 2, we reported an annualized excess return of over 4% for the contrarian strategy based on BTM. In the first sub-period, the annualized return is 12.72% and in the last period –7.38%. There are several other examples of this kind of differences in return between the sub-periods.

---

[11] The Wald-test is specified as follows: $(\underline{x}-\underline{m})'\,\underline{\Sigma}^{-1}(\underline{x}-\underline{m})$ where $\underline{x}$ is the vector of means of the sub-samples, $\underline{m}$ is the population mean multiplied by the unit vector, $\underline{\Sigma}$ is the variance matrix of the time-series for the sub-periods. Because we do not know the population mean, we use the mean of the entire sample instead. Therefore, the Wald test statistic has a c^2 distribution with t-1 degrees of freedom, where t is the number of sub-samples.

Table 4

Annualized return for a 45 month investment in a contrarian strategy and the Wald test for equality of means

In table 4, the first three lines of each panel shows the annualized return (in %) of applying a zero-cost contrarian strategy. The fourth line shows the value of the Wald test statistic testing for equality of the mean monthly return in the three sub-periods. The test statistic has a χ^2 distribution with t-1 degrees of freedom. At the 5% level, the critical value is 5.99. The shaded areas in the table indicate whether the Wald test statistic implies that the means are not equal over the sub-periods.

	BANK	BMAT	CHEM	CYCC	CYCS	FINS	INDU	INSU	NCYC	PHAR	RESO	TECH	TELE	UTIL
Contrarian strategy using earnings yield forecasts as a ranking variable														
87-91	13.75	-13.31	1.75	1.53	5.11	2.75	4.80	15.25	8.58	-1.05	4.31	0.94	3.90	11.51
91-95	-1.77	-1.03	-1.21	5.15	4.11	12.43	-1.31	12.16	-2.27	-2.05	12.01	4.22	15.51	14.46
95-98	11.47	-5.64	5.95	-1.82	-5.93	8.77	-2.94	11.05	-4.08	4.99	1.17	-11.44	-2.61	12.70
Wald	6.02	2.54	0.80	1.17	7.20	0.95	2.15	0.27	5.59	0.29	1.55	7.70	1.58	0.04
Contrarian strategy using BTM as a ranking variable														
87-91	13.53	8.60	15.39	8.73	12.72	14.39	3.24	6.81	7.24	6.73	7.77	12.42	8.22	17.70
91-95	8.25	19.48	12.57	15.36	8.51	4.79	9.77	16.10	9.52	-19.09	30.95	8.77	16.41	4.33
95-98	-1.72	3.77	8.04	1.39	-7.38	1.37	1.59	6.38	-1.74	-0.87	1.53	-6.47	0.51	12.11
Wald	3.43	3.10	0.52	2.58	25.75	7.32	1.40	1.93	5.12	3.50	7.43	11.09	1.61	2.05
Contrarian strategy using earnings yield forecasts and BTM as ranking variables														
87-91	7.61	-16.58	2.79	-14.97	1.94	13.42	4.61	6.13	5.67	3.42	5.46	-4.47	20.75	0.11
91-95	0.64	9.01	2.34	4.69	3.54	9.87	-4.40	6.00	-1.95	-4.80	12.43	12.97	5.08	5.08
95-98	7.98	-3.59	-2.11	-4.03	-5.83	5.67	-2.83	11.41	-6.87	-3.90	6.38	-12.22	-18.55	4.31
Wald	2.77	12.37	0.46	7.19	6.70	1.45	0.61	0.61	7.05	0.92	0.47	18.28	3.42	1.26

We also find statistically significant differences in returns from a strategy within a sector. In table 4, there are 12 cases where the Wald test indicates that the mean monthly return is different over the sub-periods. For the sectors of cyclical services and the technology sector (both sectors where a large number of stocks are analyzed) we find statistically significant differences in return for the three types of contrarian strategies.

All this indicates that value-growth strategies cannot unconditionally be applied in all European sectors. Moreover, this economical and statistical evidence suggests that, in most cases, returns from value-growth strategies are time varying. If a market anomaly such as the existence of a value premium can be explained by irrational behavior of market participants, the time-varying returns imply that the degree of irrationalities also vary over time. It is difficult to analyze this statement. Furthermore, given the theoretical link between the accounting fundamentals and returns outlined in section 2, we will assess the question whether these time-varying returns can be explained by risk factors. We analyze this second hypothesis in the next sections.

5.2 Factor models

Past literature has given a lot of attention to the explanation of patterns in the cross-section of international stock returns [Heston et al., 1995, Fama and French, 1998]. In order to test the cross-section of expected stock returns for this sample, we use the approach described by Fama and French [1998], assuming that factor loadings and risk premia are not time varying. We suggest that two asset-pricing models are relevant for the description of the expected stock returns[12]. Following the results from chapter II with respect to the exact factor pricing relation for European sector portfolios, we select a two-factor model and a static ICAPM where we account for the return on human capital in the description of the wealth portfolio. Hence, this wealth portfolio and the linear combination of the two factor portfolios are mean-variance efficient if these factor models explain the cross-section of expected returns.

We estimate [using OLS as in Fama and French, 1998] two factor models for the 28 sector portfolios for each of the three contrarian strategies. The first factor model is a two-factor model with the market portfolio and the momentum factor portfolio (with LMOM as the momentum factor portfolio as described in part II).

[7] $$R - F = a + b\,(M - F) + g\,LMOM$$

The second model is a static ICAPM described by a wealth portfolio where capital income and labor income (R_{LI}) are included.

[12] We also estimated the Fama-French [1998] two-factor model for global value-growth portfolios. Results are comparable to the estimations for the wealth portfolio in this paper. The p-value for the total period is 0.044.

$$R - F = a + b(M - F) + d\,R_{LI} \quad [8]$$

If a linear combination of one of the two sets of explanatory factor portfolios is mean-variance efficient, the vector of intercepts should be zero. Therefore we apply the Gibbons-Ross-Shanken multivariate test (GRS-test) described in chapter II. Table 5 shows the results of these tests for the whole period (1987:07-1998:12) and for two sub-periods (1987:07-1993:03 and 1993:04-1998:12). We look at two sub-periods as well, because the results for exact factor pricing for sector portfolios showed that there is a possibility that there are time-varying factor loadings and risk premia.

Table 5.

Statistics of the GRS-test on European sector value and growth portfolios

Average $|a|$ denotes the average absolute alpha for the 28 regressions and NCP denotes the non-centrality parameter. F-stat stands for the F-statistic for the GRS-test, with degrees of freedom 28;108 for the whole period and 28;39 for the sub-periods. The shaded areas indicate whether the ex ante mean-variance efficiency of the factor portfolio cannot be rejected (using the asymptotic chi-square (critical value of 41.34) test for the NCP and the finite-sample test reported by the p-value.

Contrarian strategy using earnings yield forecasts as a ranking variable					
		average $\|a\|$	NCP	F-stat	p-value
M-F, LMOM	87-98	.0065	102.95	2.920	.000
	87-93	.0060	130.00	2.673	.003
	93-98	.0070	113.49	2.384	.006
M-F, R_{LI}	87-98	.0064	102.08	2.895	.000
	87-93	.0064	152.83	3.143	.001
	93-98	.0071	101.89	2.141	.014
Contrarian strategy using BTM as a ranking variable					
		average $\|a\|$	NCP	F-stat	p-value
M-F, LMOM	87-98	.0075	107.62	3.052	.000
	87-93	.0076	140.97	2.899	.001
	93-98	.0070	106.83	2.224	.010
M-F, R_{LI}	87-98	.0072	71.33	2.023	.005
	87-93	.0076	69.80	1.435	.149
	93-98	.0072	102.17	2.146	.013
Contrarian strategy using earnings yield forecasts and BTM as ranking variables					
		average $\|a\|$	NCP	F-stat	p-value
M-F, LMOM	87-98	.0025	74.34	2.089	0.004
	87-93	.0035	97.69	1.979	0.025
	93-98	.0030	68.08	1.379	0.177
M-F, R_{LI}	87-98	.0025	64.85	1.822	0.015
	87-93	.0039	78.03	1.575	0.094
	93-98	.0037	62.84	1.269	0.242

Table 5 shows that for the entire period, we find evidence against mean-variance efficiency of a linear combination of the factor portfolios for both models (p-values between 0.000 and 0.015) and for the three strategies. Using earnings yield as a ranking variable, we find evidence against ex ante mean-variance efficiency of a linear combination of the factors in all cases. However, for the two other strategies, we find no evidence against ex-ante mean variance efficiency in three of the four sub-periods using the finite-sample test (p-values ranging between 0.013 and 0.242). This indicates that the returns on value-growth strategies are partially priced by these factor models, especially by the static CAPM using the wealth portfolio. But for the asymptotic test, we never find evidence of the ex ante efficiency of the factor portfolio. Combined with the results for the entire period, we get an indication that it is possible that there are time-varying components to this relation. Therefore, we conduct a second fundamental risk analysis to further explore the hypothesis that differences in expected returns exist because of fundamental risk.

5.3 A consumption-based approach

We conduct the second analysis using the dynamics of a the standard consumption-based asset pricing model. According to this model, a representative risk-averse investor will try to smooth consumption. Hence, the prediction is that this investor will prefer low risk stocks in bad states of the world (as was outlined in section 2).

Most empirical verifications of this standard model have, however, failed to corroborate this prediction. The main reason is that the covariance between stock returns and consumption (the quantity of risk) is empirically found to be low. This implies that the degree of risk aversion (price of risk) has to be very high if the standard model would hold. This phenomenon is described in the literature as the equity premium puzzle [Mehra and Prescott, 1985]. However, recent evidence indicates that the standard model cannot be rejected when consumption is measured as the deviation from the investor's consumption habits. Campbell and Cochrane [1999] find that, as consumption comes closer to habit in periods of recession, the prices of high-risk stocks will decline and hence expected returns will rise. As a consequence, stocks with a high expected return will be regarded as high-risk stocks. The instantaneous returns on these high-risk stocks should be lower in bad states of the world if this risk story holds [Lakonishok et al., 1994].

According to the standard model, these stocks will tend to make it difficult to smooth consumption in bad states of the world and will yield lower returns in those periods. A testable implication of this hypothesis is whether or not the observed return spreads for different sectors coincide with changes in the perception of consumption and indirectly, with the marginal utility of wealth. Returns for high-risk stocks should be found to be lower when this marginal utility of wealth is high. An example to clarify this statement in this chapter is that since we find that value stocks generally outperform growth stocks in 17 of the 42 analyzed cases (as outlined in section 5.1), the return on these zero-cost strategies should be lower in the states of the world where the marginal utility of wealth is high. Campbell and Cochrane [1999] describe the relation between this relationship between macro–economic variables and financial markets. Hence, if these macro-economic variables indicate a state where the marginal utility of wealth is high, returns on the strategy should be lower.

In the first place, if consumption declines towards the habit in a business cycle trough, risky stock prices will fall and expected return will rise. Second, Campbell and Cochrane specify different characteristics for this relation. The first one is that aggregate consumption is more important for the individual habit level than the past consumption of the individual. Second, the habit level is persistent because its reaction to consumption is slow. Finally, there is a non-linear reaction of the habit level to consumption.
The model proposed by Campbell and Cochrane [2000] is an example of such a framework. Testable forms of the standard consumption-based model are obtained when the marginal utilities are modeled to depend on economic variables directly. Campbell and Cochrane [2000] propose the following model for the stochastic discount factor:

[9] $$m_{t+1} = \ln(d) - g * (\Delta s_{t+1} + \Delta c_{t+1})$$

In equation 9, m is the stochastic discount factor, d is the time discount factor, Δs is the log change in the surplus consumption to habit consumption ratio and Δc is the log change in consumption. Given the relation between the discount factor and returns of any portfolio on the mean-variance efficient frontier, we use this model to test the consumption model for European data.

We examine this pricing kernel empirically by assuming that g is a linear function of macro-economic variables. This requires data that reflects the consumer sentiment as a deviation from a habitual spending pattern (s in equation 9). We use two measures to capture this variable. The first is the European Commission consumption survey data measuring the seasonally adjusted major purchases (CONS). The survey is conducted with EU residents and reflects whether the present consumption level is judged to be above or below habit consumption, with a zero-level reflecting a neutral position. Hence, recorded changes in this measure capture changes in consumer sentiment. If consumer sentiment improves, an increase of ΔCONS should covary positively with the return spread, favoring high-risk stocks.
The second variable measures expectations with respect to the business cycle. We use the log change of the composite leading indicator for the European Union (ΔCLI). An increase of ΔCLI indicates that the average European consumer

expects that future economic conditions will improve. Hence, the consumption surplus is expected to be higher in the future than the currently observed level. In this case, the current return on high-risk stocks will be lower. As a consequence, a positive change in this variable should be negatively related to the present return spread recorded for a zero-cost investment in high-risk relative to low-risk stocks.

The second part of equation 9 (Δc) is approximated by the log change in aggregate volume of retail sales in the European Union (ΔRETAIL). Increasing contemporary consumption should be correlated positively with the return spread.

Of course, in this paper, the question is more general than explaining the return spread between high- and low-risk stocks. In 17 cases, the excess return from a value-growth strategy is always positive and hence we can test whether this return can be attributed to risk. In the other cases, the return on the contrarian strategies varies over time and the question here becomes whether these factors can explain this time-varying behavior. In other words, we analyze whether the perception of risk is different over the sectors for the different ranking variables. An investor can judge the high earnings yield forecast stocks to be risky when the sentiment is that consumption is close to habit. At the same time, this investor can judge (for the same level of perceived surplus consumption) that the high earnings yield forecast stocks are not risky in a different sector.

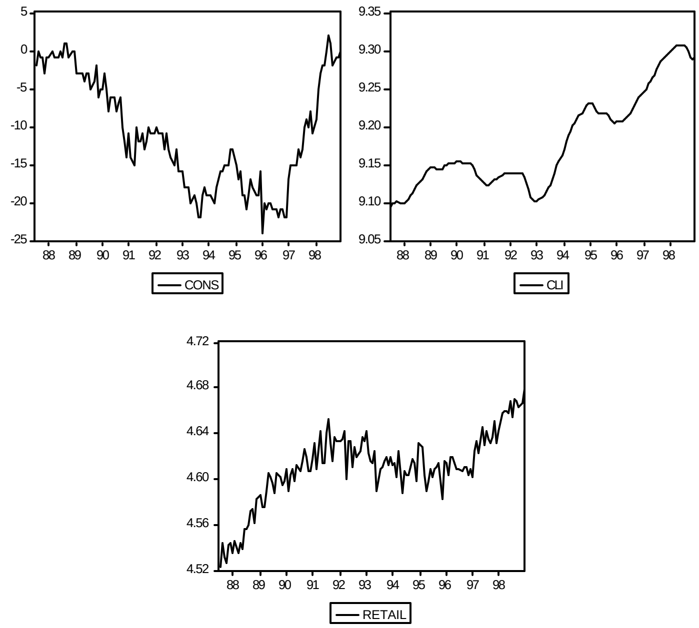
Macro-economic variables

We will first present some characteristics of the three variables used in this analysis. Figure 1 displays the three variables. In the left upper panel, the survey data for the sentiment of the level of current consumption is displayed. It is important to see that 0 is not the absolute figure for habit consumption. From this panel, we see that the consumer sentiment changes frequently and to a large extent in this period. It is high from 1987 to 1990 and again high at the end of the sample period (1998). The right upper panel shows the leading indicator for Europe. We see that in the period under study, there are both periods of good and bad prospects. From 1990 till 1993, this leading indicator declines announcing bad states of the world. The bottom panel displays the evolution of the retail consumption volume. From 1987 till 1992, this aggregate volume of retail consumption increases before it stabilizes and even slightly decreases till 1997. In the last two years of the sample period, aggregate retail consumption again increases.

In order to see whether risk can explain these returns on contrarian strategies, we estimate the short-term dynamics and the long-run equilibrium between returns and macro-economic variables for the fourteen sectors and the three value-growth strategies.

Figure 1

Macro-economic variables



The econometric analysis is conducted in the following way. In a first step, we investigate the co-integration properties of the cumulative return spread (CSP) of the high/low earnings yield strategy and the levels of the three consumption variables. Since we use the cumulative return spread (CSP) to test for integration, the analysis of the first differences in a vector error correction model (VECM) boils down to an investigation of the return spread (SP). First, we consider the lag structure of the model by an unrestricted VAR analysis using the Schwarz information criterion. The minimal optimal lag we impose is three months, or one quarter. Next, error terms are checked for autocorrelation and approximate normality (using autocorrelation tests, Breusch-Godfrey Lagrange multiplier tests, ARCH-LM tests and quantile-quantile plots for the normal distribution). In most sectors, the errors of the estimated equations are indistinguishable from white noise at 5 lags. There are some exceptions for which the residual VAR matrices satisfy the statistical conditions at 6 lags. Consequently, we estimate the basic model with 6 lags. Using six lags instead of five does not change the estimated coefficients to an extent that conclusions alter in the cases where 5 lags were sufficient.

The following model is estimated:

$$SP_t = g_1(c + 1CSP_{t-1} - b_1 CONS_{t-1} - b_2 CLI_{t-1} - b_3 RETAIL_{t-1}) + \sum a_{11}^i SP_{t-i} + \sum a_{12}^i \Delta CONS_{t-i} + \sum a_{13}^i \Delta CLI_{t-i} + \sum a_{14}^i \Delta RETAIL_{t-i}$$

$$\Delta CONS_t = g_2(c + 1CSP_{t-1} - b_1 CONS_{t-1} - b_2 CLI_{t-1} - b_3 RETAIL_{t-1}) + \sum a_{21}^i SP_{t-i} + \sum a_{22}^i \Delta CONS_{t-i} + \sum a_{23}^i \Delta CLI_{t-i} + \sum a_{24}^i \Delta RETAIL_{t-i}$$

$$\Delta CLI_t = g_3(c+1CSP_{t-1} - b_1 CONS_{t-1} - b_2 CLI_{t-1} - b_3 RETAIL_{t-1}) + \sum a^i_{31} SP_{t-i} + \sum a^i_{32} \Delta CONS_{t-i} + \sum a^i_{33} \Delta CLI_{t-i} + \sum a^i_{34} \Delta RETAIL_{t-i}$$

$$\Delta RETAIL_t = g_4(c+1CSP_{t-1} - b_1 CONS_{t-1} - b_2 CLI_{t-1} - b_3 RETAIL_{t-1}) + \sum a^i_{41} SP_{t-i} + \sum a^i_{42} \Delta CONS_{t-i} + \sum a^i_{43} \Delta CLI_{t-i} + \sum a^i_{44} \Delta RETAIL_{t-i}$$

Table 6 reports the most important findings of this analysis. We find at least one cointegration vector in all sectors at the 5% level using both earnings yield and BTM as ranking variables (third panel). For the other contrarian strategies this is the case in 18 out of 28 sectors. This finding suggests that there exists a long-run relationship between return spreads for high versus low earnings yield portfolios and consumption data. However, the interaction between the return spreads and the consumption data appears to be quite complex in certain sectors. The most important parameter in table 4 is the γ_1 coefficient, which measures the reaction of the return spread in a specific sector to deviations from the long-run equilibrium. A large and significant coefficient indicates that the speed of adjustment to the long-run equilibrium is relatively high. This coefficient is significant in eleven out of fourteen sectors using both earnings yield and BTM (third panel), implying that there is a response of return differences produced by contrarian portfolios to deviations in the previous period from the long-run path. In the three sectors that do not exhibit a significant γ_1, error-correction can be accepted since at least one of the other gamma's is significant, signaling the existence of a long-run relationship. The interim conclusion is that we find evidence that the difference in return between high/low earnings yield portfolios is reasonably explained by consumption and business cycle data for all sectors. All sectors either have a significant speed of adjustment parameter γ_1 or are influenced by the short-term dynamics of at least one factor (α^i).

In the seventeen sector strategies were we find overall profitability for the contrarian strategies, we find a cointegration relation in fourteen cases. Also in fourteen cases, we find that short-term dynamics explain some of the return spread to (significant alphas). Note that in most cases where γ_1 is significant, we find a negative sign. This implies that where there is a positive sign for the long-run relationship, which can be interpreted as states of the world where marginal utility is high, the return spread decreases as expected.

Nevertheless, if the specification of the VECM is correct, we find evidence that return spreads in only two sectors strategies (out of 42) are not influenced by one of the macro-economic factors. This is the case when both γ_1 and all α^is for that factor are statistically equal to zero. In only seven out of 42 sector strategies recent changes in short-term dynamics are no part of the explanation of the return-spread in that sector (no significant alphas). The influences however occur at different lags.

Table 6

A vector-error correction model for the relation between returns and the states of the world

In table 6, CI integrates whether or not there is a cointegration relation between returns and the macro-economic variables. If such a relation is found, the level of statistical significance is indicated as 1% or 5%. No indicates that no cointegation relation is found. g_1 is the coefficient denoting speed of adjustment to the long-run equilibrium. The two-sided 5% significance for a coefficient different then zero is indicated by **, the 10% by *. The three betas for expressing the long-run relation between returns and the variables are displayed as well, with * indicating whteher this coefficient is significantly different from zero. Finally, the R^2 is shown for the first regression in the model, explaining the return spreads.

	BANK	BMAT	CHEM	CYCC	CYCS	FINS	INDU	INSU	NCYC	PHAR	RESO	TECH	TELE	UTIL
Contrarian strategy using earnings yield forecasts as a ranking variable														
CI	1%	1%	NO	NO	NO	5%	5%	1%	5%	5%	5%	NO	5%	5%
g_1	-0.029 **	-0.062 **	-0.037	-0.019	0.001	-0.000	-0.200 *	-0.094 *	-0.132 **	-0.101 **	0.012	-0.018 **	-0.084 **	0.012 **
b_1	0.04 *	0.02	-0.01 *	-0.01	-0.01	-0.01	0.28	0.01	0.01 *	-0.03 *	-0.02	0.04	0.04 *	-0.10
b_2	-2.00	2.32	-0.90	-2.26 *	-2.24 *	-5.95 *	2.89 *	-6.06 *	1.22 *	-0.93	-9.20	0.21	-8.26 *	-17.39
b_3	1.92	1.40	-1.57	0.80	2.73	0.74	0.58	2.52	-1.90 *	0.34	7.89	-7.17	3.22	-6.85
R^2	28%	27%	14%	10%	21%	25%	22%	16%	24%	29%	23%	29%	27%	13%
Contrarian strategy using BTM as a ranking variable														
CI	5%	NO	5%	5%	NO	1%	NO	5%	NO	NO	5%	1%	5%	NO
g_1	-0.081 **	-0.007	-0.056 *	-0.003	-0.031	-0.096 *	0.005	-0.031	0.011	-0.014	0.021 **	-0.044 **	-0.002	0.000
b_1	-0.01	0.02 *	0.00	0.00	0.01 *	0.01 *	0.00	0.01 *	0.01	-0.09 *	0.00	0.00	-0.04	0.01
b_2	-0.69	-8.44 *	-5.02 *	-10.18 *	-2.54 *	-1.82 *	-3.37 *	-5.99 *	-3.81 *	-11.96	-13.54 *	-0.25	1.52	5.04
b_3	-3.33 *	2.85 *	-2.34 *	1.64	0.52	-2.01 *	0.28	-0.63	0.81	0.60	1.92	1.28	25.48	16.15
R^2	27%	23%	16%	26%	28%	25%	35%	19%	23%	21%	19%	34%	23%	14%

Table 6 (continued)

A vector-error correction model for the relation between returns and the states of the world

In table 6, CI integrates whether or not there is a cointegration relation between returns and the macro-economic variables. If such a relation is found, the level of statistical significance is indicated as 1% or 5%. No indicates that no cointegation relation is found. g_1 is the coefficient denoting speed of adjustment to the long-run equilibrium. The two-sided 5% significance for a coefficient different then zero is indicated by **, the 10% by *. The three betas for expressing the long-run relation between returns and the variables are displayed as well, with * indicating whteher this coefficient is significantly different from zero. Finally, the R^2 is shown for the first regression in the model, explaining the return spreads.

	BANK	BMAT	CHEM	CYCC	CYCS	FINS	INDU	INSU	NCYC	PHAR	RESO	TECH	TELE	UTIL
Contrarian strategy using earnings yield forecasts and BTM as ranking variables														
CI	1%	1%	5%	5%	5%	5%	5%	1%	5%	5%	5%	5%	5%	5%
g_1	0.034 **	-0.025 **	-0.033 *	-0.048 *	-0.018 *	-0.035 **	-0.088 **	0.034	-0.087 **	-0.047 *	0.003	-0.003 **	-0.200 **	0.017
b_1	-0.02	0.05	0.02 *	-0.02 *	0.01	0.04 *	0.01 *	0.00	0.01 *	0.00	0.00	0.18	0.04 *	-0.03
b_2	-2.29 *	-0.63	1.35	-2.26 *	-1.38	1.76	1.61 *	-3.19 *	1.40 *	1.47	-5.98 *	2.57	-0.99	-7.82 *
b_3	-8.92 *	-4.08	0.30	5.04*	-3.26	-4.27 *	-1.53 *	-3.88 *	-0.66	0.24	5.18	-57.44	-0.35	2.50
R^2	24%	36%	24%	17%	25%	30%	17%	21%	26%	13%	19%	28%	32%	11%

Summarizing, we find that the relation between the return spreads and the perceived risk is very complex and differs across sector strategies. This is perhaps why testing the difference in returns over booms and downturns in financial markets and economic recessions and periods of expansions does not explain the value-growth anomaly. On the other hand, we find a lot of indications that changes in the macro-economic variables explain to some extent the difference in returns between value and growth stocks.

6. Summary and Conclusions

We find that the hypothesis of different performance of value stocks portfolios relative to growth stocks portfolios differs across European sectors. We empirically assess this problem by showing that the one-year analysts' price-earnings forecasts or BTM under some assumptions can be used as a proxy for expected return. A theoretical restatement of the dividend discount model shows that under reasonable assumptions this identity can be derived using the residual income valuation model. More specifically, we find that in the sectors of financial services, insurances, resources and utilities, the high yield stock portfolio overall outperforms the low yield stock portfolio on average as well as in most months. The maximum annually estimated outperformance was found in the sector of the insurances and resources using BTM as a ranking variable. This outperformance is estimated to be about 13% per year. In other sectors we find indications that on the contrary low yield portfolios outperform high yield portfolios. Moreover, we find evidence of time-varying return spreads from the zero-cost investment strategy based on earnings yield and BTM. The zero-cost strategy based on earnings yield forecasts can be beneficial. But in European sector samples, this is not at all a stylized finding as it is in many cross-section samples that were studied in the past literature. Hence, such an unconditional investment strategy for sector funds is statistically unreliable. This strategy would only have been beneficial in the four mentioned sectors for the thirteen-year period under study.

The second hypothesis states that the difference in performance of high versus low yield portfolios is determined by the state of the world and hence embodies fundamental risk. This hypothesis is explained by asset-pricing models and in a standard consumption-based setting, and is also tested empirically. First, we test the mean-variance efficiency of two asset-pricing models for these dependent portfolios. We find evidence against an ex ante mean-variance efficiency of the factor portfolio for the entire period (1987-1998) for all contrarian strategies. The wealth portfolio is mean-variance efficient in the sub-periods for some strategies. Hence, part of the cross-section of returns can be explained by systematic risk but this evidence is weak.

We also test a long-term relation between, the return spreads and the states of the world. Therefore, we use monthly European consumption data. We find evidence that the difference in performance between yield portfolios in European sectors can partially be explained by consumption data and consumer sentiment data. On the one hand, we find that there is a long-run relationship between returns and consumption data and on the other hand we find that there are short-run dynamics between returns and consumption. Contrary to most of the past literature on this topic, this evidence indicates that there may be a fundamental risk story to the different performance of portfolios of stocks. Some stocks can indeed yield a higher return in states of the world where marginal utility is high.

A reason why it is hard to determine a relationship between returns and consumption data is that the dynamics between this relationship is complex. In this analysis, sector spreads are often influenced by the same factor but at different lags and with different signs. This is not at all illogical if we consider the nature of the data. As expected, different sectors react in different ways to expected changes in the business cycle or in large or small deviations from habit consumption. Secondly, consumers apparently consider that sectors have specific risk characteristics. Investing in the technology sector will for various reasons be considered as more risky than for example the utilities sector. As a consequence stocks with high yield properties will be treated differently in one sector to another and perhaps even more important, with a different speed, if the agent observes, for example, that his consumption will return to his habit consumption in the future.

We think that the implications of these findings are twofold. On the one hand, this framework makes it possible to investors to use the analysts' price-earnings forecast and BTM to get an indication of the individual stock's expected return because the relation within sectors seems to depend on fundamental risk. Complex models to estimate the expected return or the cost-of-equity are not always accessible or reliable. It also implies that an unconditional application of a value-growth strategy on a sector basis is unreliable. The second implication is that it is possible to determine specific periods of attractiveness for the investor for different types of stocks. From a consumption-based perspective, state variables can give information about the period in which a certain type of stock is more suitable to smooth consumption and is in that sense less risky. In terms of asset allocation, this means that an investor can overweight high yield stocks in one sector when there is a high probability of being in a good state and the expected return is lower than low earnings yield stocks, yielding a higher payoff. At the same point in time, the investor can choose to overweight low yield stocks because the probability of being in a good state can be low for that sector.

Chapter IV

Bayesian Forecasting and Financial Markets

IV a

Momentum, Rational Agents and Efficient Markets

Momentum, Rational Agents and Efficient Markets

Abstract

Descriptive behavioral models explain the momentum anomaly by assuming that financial agents are irrational. However, investors are not tested to be susceptible to the cognitive failures that were observed in psychological experiments. We suggest an environment were financial agents are rational, markets are efficient defined by the Grossman-Stiglitz [1980] efficiency and there are market imperfections in the information market. Based on a simulation experiment, we find that in this environment, returns on momentum strategies can still exist because of the noise in expert information. We empirically find that even in a sample of large and liquid stocks, this noise is still observable and hence, momentum can be empirically found for these samples even in the case where agents are rational.

1.Introduction

A market anomaly that is difficult to explain is the existence of momentum in stock prices. Past literature explains momentum by descriptive behavioral models [Daniel et al., 1998, Barberis et al., 1998] where the assumption is made that financial agents are irrational. We suggest a benchmarking of these irrationalities by studying rational judgments in financial markets. Rational behavior is often assumed in economic models. As De Bondt and Thaler [1995] state, this principle of assumed optimality entails that actual decision-making is not often studied, assuming that decision processes do not affect outcomes. This is certainly true for finance. Descriptive behavioral models start from the assumption that generally observed (especially in psychological experiments) cognitive failures[1] are true for financial agents as well. The problems with this assumption are twofold. First, we do not know to what extent financial agents are susceptible to cognitive failures that are observed in other disciplines such as psychology. Second, as Fama [1998] notices, "*descriptive behavioral models work well on the anomalies they are designed to explain*" and these models rarely come up with an alternative for the efficient market hypothesis (further: EMH). We summarize these two statements by the following argument: "*To consider the human judgment as sub-optimal without the discussion of the limitations of optimal models is naive.*"[Einhorn and Hogarth, 1988].

In this paper, we study rational judgments by financial agents in an environment that maintains its realism. The hypothesis we evaluate is whether a market anomaly such as momentum can be empirically observed when financial agents are rational. The purpose is to evaluate the hypothesis in this paper in a setting that is realistic for actual decisions in financial markets. If momentum can be explained when investors are rational, this is a motivation to document the relevance of the assumptions of cognitive failures in descriptive behavioral models.

The momentum anomaly

Stocks with a strong past performance continue to outperform stocks with a poor past performance in the next period with an average excess return of bout 1% per month [Jegadeesh and Titman, 1993, 1999]. Barberis et al. [1998] ascribe the existence of momentum because of an underreaction to news signals. The news is not quickly reflected in the price and continues to have an impact in the subsequent periods. The Fama-French [1993] three-factor model does not explain these momentum returns. Cochrane [2001] illustrates that the existence of momentum

[1] Examples of cognitive failures or irrationalities are overconfidence (the phenomenon that people overestimate the reliability of their own knowledge, [De Bondt and Thaler, 1995]) and conservatism (the slow updating of beliefs relative to new evidence, Edwards [1968]).

possibly is merely a restatement of the finding that the predictability of monthly individual stock returns is low (with an R^2 of 0.0025). He argues that this low predictability is sufficient to explain the momentum anomaly because of the large standard deviations of individual return series (40% on an annual basis). Furthermore, Carhart [1997] finds that momentum strategies are not profitable after accounting for transaction costs. Moskowitz and Grinblatt [1999] find that the portfolio at the sell-side of the strategy consists mainly of small stocks that are susceptible to infrequent trading. However, Cochrane [2001] suggests that there is a possibility that there is a small autocorrelation in individual stock return series that is related to risk.

Decision-making and finance

We benchmark the existence of cognitive failures to rational judgments in financial markets. Rational judgments are formed by Bayesian forecasting. This is the normative version of behavior in financial markets: financial agents *should* act as Bayesian forecasters if they are said to make ex-ante rational judgments. There are several reasons to evaluate this normative behavior [Kahneman and Tversky, 1988]. Firstly, agents are thought to be effective in pursuing their objectives. The objective of an investor is to make a good judgment about the expected return.

Secondly, optimal decisions are made to maximize expected utility and are hence made with the objective to survive in a competitive environment. The first-order condition of the consumption-based asset-pricing model where an agent intertemporally maximizes wealth is an example of this statement (see Part I of this dissertation). Odean [1998a] argues that overconfidence decreases expected utility for financial agents.

Thirdly, the rationality axiom provides theories that explain choice behavior. Hence, studying the rational decision-making process for financial agents can explain the problems of the assumptions that are made. For example, we can find - using transaction data - that investors tend to sell winner stocks and hold loser stocks [as in Odean, 1998b]. According to the momentum anomaly, the performance of loser stocks is lower in the next period and hence this seems an irrational choice. However, evaluating rational behavior can explain why this choice is made, and equally important, in what environment these choices are made. These three arguments motivate the importance of the evaluation of normative behavior. The main objective in this paper is to see whether momentum can exist when financial agents make rational judgments in a plausible environment. If this is the case, it is an important observation for the descriptive behavioral models that explain momentum.

2. Expected returns and momentum

We will first give a description of the financial environment in which decisions are made. An investor buys and sells assets until the marginal cost in the first-order condition of the inter-temporal asset-pricing model equals the marginal benefit. In a financial market, the decision about the value of the traded assets, and hence about the expected returns, is complex. Acquiring information about these assets is competitive so making easy profits is impossible. In this framework, we assume that small market imperfections exist [as in Grossman and Stiglitz, 1980].
A competitive equilibrium according to the EMH means that prices are such that arbitrage profits will be eliminated. This model cannot always be in equilibrium because arbitrageurs will not be able to make profits from their costly strategy and hence have no incentive for arbitrage. When arbitrage is costly, the assumption that all markets are in equilibrium, *also the market for information*, is inconsistent [Grossman and Stiglitz, 1980]. In this case, prices of assets can differ even when the expected returns are more or less the same. As an example, Thompson [1978] observed that prices of baskets of individual securities are lower than the sum of the prices of the individual assets. As Grossman and Stiglitz mention, this is not a falsification of the EMH, simply a redefinition. In this redefined EMH, prices partially reflect information of informed market participants.

First, we theoretically relate news on expected returns to prices. Suppose the relation between cash flows and expected returns is presented by a vector-autoregressive process [VAR, Cochrane, 2001]. In this framework, expected returns are driven by a state variable x_t. This state-variable is moving slowly. Furthermore, dividend growth (Δd) is unpredictable and dividend-yields are highly persistent (autocorrelation-parameter r is close to 1, see equation 2). The VAR representation takes the following form [Cochrane, 2001]:

[1]
$$\begin{aligned} x_t &= b x_{t-1} + f_t \ , \\ E(r_{t+1}) &= x_t + e_{rt+1} \ , \\ \Delta d_{t+1} &= e_{dt+1} \ , \end{aligned}$$

where f_t is the error term in the process of the state variable, e_r is the part of the expected return that is not explained by the state variable, dividend growth is not predictable and is equal to a white noise term e_d.
The VAR representation for the dividend-yield ratio takes the following form:

[2]
$$(d_{t+1} - p_{t+1}) = b(d_t - p_t) + \frac{f_{t+1}}{1 - rb} \, .$$

From this dividend-yield equation, the volatility of the dividend-price ratio can be related to the persistence and volatility of expected returns [Cochrane, 2001]:

[3] $$s\,(d_t - p_t) = \frac{1}{1 - r\,b}\, s\,(x_t)\,.$$

If b is high (slowly moving state variable) and r is high (persistent dividend-yield), a small variation in expected return translates into a large price variation. The same relation can be observed in standard valuation models. Suppose the Gordon growth model ($P = D/(r - g)$) holds. A price-to-dividend ratio of 25 means that the expected return r minus the growth rate g equals 4%. One percentage point change in expected return translates into a 25 percentage point change in price [example taken from Cochrane, 2001]. The same example can be applied for the Ohlson's [1995] residual income model (RIM, see III a). Suppose the following price dynamics for the abnormal returns: abnormal returns are persistent and other information is not. The RIM takes the following form: $P = e_{t+1}/r$ [see Dechow et al., 1999], where e_{t+1} is the one-year analysts' earnings forecast. Suppose that the earnings yield forecast is 5%. One percentage point change in the expected return causes a change of 17 percentage points in the price.

Cochrane [2001] allows for the possibility that there is a small positive autocorrelation in returns, as is empirically confirmed by an R^2 of .0025 for monthly data. However, he mentions that this small autocorrelation does not allow us to generate positive autocorrelation in future payoffs or realized returns! If there is a slow and persistent variation in expected returns we saw that from the VAR framework, there is a natural generation of negative autocorrelation in returns. News that expected returns are higher (an increase in risk) causes future dividends to be discounted at a higher rate. Hence, today's price will fall and today's payoff or return will also decline. Cochrane [2001] further states that the previous prediction will only be false in the case that shifts in expected return will be positively correlated with shocks to dividend growth. However, he argues that there is no evidence for such a correlation. The second explanation, and the one that is assessed in this paper, is that the financial market system gives information from experts, but does it imperfectly. The reason for this market imperfection is that there is noise in the precision of the expert information [as in Grossman and Stiglitz, 1980].

3.Bayesian forecast for the judgment of the next period return: the environment

We first define the environment in which the financial agent has to make rational judgments.

[as.1] There is asymmetric information in financial markets.

The first assumption is that there is asymmetric information. This assumption implies that there is information that has to be acquired against a cost. Part of the public information is costly, but it does not only concern fundamental firm information. According to decision-making theory, when the task is difficult, experts do also acquire additional information because they use models and theory to improve their judgment [Griffin and Tversky, 1992]. With respect to financial markets, this seems an important extrapolation of costly public information. Investment banks and brokers, for example, invest a lot in the search for better models and theories.

[as.2] Financial markets are efficient according to the Grossman-Stiglitz fromulation of efficiency.

The second assumption is that markets are efficient as in the Grossman-Stiglitz formulation (as described in section 2), and that there exist market imperfections because of the noise in the precision of the information[2] from informed market participants. Notice that Griffin and Tversky [1992] report that the level of dispersion in this information (or the inverse of the strength of the evidence) will be more important in the weighing of evidence by a Bayesian forecaster than the number of market participants (the weight of the evidence) who give information when the decision problem is difficult. Imagine a case where two experts communicate their opinion about the future return of a stock. If they communicate the same prediction, this will be to a higher extent influential to the decision maker than the fact that there are only two experts.
The market for information is competitive, but following from the second assumption, this market is imperfect. Together with assumption one, this implies that prices do fully reflect all costless public information. Other public information is costly.

[2] The noise in the precision of the information is also denoted as the strength of the evidence and stands in this paper for the dispersion in the expert opinion.

[as.3] There exists a theoretical relation between news on expected returns and prices.

Third, there is a theoretical relation between news on expected returns and prices. If we combine assumption 3 with assumption 2, we see the importance of the precision of this type of news. In this framework, news comes from experts who have costly public information and these experts do not necessarily trade. This is relevant for financial markets. For example, analysts make predictions of the next period's earnings but without trading on their information.

The decision problem in this paper is the judgment about the next payoff or the return. If the news is good, prices will rise, and accordingly the payoff will rise and the expected return declines. The setting in this environment is the following. A financial agent makes a judgment about the next period's return based on his information set. According to the agent defined in the Hong-Stein model [1999], this agent only uses historical price changes and makes linear forecasts. If the historical price changes are the likelihood (or the probability distribution of the data) for the investor and returns are normally distributed, the forecast of this agent about the next period's payoff will be the sample mean. We only model the rational judgment for this agent because our research question is whether momentum can exist if the investor or the financial agent is rational.
This financial agent is aware that the decision problem is complex and seeks aid from experts. He uses the prediction of the next period's payoff by the experts as prior evidence or his base rates. Using all this evidence, the investor makes a rational forecast about the next period's payoff applying Bayes'rule. As described, the evidence the agent uses in order to make a judgment about the future price change is the historical price changes (B) and the expert opinion about the future price change (A).

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)},$$

where A denotes the expert opinion and B denotes the data.

It is important to explicitly describe the difference with the Hong-Stein model [1999]. First, we only evaluate the rational decision of one financial agent. This financial agent has access to expert information from an expert who does not trade. Hence, the 'newswatcher' from the Hong-Stein model only has an impact on the price by communicating an opinion about the next period. Second, the costly information is accessible for this financial agent through the communicated expert opinion. The expert takes this information into account and uses more complex models to make a judgment about the decision problem. The fact that the financial agent has access to the newswatcher's opinion is hence a second difference with the Hong-Stein model.

This environment maintains its realism. There are expert in the markets (such as analysts) who do not trade and communicate their opinion (earnings forecasts, buy and sell signals). In the processing of the expert opinion, they use costly information (they visit the company, contact the management, use complex models) that are less accessible for the other agents (the cost is higher).

4. A normative model of judgment in financial markets

4.1 Aspects of the framework

If Bayes' rule is the appropriate language[3] for investors to form their beliefs it is worthwhile studying it. Already in 1952, Markowitz mentioned that in order to form a belief about the parameters needed to apply his expected return-variance rule, it is should be interesting to combine signals from statistical parameters and the judgment of practical men or experts. In doing so, one could formalize his probabilities by applying Bayes' rule. The framework in this paper suggests how agents should form their beliefs in order to obtain a rational prediction for the price change of the assets of interest for the next period. In order to test this setting, we simulate the rational weighing of the evidence. The advantage of an analysis of this simulated judgment is that it allows us to evaluate the actual weight function applied by a rational investor and its implications for the behavior of the financial agent.

In this paper, we assess this problem for the judgment about the next period's return. Modeling normative beliefs for future return alone is valid for different applications in finance, such as stock selection, asset allocation and fund selection. A major advantage of the framework we use, is that individual priors can be used for each individual decision in order to reduce the loss of information. In a lot of financial analyses, the problem of the use of proper priors is probably the main reason why the Bayesian approach has not always been successful for financial applications. Only in a few cases analytical solutions, using conjugate priors, are available.

An important issue is that this framework does not intend to be descriptive as is the case for most models in recent literature. In other words, we do not try to model how agents in financial markets *do* set their probabilities or *do* make decisions, i.e. how

[3] We denote a language as the calculus used to translate evidence into a judgment.

investors behave. This is important because we do not make an attempt to explain which cognitive biases in the judgment under uncertainty of individual investors are relevant for investor behavior. But more importantly, this evaluation of normative behavior allows us to analyze the cognitive failures with respect to the financial decision-making in financial markets. These suggested cognitive failures that were found in psychological experiments are assumed to be true for financial agents as well and this framework allows us to evaluate this assumption.

Model specifications

In order to simulate the judgments given the information available to the financial agent in the same way as a Bayesian forecaster would proceed, we apply the Bayesian Bootstrap regression procedure (BBR). The main reason to apply this method is the flexibility of the prior choice, which makes the possibilities much larger than the ones implied by conjugate priors or analytically tractable solutions. In this paper, the non-parametric integration method allows a flexible definition for the prior. Computational power and further development of sampling procedures in recent years makes this less stringent approach of Bayesian analysis attractive.

The basics of Bayesian sampling analysis can be found in the paper by Kloek and Van Dijk [1978]. But the foundation of the BBR approach itself is due to the development of a solution to the constrained normal linear regression model by Geweke [1986]. The procedure itself generates outcomes from the likelihood that can be used for integration procedures of expectations with respect to the posterior distribution [Heckelei and Mittelhammer, 1996]. The major difference of the procedure in this paper compared with the one suggested by Geweke [1986] lies in the fact that the parametric assumptions made in his model are not made in this approach. The application of the BBR procedure is to use it as an engine to develop Bayesian forecasts with respect to the decision problem of choosing the ex-ante optimal expected return or payoff for an asset or an asset class. This engine represents the human mind weighing the evidence as a pan balance in the case the agent is rational.

Suppose that asset prices follow a random walk. In that case, any linear forecasting rule for the future price change based on historical price changes alone is not efficient [Campbell et al., 1997]. Indeed, empirical analysis showed that returns are hard to predict at short horizons. Since we start from a decision problem about expected returns that is solved at a relatively high frequency (e.g. monthly), it is a reasonable assumption that, in reality, historical prices convey little (remind the R^2 of 0.0025 in section 2) or no information on future price changes. Moreover, this high

frequency of decisions is relevant with respect to the real environment where for example a monthly rebalancing of portfolios is not unusual.
Suppose that p_t denotes the observed time-series logarithmic prices of an asset at time *t*. If prices follow a random walk with drift, the price process can be described by equation 4.

[4] $$P_t = m + P_{t-1} + e_t \ .$$

In addition, assume that the error term is normally distributed with zero mean and variance s^2. Changing the log price in the previous equation to the left side of the equation shows that the time-series of returns of the asset (r_t) is described by a constant m and a white-noise error term. In fact this expression is the same as a normal linear regression model of the returns on a unit vector (i_t) (equation 5). The parameter m denotes the expected return.

[5] $$r_t = m i_t + e_t$$

This regression model is the point we start from to apply the BBR methodology. Such linear forecasts are inefficient if prices follow a random walk, but it is realistic to think of a financial agent who uses the first moments of the sample distribution to make his judgments about the future return [as in Hong and Stein, 1999].

Starting from this normal linear regression, we apply the BBR as described by Heckelei and Mittelhammer [1996]. Classical inference methods using a likelihood function are often difficult to evaluate. If the problem is solved in a Bayesian framework, analytical solutions are scarce [Geweke, 1989]. In most cases priors and functions of interest are not tractable. In finance, this problem is important. Often, uncertainty and estimation risk are reduced applying Bayesian methods with a common shrinkage parameter [Jorion, 1991]. For example, in portfolio selection, the return on the minimum variance portfolio is used as the prior parameter for all assets in the Bayes-Stein estimator. As Bawa et al. [1979] say, the use of a common shrinkage parameter implies a loss of information about individual assets or asset classes.

In finance, depending on the decision problem that is adressed, there are several possible priors that are relevant (such as earnings yield, expert surveys, ...) but a choice of a prior is also one of the limitations of each tested model [Lenk and Wedel, 2001]. A discussion of each prior is vital for the evaluation of each model. There are several examples of possible applications in decision making in finance. A first one is fund selection. An example of such an application is given by Baks et al. [2001] who evaluate the quality of the fund (manager) imposing specific priors on the estimation of the fund's alpha. The estimation of predictive regressions is a second example. Mostly, however, there is a restricted prior choice because analytical solutions are required to solve the problem [MacKinlay and Pastor, 1999, among others].

In this paper, we only assess the question of judgment about the next period's stock price change on a short horizon. This application is relevant both for stock selection procedures and asset allocation decisions, both a difficult and much studied topic. Because there is a lot of noise in the data, it is very difficult to make relevant forecasts. Normative forecasts, such as Bayesian forecasts, can be studied to better understand this decision process. Also, in asset allocation, improving the estimation procedure or the decision process for the mean vector as an input parameter for the expected return – variance rule is of major interest.

Bayesian bootstrap regression

Starting from the normal linear regression model in [5], we denote the return series or the series of historical price changes as the likelihood $f(r|m,s) = L(m,s|r)$. From here, we drop the subscript t from r_t for reasons of notational simplicity, with r stilll denoting the vector of historical price changes. In the Bayesian bootstrap approach, the residual series in [5] follows, in non-parametrical form, an unknown empirical distribution $h(e|s)$ determining the parameter s in the likelihood. First, we develop the general solution to the problem. Next, we extend the application to its non-parametrical form. This extension is used for the estimation of the expected return in this paper.

Using Bayes'rule as the rational calculus to make a judgment for the financial agent, we define the posterior density for the parameters (equation [6]).

[6]
$$p(m,s|r) = \frac{p(m,s)f(r|m,s)}{p(r)}$$

$$\propto p(m,s)f(r|m,s)$$

In [6], $p(m,s|r)$ is the posterior density, $p(m,s)$ denotes the prior information on the parameters and $p(r)$ denotes the data distribution. Since we want to model decisions on the expected return (m), we concentrate on the marginal posterior distribution for m (equation 7). Throughout the remainder of the text, independence between the prior parameters m and s is assumed.

[7]
$$p(m|r) = \int_0^\infty p(m,s|r)ds$$

$$\propto \int_0^{\infty} p(m,s)\, f(r|m,s)\,ds$$

$$\propto p(m)\int_0^{\infty} p(s)\, f(r|m,s)\,ds$$

In order to solve [7], the likelihood function is weighed by the prior evidence on s and the parameter s is integrated out. Next, the expression is normalized to unit total mass, dividing by the integral with respect to m. This reformulation is made in order to make it a proper density [Heckelei and Mittelhammer, 1996].

[8] $$p(m|r) \propto p(m)\frac{\int_0^{\infty} p(s)\, f(r|m,s)\,ds}{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\left[\int_0^{\infty} p(s)\, f(r|m,s)\,ds\right]dm}$$

$$\propto p(m)\, f_s(r|m),$$

where $f_s(r|m)$ can be regarded as the posterior distribution of m in the case when $p(m)$ equals 1. This is the same as assuming prior ignorance on m [Heckelei and Mittelhammer, 1996].

In this case, it is possible to formalize posterior expectations of any function of m as an integral involving this function of interest and the marginal posterior [equation 9, Geweke, 1989]. We denote $g(m)$ as the function of interest on the parameter m. $\overline{g}_n$ denotes the average of the simulated functions of the parameter of interest.

[9] $$\overline{g}_n \rightarrow E\left[g(m)\right] = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(m)p(m|r)dm$$

Tractability of the procedure and the non-parametric solution

The tractability of the expectation of the function of the parameter m depends in general on the choice of the prior, the likelihood function and the dimensionality of the problem. Geweke [1986] mentions that the expectation is solvable if the integral exists, the distribution $p(m|r)$ is bounded, the function of interest $g(m)$ is bounded and finally, there is an integrable likelihood function. As Geweke specifies, this last aspect is solvable for the normal linear model using a multivariate t-distribution. One reason to step to numerical integration procedures in order to solve the problem is that the choice of the prior is flexible [Geweke, 1999].

The parametric solution to the sampling procedure starts with sampling n i.i.d. outcomes from $p(m|r)$ [as first in Kloek and van Dijk, 1978]. This leads to n samples m_i^*, with i=1,2,...,n. A further assumption is that the expectation of the function of interest is not infinite. This is : $E[g(m)] = c < \infty$. Applying the strong law of large numbers, this means that $1/n \sum g(m_i^*) \underset{\rightarrow}{as} E[g(m)]$.

We have stressed the importance of a flexible choice of the prior for Bayesian problems. However, this flexible choice often makes sampling from the posterior distribution of m intractable. The solution to this problem is the introduction of an importance function $I(m)$ replacing $p(m|r)$. This is the systematic approach as suggested by Kloek and Van Dijk [1978] and Geweke [1986, 1989]. The most important feature of the choice of this importance function is that its tails do not decay more quickly than the tails of $p(m)$ [Geweke, 1986, 1989]. In other words, the importance function must support the posterior distribution, and the second condition is that this importance function should be easy to sample from. Introducing this importance function allows to rewrite equation [9] in the following form :

[10] $$E[g(m)] = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{g(m)\, p(m|r)}{I(m)} I(m) dm$$

The approximation of [10] is made by sampling n i.i.d. outcomes from $I(m)$ [Heckelei and Mittelhammer, 1996].

[11] $$E[g(m)] \simeq \frac{1}{n} \sum_{i}^{n} \frac{g(m_i^*)\, p(m_i^*|r)}{I(m_i^*)}$$

The rate of convergence of this procedure is determined by two factors [Geweke, 1986]. In the first place, this covergence is determined by the variability of $g(m)$. Secondly, it is determined by the ratio of density values, this is the ratio of the values of the posterior distribution and the importance function. More formal assumptions to ensure a proper simulation procedure are given by Geweke [1989].

Assumption 1 : the product of the prior density and the likelihood function is proportional to a proper probability density function defined on the parameter space.

Assumption 2 : $\{m_i^*\}_{i=1}^{\infty}$ is a sequence of i.i.d. random samples of a common distribution, having a probability density function $I(m)$.

Assumption 3 : the support of $I(m)$ includes the parameter space.

Assumption 4 : $E[g(m)]$ exists and is finite.

Geweke [1986] proposes a standard ignorance prior in order to solve the linear normal constrained regression model or the product of the uninformative prior and

an indicator function representing the constraints. As mentioned, in that case, $f_s(r|m)$ can be represented by a multivariate t-distribution with *n-k* degrees of freedom [see Heckelei and Mittelhammer , 1996].
Heckelei and Mittelhammer [1996] suggest to generalize this approach especially with respect to the form of the prior and suggest a non-parametrical solution. In order to obtain the expecation of the function of interest (the posterior distribution with respect to m as in equation [9]) this posterior distribution is normalized to unit mass in order to define a proper posterior $p(m|r)$.

[12] $$E\left[g(m)\right]=\frac{\int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} g(m)\,p(m)\,f_s(r|m)dm}{\int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} p(m)\,f_s(r|m)dm}$$

Next, the distribution of $f_s(r|m)$ is substituted by the importance function $I(m)$. In equation [12], $p(m)f_s(r|m)$ describes the weight function in the Bayesian procedure (Geweke, 1989). The practicability of this procedure depends on the expense of evaluating $g(m)f_s(r|m)$ and $I(m)$ in equation [10]. Secondly, the number of replications given m and the number of cancellations that occur in the ratio $p(m)f_s(r|m)$ on $I(m)$ in equation [10] also determine the practicability. If there is slow convergence - observable when $E\left[g(m)\right]$ is not stable as the number of replications increases – there is an indication that the choice of $I(m)$ is poor.

The numerator and denominator in equation [12] are expectations with respect to the distribution $f_s(r|m)$ (equation 13).

[13] $$\frac{E[g(m)p(m)]}{E[p(m)]}$$

If in this case, m_i^* with i=1,…,n are i.i.d. outcomes from $f_s(r|m)$, $E\left[g(m)\right]$ can be approximated by a prior weighted average.

[14] $$E\left[g(m)\right]=\frac{1/n\sum g(m_i^*)p(m_i^*)}{1/n\sum p(m_i^*)}$$

The above assumptions describe the characteristics of the sampling procedure that have to be evaluated in order to see whether the simulation behaves well. As mentioned, an important characteristic is that the expectation varies as the number of replications increases. In that case, there is no stability of the estimated expectation. Also, if extremely large values occur in the weighing function, it indicates that this function is not bounded [Geweke, 1989, 1999]. In order to test the degree of inefficient sampling, Geweke suggests to compare the numerical standard error of the procedure to the obtained posterior variance. If the ratio of the two approaches one, the posterior distribution is very well covered. A ratio smaller than 0.1 indicates a poor simulation procedure. This aspect is formalized by Geweke [1989]. If $\overline{g}$ denotes the expectation, then the following characteristic is obtained for the simulation procedure : $n^{1/2}(\overline{g}_n - \overline{g}) \Rightarrow N(0, s^2)$. This implies $ns_n^2 \to s^2$, where s_n^2 is the numerical standard error of $\overline{g}_n$. This numerical standard error is adversely influenced by the variance of $g(m)$ and by extreme relative observations in the weight function. If the simulation procedure behaves well, the numerical standard error is close to the posterior variance.

Finally, from equation [14], we can estimate the necessary expectations of the function of interest. In [15], we show the expectation of *(a)* the posterior mean and *(b)* the numerical standard error (nse) using the Bayesian bootstrap regression (with N denoting the number of replications)[Heckelei and Mittelhammer, 1996].

[15] $$(a)\quad \hat{E}(m) = \frac{\sum_{i=1}^{N} m_i^* \, p(m_i^*)}{\sum_{i=1}^{N} p(m_i^*)}$$

$$(b)\quad nse = \frac{\sum_{i=1}^{N} (m_i^* - \hat{E}(m))^2 \, p(m_i^*)^2}{\left[\sum_{i=1}^{N} p(m_i^*)\right]^2}$$

4.2 The case of expert evidence

Before we assess the hypothesis in this chapter, we examine a practical case of expert evidence based on the environment as presented in section 2. What we focus on is the strength of the expert evidence for a sample of stocks that can be considered for investment purposes.

Carhart [1997] finds that after accounting for transaction costs, he no longer finds evidence for beneficial momentum investment strategies. This is rather confusing because Rouwenhorst [1998] finds a return on a European momentum strategy of 1% per month. These findings are based on return data for the period 1980-1995 and stocks in the sample cover about 60% to 90% of the included European countries. Apparently, no really small stocks are included and the individual stocks should not be susceptible to infrequent trading. Nevertheless, momentum is observed.

This implies that the question whether there are differences in the strength of the expert evidence is important. Therefore, we investigate this question for a European database of liquid stocks. We collect relevant information about all European stocks in the MSCI Europe index between March 1992 and August 2000 from the IBES database. From the IBES database, we collect the mean consensus forecast, the highest and lowest forecast and the number of analysts covering the stock. We use the forecasted earnings yield as the expert opinion about the next period's cash flow (see section 2 of this chapter).

The final database we use is an intersection database between the IBES data and the Datastream dataset of market values and returns for the stocks in the MSCI Europe index. The Datastream data is collected from January 1987 to August 2000. The average number of stocks in the index during this period is 588. The average number of MSCI Europe index stocks that were covered by analysts is 471. Between 1992 and 2000, the coverage of index stocks in Europe increased on a steady basis. In 1992, an average of 16 analysts announced a one-year ahead forecast on a stock. This increased to a maximum of 21 analysts in 1997, and in 2000 this number is 19. Given the increasing coverage of stocks, this implies that at the end of the sample more stocks were covered by, on average, less analysts. A last important feature of the dataset is that Portuguese stocks entered the index only at the end of 1997. There are almost no stocks that are covered by only 1 analyst (in that case, τ is set to be equal to two times the standard deviation of the market portfolio).

For all stocks, we calculated a proxy for the strength of the expert evidence. Kinney et al. [1999] suggest to use the range of expert opinions or the standard deviation of opinions as a measure of precision. Comparable to the latter measure, we used the highest (H) and lowest (L) earnings yield forecasts on a stock to compute the proxy for the strength of evidence (t). The variance of the expert opinion is calculated from the Parkinson measure [1980] for extremes. The variance of the expert opinions is given by $0.361*(H-L)^2$. Furthermore, using the range of expert opinions and not taking into account the number of analysts making a forecast is defended both by findings in the psychology literature and the finance literature.

Griffin and Tversky [1992] suggest that the strength of the evidence (the level of divergence in the expert opinions) dominates the weight of the evidence (the number of analysts making a forecast). Hong et al. [2000] find that the diffusion of information is lower for small caps. Information diffusion is proxied by the number of analysts covering a firm. This number is best explained by the size of the firm. Other variables such as book-to-market or residual analyst coverage (the impact of one additional analyst) explain few of the total analyst coverage. Anticipating table 1, we find that the strength of the evidence is strongly related to the size of the firm as well. Hence, using the Parkinson measure as a proxy for the strength of the evidence measures the disperion in the expert opinion but also indirectly captures the number of analysts making a forecast because of the relation between the Parkinson measure and the firm size.

Table 1.

Stock characteristics and signal strength.

Table 1 shows the average characteristics for firms in decile portfolios ranked on market capitalization (MV) or beta. For the MV deciles, the Parkinson measure (τ) is displayed as well as the average beta for the decile. For the beta deciles, only the Parkinson measure is given. The Q1 portfolios are the portfolios with low values for the ranking variable.

	ranking variable: MV reported variable: $\overline{t}$	ranking variable: beta reported variable: $\overline{t}$
Q1	.02914	.00169
Q2	.02801	.00158
Q3	.00331	.00077
Q4	.00219	.00047
Q5	.00085	.00077
Q6	.00040	.00077
Q7	.00042	.00023
Q8	.00036	.00093
Q9	.00028	.00108
Q10	.00008	.00526

The Parkinson measure is computed for all stocks that are covered by analysts forecasts. All stocks are ranked according to their market capitalization (MV) and their beta (computed using the CAPM and a market value weighted portfolio) and ten deciles (of on average 50 stocks) are formed for each ranking parameter. For each decile and for each month, the equally weighted strength of evidence, $\overline{t}$, is computed. Notice that we are only interested in the difference of the signal strength across deciles for one ranking variable. We use mean absolute deviation to get an estimate of the sample mean of $\overline{t}$. This correction is made because of the presence of extreme outliers in this data. The consequence is that the average level $\bar{t}$ is not comparable between the two ranking procedures, but this has no impact on the analysis.

We find that even in a sample of liquid stocks, there are large differences in the strength of expert evidence. In the Q1 and Q2 portfolio based on size (smallest stocks), the precision of the expert information is extremely low. The variance of the information is around 2.85% (standard deviation of 17%). This implies that there are stocks in these deciles that are surrounded by noisy expert signals, even after accounting for outliers. The noise of the expert opinion decreases with the size of the firm. For the decile of most liquid stocks, the variance of the expert evidence is 0.0008 (standard deviation of 0.9%). These figures confirm the finding of Hong et al. [2000] that news diffuses at a higher speed for larger firms. This implies that news from experts are quickly reflected in the stock prices of large firms.
Expert evidence is also weak on average for low beta stocks and high beta stocks. The variance of the Q1 beta decile is 0.00169 (standard deviation of 4%) and for the high beta decile 0.00526 (standard deviation of 7%). The conclusion of this analysis is that even in a liquid sample of stocks, momentum effects can be found if financial agents are rational. Kinney et al. [1999] for U.S. data report that the accuracy of the expert opinion increases in the nineties. Also the precision of the communicated evidence increases. If the market imperfection we assumed reduces, and agents are indeed rational, we can argument that momentum strategies will be less beneficial in the future.

4.3 Scenarios and weighing functions

Scenarios
We evaluate three scenarios. In each scenario, the consensus expert opinion is that the return in the next period will be 4%. This implies that an imaginary stock price of 100 will rise to 104. This expert opinion is communicated with a different strength. For the three scenarios, we estimate the Bayesian forecasts using the BBR. Appendix 1 shows the general algorithm presented by Heckelei and Mittelhammer [1996] in order to explain how the Bayesian bootstrap regression is conducted. We use 5000 bootstraps to mimic the weighing of evidence by the human mind using the BBR.
The likelihood used to make a decision consists of 60 months of historical price changes on the Belgian stock market (it could be any real dataseries of stock returns). Returns are computed as the monthly percentage of change in the return index, are taken from Datastream using the total market index for Belgium. These returns are the monthly returns for the period 1996 till 2000. The difference in the

strength of the evidence is represented by the figures in table 1 for size deciles. The high strength of evidence is givezn by the Q10 number: 0.00008 or a standard deviation of 0.8944%. The medium strength is given by the Q5 number (0.00085 or a standard deviation of 2.9155%) and the weak evidence is given by the Q3 number (0.00331 or a standard deviation of 5.7533%). We do not use the Q1-Q2 numbers because of their extreme nature. In the analysis in table 1, negative earnings forecasts are not excluded and explain the extreme observations in the Q1-Q2 deciles. We do not want to take the negative earnings yield forecasts into account and hence use the Q3 number.

Table 2 presents the three scenarios and the most important characteristics of the judgment. The first column of table 1 displays the strength of the evidence (we denote this as the scale t). The assumption is that the financial agent obtains a consensus opinion about the expected price change for the next period with a level of disagreement displayed by the experts. The expert opinion is that given all costless and costly public information, the stock price has to rise 4%. We assume that this consensus opinion is true meaning that given all information the valuation of the imaginary stock price is 104. Disagreement amongst experts can exist because the experts use different models to make a forecast and have different relations with the firm. We suggest that this is a true representation of the financial environment.

Table 2.

Scenarios

The expert opinion about the next period's return is 4% in all cases. The historical sample estimate for the next period's return is 1.668% (std.dev. of 4.17%). For each opinion, a signal-strength (scale) is given in the first column. the second coums shows the square root of ther strength of the evidence or the standard deviation of expert opinions. Columns three and four show the highest and lowest forecast based on the strength of the evidence when the two opinions are symmetrical around 4%. The last column shows the Bayesian forecast for the next period's return for the three scenarios. Numbers in the first column are expressed in percentage points.

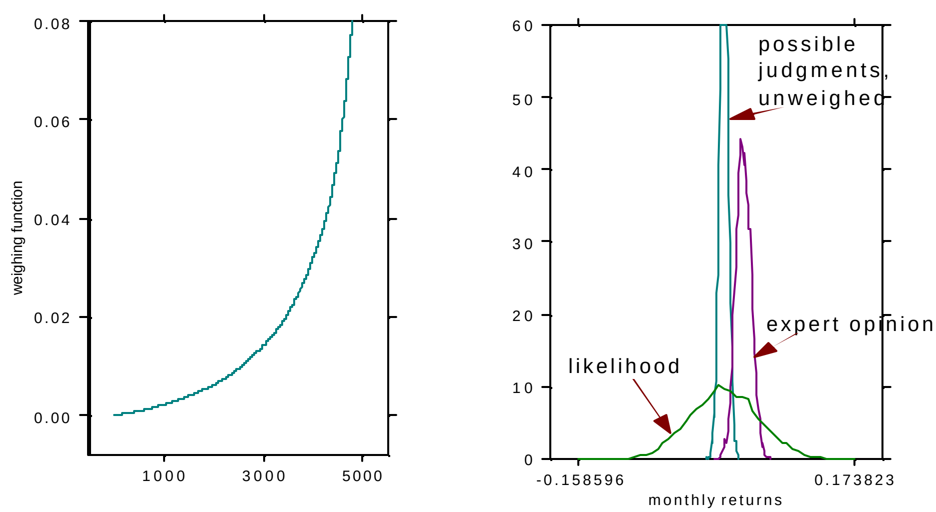
strength of the evidence t	standard deviation $\sqrt{t}$	Highest expert forecast	Lowest expert forecast	Bayesian forecast
0.00008	0.89944%	4.744%	3.256%	2.262%
0.00085	2.91550%	6.426%	1.574%	1.728%
0.00331	5.75330%	8.788%	-0.008%	1.680%

In the second column, the square root of this scale or the standard deviation of the expert opinions is given. Hence, when experts disagree, the standard deviation of their opinions is as least as large as the standard deviation of the sample returns (in this example 5.75% compared to 4.17% for the historical sample). In section 3, we denoted this strength of evidence as the precision of the information from informed

market participants. It is the noise in this signal that causes the small market imperfections in the Grossman-Stiglitz model [1980].
Columns three and four give the highest and lowest forecast as communicated by the experts. It is calculated from the definition of the Parkinson measure and assumed to be symmetric around 4%. The last column shows the Bayesian forecast made by a rational financial agent who observes the historical price changes, the expert opinion and the level of disagreement amongst analysts.

If there is no disagreement among experts, the financial agent's Bayesian forecast will be 4% (all the weight in the prior is on the 4%). In the case where the level of disagreement amongst experts is low ($t = 0.00008$), the financial agent expects that the return on the stock will be 2.262%. Looking at table 1, this is the case of the large stocks. Figure 1 shows the environment for the financial agent we model and his weighing function. In the left panel, the weighing function is shown for the case where the precision of the information is high. In the right panel, the likelihood is shown as well as the expert opinion and the distribution of possible outcomes. These possible outcomes are thought of as the possible returns the financial agent has in mind based on the historical price changes. In the simulation, these possible outcomes are the bootstraps for the mean. The weights of the weighing function are ranked according to the 5000 possible outcomes for the next period's payoff and are accordingly ranked on the X-axis. The left part of the weighing function denotes the weights assigned to the lowest possible outcomes by a rational investor.

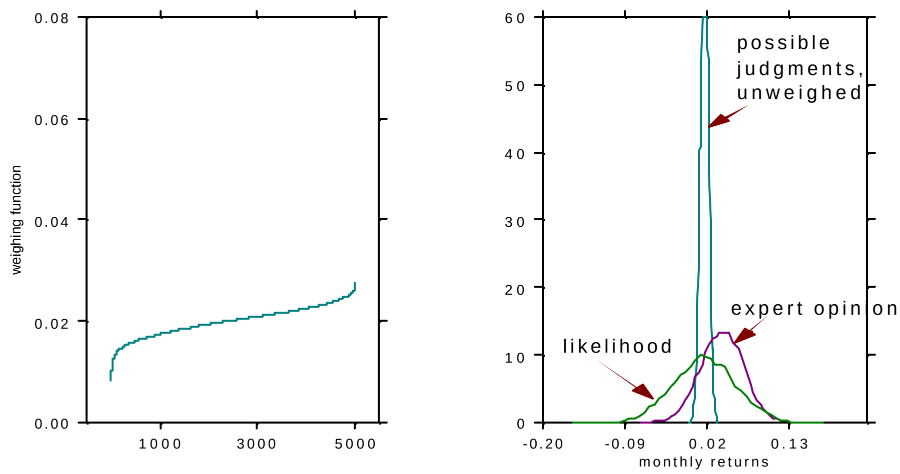
Figure 1: low level of disagreement among experts



From figure 1, some interesting features emerge. In the case where the information signal is strong (or in other words, if the strength of the evidence is high), the weighing function of the agent is to a large extent influenced by the optimistic expert opinion. Most of the decision weight goes to the larger possible outcomes for the next period's return. In this case, the Bayesian forecaster sets his expectation about the next period's return at 2.262%, which is above the expectation of the financial agent who does not use expert information (1.668%). Also, in the right tail of the possible judgment density, the weights sharply increase. The deviation of the updated judgment from to the mean is in this case economically important (.594% per month).

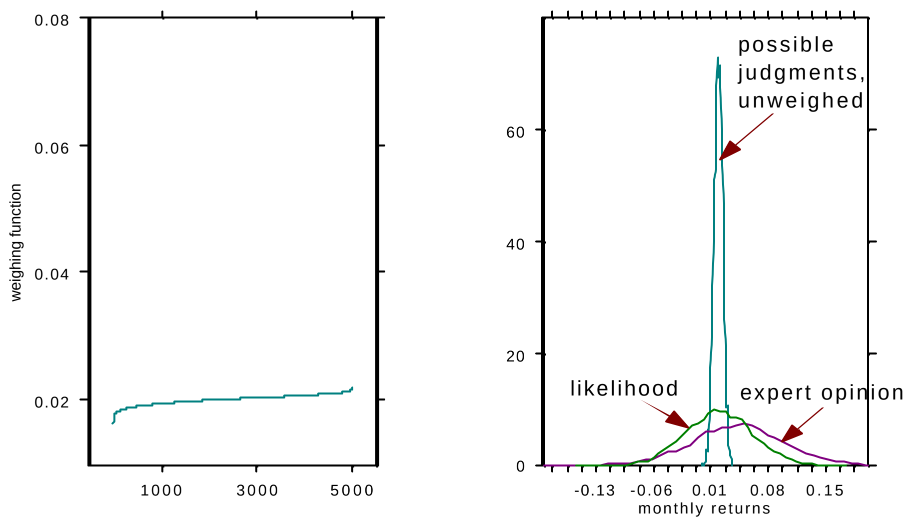
If the expert signal is weaker however (bottom panel), the weighing function is far more conservative, even with an optimistic expert opinion of 4%. In this case, the standard deviation of expert opinions is still lower than the standard deviation of historical returns (2.92% compared to 4.17%). We see that weights are only slightly increasing for the larger possible judgments (figure 2). On the lower tail, the agent displays a not much belief in a bad outcome of the return. What is surprising is that the difference in the steepness of the weighing function is large. This means that the rational agent already puts more weight on the statistical evidence. The expected return in that case is above the price trader's expectation (1.728% or 0.060% per month).

Figure 2: moderate level of disagreement among experts



When there is a lot of disagreement among experts, almost all the weight that a rational financial gives to the evidence goes to the historical observations of returns (figure 3). The weighing function is rather flat, indicating that even in the case of an optimistic expert opinion of 4%, the rational financial agent will have no strong preference for any of the possible outcomes her has in mind. The only effect the optimistic expert opinion has on the judgment of the rational financial agent is that he thinks that a bad outcome is a bit less probable than a good outcome. The Bayesian forecast of the return is 1.680% or 0.012% per month above the historical mean.

Figure 3: high level of disagreement among experts



Is there a case for momentum effects?

What does this imply for the existence of momentum? In the scenario where the strength of the evidence is high, the rational financial agent expects that the price of the asset will go up and hence bids up its price. In the second and third scenario, the rational agent does not give a lot of weight to the expert opinion because the strength of the evidence is low. In this case, the expected return will not change much (the financial agents expects that the price will not increase much, in contrast with the optimistic expert opinion). In the case where the expert evidence is strong, the precision of the information will be high and the prices of these stocks will be bid up. Hence, new information will quickly be reflected in the price and autocorrelation in returns will not be observed.

The next question is to what extent the diffusion of information induces momentum effects. In order to answer this question, we conducted the following simulation. In all three scenarios, the rational financial agent makes a judgment about the expected return. With respect to his consumption behavior, he will accordingly buy additional units of the asset based on this judgment. We assume that there is no additional information about the asset in the next periods and hence, the expert continues to value the price of the asset at 104. We simulate the return after two periods when prices are bid up relative to the expected return. If the financial agent bids up the price of the stock when the strength of the expert evidence is high to 102.262 (expected return of 2.262%, the experts still expects the stock price to rise 1.699%. In table 3, we show the simulated forecasts for the three scenarios assuming that there is no additional information in these consecutive periods and the expert's valuation remains the same. Additionally, we assume that the financial agent observes that the expert consensus remains the same and hence puts more weight on this opinion in the next period. We formalize this increase in weight by doubling the strength of the expert evidence.

Table 3.

Bayesian forecasts in the consecutive periods.

Column 1 of table 3 displays the stock price at the beginning of t=2. The second column shows the expert opinion at the end of t=1 in the case when there is no additional information. Column three gives the strength of the evidence, which is the double of the strength of the evidence in the previous period. The fourth column shows the Bayesian forecast for the second period. Column five shows the price at the end of t=2

P at t=1	expert forecast	*t*	Bayesian forecast at *t=2*	Price at the end of *t=2*
102.262	1.699%	0.00004	1.671%	103.97
101.728	2.223%	0.00043	1.693%	103.45
101.680	2.281%	0.00166	1.671%	103.38

Table three shows the case of no additional public information, and the case of an increase in the weight a financial agent gives to the expert opinion when this opinion remains the same as in the previous period. In this case, the stock price will only rise to its true value of 104 in the first scenario after two periods. In the other two scenarios (where the strength of the evidence is lower), the true stock price (104) will not be attained after two periods. This means that differences in the information diffusion can account for the existence of momentum.

The cases studied here are intermediate cases. Remind that the stock price will rise to its true value in the first period when there is no disagreement among experts. On the other hand, we observed cases of higher levels of disagreement than the ones tested here (the Q1 and Q2 decile in table 1). Hence, in the imaginary case that the expert forecast is the true forecast and in the case where financial agents are rational, we can still find that momentum effects can exist because of frictions in the market of information. Here we defined these frictions as the difference in the expert opinion because of the use of different complex models and a different relation between the experts and the firms [as documented by Lim, 2001].

Further applications of this Bayesian method

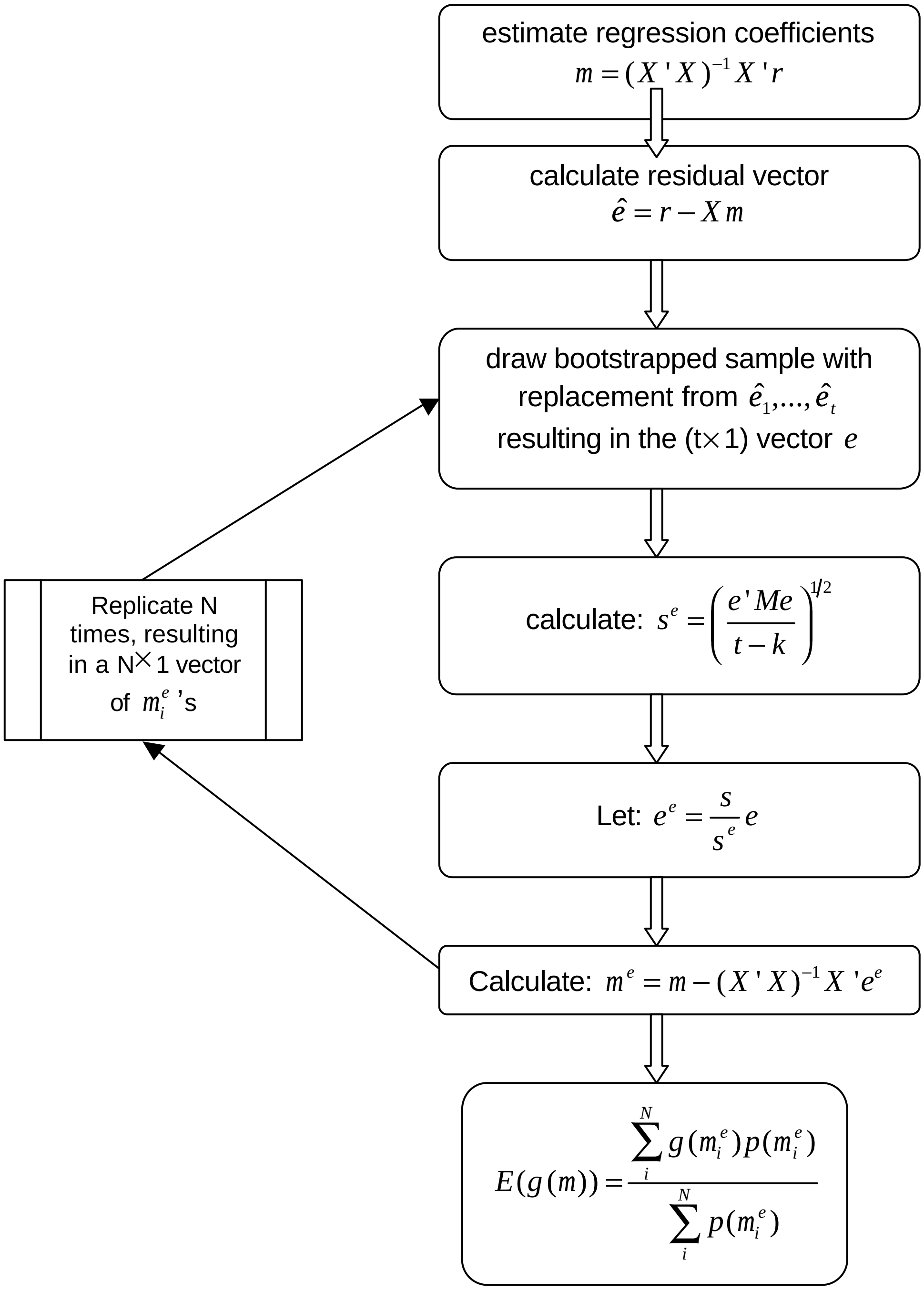
Applying this BBR methodology in decision problems in finance is an example of one of the challenges Geweke (2001) indicates as a possible application of Bayesian sampling methods. He mentions that applications that are relevant for real cases can prove their relevance to probability forecasting and decision-making. Real cases in finance are the estimation of the input parameters for the expected return-variance rule, fund selection, evaluation of projects, the determination of th e initial price of an IPO.

In a comment to this paper, Lenk and Wedel [2001] also recognize that a shift in the use of Bayesian methods from inference to entire decision problems is a challenge for Bayesian analysis. They state that making better decisions implies that there is a need for research that develops better algorithms and frameworks to clarify objective information and to test the impact on decision problems. Moreover, a part of this problem concerns the choice of appropriate priors. Geweke [2001] recognizes this as being one of the major problems practitioners are confronted with. This is an interesting topic that is open for future research, especially in financial markets [Lenk and Wedel, 2001]. Relevant choices of prior evidence or expert opinions are not always available. Because the simulation-based Bayesian methods become more and more available and because computational power increases, there is a possibility for further evaluation of this topic.

Conclusion

Descriptive behavioral models provide explanations for observed anomalies assuming that financial agents are irrational. The violation of the rationality axiom is taken to be true for financial agents relying on general findings from psychology. In this paper, we explore whether the momentum anomaly (that cannot be explained by risk) can also empirically be found when markets are efficient and agents are rational. We assume that there is a market imperfection in that the strength of the expert evidence is noisy. Hence, the costly public information they reflect through the forecasts is slowly diffused in the markets and prices do not fully reflect all costly public information. According to the Bayesian forecasts and the weighing functions for the evidence we computed, momentum or underreaction can exist in efficient markets where agents are rational. This is even true for samples of liquid stocks. It also implies that when the strength of the evidence will increase, momentum strategies will be less beneficial. Hence, this study of normative behavior also implies that an improvement of this segment of the market microstructure will increase market efficiency. More concrete, experts should be able to communicate fully independent opinions, regardless of their reputation. If that is the case and they have a strong consensus about the stock's value or payoff, prices will better reflect the available information.

Appendix 1: The general Bayesian bootstrap regression algorithm.



Note: N denotes the number of replications and the superscript ^{e} denotes a drawing from the empirical distribution.

IV b

Bayesian Forecasting and Stock Selection

Bayesian Forecasting and Stock Selection

Abstract

Investor behavior can explain to some extent the stock market anomalies from a psychological viewpoint. A lot of models are presented in the recent literature without testing predictability implied by the models and without a discussion of the limitations that are implied by the design. Mostly, these models are descriptive. In these designs, the question about relevant normative models is left aside. In this paper we propose a normative model that allows empirical testing of whether the way investors should behave given the information is useful in making judgments in financial markets. Contrary to most papers, we apply individual priors to form a judgment about the future price change of each asset at each point in time. These priors are the expert opinion and are given by the one-year consensus forecast of earnings yield as provided by analysts. This design allows tests of the predictions for a normative setting using actual market data. Comparing Bayes' rule to decisions by an uninformed agent, we find that economic loss is lower for the uninformed agent than for the Bayesian trader under several specifications. However, using expert information in the Bayes' rule leads to better predictions for stocks that do not have high-risk characteristics.

1.Introduction

In the previous part of the dissertation we have shown that modeling Bayesian forecasting (or in other words studying normative behavior of financial agents) leads to additional insights in the dynamics behind expected returns. In this part and the following part of chapter four, we will analyze Bayesian forecasts of returns for stock selection and asset allocation purposes.
Studying normative models can give more insight about the optimality of the behavior that is considered as being correct. Forecasting returns for stock selection purposes is a difficult task. Especially where the environment is complicated, it is rather hard to judge whether choices are optimal [Einhorn and Hogarth, 1988]. The environment that we focus on in this part of chapter four is the following: a stock selector makes forecasts about the future returns for the assets in his sample and buys the stocks that have the highest forecasts. The main question we want to address with this empirical study is whether ex-ante rational forecasts are also optimal in an ex-post evaluation of a stock selection. Second, we analyze the environment in which the stock selection is conducted questioning how the environment influences this kind of decisions in financial markets.

In this paper we empirically explore the properties of rational judgments or Bayesian forecasts of future returns compared to other types of forecasts when the objective is to perform a stock selection. Different aspects are empirically studied. In the first place, we compare the optimality of rationally forecast returns as opposed to simple forecasts (using the sample mean of past price changes as a linear forecast based on past price changes) and as opposed to an expert opinion of the forecast with respect to stock selection. Secondly, the procedure described in the previous part of the dissertation and also used in this part allows us to analyze the characteristics of the environment in which the forecast is made. The most important characteristic of this environment is the noise in the information that financial agents observe [as in Kinney et al., 1999].

Forecasting individual stock returns for stock selection purposes is a difficult task. In this part of chapter four, we compare the results from a monthly stock selection based on rational forecasts to a stock selection based on the historical mean and a stock selection based on an expert forecast of returns. A first simple criterion to evaluate the ex post optimality of the stock selection is the comparison of the returns from for the three types of forecasts. Next to these returns, we also evaluate the environmental characteristics of the decisions made. This implies that we also look at the size of the individual firms and the strength of the expert evidence when an expert forecast is used to make a decision about the return. The reason why we also

focus on these environmental characteristics is because of the growing efforts from behavioral finance literature to explain market anomalies [as in Barberis et al. 1998 among many others] and because of the growing attention of the importance of information diffusion for the stock price process [as in Hong et al., 2000]. We analyze these questions for ten portfolios that are formed based on the forecasts of returns.
Next, we extend this analysis with additional investigations about the ex post optimality of the performed stock selection. First, we analyze whether the stock selection algorithms are optimal when we correct for risk. Risk is modeled by means of the CAPM beta because of its universality and because in chapter two, we found no evidence against the CAPM fro European stock markets using country portfolios. Second, we also evaluate the forecasts based on a loss function analysis. Because it is difficult to analyze whether an investor will ex post evaluate a forecast as good or bad, we apply different types of loss functions to answer this question.

Three different forecasts of future returns are studied for their relevance for stock selection purposes. We study an investor who is convinced that he can capture the complexity of the environment by looking at past price changes alone in order to form his predictions. We denote this financial agent as the uninformed agent. As in Hong and Stein [1999] this investor is limited in the way he sets his forecasts. These predictions will be simple functions of the observed past price changes. We apply the same definition of 'uninformed' as in the previous part of the dissertation. Markets are efficient according to the Grossman-Stiglitz definition, and there are small imperfections in the information market. The uninformed agent only uses costless public information that is reflected in the prices.
Next, we will focus on a second investor who is aware that the expert opinion is valuable for his decision-making process. This agent is denoted as the rational agent. We specify that this investor takes into account the consensus forecasts as communicated by analysts who are the experts in this chapter. The rational agent does not rely on the opinion of the experts alone because there is noise in the pool of information coming from different experts. The choice of this information as prior information is crucial for this analysis and will be extensively discussed in section 3 of this chapter. This agent will process this evidence in accordance with Bayes' rule (as described in the previous chapter). In this way, he processes the costly public information that is reflected in the expert opinion. This is different from a setting with heterogeneous[1] agents [as in Hong and Stein, 1999] and the most important is that we do not assume that uninformed agents do not have access to the expert

[1] The agents in the Hong-Stein [1999] model are heterogeneous because there are information asymmetries. In this paper, there are also information asymmetries, but we only study the stock selection of the rational agent without assuming that there is an interaction between heterogeneous agents. In other words, heterogeneous agents in this paper are defined by the information they hold, not by the information they trade upon. The experts themselves do not trade.

information or the opinion of newswatchers who set their beliefs based on the stock price fundamentals and use complex models.
The third agent in this paper making a stock selection is an agent who only takes into account the forecasts that are communicated by experts. This agent observes the expert opinion and disregards the fact that there is noise in this information, based on the idea that the expert forecast is more accurate since analysts hold costly public information. We denote this agent as the expert agent.
In this framework, we are able to compare the stock selection abilities of an uninformed agent, a Bayesian trader who also observes the expert opinion and, thirdly, the expert opinion itself.

The remainder of the paper is organized as follows. Section 2 outlines the theoretical background of the Bayesian simulation procedure used in this paper to forecast returns using Bayes' rule. In section 3, we motivate the choice of the prior. We motivate the choice both from a theoretical viewpoint and from the fact that the prior data indeed is a good reflection of an expert opinion of future return. Section 4 describes the data and the practical issues of the computation. Section 5 analyzes the results. Evaluating whether a judgment in finance is an appropriate one is not straightforward. To repeat the problem: it is difficult to evaluate choice as being rational ex ante versus optimal ex post. Therefore we evaluate the predictions in ways that seem important in the decision process in which the estimated parameters are used in finance. Section 6 concludes.

2.A Bayesian Bootstrap Procedure using individual prior information

An important consideration in this paper is the relevance and the tractability of the behavior according to Bayes' rule. In order to weigh the evidence of the second investor, we model his prediction as a weighing of the evidence at hand using Bayes' rule using the same algorithm as described in the previous part of this dissertation. This means that we do not discuss if the Bayesian language is the appropriate one compared to others. However, we are able to study the relevance of the Bayesian language for financial agents for stock selection purposes. Does rational forecasting improve stock selection? What we in fact do using a Bayesian setting, is measuring the strength of the available evidence by its relative importance to the prior scale. This means that there are some requirements to develop such a framework, which we think are mostly fulfilled in a finance setting. There is a numerical scale, there are canonical examples for the probabilities, and there exist calculus frameworks to obtain the complex judgments [Shafer and Tversky, 1988].

Let us describe the objective in this paper using the framework from part IV a. The first specification is that, under the assumption that the historical estimate of the covariance matrix is the true population matrix, we only make inferences about m, the expected return. By using the expert opinion, common factors across stocks (such as industry membership) are captured by the Bayesian investor when he uses the evidence form the expert forecast. The expert uses models and theory in order to account for all available information. If there is predictability in the monthly return series, we assume the expert will exploit this possibility to make his forecast more accurate. Secondly, to reduce the loss of information implied by a common prior, we assess this problem for each individual stock *i*, forecasting m_i based on specific prior information $p_i(.)$. In order to apply the first specification, we use a Bayesian simulation method. For each individual stock, we interpret the estimation of the maximum likelihood m as a regression of historical observations r_t on the unit vector under the assumption of a normal linear model. This means that $r_t = mi_t + e_t$, and the estimated expected return m is evaluated as a forecast of future return for the first type of agent looking at historical price changes alone. Also, this setting implies that $s_r^2 = s_\varepsilon^2$. In order to estimate the expected return for an agent dealing with both types of information and applying Bayes' rule, we use a Bayesian regression as proposed by Geweke [1986], but in a non-parametric setting. The actual simulation is based on the Bayesian bootstrap regression that was described in the previous chapter [Heckelei and Mittelhammer, 1996]. In what follows, we specify $m = m_i$, the next period return of an individual stock we want to forecast. Henceforth, the subscript *i* is dropped for notational simplicity and the model development is written in a general form. Also, for notational simplicity, the subscript *t* is dropped from r_t, with r denoting the time-series of historical price changes.

The joint posterior probability density function of the parameters of interest is calculated to make inferences (eq. [1]). The random vector of observations has a density $f(r|m,s)$ which is the likelihood function. Prior information about the parameters of interest is described by a prior density function *p* of the form *p(m,s)*. The joint posterior is then :

$$[1] \qquad p(m,s\,|r) \propto p(m,s)\,f(r|m,s)$$

To make inferences about one of the parameters described in the posterior density, it is necessary to calculate the marginal posterior distribution of interest. Studying the mean vector alone, we illustrate this estimation for the expected return. In part IV a, we illustrated this Bayesian Bootstrap Regression and showed that expectations of a function of the parameter of interest can be sampled in the following form :

[2] $$E\left[g(m)\right]=\frac{\sum_{n=1}^{N} g(m_i^*)\,p(m_i^*)}{\sum_{n=1}^{N} p(m_i^*)}$$

Using [2] allows us to estimate any expectation of a function on a parameter of interest as there are: the posterior mean, the posterior variance and the numerical standard errors.

3. The choice of analysts' forecasts as the prior information

Earnings yield and return

In this paper, we apply a Bayesian simulation method to estimate the stock's future return in order to evaluate the forecasts of an uninformed agent that takes into account costly public news from experts compared to the forecasts formulated by an uninformed agent alone. It is common practice to use one shrinkage parameter for all stocks in the market. This induces loss of information by assuming that the prior is the same for all stocks. The problem at hand, however, is that for individual priors, a decent prior specification for each individual stock is not straightforward. As Geweke [2001] notes : *« A research challenge in Bayesian forecasting is inventing new models, with one eye fixed on the demand of practical forecasters and decision makers, and the other on the rich opportunities opened up by advances in simulation methods for Bayesian inference. »*. We think this application is one example of such a challenge. Given the very complex environment, and the pay-offs and losses that are linked to a judgment about expected return, this Bayesian forecasting application is relevant. This section describes a possible prior choice that is both motivated by the behavior of practitioners and, under some assumptions, by theory. This section refers to part III a, where the relationship between earnings yield forecasts and expected returns is outlined.

Starting from the dividend discount model (DDM) under the assumption that the dividend pay-out ratio is one, the earnings yield ratio equals the required return minus the growth rate [Dechow and Sloan, 1997]. Otherwise formulated, the earning's yield and the growth rate together, under this assumption, provide the required return. In equation [3], *E/P* is the earnings yield ratio, m is the required rate of return and g is earnings growth.

[3] $$E/P = m - g$$

It is common practice that practitioners compare this level of required return to the expected return from asset pricing models. The information obtained from this difference is for instance useful to investment processes. The drawback in using present earnings yield as a prior is that present earnings do not entirely reflect consensus expectations derived from fundamental values. Hence, earnings yield forecasts by analysts are a better approximation [see Dechow et al., 1999 and chapter III of this thesis].

[4] $$\frac{E_{t+1}}{P_t} = m$$

Under the described assumptions, we propose to use one-year consensus analysts' forecasts as prior information or expert information or otherwise said: "the judgment of practical men" as Markowitz [1952] calls it.

What is the actual information in the analysts' earnings forecast? An analyst announces a one-year earnings forecast. By this announcement, he expresses an expected price change or an expected future return. Suppose that the present value of the stock is 10. The expected return is 5%. At t=0, the implicit earnings per share forecast is 0.5 (we use the same relation as described by Cochrane [2001] and outlined in part IV a). If the consensus of the experts changes to 0.6, the stock price is expected to be 12. Dechow et al. [1999] observe that this one-year earnings forecast is a good predictor of firm value. Hence, an announced expert opinion implies a change in the firm's stock price and hence a change in the same direction of the expected payoff. In the above example, the price shifts from 10 to 12 and the expected payoff is 20%. The rational investor will weigh this expert evidence according to Bayes'rule.

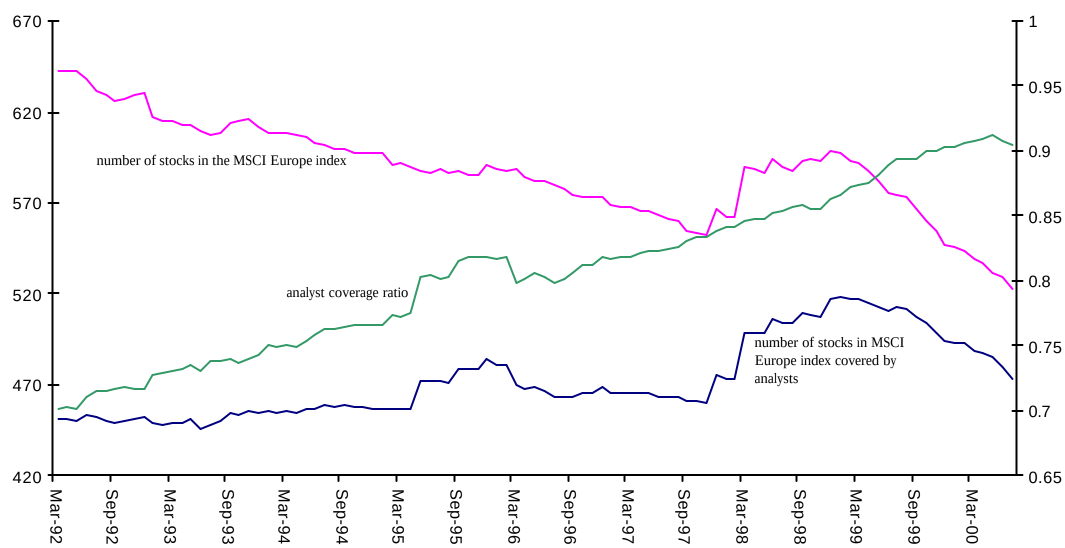
The expert opinion data

In order to assess the question of Bayesian forecasting in finance, we use a set of European stock data. We collect all relevant information (returns as the percentage change in the return index and analyst data) about all European stocks in the MSCI Europe index between March 1992 and August 2000 from the IBES database. Coverage of European stocks by analysts before that period was rather limited. From the IBES database, we collect the mean consensus forecast, the highest and lowest forecast and the number of analysts covering the stock. Because our database is prone to a selection bias, some details are required. The first question is of course why only stocks with a large market capitalization should be used in order to test the Bayesian forecasting in the European market. Firstly, this set of stocks is the most selected set among portfolio managers. For reasons of liquidity, institutional investors are often restricted to this sample. Secondly, because this analysis is as strong as the number of stocks that can be used being covered by analysts, we are not able to look at the updating of expected return for stocks that are not covered by analysts.

The final database we use is an intersection database between the IBES data and the Datastream dataset of market values and returns for the stocks in the MSCI Europe index. All stocks that were in the MSCI Europe index are available in the Datastream database, but only the ones having analyst forecast data in IBES are used in this analysis. The Datastream data is collected from January 1987 to August 2000.

The average number of stocks in the index during this period is 588. The average number of MSCI Europe index stocks that were covered by analysts is 471. Figure 1 shows that between 1992 and 2000, the coverage of index stocks in Europe increased on a steady basis. In 1992, an average of 16 analysts announced a one-year ahead forecast on a stock. This increased to a maximum of 21 analysts in 1997, and in 2000 this number is 19. Given the increasing coverage of stocks, this implies that at the end of the sample more stocks were covered by, on average, less analysts. A last important feature of the dataset is that Portuguese stocks entered the index only at the end of 1997.

Figure 1.
The MSCI Europe and analyst coverage.



Information diffusion

Recent papers stress the importance of the information diffusion in financial markets [Hong et al., 2000]. Not only the direct information on prices and fundamentals is important but also the way this information diffuses into the market. Because we stress market imperfections in the information market, the noise in the information in the analysts' forecasts is important. Many papers document that analysts' forecasts are overoptimistic [Lim, 2001]. There are several reasons why this bias may exist. But Lim argues that the idea that this bias is possibly due to irrationality of analysts is not necessarily correct because the judgment about the forecast they want to communicate is rational given the environment of the analyst (depending on the size of the brokerage firm, the reputation of the analysts among other things). In order to make Bayesian forecasts using analysts' forecasts as prior information, the interpretation of the strength of the evidence[2] given by experts or analysts is important as well [Griffin and Tversky, 1992]. In terms of decision-making, the question is raised how much weight a financial agent will attach to the expert opinion or the consensus forecast. In part IV a, we showed that the strength of the signal is important for the eventual weight function.

In this view, the modeling of the strength of the evidence of the expert information becomes very important. Hong et al. [2000] suggest that firm size and residual analyst coverage are a good proxy for the speed of information diffusion. Firm size explains most of the number of analysts covering a stock. If information is released into the market at high speed, agents will have low uncertainty about the consensus expectation. Following the findings by Hong et al. [2000] this would imply that agents are confident about expectations for large firms. In our setting, this would be an assumption that oversimplifies the market behavior since the dataset consists of stocks that are in a large index. Hence, we apply a different methodology to formalize the uncertainty about the fundamental signal. In line with the measures used by Kinney et al. [1999], we use the Parkinson measure[3] [1980] to model this parameter based on the highest and lowest consensus forecast of earnings yield. Suppose there is a consensus earnings yield of 3%, and the highest forecast is 5.3%, the lowest 1.6%. In that case, the Parkinson measure is 0.0004942, which can be regarded as the variance for the expert opinion or the strength of the signal. This number is equivalent to a standard deviation of 2.22%. The idea is that if analysts disagree about the consensus, the highest and lowest forecast are far apart and the investor gets a weak signal from the expert opinion. We further denote this Parkinson measure as τ. We repeat the Griffin and Tversky [1992] finding, that

[2] By the strength of the evidence we mean the degree of consensus about the forecast. If the dispersion of analysts' forecasts is high, the strength of the evidence is low.

[3] variance or dispersion of the forecasts = 0.361*(highest forecast – lowest forecast)²

when agents are confronted with a difficult task, which clearly is the case here, most attention goes out to the strength of the evidence. So in order to condition the tests on environmental realism, we focus on this strength of the evidence rather than on the weight of the evidence[4].

The one drawback this measure has is when there is only one analyst covering the firm. In that case, we set the prior information to be diffuse[5]. For our sample of large capitalization stocks in the MSCI Europe index, we find that this case hardly ever occurs. If there is coverage of a stock by analysts, in almost all cases there is more than one analyst. Another problem can arise when there are few analysts, where the agents follow an analyst with a better reputation to set his expectations [see Lim, 2001 for an explanation of the phenomenon] (difference between highest and lowest estimate is very small). No adjustments are made for this case, since this concerns one of the main questions in this paper. Suppose that the number of analysts is low and the opinions are close together, how informative is the expert opinion? Furthermore, this modeling of uncertainty of the expert information stream allows for small stocks to provide confident consensus expectations.

Also in this sample, there are more analysts who cover larger firms [as in Hong et al. 2000], the question is to what extent our measure of dispersion formed by the difference between the highest forecast earnings yield and the lowest forecast earnings yield correlates with the analyst coverage. This is the same as asking the question whether there is a correlation between the strength of the prior evidence and the weight of the prior evidence. We expect that firms covered by more analysts have a high speed of information diffusion and hence we expect a negative correlation with the Parkinson measure. We find this indeed is the case, but correlations are low and in almost every case not significant. The highest correlation in absolute terms .13. This is somewhat surprising because it implies that the strength of the signal on evidence is on average not higher for firms covered by more analysts. Moreover, we think this Parkinson measure is a more robust measure of news diffusion for our sample given the decision problem at hand [Griffin and Tversky, 1992].

Because no really small stocks are included in this analysis, news can be diffused at high speed for firms with a small number of analysts for example in the case where the management has implemented a transparent governance structure. In this case, even with a small number of analysts, the signal coming from the experts will be evaluated as a strong signal.

---

[4] In this paper, the weight of the evidence is given by the number of analysts announcing a forecast

[5] In this study, diffuse means that we use two times the standard deviation of the returns of the MSCI Europe as a scale parameter.

In table 1, we show the measure representing the sample average of the strength of the signal ($\bar{t}$) for firm size deciles and for beta deciles (betas for all stocks are re-estimated monthly using the single-index market model and 60 months of historical returns). The average beta for market capitalization deciles is reported as well. By analyzing these figures, we get an insight on the link between the signal strength and the size of the firm. Also, by reporting the beta characteristics, we analyze whether riskier stocks are accompanied by weaker evidence or not (Most of these figures were already shown in table 1 in part IV a).

Table 1.

Stock characteristics and signal strength.

Table 1 shows the average characteristics for firms in decile portfolios ranked on market capitalization (MV) or beta. For the MV deciles, the Parkinson measure ($\bar{t}$) is displayed as well as the average beta for the decile. For the beta deciles, only the Parkinson measure is given. The Q1 portfolios are the portfolios with low values for the ranking variable.

	ranking variable: MV reported variable: $\bar{t}$	ranking variable: MV reported variable: beta	ranking variable: beta reported variable: $\bar{t}$
Q1	.02914	0.914	.00169
Q2	.02801	0.997	.00158
Q3	.00331	1.001	.00077
Q4	.00219	1.034	.00047
Q5	.00085	1.033	.00077
Q6	.00040	1.025	.00077
Q7	.00042	1.003	.00023
Q8	.00036	1.072	.00093
Q9	.00028	1.063	.00108
Q10	.00008	1.048	.00526

From table 1 we see that although the number of analysts is related to size and not related to the signal strength $\bar{t}$, there is a distinct relation between size and the sample average of τ. The level of agreement around a consensus forecast of a large firm is much higher than for a smaller firm and this relation is almost linear. Notice that we are only interested in the difference of the signal strength in the deciles, not in a correct estimation of the actual $\bar{t}$ for each deciles. Therefore we use the mean absolute deviation to get an estimate of $\bar{t}$. This correction is made because of the presence of extreme outliers in the data. The consequence is that the average level of

τ is not comparable across decile rankings, but as mentioned, this has no impact on the analysis.

In the second column, we see that the smaller firms in this sample have somewhat lower betas than the larger firms (consistent with the findings in Part II and conflicting with the findings on U.S. data). This relation is however relatively flat. But important is that this finding implies that larger stocks, that are somewhat prone to more market risk and are riskier if the CAPM holds, get stronger information signals. Finally, there is a weaker relationship between τ and beta deciles, but it is an interesting observation that the low and high beta deciles do have the weakest information signals (larger $\bar{t}$). This looks somewhat puzzling because the large caps have greater betas, stronger signals, but at the same time, larger beta stocks exhibit the weakest signals. The explanation for this finding is that the relation between size and beta is rather flat. Stocks with different betas are present in all size deciles. If we look at average betas for the beta deciles, low and high betas are far away from 1.

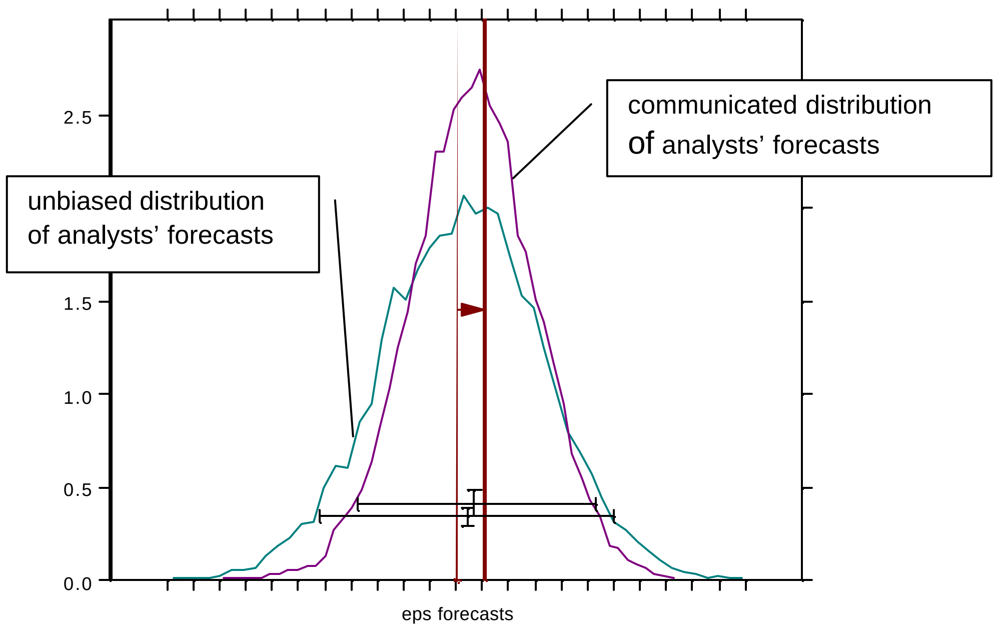
Herding behavior and over-optimism.

Past literature documents two possible biases in analysts' earnings forecasts. The two are related. First, different authors report the existence of herding behavior [Olsen, 1996, Cote and Sanders, 1995]. Depending on the reputation of the analyst and because of managerial incentives and relations between investment banks and firms, analysts tend to formulate opinions that are close together. Also because of managerial incentives and relations between investment banks and firms, analysts' forecasts are sometimes overoptimistic because analysts communicate favorable forecast in order to continue to receive as much information about the firm as possible [Michaely and Womack, 1999]. As mentioned, Lim [2001] argues that these biases do not necessarily imply that analysts' forecasts are irrational.

What are the implications for the rational investor using this relation? In figure 2, we present a graphical interpretation of the impact of both possible biases [as given by Olsen, 1996]. The distributions are simulated for the parameters from observed expert evidence of a randomly chosen individual stock. The unbiased distribution uses 0.75 times the first moment of the communicated distribution and 0.75 times the dispersion of the communicated distribution, in the assumption that the expert evidence for this stock is susceptible to both herding behavior and over-optimism.

A combination of herding behavior and over-optimism causes the level of the consensus expert opinion and the strength of the evidence to be too high. If the strength of the evidence is sufficient for the rational trader to give substantial weight to the expert opinion, the rational forecast using analysts' forecasts will be too high and there is a possibility of overreaction in the price and hence a negative momentum can be observed. It should be noted that given the evidence in table 1, the evidence is not always strong and the biases will not always cause overreaction.

Figure 2.
The impact of herding behavior and over-optimism.



4.Updating expected returns: the data and computation

We estimate the expected future return of all stocks with sufficient data for all months. We use 60 months of historical returns as known at t-1 and the analysts' forecasts at the end of t-1 in order to forecast the next period t return. Notice that even stocks with a short history of past returns are included in the analysis. This means that if a stock has a shorter history than 60 months, the uninformed agent will be able to evaluate the stock price based on the available price changes. Also notice that the forecast of the next period return is completely out-of-sample. For all available stocks and for all months (101) we obtain 47610 forecast returns using the Bayesian bootstrap regression procedure outlined in the previous chapter[6].

For the simulation procedure, we approximate the likelihood by N bootstrapped replications for this regression structure (see part a of chapter four). We use the empirical distribution for the residual vector. So in the regression model, we first estimate the residual vector e, with m as the maximum likelihood estimate of the

[6] All estimations are done in S-plus

expected return. Next, N bootstrapped samples with a size equal to the length of the vector of available historical returns are drawn, resulting in N new residual vectors. Inserting these new residual vectors into the equation results in N new estimates of the parameter of interest (eq.[6]).

[5] $$e = r - X\,m$$

[6] $$m^{n} = m - (X^{'}X)^{-1}X^{'}e^{n}$$

As explained by Heckelei and Mittelhammer [1996], a bootstrapped drawing from the original residual vector would imply a prior knowledge of the scale factor of the residual distribution. Therefore, the N drawn residual vectors have to be rescaled, implying ignorance about the scale factor of the original error distribution in [5]. In [7], s is the estimated scale factor of the error distribution and s^n is the estimated scale factor of a bootstrapped residual vector. Eventually, the posterior functional of m can be calculated as a weighted function of the m^ns (equation 8).

[7] $$m^{n} = m - (X^{'}X)^{-1}X^{'}e^{n} * \frac{s}{s^{n}}$$

[8] $$g(m) = \frac{\sum_{n=1}^{N} m_i^{*}\, p(m_i^{*})}{\sum_{n=1}^{N} p(m_i^{*})}$$

The weights in equation 8 are based on the substantial information from the prior. In the framework described in section 2, this density can be formulated as the normal density. The central parameter of this density is the earnings yield forecast for the particular stock. The uncertainty about this parameter is formalized by the Parkinson measure τ.

A final important issue in the computation is the numerical accuracy of the procedure. This is calculated by the numerical standard error (NSE, equation 9) of the procedure.

[9] $$NSE = \frac{\sum_{n=1}^{N} (m_i^{*} - \hat{E}m)^2\, p(m_i^{*})^2}{\left[\sum_{n=1}^{N} p(m_i^{*})\right]^2}$$

A low NSE corresponds to a reasonable choice for the importance function, which is of major importance. The mean NSE for the 50 randomly chosen assets for 1000 replications is .0029, which is reasonably low. Even more important is the check of the practicability of the procedure. The ratio of the numerical standard to the posterior variance is a good indicator of this practicability. In the estimation of the Bayesian expected returns, this ratio varies but in most cases is far above 0.1, which is acceptable [Geweke, 1989].

Figure 3.

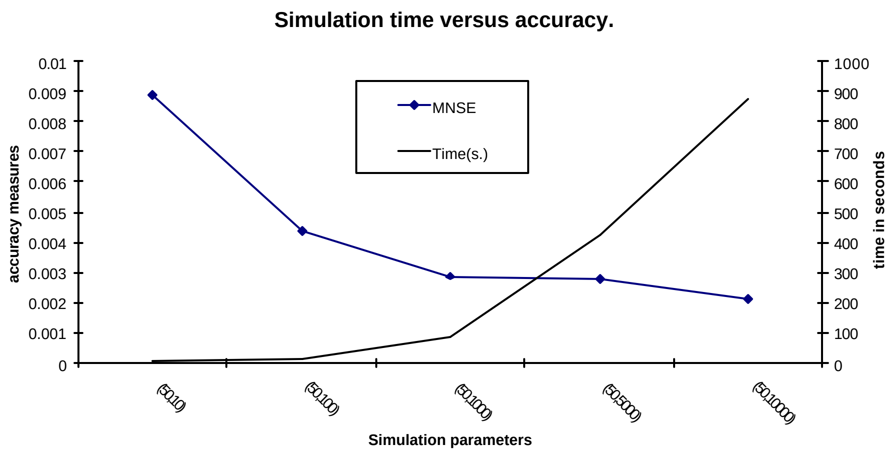


Figure 3 plots the mean NSE versus the average computing time (in seconds) for 50 randomly chosen assets. The simulation parameters on the X-axis denote the number of replications (N) used for these 50 stocks. From figure 3, we can see that the accuracy increases when the number of replications goes up from 100 to 1000. At the same time, there is an increase in computing time when the number of replications further increases. Therefore, all simulations are done using 1000 replications for each generation. In accordance with this, and perhaps even more important, is the mimicking capacity of the importance density for the posterior density [see Bauwens and Richard, 1985]. Geweke [1989] mentions that poor tail coverage is mostly observed when there are substantial fluctuations of the posterior mean for several thousands of replications. This is also checked for the randomly chosen sample described above and found to be solid. For the random sample, also the weight vector $p(m_i^*)$ is checked. Most weight vectors are well behaved in a sense that, in most cases, the mass is not on a few observations only.

5. *Empirical analysis of the return forecasts*

5.1 stock selection

An ex-post evaluation of rational expectations is a difficult task. Judgments that can be characterized as rational ex-ante are not necessarily the best ones ex-post. As an example, Lim [2001] arguments that analysts' forecasts are often evaluated to be too optimistic ex-post but are, according to Lim, ex-ante rational judgments. We evaluate the expected returns in this paper in several ways. At first, we perform a stock selection based on the expectations. Portfolios are formed based on the forecasts of the uninformed agent, the forecasts from the rational agent that are estimated by Bayesian updating and the forecasts of the expert agent. The three types of agents make a forecast of the future return for all stocks were the information is complete (historical data and an expert opinion) and form 10 portfolios (Q1 to Q10) by ranking the stocks from the lowest forecast to the highest forecast.

What we test for is the applicability of these forecasts for stock selection purposes. Next, we also analyze the kind of information that is communicated, asking the question whether different types of forecasts are communicated with a lot of certainty in the expert information or not. In other words, we evaluate whether the strength of the prior evidence has an impact on the forecasts. Also, we evaluate the forecasts relative to the size of the firm, asking the question whether optimistic forecasts are related to larger or smaller firms.

Table 2 reports the results of this analysis. In table 2, deciles are formed, ranking all available stocks based on the three types of forecasts in each month. Both market capitalization weighted (m.c.w.) and equally weighted (e.w.) portfolio returns are calculated for the whole period for the three types of agents and for the ten deciles. In table 3, the sample mean of the e.w. time-series (θ_e) is reported as well as the m.c.w. sample mean (θ_m). The standard deviation of the return series for the decile samples is reported as well. Next, the average market value (MV, in 100.000 DEM) for each decile is reported. For the expert agent forecast and the rational forecast, the signal strength parameter $\bar{t}$ is given as well.

Table 2.

In table 2, deciles are formed based on the monthly ranking of all stocks based on the expected future returns. In the first row of each panel, market capitalization weighted average portfolio returns (θ_m) and equally weighted average portfolio returns (θ_e) are given. For both return series, also the standard deviations are reported (σ_m,σ_e). MV is the average market capitalization of a stock in a decile portfolio. For the second and third panel, also the sample mean of the Parkinson measure ($\overline{t}$) reflecting uncertainty about the information is reported. All return figures are reported in percentages. The Parkinson measure is reported in decimals.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Panel 1 the uninformed agent										
θ_m	2.064	1.765	1.615	1.379	1.692	1.569	1.532	1.429	1.306	1.950
σ_m	6.717	6.013	5.435	4.716	5.026	4.633	4.539	4.188	4.826	5.826
MV	2002	3609	5774	7038	9906	12583	15361	16592	17809	16152
θ_e	1.646	1.154	1.344	1.121	1.358	1.280	1.353	1.243	1.373	1.885
σ_e	6.213	5.511	5.018	4.716	4.617	4.398	4.405	4.178	4.721	5.489
Panel 2 the expert agent										
θ_m	1.923	1.550	1.343	1.516	1.428	1.472	1.466	1.730	1.712	2.364
σ_m	7.838	5.571	4.707	4.608	4.617	5.036	4.884	5.101	5.415	5.527
MV	1958	18938	17599	16261	13771	13269	11112	9393	7880	4716
θ_e	1.615	1.266	1.220	1.253	1.235	1.179	1.030	1.326	1.310	1.920
σ_e	6.494	5.314	4.391	4.193	4.507	4.434	4.576	4.409	4.857	5.196
$\overline{t}$	.11424	.00060	.00024	.00013	.00016	.00020	.00023	.00035	.00046	.00205
Panel 3 the rational agent										
θ_m	2.346	1.768	2.015	1.728	1.590	1.415	1.372	1.529	1.584	1.337
σ_m	7.271	5.990	4.883	4.959	4.568	4.712	4.508	4.683	4.920	5.495
MV	2804	6851	10590	12770	13833	14407	12974	12136	10633	9659
θ_e	1.507	1.447	1.260	1.327	1.451	1.325	1.363	1.374	1.185	1.531
σ_e	6.23	5.105	4.728	4.718	4.309	4.578	4.502	4.501	4.456	4.875
$\overline{t}$	.04451	.001427	.00068	.00029	.00012	.00011	.00008	.00007	.00007	.00010

The first panel of table 2 shows the parameters for the portfolio deciles formed based on the forecasts by the uninformed agent (i.e. using the historical mean). These figures learn that there is only a weak link between the forecasts and the realized returns. Indeed, the decile of highest forecasts performs well ($q_{\mathrm{m}}^{\mathrm{Q10}} = 1.95\%$, $q_{\mathrm{e}}^{\mathrm{Q10}} = 1.89\%$, compared to a m.c.w. sample mean of 1.67% and compared to an e.w. sample mean of 1.57%), but more troubling is the fact that also the decile of lowest forecasts performs well ($q_{\mathrm{m}}^{\mathrm{Q1}} = 2.06\%$, $q_{\mathrm{e}}^{\mathrm{Q1}} = 1.65\%$). Also, there often is a difference in return between the equally weighted and market capitalization weighted return. The stocks with a larger market capitalization performed better in general in this period under study[7]. Other important observations are in the first place that the lowest simple forecasts are on average riskier stocks (with a standard deviation for the m.c.w. Q1 portfolio of 6.72% compared to 4.17% for the MSCI Europe index) and also smaller stocks on average. Stocks with higher simple forecasts are on average found for larger stocks and for less volatile stocks (Q7-Q9). The Q10 decile also consists of large caps, but more volatile stocks (standard deviation of 5.83%). Hence from the first panel, the main problem is that the lowest forecast stocks perform very well, and between the lowest and highest decile, the relation between forecasts and realized returns is quite flat.

The second panel reports the characteristics of the decile portfolios for expert forecasts. The same problem as in panel 1 is observed. Low forecast stocks perform well ($q_{\mathrm{m}}^{\mathrm{Q1}} = 1.92\%$, $q_{\mathrm{e}}^{\mathrm{Q1}} = 1.62\%$). It should be noted that from table 1, the portfolio of high expert forecasts (Q10) performs best ($q_{\mathrm{m}}^{\mathrm{Q1}} = 2.36\%$, $q_{\mathrm{e}}^{\mathrm{Q1}} = 1.92\%$). The low forecast portfolio consists of the small stocks in this sample compared to the other deciles. Note that in this analysis, negative forecasts (of earnings yield) are not excluded. The most extreme expert opinions are given for the smallest stocks. Both for the Q1 and the Q10 portfolio, average MV is small (1958 and 4716). Except for Q1, the lowest forecasts are on average found for the larger stocks in the sample (the average size of a Q2 firm is about 10 times greater than in Q1 and 5 times greater than in Q10). Notice that the relationship between size and the level of the average forecast is very different with the sim ecast where the higher forecasts were found for the larger stocks. It also denotes that the herding behavior and overoptimistic forecasts of the experts do not distort the rational agent's forecasts to a large extent. First, we saw that the expert forecasts for large stocks are given as strong evidence (low τ in table 1). Second, we see in table 2 that the forecasts for larger stocks are rather conservative compared to the other stocks. Together with the observation made by Kinney et al. [1999] that expert earnings forecasts are far

[7] The market-cap weighted average monthly return for the available dataset in this chapter is 1.670%, while the equally weighted average monthly return is 1.568%. The average monthly return on the MSCI Europe index is 1.590%.

more accurate in the nineties than in the periods before, news about large firms is expected to be quickly reflected in the stock price.

The parameter τ denoting the strength of the forecast is informative as well. The low forecast portfolio consists of stocks which are forecast with a lot of uncertainty (τ=.11424). The main reason why the consensus around these forecasts is low is because the negative forecasts are included in this analysis and those are communicated with a lot of disagreement among experts. The rational agent will give a very small weight to the expert opinion. This is beneficial to the rational investor, because the expert forecast for the Q1 portfolio is inaccurate. The strongest evidence is given in the Q4-Q5-Q6 portfolios (low τ). However, these forecasts seem to be the most neutral ones and the portfolios including the stocks that have a strong expert opinion performing worse than average.

More troubling is that the portfolio of high forecasts contains stocks that are forecast with a lot of uncertainty among experts as well (τ=.00205). In Part IV a, we have shown that updating with this kind of weak signals is not beneficial to a rational investor because the weight function is flat in that case. In other words, experts provide good forecasts at the upper side of the sample, but the expert opinion is not informative about it and the rational agent put a low weight on the expert opinion in his judgment. This means that for the Q10 stocks, good news will be slowly diffused into the market, and for this decile, autocorrelation in the individual stock returns can be observed.

Compared to the past literature [Lakonishok et al., 1994, Bakshi and Chan, 2000], it is surprising that in this sample the relation between forecast earnings yields and realized returns is U-shaped. Of course this analysis is conducted on a monthly basis, but forecast earnings yields are relatively persistent, meaning that the change of a stock from one decile to a totally different decile does not occur at a high frequency. Also notice that in the previously mentioned studies, stocks having negative earnings yields are excluded from the analysis.

Panel three reports the characteristics for the Bayesian forecasts computed by the BBR. The equally weighted returns suggest that this rational updating of expectations flattens out the returns across deciles. The two highest realized returns are still observed in the Q1 and Q10 portfolio, but they are less outspoken compared to the previous panels. The market capitalization weighted portfolios show that the relation becomes worse. For the Q10 portfolio, the m.c.w. return is 1.34% and the e;w. return is 1.53%, both below the sample mean. The only observable relation between forecasts and realized returns is an inverse one.

Based on the findings of the previous two panels, this means that the rational agent will select stocks from the Q7-Q8-Q9 portfolios and not from the Q10 portfolio in panel 2. We showed in the previous part that the higher the measure of

disagreement among experts is, the lower weight the agent will attach to these kind of opinions. Hence, when a rational agent selects the stocks with the highest expected return from the Q7-Q8-Q9 portfolio in panel 2, because in that case the strength of the evidence is high, it will lower the average realized return as observed in panel 3.
With respect to the strength of the prior evidence, the analysis reveals that a bad stock selection (in the case the agent buys the Q10 stocks) based on rational forecasts is caused by the way the expert opinion is communicated to the market. Still, the Q1 portfolio contains the stocks with the noisy signals. But on average, this signal is not as weak as for the expert forecasts, and the average selected stock has a larger market value. This means that the uninformative bad forecasts cause a flat weight function (based on the analysis in the previous part) and hence stocks with a strong past performance have a larger possibility of being influenced by the expert opinion. This will be observed, even in the case where a rational agent receives a pessimistic expert opinion. Panel 3 shows that this effect is positive, yielding a higher average return in the Q2 to Q5 portfolios. As mentioned, the explanation is that the rational investor gives more weight to the statistical evidence, which is beneficial because the analyst forecast is inaccurate in the Q1 decile. The impact of the uninformative pessimistic expert opinions is beneficial to the rational investor observed by a better performance of the Q2 to Q4 portfolios in panel 3 compared to the previous panels.
Also, τ declines almost linearly with the forecasts. We have shown that the rational agent has a tendency to select stocks that have an optimistic expert opinion, communicated with a strong signal. This is problematic because of the ex-post empirical finding that analysts' forecasts are optimistically biased on average implying that optimistically biased forecasts with a small τ are most likely to be selected by a rational investor. According to table 2, this is unattractive for the rational stock selector, resulting in a lower return for the Q10 portfolio in panel 3 compared to the previous panels. Panel 3 shows that the on average stocks with a strong expert consensus (low τ) are selected in the Q7 to Q10 portfolios.
Finally, we observe that in panel 3, the e.w. return is much lower than the m.c.w. return for the Q1 portfolio. The e.w. return for the Q1 portfolio in panel 3 is lower than in the previous panels. This means that, given the findings in the previous panels, especially the large capitalization stocks with a pessimistic expert opinion perform well. In table 1 we saw that there is a strong consensus around the expert opinion for the largest stocks. So the wrong pessimistic expert information on the largest stocks causes the very high return for the m.c.w. Q1 portfolio in panel 3.

Summarizing, we find that simple stock selection procedures are not beneficial for the rational investor. The main cause is the fact that experts disagree about their most optimistic forecasts that are accurate in an ex-post evaluation. Because the experts disagree, the rational agent will not take this opinion into account, causing the poor stock selection. False pessimistic opinions about the largest stocks cause the high realized m.c.w. return in the Q1 portfolio for rational investors.

5.2 Risk-based stock selection.

Table 1 showed that the strength of the signal depends on the stock's beta, the lowest-beta stocks and especially the highest-beta stocks are forecast by experts with the largest disagreement among experts. In table 3, we select portfolios based on the forecasts conditional on the stock's beta. We select three beta portfolios (of low, medium and high betas) for each of which three portfolios are selected based on the forecasts (low, medium and high forecasts), eventually leading to 9 portfolios. This is done in order to keep the number of stocks per portfolio comparable to table 2. Table 3 shows that for the simple forecasts, the stock selection for medium beta portfolios is best. The annualized premium for a long strategy in high forecast-stocks and a short position in low forecast-stocks is 6.80%. Also, for medium beta-portfolios, the relation between forecasts and realized returns is correct, meaning that the stocks with the highest forecast perform best ex-post. For low beta stock portfolios, the realized return for the low forecast-portfolio is greater than for the mean forecast-portfolio. For the high beta stocks, the relation between expected returns and realized returns is completely the inverse.

The accuracy of the expert opinion in the three beta classes is very different. Only for the low beta stocks, the expert forecasts are accurate. In two thirds of the months, the high forecast portfolio performs better than the low forecast portfolio ($\pi_{(H>L)}$ is 65.4%). Especially for the medium beta-stocks, where the expert information is communicated with the strongest signal (see table 1) on average, the forecasts perform poorly (with the highest ex-post return – 1.73% - for the stocks with the lowest forecasts). This is informative with respect to other inferences: the expert opinion is only accurate for the low beta stocks, for which the level of disagreement is relatively high.

In the last panel of table 3, the Bayesian forecasts are analyzed relative to the beta of the stock. While in the case of the unconditional stock selection the performance was bad (in table 2), conditioning on the beta of the stock improves the stock selection a lot. Both for low betas and medium betas, the relation between forecasts and realized returns is correct, yielding on average higher realized returns for stocks with higher forecasts. Moreover, where the performance was worst for high beta stocks in the previous cases, the relation between forecasts and realized returns is still perverse, but better than in the previous cases.

Table 3.

For each type of expected return, nine portfolios are composed. First all stocks are ranked monthly based on their beta (β). Next, each subgroup of beta stocks is ranked according to their expected return (μ). For each of the nine portfolios, the average monthly return is reported as well as the standard deviation of the return series (between brackets). Next, an annualized return spread is reported between the high and low expected return portfolio in each beta category ($A_{(H>L)}$). In the bottom row, the proportion of months in which the realized return of the high expected return portfolio was higher than the realized return of the low expected return portfolio is reported for each beta category ($\pi_{(H>L)}$). All return numbers are reported in decimals.

	Uninformed agent			Expert agent			Rational agent		
	Low β	Medium β	High β	Low β	Medium β	High β	Low β	Medium β	High β
Low μ	.0137	.0099	.0154	.0099	.0173	.0159	.0120	.0103	.0146
	(.043)	(.053)	(.062)	(.039)	(.050)	(.063)	(.041)	(.051)	(.058)
Medium μ	.0124	.0134	.0137	.0133	.0127	.0125	.0131	.0143	.0131
	(.039)	(.043)	(.052)	(.039)	(.045)	(.055)	(.038)	(.042)	(.054)
High μ	.0152	.0154	.0127	.0159	.0135	.0145	.0144	.0152	.0132
	(.041)	(.045)	(.053)	(.044)	(.049)	(.055)	(.043)	(.045)	(.052)
$A_{(H>L)}$	.0182	.0680	-.0319	.0744	-.0447	-.0167	.0292	.0604	-.0167
$\pi_{(H>L)}$	.545	.555	.515	.654	.455	.515	.545	.535	.555

All this indicates that when rational investors receive accurate signals, both with respect to the level of the forecast and the strength of the prior evidence, they are able to make an accurate selection. In table 1, we have seen that the expert opinion is uninformative both for the low beta stocks (Q1-Q2) and the high beta stocks (Q8-Q9-Q10). This implies that giving more weight to the statistical evidence for low and beta stocks and slightly adjusting to the expert opinion is ex-post beneficial. For high beta stocks, both the statistical evidence and the expert opinion provide wrong evidence. This makes stock selection for high volatility stocks difficult for a rational investor. Notice that this analysis is made for stocks that are part of a large and important index and it is remarkable that already in this sample this is the case.

5.3 A loss differential analysis

We started of with the remark that the ex-post evaluation is a difficult task. Therefore, we perform an additional analysis on the quality of the forecasts. We have seen that for stock selection purposes, rational investors are not better of ex-post if they do not condition on the beta of the stock. In this section, we evaluate the quality of the forecasts by means of loss functions. Forecasts are not evaluated for their stock selection purposes, but with respect to different possible objectives the investor has in mind when he makes the forecast.

There are different possible criteria for an investor to evaluate these forecasts ex-post (e.g. with respect to loss aversion). This analysis is performed in terms of loss functions and the evaluation is made for several realistic cases (e.g. an investor gives more weight in the loss function for forecasts that are too high), in which the investor can evaluate his forecasts ex-post. The evaluation of expected returns requires a different evaluation approach than for example macro-economic forecasts. Moreover, a statistical evaluation of the forecasts under study does not always imply that the forecasts are economically relevant [Diebold and Mariano, 1994]. In this section, we evaluate the forecasts based on different loss functions reflecting representative evaluation of the forecast error.

Denote by $\{m_{it}\}$ the series of realized returns in period t on asset i. We only compare the statistical forecasts to the Bayesian forecasts. Hence the question is whether an investor evaluates his forecasts positively ex-post when he collects information from experts. The series of forecasts by the Bayesian procedure are denoted by $\{\hat{m}_{it}^{B}\}$ and the series of forecasts by maximum likelihood are $\{\hat{m}_{it}^{M}\}$. Recall that these forecasts are completely ex-ante, based on information up to t-1. The forecast errors associated with both generated series are $\{e_{it}^{B}\}$ and $\{e_{it}^{M}\}$. In general, the loss function associated with the errors is $g(e)$. The difference between the losses from the two series is the loss differential $d_t(B,M) = g(e_{it}^{B}) - g(e_{it}^{M})$. Based on this loss differential, we can test the null hypothesis of equal accuracy of the forecasts or $E_t[d(B,M)] = 0$ [Diebold and Mariano, 1994]. If the accuracy of the forecasts by the uninformed agents alone is higher than the case where the uninformed agents also look at the newswatchers (or the experts), there is no reason for uninformed agents to update their expectations based on expert information (in this case analysts' forecasts).

We specify a general class of loss functions and explicitly test some of them. The essential idea behind these functions is that investors will evaluate the forecast error in several possible forms. One example: a particular investor could, a priori, try to use forecasts that are not too extreme, i.e., forecasts that are not higher than the realized return of a particular asset. When the forecast is indeed too high ex-post, the investor can evaluate the loss from being overweighted in that asset as important being indifferent about other losses at the same time.

Equation 9 specifies the general form of the loss function. In equation 9, only absolute loss is tested. Firstly, the investor is really worried about real gains and losses. Secondly, we also test quadratic loss, but there are few differences in interpretation under the specified loss functions.

[10] $$d_t(B,M) = \Upsilon_1 \times \Upsilon_2 \times \Upsilon_3 \times \left|\hat{m}_{it}^{B} - m_{it}\right| - \Upsilon_1 \times \Upsilon_2 \times \Upsilon_3 \times \left|\hat{m}_{it}^{M} - m_{it}\right|$$

$$\text{with } \Upsilon_1 = \begin{cases} c_1 \text{ if true} \\ 1 \text{ else} \end{cases}$$

$$\Upsilon_2 = \begin{cases} c_2 \text{ if true} \\ 1 \text{ else} \end{cases}$$

$$\Upsilon_3 = \begin{cases} c_3 \text{ if true} \\ 1 \text{ else} \end{cases}$$

The Υs are indicator functions, denoting the importance of a certain loss to the investor. Υ_1 denotes the aversion for forecasts that are too high ($\hat{m}_{it} > m_{it}$). The indicator Υ_2 represents aversion for bad market timing when the assets return turns out to be negative ex-post. The indicator Υ_2 is true when ($\hat{m}_{it} > 0$ & $m_{it} < 0$). The last indicator, Υ_3, is true when the forecast return of a particular asset is larger than the average forecast return while the realized return for that asset is below market return. In other words, Υ_3 is true when ($\hat{m}_{it} - \overline{\hat{m}}_{it} > 0$ & $m_{it} - \overline{m}_{it} < 0$). In this case, the forecast indicates a better than average performance of the stock while ex-post, the performance of the stock was below average.

Under this general loss function, seven specifications, defined by the form on the loss function as constructed based on the c_1,c_2 and c_3 weights shown in table 4 of the loss differential are obtained. For these loss differentials the null hypothesis of equal accuracy of the forecasts can be tested. This test evaluates the formulated hypothesis that agents that make rational forecasts perceive that they are better off when they make an ex-post analysis about the forecasts. In order to do so, we apply an exact finite-sample test. The Wilcoxon signed-rank (WSR) test is powerful for this purpose in a sense that it takes into account the amount of loss associated with a forecast [Diebold and Mariano, 1994]. The only caveat in applying this test is that it assumes symmetry of the loss differential. Equation 10 reports the test statistic for the WSR test and its associated test. In its Studentized version, this test is asymptotically standard normal.

[11]
$$S = \sum_{t=1}^{T} \mathrm{I}_{+}(d_t(B,M))Rank(|d_t(B,M)|)$$
$$Test = \frac{S - T(T+1)/4}{(T(T+1)(2T+1)/24)^{1/2}} \sim N(0,1)$$

Table 4 reports the parameters of the loss evaluation based on the forecasts. The table reports results both on the realized returns in the next month – which is straightforward because the analysis is done on a monthly basis – and on the average monthly return over the next year. The table reports results for different specifications of the indicator functions, or in other words, different ways of ex-post evaluation by the investor.

In table 4, different specifications of the loss function determining the loss differential are made. If an indicator function is true, then an investor evaluates this loss at x%, if not, the loss function is evaluated at 100%. An example: the loss function L_2 evaluates the loss if indicator function Υ_1 is true at 150% (in that case, c_1 = 1.5). There is one more remark to make. In the evaluation of the loss functions (6) and (7) in table 4, the loss is considered of the absolute value of the excess forecast minus the excess return on an equally weighted market portfolio (MSCI Europe). This is a relevant test since the investor can possibly only be interested in the relative performance of the stock.

For the one-month ex-post returns, we evaluate 711 time-series of forecasts. For the average monthly return over the next year, there are 675 series. Note that all series have different length. This depends both on their presence in the index and the analyst coverage. The fact that the results are evaluated by this exact finite-sample test is a good solution for possible small sample problems. In what follows we will denote the forecast formulated by the uninformed agent as the maximum likelihood forecast and the forecast of an investor who combines expert information and historical price changes as the Bayesian forecast.

Loss function L_1 evaluates the absolute loss differential. All other loss differentials put more weight on specific events ex-post. The first line of the table indicates that the minimum loss differential among the 711 time-series is –2.4%. This means that, for that particular stock, the Bayesian forecast has an error that is about 2.4% less than the maximum likelihood forecast error.

Table 4.

Parameters for the loss differential analysis

All numbers are in percentages, except the column of the mean test statistics (Mean(t)). The first column displays the minimum observed average loss differential for a stock, the second column reports the average loss differential for the sample of stocks and the third column reports the maximum observed average loss differential for a stock. Column 4 displays the average test statistic for the Wilcoxon signed-rank test. The fifth column reports the percentage of cases the test statistic was negative, indicating a superior performance of the Bayesian forecasts (B>M). The next column presents the number significant negative test statistics (B>M*). The last column reports the number of significant positive test statistics (B<M*). The first panel reports the parameters for the one-month ex-post returns and in the second panel, the parameters for the average monthly one-year ex-post returns are reported. The seven reported loss functions (L[c_1,c_2,c_3]) used indicate the importance a representative investor attaches to a certain loss. If the indicator function is true, the number between brackets indicates the weight given to that loss. If the indicator is false, the weight is 1 or 100%. The L1 to L5 loss functions evaluate the absolute returns, the L6 and L7 loss functions evaluate the relative returns (relative to the market return).

	Min d_t	Mean d_t	Max d_t	Mean(t)	B>M	B>M*	B<M*
1-month ex-post							
L_1[1,1,1]	-2.40	0.23	4.00	0.81	21.4	1.8	12.0
L_2[1.5,1,1]	-4.56	0.63	6.45	1.93	13.9	3.0	52.9
L_3[1,1.5,1]	-8.77	0.55	5.52	1.40	22.1	5.9	42.3
L_4[1,1,1.5]	-5.68	0.47	6.63	0.83	31.5	7.2	27.4
L_5[1.25,1.25,1.25]	-7.10	0.71	8.86	1.34	27.7	8.3	42.9
L_6[1,1,1.5]	-5.72	0.13	5.51	0.47	38.0	5.5	15.6
L_7[1.25,1.25,1.25]	-6.80	0.07	6.88	0.55	37.8	7.7	21.9
1 year ex-post							
L_1[1,1,1]	-1.61	0.62	4.96	2.43	13.6	3.1	61.2
L_2[1.5,1,1]	-3.16	1.11	7.46	3.22	9.3	2.4	73.3
L_3[1,1.5,1]	-2.29	0.80	6.89	2.49	14.8	5.6	63.1
L_4[1,1,1.5]	-2.38	0.91	7.29	2.42	15.7	4.6	60.9
L_5[1.25,1.25,1.25]	-3.41	1.19	9.29	2.79	14.4	5.2	65.8
L_6[1,1,1.5]	-3.38	0.48	5.62	2.11	17.8	5.5	53.9
L_7[1.25,1.25,1.25]	-4.12	0.55	6.79	2.28	17.0	5.9	57.0

However, the average absolute loss differential for all stocks is positive. This means that on average, we get a first indication that the maximum likelihood forecast is a better forecast in economic terms than the Bayesian forecast in a cross-section of stocks. In other words, taking into account the expert opinion does not provide a better return forecast. The maximum observed loss differential is +4%, indicating for that stock that in absolute terms, the Bayesian forecast error is 4% higher for that period than the ML forecast. More robust findings are reported in the remaining columns. The next column reports the average test statistic provided by the Wilcoxon signed-rank test as described by equation 20. For the L_1 loss function, the average t-statistic for all 711 time-series is 0.81.

In the following column, the number of stocks for which the Bayesian forecast did better than the ML estimator is reported. For the absolute loss, the Bayesian forecast does a better job 21.4% of the tested stocks than the ML forecast. In only 1.8% of those cases, this difference was significant, based on a two-sided 95% interval. In fact, this column indicates the test value of the formulated hypothesis. The hypothesis stated that agents applying a Bayes' rule will obtain better forecasts than agents who rely on historical price changes alone. For the first loss function, this is only true in 1.8% of all cases. The last column reports the cases were the ML estimator led to a significant smaller loss. For the L_1 function, this is the case in 12% of all observations. Summarizing the remaining results for the one-month ex-post evaluations, there are a lot of indications that analyst information does not improve the economic value of a return forecast measured in terms of absolute loss.

The loss functions that take into account that the investor evaluates forecasts in different ways in terms of economic loss, do not affect this finding much. There are however some interesting observations. The second loss function, punishing the loss when the forecast is too high clearly shows a shift to the right for the overall loss differential. Not only is the mean loss differential (0.63) higher than the L_1 loss differential, but the number of significant cases where the ML estimate outperforms the Bayesian forecast increases substantially (from 12% to 52.9%). This finding is in line with previous literature documenting that analyst' forecasts are generally too extreme.

Furthermore, results improve a lot when the investor evaluates the forecast relative to the market forecast (L_4). The best results however are found when the investor calculates his loss relative to the benchmark performance (L_6 and L_7). In both cases, the average loss differential is close to zero (0.13% and 0.07%). There are more significant better forecasts for the Bayesian method than in L_1 (in respectively 5.5% and 7.7% of all cases). Even in this case, the ML estimator still seems to be a better forecast than the Bayesian forecast. The overall conclusion is that, based on this model, taking analyst information into account does not reduce the economic loss induced by forecast errors. In terms of the hypothesis, we find that the Bayesian' rule using also expert information does not provide better forecasts than a forecast based on historical price changes.

A possible remark on this finding can be that analysts' forecasts reflect longer-term expectations than the one-month foresight applied here. Therefore, the average monthly return over the next year is evaluated in a similar fashion. The previous finding also holds in this evaluation. Results are even more pronounced for longer-term realizations. In about 60% of the cases, ML estimators lead to a significantly lower loss than the Bayesian forecasts. Furthermore, all average t-values for the WSR test are above 1.96. Finally, the Wilcoxon sign test gives the same indications although, for this test, there are less significant differences found between the two forecast series. This means that the assumption of symmetry of the distribution when applying the Wilcoxon signed-rank test is not crucial for the conclusions.

One last observation we make based on the 47.610 forecasts is in favor of the Bayesian methodology. When we look at the market timing implied by the forecasts, we find that the estimated probability of a correct positive forecast, conditional on positive realized returns is higher for the Bayesian forecasts. This conditional probability for the Bayesian forecasts is .523, while for the ML estimates this probability is .475. This indicates that there is somewhat a better ability to forecast positive returns when realized returns are positive ex-post in the Bayesian case.

6.Concluding remarks and future research

Much of the literature these days is devoted to the modeling of investor behavior. It is however difficult to evaluate a lot of these descriptive models that are based on irrationalities in the investor behavior, because they are designed with the hypothesis of cognitive failures in mind. Hence it is useful to study normative behavior as well to see the implications of rational behavior on realistic problems, such as stock selection.

In this paper we test the hypothesis that Bayes' rule based on analysts' forecasts and historical price changes leads to a better return forecast than the forecast based on price changes alone. The design is such that we evaluate the usefulness of the Bayes' rule as a normative language for financial agents. A novelty in this paper is the way in which agents who have access to more than one type of information form their expectations. In practice, we apply a Bayesian simulation procedure to formalize the Bayes' rule. Prior knowledge arrives in the market through analysts who announce an expectation about return by their one-year analysts' earnings yield forecast. In forming their prediction, rational agents evaluate how valuable this information is next to their knowledge of historical price changes. The second novelty in this paper is that, in order to reduce the loss of information, we use individual priors for each asset.

We find that in the cross-section of all tested stocks there is no evidence in line with our hypothesis. The forecasts are evaluated in terms of economic loss for different specifications of the loss function. Next, we also look at the returns from investment strategies based on the forecasts. For both tests we find that forecasts based on historical price changes alone are superior to the ones exploiting Bayes' rule for the cross-section of stocks. In a next part we looked at three sub-samples of the cross-section, based on the level of systematic risk. For these sub-samples, the conclusion is the inverse of the previous one. Forecasts based on Bayes' rule are superior to the uninformed agent's forecasts and the expert agent's forecasts. This is especially the case for low and medium systematic risk-stocks implying that agents form a better forecast for lower risk stocks than for stocks that are harder to forecast.

We also find that the strength of the expert evidence causes a poor performing cross-sectional stock selection. The highest expert forecasts are accurate but uninformative. A rational agent will not select these stocks. Low expert forecasts are both inaccurate and uninformative. This leads to a better selection of stocks by the rational agents among the low forecast stocks.

This line of research is open to improvements as already pointed out in the paper. More and more models do emerge trying to model investor behavior and trying to explain market anomalies. In our specification we find that momentum exists even when agents are rational because good news about firms (stocks with high expert forecasts) is not quickly reflected in the stock price because experts communicated this good news as a weak signal.

Some of the specifications of the simulation procedure mimicking Bayes' rule are, as previously mentioned, debatable. Hence, future research can take that route in the empirical evaluation of these types of models. In the first place, the choice of a prior is a delicate one. There is a wide choice of priors available, but little priors will embody sufficient information about consensus expectations that are based on fundamentals. Secondly, the measure that expresses the uncertainty that an investor attaches to an information stream can be as important as the central parameter in the model itself. This also should be explored further. Finally, the same type of model could be applied to a portfolio manager forming expectations about asset classes in order to conduct portfolio formation. The advantage this application of the model has is that the effects of the Bayesian' rule on optimization can directly be tested.

Finally, finding that the Bayes' rule using expert opinion to predict future price changes is only applicable to stocks that do not have high systematic risk characteristics, is not sufficient to call the Bayesian language the right one. It would be useful to test different languages that are carefully chosen with respect to the environment and the complexity of the problem at hand.

Part IV c

Bayesian Forecasts of the Mean Vector for Portfolio Analysis

Bayesian Forecasts of the Mean Vector for Portfolio Analysis *

Abstract

This paper evaluates the characteristics of actively managed portfolios for different estimators of the mean vector. Additional to traditional estimators such as the historical mean estimator, the Bayes-Stein estimator and the CAPM estimator, we introduce an estimator for the mean vector forecasted by an investor that is defined as rational. This investor makes Bayesian forecasts of the elements of the mean vector. Portfolios are optimized with and without short sales and portfolio performance is measured before and after transaction costs. We find that the mean vector that consist of Bayesian forecasts for the individual means performs better both statistically and economically than the CAPM estimator and the Bayes-Stein estimator. However, the historical mean estimator is better in economic terms. There are indications that both the level of the prior evidence and the strength of the prior evidence are the reason why the ex-ante rational estimator does not perform better than the historical mean estimator.

* Co-authored by Wim Van Hyfte

1. Introduction

In part b of chapter four, we analyzed the use of Bayesian forecasts for stock selection purposes. In this part, we compare rational forecasts of the mean vector to more traditional forecasts. The mean vector is used as an input parameter for the expected return-variance rule and hence used for asset allocation purposes or to determine the portfolio weights of asset classes. A problem often met by practitioners is that small changes in the estimation of the mean vector lead to large adjustments in optimal portfolio weights, which motivates a more profound study of the subject [first in Jobson and Korkie, 1980].

The main objective of this chapter is to test the implications of rational judgments for portfolio selection. Bayesian forecasts or rational judgments about the mean vector are compared with more traditional estimates of the mean vector [as in Jorion, 1991]. This chapter is an extension of the Jorion paper because it adds a test of the accuracy of the rational forecasts for the mean vector. The objective is the same: a test of the forecast accuracy of different estimation methods for the mean vector based on actual data. We allow for additional extensions of the Jorion paper by analyzing portfolios that are optimized both with and without short sales, and secondly, by evaluating the ex post performance of active strategies both before and after transaction costs. Moreover, we empirically evaluate the properties of optimized portfolios using Bayesian forecasts against the properties of optimized portfolios using the traditional estimators for the mean vector. We only assess the problem of judgments about the mean vector. The estimation of the covariance matrix has been documented in the past to be easier and to have less of an impact on active portfolio strategies [Grauer and Hakansson, 1995]. A final extension of the Jorion paper is that we perform the analysis for both country and sector asset classes instead of sector classes alone.

In his 1952 paper, Markowitz suggested that in order to test the expected return – variance rule, it is important to combine statistical parameters and the judgment of practical men. This is exactly what this paper does for the expected return vector. In this paper, the judgment problem about the price change in the next period is solved for a portfolio manager who looks at historical price changes in order to make a judgment about the next period's price change. In his decision-making process, the possible pay-offs as well as the possible losses are large enough to motivate the agent to seek aid from an expert [Edwards, 1975]. The portfolio manager will collect information from experts and will weigh all this evidence. He uses a Bayesian language to transform all his evidence into a judgment about the next period's price change. Hence, he is rational in the weighing of his evidence, conditional on the simplification of the environment. Bayesian solutions for the judgment about the returns are not new in finance (for example: the Black-Litterman procedure, predictive regressions). However, in this paper, we explicitly use expert forecasts as prior base rates. For each individual asset class, an expert opinion that is observable in the market is given.

The hypotheses

By comparing Bayesian forecasts of expected returns to traditional measures, we evaluate three hypotheses. First, we empirically test the hypothesis that ex ante rational opinions about input parameters, used in portfolio selection, are also optimal ex post. In order to evaluate this question, characteristics of optimized portfolios are compared using Bayesian forecasts as well as traditional estimators for the mean vector. For the covariance matrix we assume certainty equivalence (i.e., we assume that the sample estimates of the second moments of the distribution are assumed to be the true parameters) Hence, covariance matrices are identical for a portfolio manager who uses one of the techniques to estimate the mean vector.

Second, by comparing the characteristics of the optimized portfolios of a rational portfolio manager and a manager using a different technique to estimate the mean vector, we can test the hypothesis that judgments about the mean vector that are ex ante not determined to be rational cause unattractive portfolio characteristics. One first example of this type of question is whether transaction costs are higher when we use traditional measures to estimate the mean vector. A second example of this question is whether portfolio rebalancing is large using traditional estimators.

Third, as an additional hypothesis, and keeping in mind the Griffin and Tversky [1992] comments, we will also empirically assess the question whether the strength of the expert information is important in the accuracy of the input parameters in order to conduct portfolio selection.

Finally, because this paper is an extension of the Jorion [1991] paper, we also make conclusions about the impact of transaction costs, the short-sales restriction and the difference between sector allocation strategies and country allocation strategies.

The remainder of this chapter is organized as follows. Section 2 describes the estimators for the mean vector that are used in this paper. This section also repeats a brief description of the algorithm used by a Bayesian forecaster to estimate the judgment about the future return. Section 3 outlines a detailed description of the choice of the expert opinion as prior information. Section 4 describes the data. Section 5 evaluates the empirical findings of the portfolio formation. Section 6 concludes.

2. Portfolio optimization and estimators for the mean vector

Portfolio optimization and traditional estimators for the mean vector

In the existence of a risk-free asset, portfolio optimization consists of the maximization of the Sharpe ratio (the ratio of average realized excess return to the portfolio standard deviation) [Jorion, 1991]. This optimization procedure takes the following form:

[1] $$v = \frac{\Sigma^{-1} E\left[m|r\right]}{i' \Sigma^{-1} E\left[m|r\right]}$$

with v the vector of optimal weights, Σ the sample covariance matrix,
i de unit vector and $\mathrm{E}\left[m|r\right]$ the estimator of the mean vector.

The constraint is that the sum of the optimal weights equals 1. Furthermore, in this formulation, there is no short sales constraint. All portfolios are also optimized following the Markowitz procedure with no short sales allowed. This type of active investment strategies is quite realistic with respect to the practitioner's situation.

First we give an overview of the traditional estimators for the mean vector [see Jorion, 1991]. If returns are i.i.d., the average return ($\bar{r}$) of the historical returns (r) of each asset class form the historical mean estimator of the mean vector.

[2] $$E\left[m|r\right] = \bar{r}$$

A second traditional estimator is based on the CAPM. In this case, the expected return of each asset class is only related to systematic risk. The expected return is determined by the individual beta of each asset class, and the expected risk premium on the market portfolio. Betas are estimated using 60 months of historical price changes. Also, the expected excess return on the market portfolio can be estimated from the historical observations (average excess returns on the market capitalization weighted market portfolio). Equation 3 shows the mean vector based on the CAPM.

[3] $$E\left[m|r\right] = b \; r_m^e$$

with b the vector of betas for each asset class, and r_m^e the expected return on the market portfolio. Based on the assumption of risk aversion, this expected risk premium must be positive [Jorion, 1991]. So, $r_m^e = \max(0, r_m^e)$.

The third and final traditional estimator is the Bayes-Stein estimator suggested by Jorion [1986]. The purpose of the use of this estimator is to statistically reduce estimation risk[1]. All expected returns are shrunk to a common shrinkage parameter. The common shrinkage return in the Bayes-Stein estimator is the return on the minimum variance portfolio (r^*), assuming that short sales are allowed. Stein [1955] first suggested this estimator, based on his finding that the estimator of simple sample means was inadmissible[2]. Note that this estimator uses the information in the sample covariance matrix. Following Jorion [1991], the following notation for the Bayes-Stein estimator of the mean vector is obtained:

[4] $$E\left[m \middle| r\right] = (1-f)\overline{r} + f\ r^* i \quad ,$$

$$r^* = (i'\Sigma^{-1}\overline{r}) \big/ (i'\Sigma^{-1} i) \ ,$$

$$f = l \big/ (\ l + T)\,,$$

$$l\ = (N+2)(T-1) \Big/ \left[(\overline{r} - r^* i)'\Sigma^{-1}(\overline{r} - r^* i)(T - N - 2) \right],$$

with N the number of asset classes and T the number of time periods.

Bayesian simulation

Assuming that the underlying moments of the distribution of returns are known (certainty equivalence), the investor maximizes his utility of the portfolio return in setting optimal portfolio weights ω, where ω_i denotes the weight assigned to an asset class. Certainty equivalence implies that the vector of parameters (m, s) estimated from the historical return data matrix r**,** is assumed to be the true parameter vector of the population distribution. This implies that the vector r_i is a vector of returns for one specific asset class. Relaxing this assumption and introducing uncertainty about the estimation of the parameters makes the problem a Bayesian one. The optimal portfolio choice ω is now formulated in terms of its predictive density function. With m the vector of future returns, Π is the maximization of the expected utility of the portfolio wealth and is formalized by finding an optimal $\omega' r$ [Jorion, 1986]. The density function is obtained by integrating out unknown parameters (m,s). This procedure allows a correction for parameter uncertainty [see Jorion,

[1] estimation risk refers to the uncertainty about the parameters of the return process

[2] a solution to a decision problem about m is inadmissible if the expected utility for m is lower than an alternative specification of m.

1986]. This paper uses the certainty equivalence assumption for the estimation of the covariance matrix for reasons described above.

As described in the previous chapters, we use the estimator of the marginal posterior distribution. Focussing only on the procedure of a Bayesian forecaster to make his judgments about the expected returns, the mean vector m is decomposed into i expected returns for i asset classes. So the Bayesian forecast procedure will estimate each m_i separately. The joint posterior of each asset class is described by equation 5.

$$[5] \qquad p(m_i, s_i | r_i) \propto p(m_i, s_i) f(r_i | m_i, s_i)$$

In equation 5, $f(r_i | m_i, s_i)$ is the likelihood function and $p(m_i, s_i)$ is the prior density function, and r_i denotes the historical return vector of asset class i. The assessment of this Bayesian problem has not always been regarded as attractive in finance. A first important reason why this is so is the problem of the prior specification. Given that the environment is so complex, the choice of a relevant prior is very difficult and, also, this choice influences the inferences drawn from the algorithm. This problem will be discussed in the next section.

The second problem in this kind of setting is that conjugate priors are only available in a few cases making analytical solutions hard to calculate. In recent years however, computational power makes it possible to solve the integrals by means of numerical algorithms. This is what we decided to use because it makes the choice of the prior less crucial for the tractability of the results and the calculation of the integrals.

We use the Bayesian bootstrap technique described in the previous chapters in order to estimate the Bayesian forecast of the expected return. Using the description of this estimator from the previous chapters, the final estimator of the mean vector in this paper is the following:

$$[6] \qquad E\left[m | r\right] = \left[\frac{\sum_{i=1}^{R} m_i^* p(m_i^*)}{\sum_{i=1}^{R} p(m_i^*)} \right]^N ,$$

where R denotes the number of replications, m_i^* are the bootstrapped parameter of interest and $[\ ]^N$ denotes the vector of means for N asset classes.

Model specifications

Some important remarks about the estimation of the mean vectors have to be made to clearly picture the scope of this analysis compared to the Jorion [1991] paper. An important drawback of this extension of the Jorion [1991] paper is that because of the available data, the period studied is reduced implicating that it is more difficult to obtain statistical significance. However, we study a period of more than nine years, which is an investment horizon of reasonable importance.

Secondly, we extend the study from industry classes for one country to both sectors and countries for international data. This implies that we assume that the European stock markets are not segmented. Since we extend the analysis to a group of countries, there is an additional assumption that there are no investment barriers All data is collected in Deutschmark. Hence, we study the problem for a German investor. With respect to exchange rate risk in Europe, this is acceptable because in the period under study, more than one currency was linked to the Deutschmark and hence the analysis is extendable to an investor in a wider European area than Germany alone.

Furthermore, portfolios in this paper are actively managed, optimizing with and without short sales. The Jorion paper only reports results for the optimization when short sales are allowed. This implies that an active portfolio manager does not know to what extent the different measures for the mean vector are applicable in practice. As an additional extension, we also analyze the performance of the portfolios before and after transaction costs. Both optimizing with and without short sales, and evaluating actively managed portfolio before and after transaction costs, make it possible to evaluate the practicability of one estimator relative to another with respect to the rebalancing of the actively managed portfolios.

3. The expert opinion as prior information

The difficulty to find proper prior information is one of the reasons that Bayesian estimation using individual prior information in asset allocation procedures has not been successful in the past. One could come up with different suggestions for the prior choice but each choice is debatable which is no different for this paper. Prior information should reflect an appropriate parameter space with respect to the problem. In this case, we look for a prior reflecting expert opinion about expected returns. The prior information in this paper is given by the one-year analyst forecast of earnings yield. The criteria for the choice of a prior we used in this paper are threefold. First, the prior should reflect actual forecasts for the subject of the decision problem. Since forecasts for future price changes are not often made directly by experts, we choose to apply a proxy (see the previous chapters). The assumptions required to use earnings yield as a proxy for expected return has been widely documented in the past [see Dechow et al., 1999, and the previous chapters]. Second, the expert opinion or prior evidence should be observable in a clearly communicated quantifiable number. Expert opinions in stock markets are also available for categorical quantities (buy and sell signals), but this is only indirectly quantifiable. Finally, we aim to apply a purely Bayesian procedure. Hence, the expert opinion is not drawn from the historical data (as is the case for the Bayes-Stein estimator, which is in fact an empirical Bayesian procedure). Moreover, the empirical finding [for most countries, see Bakshi and Chan, 2000] that high forecast earnings yield stocks earn a greater return ex-post is an interesting feature to use this multiple as a prior proxy for expected return.

We start from the idea that in a decision-making question in finance, which is the problem addressed here, it is important to consider a realistic environment in which an agent (here a portfolio manager) forms his judgments. Therefore, we suggest testing an environment where the portfolio manager evaluates his asset classes by looking at the historical price changes (a uninformed agent who seeks aid from an expert (a “newswatcher”)). These two types of agents in the market are very similar to the specification by Hong and Stein [1999]. The major difference here is that the uninformed agent actually observes the expert opinion (as explained in Part IV b). We evaluate the estimator of the uninformed agent. This is the historical mean estimator. Next, we also evaluate the estimator of a rational agent. This agent uses the Bayesian forecasts to estimate the mean vector. This setting outlines the environment in which the decision is made and seems a reasonable simplification of the complex environment in which the portfolio manager makes his judgments.

For our dataset, we use the mean consensus forecast as well as the highest and lowest forecast recorded by IBES, to determine the prior information. The analyst forecasts used are the ones at the end of previous month relative to the month concerning the judgment. In this way, the forecast is observable by the portfolio managers before the judgment is made. The portfolio manager knows the expert opinion at the time he makes his decision. The discussion that the expert opinion changes because of the latest published accounting data of the firm does not seem relevant here. It is our purpose to mimic a realistic decision problem according to the environment the decision maker observes at that point in time. If the agent observes a forecast at the time he makes a decision, he will use that forecast to make a decision.

In order to determine a scale factor for the prior information, we rely on the extreme forecasts using Parkinson's measure[3] [1980] as we explained for reasons documented by Griffin and Tversky [1992]. This measure reflects the uncertainty the portfolio manager will assign to the observed expert opinion. If the opinions from different experts are far apart, this will make the financial decision maker doubtful about the relevance of the consensus forecast.

These forecasts are collected for all individual firms in one asset class (the asset classes are the same as defined in appendix 1 of chapter 2 using country and sector classes). We use the equally weighted average of all forecasts in one class to reflect the expert opinion about the asset class. Again, equally weighing all expert opinions is relevant for the environment of the portfolio manager. What he is in fact looking for is a common sentiment in a certain asset class irrespective of the distribution of the size of the assets that are in that class. A pessimistic expert opinion on all the small caps in an asset class is, in our view, as informative as a pessimistic opinion on the large caps of that asset class. Hence, it is acceptable that the investor will be as sensitive to expectations about smaller firms as to expectations about larger firms.

Because we start from the forecasts of individual firms, outliers are observed at both sides of the range of analysts' forecasts. We apply a robust methodology (MAD[4]) in order to keep the range of the extremes within reasonable bounds. For each month, the consensus forecast and the range of forecasts for an asset class is calculated, obtaining a time-series of analyst's forecasts. In order to obtain a reliable estimate for each of these three components, we exclude individual outliers in each asset class, applying a mean absolute deviation algorithm to get a robust estimate. In the first place, we think that one extreme forecast from one analyst on one stock is not representative for the overall expert opinion for the whole asset class. Allowing for such an outlier induces large uncertainty about the signal, while on average, experts can be confident on their signals for that asset class. Moreover, extreme negative earnings yields cannot be regarded as reasonable proxies for expected returns. Also, the observation that analysts' forecasts are sometimes too extreme can be regarded as known by investors since it is widely documented in the literature, and hence, less weight is given to this overall asset class opinion. Again, all choices are debatable, but we want to stress again that choices are made in order to mimic a realistic environment in which the decision has to be made and a realistic representation of the strength of prior evidence.

[3] σ^2=0.361[Highest forecast – Lowest forecast]²

[4] Mean Absolute Deviation

4. The data

In order to test the impact of rational forecasts of the mean vector on asset allocation empirically, we use an intersection database for all stocks in the MSCI Europe index between March 1992 and August 2000. Analysts' forecasts are taken form the IBES database. The returns (percentage changes of the return index) are from Datastream. All stocks that comprise the index in the studied period and are covered by analysts are used in the analysis. The magnitude of the available sample is on average 471. All stocks are assigned to a country portfolio (15 portfolios) and a sector portfolio (14 portfolios). The 29 asset classes are the same as defined in chapter II (appendix 1).
For the period under study (101 months), asset class returns are calculated (market capitalization weighted) and an equally weighted analyst' forecast is calculated for each asset class for each month taking the value at the end of the previous month as was described in the previous section.

Using this intersection data and the Bayesian bootstrap regression framework as described above and in the previous chapters, we estimate a Bayesian forecast of the expected return for the next period for each asset class in each month. Given that the portfolio manager only evaluates past price changes and observes the expert opinion at the end of the previous month, these forecasts are out-of-sample. Also, the traditional measures defined in section 2 are predicted out-of-sample. The covariance matrix used to apply the expected return-variance rule is estimated under the certainty equivalence assumption. However, every 12 months, the estimate for this matrix is updated, each time using the past window of 60 months to estimate the matrix. The market portfolio used to obtain the CAPM estimator of the mean vector is the market capitalization weighted portfolio. The results reported by Jorion [1991] are in favor of the equally weighted market portfolio, but in portfolio selection, the market cap weighted portfolio is a more realistic benchmark as the mean-variance efficient portfolio if the CAPM holds. Applying the four methods to estimate the mean vector provides a forecast for each month for each asset class for the expected return.

5. Empirical evaluation of the estimators for the mean vector

Forecast accuracy

A first analysis looks at the quality of the forecasts. Ex post evaluations by the portfolio manager could attach larger weights to an overestimation of the expected return (i.e. the estimation of the expected return is higher than the realized return) relative to an underestimation. We analyze the proportion of overestimations and underestimations for the four methods and the two types of portfolio strategies (country-based and sector-based). Given the extremely low proportion of perfect predictions, the proportion of underestimated expected returns can be set equal to one minus the proportion of overestimated expected returns. Table 1 shows these proportions. Also, we evaluate the average absolute forecast error for the four estimators of the mean vector and both allocation strategies.

Table 1 shows that the proportion of overoptimistic return forecasts is more or less the same for all traditional methods (about .460 for country allocation portfolios and .440 for sector allocation portfolios) but this proportion is larger for the Bayesian forecaster (.537 for country allocation portfolios and .494 for sector allocation portfolios). Knowing that the simple forecast (HM or the statistical evidence) has a smaller proportion of overoptimistic forecasts, this finding indicates a possible observation of overconfidence of experts. Even with rational expectations, making too high a judgment about the future returns is due to the expert opinion. First, it is a common empirical finding that analysts' forecast are overoptimistic [as in Easterwood and Nutt, 1999].

Table 1.

Proportion of overoptimistic return forecasts.

All forecasts for the four methods are evaluated over 101 months. For each estimation method (Bayesian bootstrap regression [B], historical mean, [HM], CAPM [C] and Bayes-Stein [BS]) the forecasts are compared to the one-month ex-post realized returns for the country-based strategy (1515 observations) and the sector-based strategy (1414 observations). The first column of each strategy indicates the estimation methodology, the second the proportion of overestimated returns and the third the average absolute forecast error (AAFE).

Quality of return forecasts					
Country-based strategy (1515 observations per estimation method)			Sector-based strategy (1414 observations per estimation method)		
	Proportion	AAFE		Proportion	AAFE
B	.537	.0457	B	.494	.0404
HM	.467	.0455	HM	.446	.0405
C	.456	.0455	C	.433	.0409
BS	.458	.0453	BS	.448	.0404

Second, as documented in part IV a and part IV b, the strength of the evidence as provided by experts can be too high, and much of the weighing function will in this case be determined by the expert evidence [Griffin and Tversky, 1992]. In the case of asset classes, the expert opinion will generally be presented as strong evidence.

Notice that the explanation could very well be that the experts are overconfident about their opinion because of the difficulty of the task [Griffin and Tversky, 1992].

Let us explain this in more detail. It is acceptable to assume that an investor tries to obtain correct forecasts. The investor rarely succeeds in making perfect predictions, so his main objective is to be as accurate as possible conditional on his perception of the environment. Trying to make better forecasts and hence seeking aid from experts, seems to induce even more optimistic judgments (table 1), even in the case of rational expectations formed by a Bayesian forecaster. Hence, portfolio managers are not necessarily susceptible to the cognitive failures or irrationalities that are used in descriptive behavioral models. The reason could very well be that the strength of the expert evidence is too high. Because we use asset classes, the problem of the market imperfection that there is noise in the information market is the reverse compared to the one mentioned in part IV b. In that case, high expert forecasts were accurate but uninformative. Here, expert forecasts are generally overoptimistic and in most cases presented as strong evidence.

When looking at the average absolute forecast errors, we see that there is not a lot of difference between the four methods. The only interesting issue is that the errors are smaller for sector allocation parameters than for country allocation parameters. This seems reasonable because a lot of the analysts are pooled to generate forecasts based on sectors and not on countries. In other words, it is more commonplace to have an analyst who evaluates two IT stocks than an analyst who simply evaluates two Swedish stocks. Also, the overall error is large, implying that, in line with the literature, expected returns are hard to predict [as in Jorion, 1991]. With respect to the first hypothesis, it is hard to draw firm conclusions. One indication is however that the rational forecasts tend to be overly optimistic about the future return. This means that, in this setting, ex-ante optimal judgments are not optimal ex-post based on this one indicator.

Portfolio optimization

The framework allows us to evaluate the return properties of optimized portfolios for the four procedures to estimate the mean vector. It is conducted for a portfolio manager who manages an active country-based portfolio strategy and a portfolio manager who manages an active sector-based portfolio strategy. Portfolios are optimized both with and without short sales and evaluated with and without transaction costs (see section 2). Assuming that there is no restriction on short sales allows us to analyze the stability of the optimized weights without short sales with respect to transaction costs and the impact of no longer allowing for short sales. As in Jorion [1991], we mainly focus on the Sharpe ratio because of its attractive small

Table 2.

Portfolio formation properties.

Based on the estimated mean vectors using Bayesian forecasts (B), historical mean (HM), the CAPM (C) and the Bayes-Stein estimator (BS) and the covariance matrix estimated under the certainty equivalence assumption, portfolios are optimized monthly for the period April 1992 to August 2000. This optimization is performed both with (ss) and without short sales (nss) restrictions. The return and risk measures of the portfolios are calculated before (BTC) and after transaction costs (ATC, .5%). For all optimized portfolios and the equally-weighted (EW) and market capitalization-weighted (MW) benchmark, the return [R], the risk [σ] and the Sharpe ratio (R/σ) are reported. Also, for the optimized portfolios, the proportion π of months where the realized return for the optimized portfolio exceeds the benchmark return (no transaction) costs is reported. The reported numbers are calculated relative to the EW benchmark (numbers are equivalent for the MW benchmark)

	country allocation							sector allocation						
	BTC			ATC .5%				BTC			ATC .5%			
	[R]	[σ]	R/σ	[R]	[σ]	R/σ	π	[R]	[σ]	R/σ	[R]	[σ]	R/σ	π
EW	1.355	4.40	.308					1.301	4.27	.305				
MW	1.360	4.27	.319	1.348	4.27	.316		1.334	4.27	.313	1.322	4.26	.310	
nss B	1.286	4.77	.270	1.031	4.76	.217	.475	1.452	4.17	.348	1.320	4.17	.317	.545
nss HM	1.589	4.65	.342	1.456	4.65	.313	.574	1.692	4.37	.388	1.577	4.38	.360	.505
nss C	1.293	4.96	.261	1.135	4.97	.228	.465	.984	5.19	.190	.812	5.20	.156	.396
nss BS	1.081	4.35	.248	.528	4.37	.121	.446	1.751	4.16	.421	1.117	4.13	.283	.545
ss B	2.682	19.18	.140	1.117	19.02	.061	.545	2.898	10.23	.283	.928	10.31	.090	.594
ss HM	3.997	25.29	.158	2.448	25.58	.096	.564	4.751	15.41	.308	1.689	15.95	.106	.624
ss C	-1.364	25.30	<0	-8.930	37.87	<0	.505	3.353	38.00	.088	-5.948	37.44	<0	.456
ss BS	.648	9.94	.065	-4.000	12.34	<0	.436	3.205	12.90	.248	-5.047	14.12	<0	.584

sample properties. Jobson and Korkie [1981] pointed out that although the Treynor ratio is the straightforward measure to use, its small sample properties are weak. Secondly, because the active strategies are computed to maximize the Sharpe ratio ex-ante, it is logical to focus on this performance measure.

Some overall features emerge from the results in table 2. In the first place, after transaction costs, the simple forecasts (HM) outperform the other forecasts in all cases of the portfolio formation process. The Sharpe ratio is always the highest (with a maximum of .36) and the proportion of months where this active allocation strategy outperforms the passive strategy (p) is in most cases higher than for the other methodologies. Even after transaction costs, the optimized portfolio with no short sales allowed based on the HM mean vector outperforms the passive strategies (MV) for both country portfolios and sector portfolios.
Based on the Sharpe ratio, the active strategy based on the Bayesian forecasts also outperforms the passive strategy for the sector portfolio (.348, still outperforming after transaction costs). This is in line with the previously mentioned idea that aggregate analysts' forecasts for asset classes are more relevant for sector classes than for country classes. Moreover, using rational expectations from the Bayesian forecaster in the sector allocation procedure lowers the portfolio risk in all cases (before and after transaction costs, and with and without short sales allowed, right panel of table 2). One attractive characteristic of active portfolio management based on this rational estimator is that the standard deviation of the ex-post portfolio returns is in most cases lower but never much higher than for the other estimators. Finally, sector allocation strategies have in all cases better risk-return properties than country allocation strategies. The risk of a European diversified sector portfolio is in most cases lower than for a country portfolio. In most cases differences are small.

The CAPM as a model to make forecasts about returns performs poorly. The risk-return ratio for this method never exceeds the passive portfolio strategy. On top of that, when short sales are allowed, the performance based on the CAPM method and the Bayes-Stein method is severely reduced because of the transaction costs. Especially for the Bayes-Stein estimates, the problematic level of the transaction costs is even observed when short sales are not allowed. The return decreases from 1.081% to a net return of 0.528% for the country portfolio and from 1.751% to 1.117% for the sector portfolio. This is not the case for the simple forecasts (HM) and the ex-ante rational forecasts (B). This indicates that the CAPM and the Bayes-Stein estimators induce large monthly shifts in the portfolio weights and make them hard to apply in practice.

The empirical finding that the rational expectations estimator performs better in estimating the mean vector for asset allocation purposes than the CAPM or the Bayes-Stein algorithm, is interesting. With respect to the first hypothesis, it is conventional to say that active strategies perform well ex-post using the Bayesian forecasts, especially for sector portfolios. With respect to the second hypothesis, we see that if forecasts about future returns are not fully rational conditional on the environment, there is a possibility that the portfolio has undesirable properties. Especially in the case of the Bayes-Stein estimator, we see that transaction costs make this estimator hardly applicable in practice (without short sales, the return drops from 1.08% to 0.53% for the active country strategy). Also, if short sales are not allowed, a restriction often met in practice, this conclusion does not change. Notice that a lot of these findings (such as the best performing estimator, the standard deviation of portfolio returns) are in conflict with the Jorion results [1991]. First, the time period studied by Jorion was much longer. But if the conclusion from the paper is that the CAPM provides the best estimator of the mean vector, a European investor, investing between 1992 and 2000 in either country classes or sector classes based on CAPM estimates of the mean vector, would have had a bad performance. Second, we extended the strategies to both short sales and no short sales and with or without transaction costs.

Summarizing table 2, we see that both modifications (short sales and transaction costs) have a large impact on the results. Especially adding transaction costs in the analysis seems like a necessary extension given the results for the Bayes-Stein estimator. Also the stability of the results without short sales seems to be more relevant towards practitioners, given the large standard deviations that stem from the analysis where short sales are allowed.

Statistical difference in portfolio performance

Before we draw any further conclusions about the better performance of any strategy relative to another, it is interesting to see whether the difference in performance is not only economically relevant (higher Sharpe ratios), but statistically significant as well. Therefore, we test whether there is a statistical difference in portfolio performance using the Jobson and Korkie [1981] test statistic.

Table 3 reports the Jobson and Korkie performance comparison measure for all pairs of portfolios. We estimate these statistics for strategies with and without short sales and after transaction costs. The test statistics for the active strategies before transaction costs are reported in appendix 1. A positive test statistic indicates that the row portfolio strategy performs better than the column portfolio strategy.

A first conclusion from the table that is interesting for the discussion of the other results is that for both strategies, the historical mean estimator (HM) and the Bayesian forecasts (B) outperform the CAPM and the Bayes-Stein estimator significantly in the case where short sales are allowed. There is, however no statistical difference in the performance of the active strategies based on the HM estimator and the rational forecasts. When short sales are not allowed, the main conclusion is that active sector strategies based on the CAPM estimator of the mean vector perform worse than the other estimators, both before and after transaction costs (see appendix 1). Also, in appendix 1, we see that this is the only conclusion we can make about a statistical difference in performance when transaction costs are not included.

Table 3.

Statistical difference in portfolio performance

Table 3 reports the Jobson and Korkie performance comparison measure for all pairs of portfolios, both for the cases with and without short sales but only after transaction costs. C stands for the CAPM estimator, BS for the Bayes-Stein estimator, B for the Bayesian forecasts and HM for the historical mean estimator. * denotes significance at the 10% level, and ** at the 5% level. A positive test statistic indicates that the active row portfolio strategy outperforms the column portfolio strategy.

	No short sales allowed			Short sales allowed		
Sector	C	BS	B	C	BS	B
BS	2.131**			-1.445		
B	2.191**	0.553		1.852*	3.180**	
HM	2.709**	1.097	1.153	1.919*	3.104**	0.286
Country	C	BS	B	C	BS	B
BS	-1.687*			-0.621		
B	-0.172	1.396		2.194**	2.774**	
HM	1.125	2.891**	1.740*	2.521**	2.962**	0.582

The main finding with respect to the first two hypotheses is that the Bayesian forecasts are optimal ex-post compared to the CAPM and Bayes-Stein estimators. The Bayesian forecasts also reduce the unattractive properties of portfolio selection, such as frequent and large portfolio rebalancing and weak performance after transaction costs, we observed for other estimators. Using expert evidence and statistical evidence together does not improve the portfolio performance. We observe no statistical difference in performance between the active strategies based on the Bayesian forecasts and active strategies based on the HM estimator. When we look at the other parameters of performance (Sharpe ratios and realized returns), the active management based on the HM estimator performs even better. This brings us back to the question why rational forecasts dealing with expert information do not outperform the simple forecasts. One reason was already pointed out: there is a possibility that experts are overconfident when confronted with the difficult task, such as the prediction of expected returns. But secondly, as we illustrated in the previous parts, there is also a possible problem with the strength of the evidence when the task is difficult.

Strength of the evidence and portfolio selection.

We find that, especially for an active sector strategy, the rational forecasts are beneficial ex-post as well. However, it is remarkable that the simple forecasts always outperform the Bayesian forecasts. This implies that seeking aid from experts does not improve the quality of the forecasts and the performance of portfolios that are optimized based on these mean vectors. The question is what possibly causes this problem: the strength of the information or the overconfident experts conditional on the simplification of th e environment we made?

Table 1 indicated that experts are indeed overconfident. In part IV a, we showed that the rational forecaster only adjusts the expected return upwards when the prior evidence is reasonably strong and the expert opinion is optimistic. Given the findings in table 1, we have strong indication that experts are indeed overconfident about their opinion on asset classes. The problem of adjusting expectations to prior evidence is possibly enforced by the strength of the prior evidence (as documented in part IV b).

Hence we assess the third hypothesis that the strength of the prior evidence or expert information is important for the accuracy of the input parameters for portfolio selection. As we mentioned earlier, a large consensus around the expert opinion together with overoptimistic opinions induce expected future returns to be too high. The way we approach this problem is the following. If the financial agent observes the expert information he will take the level of disagreement among experts as the strength of the evidence as described before. Suppose that for reasons well described in past literature [as in Lim 2001] there is a bias in this observation of the strength of evidence and experts display more agreement than actually is the case.

In order to test this empirically, we repeat the sampling procedure and the portfolio selection for the Bayesian forecasts but with different strength parameters for the prior evidence. Instead of the Parkinson measure, we impose the disagreement among experts about the next period's return of a certain asset class to be equal to the conditional variance (i.e. volatility follows a GARCH(1,1) process) of that asset class in the previous month. The idea is that if the return is very volatile in the previous month, the level of agreement will be lower in the next month. In this case, it is even more difficult to estimate the expected future return. Using this alternative measure for the strength of prior evidence, we re-estimate the Bayesian forecasts with all the other evidence equal to the previous case for which the results are shown in table 2.

Table 4 shows the results of the active portfolio strategies using the Bayesian forecasts and the new measure of the strength of the prior evidence. From table 4, we see that for the active country strategy, performance parameters improve over the whole line. The Sharpe ratio, measured after transaction costs goes from .270 to .290 without short sales and from .140 to .173 when short sales are allowed. In this last case, the Sharpe ratio even exceeds the one for the HM estimator. For active country strategies, this indicates that we find no evidence against the third hypothesis. In some cases, realized excess returns increase while the risk is lowered.

Table 4.

Portfolio performance using Bayesian forecasts and an alternative measure of the strength of evidence

Table 4 reports the portfolio selection strategies based on Bayesian forecasts and a conditional volatility measure of strength of evidence. [R] denotes the realized portfolio excess return, [σ] the volatility and R/σ denotes the Sharpe ratio. NSS stands for the optimization when short sales are not allowed, SS when short sales are allowed. BTC and ATC stand for the performance measures before and after transaction costs.

		Country strategy		Sector strategy	
		BTC	ATC	BTC	ATC
NSS	[R]	1.328	1.181	1.490	1.390
	[σ]	4.58	4.58	4.34	4.34
	R/σ	.290	.258	.343	.320
SS	[R]	2.647	1.786	2.441	.949
	[σ]	15.33	15.41	13.93	14.31
	R/σ	0.173	.116	.175	.066

For the sector strategies, the conclusion is not that firm. If short sales are not allowed, the realized excess return increases, but also the risk increases. After transaction costs, there is a minor upgrade in the Sharpe ratio (from .317 to .320). Results are worse for the case where short sales are allowed, but the part of the return that goes to transaction costs lowers.

Overall, these findings indicate that the strength of evidence is an important feature in the objective of making rational forecasts. Together with the finding that there is a possibility that the experts are overconfident, the ex-post optimality of the rational forecasts depends very much on the quality and the clarity of the expert opinion.

6.Conclusion.

In this paper, we evaluate estimators for the mean vector and their use in active portfolio strategies. The traditional estimators as proposed by Jorion [1991] are used. We extend this analysis using an additional estimator for the mean vector. This estimator is designed to produce rational Bayesian forecasts conditional on the simplification of the environment. Furthermore, we analyze the performance of active strategies for portfolios that are optimized with and without short sales restrictions. Also, performance measures are calculated before and after transaction costs.

First, we analyze whether the ex-ante rational estimations of the mean vector using a Bayesian framework are also optimal for the portfolio manager ex-post. We find that this is indeed the case, but we find no statistical difference compared to the performance of the actively managed portfolios that are optimized using the historical mean estimator for the mean vector. Moreover, in most cases, these latter portfolios outperform the actively managed portfolios that are optimized using the rational estimator for the mean vector based on economical performance measures, such as the Sharpe ratios. There is some evidence that Bayesian forecasting using expert information is less accurate (in an ex-post evaluation) than historical mean estimations of the mean vector because of overoptimistic expert information.
We also find that extending this analysis to both with and without short sales and before and after transaction costs yields valuable insights for portfolio selection. Allowing for short sales induces large portfolio rebalancing for most estimators. Also, ex-post standard deviations for the actively managed portfolios are large when short sales are allowed. Moreover, if we make an evaluation of the performance measures for actively managed portfolios after transaction costs, we see that (especially for the Bayes-Stein estimator) performance is weak. This is not the case for the rational estimator and the historical mean estimator for the mean vector. These findings suggest that the portfolios that are managed based on the rational estimator do not have unattractive characteristics for the portfolio manager such as large portfolio rebalancing and poor performance after transaction costs.
However, we find no indications that the Bayesian forecasts are optimal compared to historical mean estimation of the elements of the mean vector. We have already mentioned that this is possibly due to the fact that experts are overoptimistic. A second possibility is that the strength of the prior expert evidence is too high and

hence, too much weight is given to this evidence. When we test this empirically, we find that when a portfolio manager does not rely on the strength of the evidence as displayed by experts, the rational estimator is more accurate. This is especially the case for actively managed country portfolios.

Overall, rational estimations of the mean vector perform well amongst other estimators. Traditional estimators that have been tested empirically in the past perform poor for this European dataset in the nineties. The one exception is the historical mean estimator. In portfolio selection procedures, the Bayesian forecasts of the elements of the mean vector are attractive because there is no statistical difference in the performance comparison with the historical mean estimator. Furthermore, using this estimator for the mean vector does not induce large portfolio rebalancing and poor performance after transaction costs. A final attractive characteristic of active portfolio management based on this rational estimator is that the variance of the ex-post portfolio returns is not higher than the ones obtained for other estimators, in all cases.

Appendix 1.

Table 3.bis

Statistical difference in portfolio performance

Table 3 reports the Jobson and Korkie performance comparison measure for all pairs of portfolios, both for the cases with and without short sales, before transaction costs. C stands for the CAPM estimator, BS for the Bayes-Stein estimator, B for the Bayesian forecasts and HM for the historical mean estimator. * denotes significance at the 10% level, and ** at the 5% level. A positive test statistic indicates that the active row portfolio strategy outperforms the column portfolio strategy.

	No short sales allowed			Short sales allowed		
Sector	C	BS	B	C	BS	B
BS	3.781**			1.107		
B	2.162**	-1.168		1.450	0.259	
HM	2.627**	-0.468	1.047	1.568	0.402	0.471
Country	C	BS	B	C	BS	B
BS	-0.199			0.918		
B	0.130	0.315		1.457	0.569	
HM	1.066	1.427	1.308	0.569	0.691	0.327

Chapter V

Conclusions,
&
Implications

Conclusions & Implications

1. The cross-section of expected returns for European stock portfolios

The estimation of expected returns and discount factors concerns both the right identification of priced factors and the correct econometrical assessment of the problem. There are several reasons why this topic is important for investors. Using the expected returns for portfolio selection, estimating the cost of equity, evaluating investment plans and evaluating funds are some of them.

In chapter II, we test whether well-described factor models, widely used for U.S. datasets, apply to a large set of European stocks. We assess this analysis for different regroupings of stocks into portfolios and for different asset pricing models (further: APMs). The main question is whether APMs that are relevant for U.S. data, in the sense that there is ex-ante mean variance efficiency for a linear combination of these factor portfolios, are also relevant for an independent European dataset.

We find that for country portfolios this indeed is the case. For the capital asset pricing model, we never find evidence against an ex ante efficiency of the factor portfolio. This holds for both the entire period and the sub-periods studied. One implication of this finding is that we find no evidence against the hypothesis that European stock markets are integrated. The power of this multivariate test is evaluated based on both a risk-based and a nonrisk-based alternative specification of the factor model. For the twenty-year period under study, the power is reasonably high. This indicates that market capitalization weighted European portfolio of stocks is ex ante mean-variance efficient for the analysis based on country portfolios. The power is also acceptable for the models including extra explanatory factor portfolios. Moreover, the power statistic using the extended models does not vary a lot compared to the one-factor model indicating that there is no loss in power when extra factors are included in the APM. All this implies that portfolio selection in European stock markets based on country allocation can be conducted based on the CAPM. The same goes for the estimation of expected returns for European country portfolios. Moreover, assumptions about the pricing of exchange rate risk factors are not required for ex ante efficiency.

The conclusion, however, is different for sector and size portfolios. For the entire period we find evidence against ex ante efficiency of the factor portfolio. This indicates that for the whole period, we find evidence against an exact factor pricing relation for these two types of portfolios for all specifications of the APM. This finding conflicts with earlier findings on U.S. data where three-factor models are found to efficiently price assets. Moreover, and more important, we find that there is an efficient pricing relation in most of the sub-periods for the different APMs for both the size and the sector portfolios, especially based on the finite-sample test. This finding indicates that factor loadings and/or risk premia could be time varying. The power that evaluates the multivariate test-statistic for an alternative specification of the APMs is for both types of portfolios sizeable. This again indicates that adding additional factors does not reduce the power of the multivariate test.

We find that based on the p-values for the test-statistic in the sub-periods, the local momentum factor portfolio (LMOM) is a better additional factor on top of the market portfolio to efficiently price the sector portfolios. For size portfolios, results are rather inconclusive. Also, the findings for sector portfolios support the argument to use a multivariate test to analyze the factor pricing relation in order to determine which factors or factor portfolios are priced. The univariate statistics indicate that the high-minus-low book-to-market factor (HML) explains a part of the cross-section of returns. However, using the multivariate test, we find indications that this is to a smaller extent the case. As mentioned, LMOM improves the exact factor pricing relation for sector portfolios in European stock markets.

We previously mentioned that the results for the sub-periods with respect to sector portfolios and especially size portfolios indicate that factor loadings or risk premia could be time varying. Because the three periods show different patterns in the short-term interest rate, we can draw some inferences with respect to this feature. Overall, p-values for the multivariate test-statistics are higher in the period where the volatility of the risk-free interest rate is lowest. This is especially the case for size portfolios. This is an indication that, for the choice to regroup stocks into dependent portfolios, it is important to evaluate the APM conditional on the monetary policy stance. We find lower p-values for the multivariate test-statistics for size portfolios in the periods where interest rates are distinctively rising or falling. However, these findings are not always confirmed by the asymptotic test. Hence, it seems easier to find evidence against an exact factor pricing relation for size portfolios in European stock markets when interest rates are less volatile.

We have to make three more comments. Firstly, the hypothesis that the wealth portfolio in the single-factor model is misspecified is also tested in this chapter. We cannot reject the ex ante efficiency of the wealth portfolio when the return on human capital is taken into account next to the market portfolio of stocks. This is the case for country and sector portfolios but not for size portfolios. We think that this line of

research in European financial markets is promising enough to deserve more attention in the future.
The second is that the power of the multivariate test-statistic against a risk-based alternative is higher than in the case of a nonrisk-based alternative. This is in conflict with findings for U.S. data and because of the different time-period studied and because of the different specification of the risk-based alternative, this topic requires more research.
Thirdly, we also conducted this entire analysis for the same dataset with the synthetic euro as the currency of denomination. This currency is an artificial one, mimicking what the euro currency would have looked like. However, in light of the single currency for the EMU, it is interesting to see that none of the previously made conclusions change when this currency is used. It is hence an interesting currency to use in future research for the asset markets in the EMU zone or for the representative investor in the EMU when the problem arises to express returns of EMU members in a common currency before 1999.

Summarizing, this chapter learns that the market portfolio of stock returns is mean-variance efficient in the case where European country portfolios are studied. This holds for the entire period under study as well as for the sub-periods. For sector and size portfolios, we find evidence against an exact factor pricing relation for the whole period, implying that suggested APMs for U.S. data are not applicable to all European subsets of assets. Evidence for the sub-periods suggests that the factor loadings and/or risk premia could be time varying. In a multivariate test, the LMOM factor portfolio provides a more appropriate extension for the sector portfolios in the sub-periods. Especially for size portfolios, we observe that the monetary policy stance could be important for the analysis of the priced risk factors. Finally, we present evidence that taking into account human capital in the wealth portfolio is an important consideration in the one-factor model.

2. Earnings yield as a proxy for expected returns for European sectors

The relationships between earnings yield, book-to-market (BTM) and stock return is a widely investigated subject. For U.S. data, annualized returns are reported for a zero-cost strategy based on both earnings yield and forecasted earnings yield (also known as contrarian strategies) of about 9%. Also for different regions this return is ascertained. We assess this question for sector portfolios, assuming that the fundamentals and market-risk characteristics for stocks in the same sector are comparable, but are different across sectors.

First, we test the hypothesis that the expected return for high accounting multiple stocks is the same as for low accounting multiple stocks in a sector. We formed portfolios based on earnings yield and BTM. Next, we also formed value and growth portfolios based on a combination of earnings yield and BTM relying on results presented in the past accounting literature.
For the period under study, we find a positive return for the zero-cost strategy for the insurance sector, the financial services sector, the utilities sector and the resources sector over the whole period as well as in all the sub-periods. However, we find that overall differences in returns are often small. A closer look reveals that a lot of these returns on contrarian strategies have time-varying characteristics. Hence, the second hypothesis emerges that differences in risk account for the difference in return for the zero-cost strategy based on earnings yield and BTM. If for example high earnings yield stocks are fundamentally riskier in bad states of the world, their prices will fall and their expected return will rise.

The two hypotheses are relevant for portfolio selection purposes. Especially with respect to the management of sector funds, but for other funds as well, the knowledge about the driving forces behind expected returns for individual stocks is highly informative. As mentioned before, we find that over the whole period we cannot conclude that a difference in return for high and low accounting multiple portfolios is a stylized fact. Compared to the findings about the country cross-section of earnings yield in the past literature, it is of importance to see that this empirical fact is not directly applicable to sub-samples of stocks that have the same fundamental values and risk characteristics and hence is not an investment strategy as such for any sample.

Finding that there are different patterns of time-varying returns for the zero-cost strategy, however, is informative with respect to the management of the portfolio. According to theory, the portfolio manager has to be aware that the prices of stocks that are fundamentally riskier will fall in bad states of the world. We assessed this question by an econometric evaluation of the long-term relationship and the short-term dynamics between the return series for the zero-cost strategy and the states of the world. The latter part is modeled relying on a consumption-based asset-pricing model. The perception of the state of the world will be determined by aggregate consumption as well as by the difference between the current consumption and the habit of consumption. We find that a part of the return of the zero-cost strategy is explained by these dynamics. This implies that the observed time-varying properties of the returns can be explained both by the long-term relation between the states of the world and the returns and by the short-term dynamics. However, the dynamics are sector dependent and complex. This is informative as well with respect to portfolio management. The actual portfolio formation should be conducted conditional on the state of the world when the portfolio manager is managing a sub-sample of all assets. Finally, notice that the indication that returns for this zero-cost strategy depend on fundamental risk for European sector data, sheds new light on this discussion. For the unconditional cross-section of U.S. data, there has been no indication that this fundamental risk story explains this anomaly.

Summarizing, we find that an investment strategy based on earnings yield forecasts is not applicable to all subsets of stocks for any horizon. We find indications that the return on a zero-cost strategy based on forecasted earnings yield is time-varying. A part of the explanation why this is the case is found in the identification of risk. Assuming equal fundamental values and risk characteristics in a sample of stocks, stocks with high fundamental risk properties will have a different reaction than stocks with low fundamental risk properties in different states of the world.

3. Bayesian forecasting and investment decisions

The last part of this thesis documents the benchmarking of cognitive failures in financial markets. The purpose of this chapter is to evaluate investment decisions that are made based on rational expectations of the investor or, in other words, when the investor acts as a Bayesian forecaster. It is not clear to what extent the cognitive biases that are found in psychological experiments are applicable to the behavior of investors, which is a motivation to do this type of investigation.

In a first part, we show that some of the assumed cognitive failures in financial markets can be explained by the environment agents observe rather than by irrationality. This is important because irrationalities form the basis of descriptive models in behavioral finance. These models are designed to explain market anomalies. However, if the assumed irrationalities are no irrationalities but market imperfections, the foundation of these models can be questioned. This motivates the explicit study of cognitive biases for investors. With respect to practice, we see a research topic for the field of contrarian investing. The existence of intuitive forecasts makes it possible to benefit from a strategy where stocks with a high expected future growth are sold. We argue that an investment strategy based on this cognitive failure of investors does not necessarily exist because investors are irrational. On the contrary, we show that the market imperfection of incorrect information diffusion creates the illusion that investors are irrational. The main cause of this problem is found to be the strength of the prior evidence. Even if investors look at the expert opinion to model their base rate, a false observation of the certainty the experts display about their opinion induces the sub-optimal forecasts. We stress that these forecasts are processed as Bayesian forecasts and hence, are rational forecasts.

This implies that if the market microstructure improves, contrarian investing and funds based on contrarian strategies will be less profitable. One example is that an increased independence of analysts, who can be considered as the experts, can improve the market of information. An opinion of an analyst that is not dependent on the companies' management to get his information can then truly reflect his belief about the value of the company. The investor in that case clearly observes both the consensus and the true level of disagreement about the intrinsic value of the firm. What we refer to is that herding behavior and over-optimism by analysts can be reduced an this can improve the market of information.

Moreover, the idea that the market microstructure with respect to financial decisions improves has been observed in past literature. It has for example been documented for U.S. data that the amount of zero earnings surprises as the difference between analysts' forecasts and announced returns has been increasing substantially over the last years.

We test the implications of ex ante rational expectations about the next month's price change for stock selection purposes. Overall, we find that ex ante rational judgments are not optimal ex post when applied in stock selection. First, the strength of the information provided by experts explains a part of the sub-optimal decisions. We find indications that the expected stock return is ex post sub-optimal because the prior evidence from experts causes sub-optimal decisions. This evidence is found for investors who make Bayesian forecasts and are hence rational. Moreover, we find that rational judgments about the next month's return are optimal compared to purely statistical forecasts and expert forecasts in the case where an investor makes a stock selection conditional on the betas of the stocks. The main problem however remains for high-beta stocks. There is no relation between forecast return and realized return for any forecasts for the sub-sample of high-beta stocks.

An analysis of the decision about the next period's return, in this case on a monthly basis, is relevant for different applications in financial markets. Hence, this chapter is informative both with respect to the ex-post optimality of Bayesian forecasts. Moreover, this chapter also provides caveats for the weighing of information. In a simplified environment, we find that the quality of the prior expert evidence accounts for a lot of the ex-post sub-optimality of Bayesian forecasts for stock selection purposes.

This information is relevant for retail investors. Past research documented that investors tend to realize capital gains and tend to keep stocks with capital losses in their portfolios. In the decision process of which stocks should be held and which stocks should be sold, the previous conclusions are informative. First, it is important to be aware that even for a sample of large stocks, low- and medium-beta stocks are easier to forecast than high-beta stocks. Second, if investors evaluate their statistical evidence and their base rates in a Bayesian way, as is theoretically evaluated as being rational, it is important to see that the base rates provided by experts should be handled with care. Strong signals from the experts are often exaggerated (without the experts being irrational, [Lim, 2001]) and hence, giving too much weight to the base rates can be sub-optimal. We gave the example of the analysts' forecasts, where opinions can be close together for reasons that are not motivated by the forecasts and in reality forecasts can diverge more than it is communicated.

The decision about the next period's return applies as well to portfolio managers. Deciding what stocks that are kept in the portfolio and what stocks are sold is the same decision problem. On an individual stock level, the rational estimation of the cost of equity is another example of this type of decision problem. Finally, we see a direct application with respect to the care that should be taken when investing based on a contrarian strategy. We argued before that in the case where the expert information does not reflect the true expectations experts have. Based on past literature, we argue that this does not mean that experts are not rational. Again, an improvement of the market's microstructure with respect to the communication of experts' forecasts can reduce the profitability of contrarian strategies because we find no evidence that investors are indeed irrational.

Finally, we also evaluate rational forecasts made by a portfolio manager in order to conduct portfolio selection. This portfolio selection is evaluated both for a country-based strategy and a sector-based strategy. The main finding is that we cannot conclude that Bayesian forecasts of the elements of the mean vector are sub-optimal to other estimators of the mean vector. We assign this conclusion based on an evaluation of the characteristics of optimized portfolios, both in the cases were short sales are allowed and are not allowed. Performance measures are evaluated before and after transaction costs.

First, we find that the Bayesian forecasts of the mean vector and the simple forecasts of the mean statistically outperform the CAPM estimator and the Bayes-Stein estimator. This conclusion is enforced by the economic interpretation of portfolio performance. The main problem for the CAPM estimator and the Bayes-

Stein estimator is that they have poor performance measures when they are evaluated after transaction costs.

Second, we find that in economic terms, and especially in terms of the Sharpe ratio, the simple mean estimator is better than the rational estimator for the mean vector. This implies that taking expert information into account does not improve the estimation of the mean vector overall. We further explore this problem and find that the problem is twofold. Firstly, expert opinions are on average too optimistic. This has been extensively documented in the past, but we repeat that this does not imply that experts are irrational. Secondly, we again observe that the strength of the expert evidence is an additional problem for the optimality of the estimator. We argue, based on the empirical findings, that expert information is less noisy for sectors than for countries. Allowing a different measure for the strength of expert evidence we observe that the performance characteristics for the country-based optimized portfolio improve.

This last part provides some caveats with respect to the estimation of the mean vector and its use in asset allocation. Past research documented the usefulness of some estimators over others. Extending the tests to the short sales restriction issue and the issue of transaction costs leads to the conclusion that findings from past research are not extendable to all horizons and to all areas. Most important we find that the Bayesian forecasts of the mean vector have reliable performance characteristics. An improvement of the market's microstructure can even improve the ex-post optimality of the Bayesian forecasts.

Summarizing, we find in this third part of the thesis that the benchmarking of irrational behavior reveals interesting information. The main finding is that a false communication by experts of the strength of their evidence leads to ex-post sub-optimal rational decisions. This finding has important implications for practitioners as well as for academics. For practitioners, it implies that they can use expert information as a base rate for their evidence, but that they should be aware of the shortcomings of this base rate in reality. This finding calls for attention to the subject of the role of expert information that is publicly made available. This segment of the market's microstructure distorts the ex-post optimality of rational decisions in financial markets. For academics, it implies that the wide range of possible assumptions about cognitive failures is not unconditionally applicable as basic assumptions for descriptive behavioral models. This finding calls for further research on the topic of cognitive failures susceptible to investors, which as we recognize is an extremely difficult task. As long as this is not achieved, it will be hard to come up with an behavioral alternative for the efficient market hypothesis.

4. Implications

Chapter II

- We find no evidence against the hypothesis that a portfolio manager can conduct a country-based portfolio strategy based on the factor loadings on the market portfolio alone. This implies that for this sub-sample of stocks, the market portfolio is ex-ante mean-variance efficient or that the CAPM is not dead. It also implies that we find empirical evidence that the European stock markets are not segmented between 1980 and 2000.

- We find a lot of evidence that factor models that are identified as being useful to price samples of U.S. assets are not directly applicable to European assets. The Fama-French three-factor model is evaluated as weak in explaining the cross-section of European stock returns.

- Especially for sector and size portfolios, there are clear indications that the identification of a correct asset-pricing model should contain time-varying factor loadings or risk premia.

- For sector portfolios we find in a multivariate test setting that a factor mimicking portfolio based on momentum is a good extension of the one-factor model. Overall, for European stock markets, a factor based on momentum better explains the cross-section of returns than a factor based on book-to-market.

- Including the return on human capital in the wealth portfolio for the single-index model provides promising results. In identifying the set of factors that are sufficient to price European assets, this aspect deserves a lot of attention in future research.

Chapter III

- Earnings yield forecasts as a proxy for expected returns is a relation that empirically holds, especially for the U.S., but also for the cross-section for a lot of other regions. However, we find that this European value-growth strategy is not applicable as an investment strategy as such. For European sector samples in the nineties, we find that this investment strategy would only have a positive pay-off for financial sectors, the resource sector and the utilities sector.

- For European assets, we find that portfolio strategies based on the book-to-market ratio are more profitable on average than strategies based on earnings yield.

- Returns on European contrarian strategies are not constant through time. It implies that an investment in these strategies is not profitable as such but depends on the time of the initial investment and the investment horizon.

- We find evidence that risk explains a part of the time-varying return premium from an investment strategy based on earnings yield and book-to-market. It means that, in contrast with past literature, we find evidence that the fundamental risk explanation holds.

Chapter IV

- Academics often assume that assumptions of irrationalities are applicable to finance. These cognitive failures are the basic assumptions for descriptive financial models. Evaluating rational behavior for investors, we find indications that market anomalies can be found when investors form rational expectations and markets are efficient. This finding questions the validity of these assumed irrationalities for investors. It furthermore implies that universal rationality as an axiom is not dead. Hence, assumptions in behavioral finance should be well documented before they can be applied.

- Practitioners who make complex investment decisions are often evaluated as irrational. However, a closer look at the rational behavior for a realistic simplification of the environment reveals that they are not irrational as such. Moreover, we find that if they make rational decisions, they should be aware that the expert opinion they use should be carefully evaluated in order to improve the ex-post optimality of their rational decisions.

- Ex ante rational judgments about the expected returns on individual stocks are only optimal ex-post for low- and medium-beta stocks. High-beta stocks are hard to forecast, based on the statistical evidence as well as based on expert evidence as well as by rational investors weighing the two.

- Rational decisions about the mean vector provide useful input parameters for portfolio selection purposes. Again, portfolio managers should carefully evaluate the expert opinions they use as base rates.

- Financial market's microstructure can be improved by setting guidelines for the way expert forecasts are disclosed to the market. Communicating the level of disagreement amongst analysts is highly informative for rational investors. This is motivated by the fact that optimal ex post rational decisions improve market efficiency.

- Reducing herding behavior among analysts and reducing over-optimism displayed by analysts are examples that improve the information market in financial markets. If this can be accomplished, we expect some market anomalies to be reduced or even to disappear. This is important knowledge for practitioners who base their investment strategy on market anomalies.

References

Adams, C.A., P. Weetmer, E. Jones and S.J. Gray, 1998, 'Measurement in UK Accounting Standards : A Case of Increasing Disharmony in Relation to US GAAP and IASs', Accounting and Business Research, Summer 1998.

Affleck-Graves, J., and B. McDonald, 1990, 'Multivariate Tests of Asset Pricing: the Comparative Power of Alternative Statistics', Journal of Financial and Quantitative Analysis 25.

Ashiq, A. and L.-S. Hwang, 2000, 'Country-Specific Factors Related to Financial Reporting and the Value Relevance of Accounting Data', Journal of Accounting Research, 38.

Baks, K.P., A. Metrick and J. Wachter, 2001, 'Should Investors Avoid all Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation', Journal of Finance, 56.

Bakshi G. and A. Chan, 2000, 'Earnings Yield, capitalization, and the Cross-Section of Holding Period Returns: A Global Perspective', Mimeo.

Banz, R. W., 1981, 'The Relationship between Return and Market Value of Common Stocks', Journal of Financial Economics 9.

Barberis, N., A. Shleifer and R. Vishny, 1998, 'A Model of Investor Sentiment', Journal of Financial Economics, 49.

Basu, S., 1977, 'Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: a test of the efficient market hypothesis', Journal of Finance 32.

Bauwens L. and P. Richard, 1985, 'A 1-1 poly-t Random Variable Generator with Application to Monte Carlo Integration', Journal of Econometrics, 29.

Bawa V.S., S.J. Brown and R.W. Klein, 1979, 'Estimation Risk and Optimal Portfolio Choice', in: 'Studies in Bayesian Econometrics', ed. Zellner, Arnold and Joseph B. Kadane, Vol.3, North-Holland.

Breen, W., 1968, 'Low Price-Earnings Ratios and Industry Relatives', Financial Analysis Journal.

Breen, W. and R. Korajczyk, 1993, 'On Selection Biases in Book-to-Market Based Tests of Asset-Pricing Models', Working Paper, 167, Northwestern University.

Campbell, J.Y., 1998, 'Asset Prices, Consumption and the Business Cycle', in Taylor, John and Michael Woodford (eds.), Handbook of Macroeconomics, Amsterdam: North-Holland, 1999.

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

Campbell, J.Y. and J.H. Cochrane, 1999, 'By Force of Habit : A Consumption-based Explanation of Stock Market Behavior', Journal of Political Economy, 107.

Campbell, J.Y. and J.H. Cochrane, 2000, 'Explaining the Poor Performance of Consumption-based Asset Pricing Models', Journal of Finance, 55.

Carhart, M.M., 1997, 'On Persistence in Mutual Fund Performance', Journal of Finance, 52.

Chan, L.K.C., N. Jegadeesh and J. Lakonishok, 1996, 'Momentum Strategies', Journal of Finance, 51.

Chen, N., Roll, R., and S.A. Ross, 1986, 'Economic Forces and the Stock Market', Journal of Business 59.

Chordia, T. and L. Shivakumar, 2000, 'Momentum, Business-Cycle and Time-Varying Expected Returns', Working Paper, Emory University and London Business School.

Cochrane, J., 2001, 'Asset Pricing', Princeton University Press.

Connor, G. and R. Korajczyk, 1988, 'Risk and Return in an Equilibrium APT: Application of a New Test Methodology', Journal of Financial Economics, 19.

Cote, J. and D. Sanders, 1995, 'Herding Behavior and Analysts' Earnings Forecast Bias', Working paper, Washington State University, Vancouver.

Daniel, K., D. Hirschleifer and A. Subrahmanyam, 1998, 'A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Over-reactions', Journal of Finance, 53.

De Bondt, W. F. M. and R. Thaler, 1995, 'Financial Decision-Making in Markets and Firms: A Behavioral Perspective', in: R.A. Jarrow, V. Maksimovic and W. Ziemba (eds.), The Handbook of Operational Research and Management Science, Elsevier Science.

Dechow P. and R.G. Sloan, 1997, ‘Returns to Contrarian Investment Strategies: Tests of Naïve Expectations Hypothesis’, Journal of Financial Economics, 43.

Dechow P., A.P. Hutton and R.G. Sloan, 1999, ‘An Empirical Assessment of the Residual Income Valuation Model’, Journal of Accounting and Economics, 26.

Diebold F.X. and R.S. Mariano, 1994, ‘Comparing Predictive Accuracy’, NBER paper, n.169.

Dimson, E. and P.R. Marsh, 1984, ‘An Analysis of Brokers’ and Analysts’ Unpublished Forecasts of UK Stock Returns’, Journal of Finance, 39.

Dreman, D.N. and E.A. Lufkin, 1997, ‘Do Contrarian Strategies Work Within Industries ?’, Journal of Investing, 6.

Dumas, B. and B. Solnik, 1995, ‘The World Price of Foreign Exchange Risk’, Journal of Finance 50.

Edwards, W., 1975, ‘Comment’, Journal of the American Statistical Association, 70.

Einhorn H. J. and R. M. Hogarth, 1988, ‘Behavioral Decision Theory: Processes of Judgment and Choice’, in: D. E. Bell, H. Raiffa and A. Tversky (eds.), ‘Decision Making, Descriptive, Normative, and Prescriptive Interactions’, Cambridge University Press.

Elton, E.J., S. Grossman and M.J. Gruber, 1986, ‘Discrete Expectational Data and Portfolio Performance’, Journal of Finance, 41.

Fama E., 1998, ‘Market Efficiency, Long-Term Returns, and Behavioral Finance’, Journal of Financial Economics, 49.

Fama, E.F. and J.D. MacBeth, 1973, ‘Risk, Return and Equilibrium: Empirical Tests’, Journal of Political Economy 81.

Fama, E.F. and K.R. French, 1992, ‘The Cross-Section of Expected Stock Returns’, Journal of Finance 47.

Fama, E.F. and K.R. French, 1993, ‘Common Risk Factors in the Returns on Stock and Bonds’, Journal of Financial Economics, 33.

Fama, E.F. and K.R. French, 1995, 'Size and Book-to-Market Factors in Earnings and Returns', Journal of Finance 50.

Fama, E.F. and K.R. French, 1996, 'Multifactor Explanations of Asset Pricing Anomalies', Journal of Finance 51.

Fama, E.F. and K.R. French, 1998, 'Value versus Growth: the International Evidence', Journal of Finance, 53.

Ferson, W.E. and C.R. Harvey, 1993, 'The Risk and Predictability of International Equity Returns', Review of Financial Studies 6.

Garcia-Ayuoso, M., J. Monterrey and C. Pineda, 1998, 'A comparative Analysis of the Value Relevance of Accounting Information in the Capital Markets of the European Union, Mimeo.

Geweke J., 1986, 'Exact Inference in the Inequality Constraint Normal Linear Regression Model', Journal of Applied Econometrics, 1.

Geweke J., 1989, 'Bayesian Inference in Econometric Models Using Monte Carlo Integration', Econometrica, 57.

Geweke, J., 1999, 'Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication', Econometric Reviews, 18.

Geweke, J., 2001, 'Bayesian Econometrics and Forecasting', Journal of Econometrics, 100.

Gibbons, M.R., 1982, 'Multivariate Tests of Financial Models: a New Approach', Journal of Financial Economics, 10.

Gibbons, M.R., Ross, S.A. and J. Shanken, 1989, 'A Test of the Efficiency of a Given Portfolio', Econometrica 57.

Grauer, R.R. and N.H. Hakansson, 1995, 'Stein and CAPM Estimators of the Means in Asset Allocation', International Review of Financial Analysis, 4.

Griffin D. and A. Tversky, 1992, 'The Weighing of Evidence and the Determinants of Confidence', Cognitive Psychology, 24.

Grossman, S. and J.E. Stiglitz, 1980, 'On the Impossibility of informationally Effcient Markets', American Economic Review, 70.

Heckelei, T. and R. C. Mittelhammer, 1996, 'Bayesian Bootstrap Inference via Regression Structure Likelihood', in: Hill R. Carter, (ed.), 'Bayesian Computational Methods and Applications', Advances in Econometrics, vol. 11. (Part A). Greenwich, Conn. and London: JAI Press.

Heston, S.L.K., G. Rouwenhorst and R.E. Wessels, 1995, 'The Structure of International Stock Returns and the Integration of Capital Markets', Journal of Empirical Finance, 2.

Hong H. and J. C. Stein, 1999, 'A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets', The Journal of Finance, 54.

Hong H., T. Lim and J. C. Stein, 2000, 'Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies', Journal of Finance, 55.

Jagannathan, R. and Z. Wang, 1996, 'The Conditional CAPM and the Cross-Section of Expected Returns', Journal of Finance 51.

Jegadeesh, N. and S. Titman, 1993, 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', Journal of Finance 48.

Jegadeesh, N. and S. Titman, 1999, 'Profitability of Momentum Strategies: an Evaluation of Alternative Explanations', NBER working paper, 7159.

Jensen, G.R., Mercer, J.M. and R.R. Johnson, 1996, 'Business Conditions, Monetary Policy and Expected Security Returns', Journal of Financial Economics 40.

Jobson, J.D. and B. Korkie, 1980, 'Estimation for Markowitz Efficient Portfolios', Journal of American Statistical Association, 75.

Jobson, J.D. and B. Korkie, 1981, 'Performance Hypothesis Testing with the Sharpe and Treynor Measures', Journal of Finance, 36.

Jorion, Ph., 1991, 'Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection', Journal of Banking and Finance, 15.

Jorion, Ph., 1986, 'Bayes-Stein Estimation for Portfolio Analysis', Journal of Financial and Quantitative Analysis, 21.

Kinney, W., D. Burgstahler and R. Martin, 1999, 'The Materiality of Earnings Surprises', Working Paper, University of Texas at Austin, University of Washington and Indiana University.

Kloek, T. and H.K. Van Dijk, 1978, 'Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo', Econometrica, 46.

Kothari, S.P., J. Shanken and R.G. Sloan, 1995, 'Another Look at the Cross-Section of Expected Stock Returns', Journal of Finance, 50.

La Porta, R., 1996, 'Expectations and the Cross-Section of Stock Returns', Journal of Finance, 51.

Lakonishok, J., A. Shleifer and R.W. Vishny, 1994, 'Contrarian Investment, Extrapolation, and Risk', Journal of Finance, 49.

Lehman, B.N. and D. Modest, 1988, 'The Empirical Foundations of Arbitrage Pricing Theory', Journal of Financial Economics, 21.

Lenk P. and M. Wedel, 2001, 'Bayesian Econometrics: A Reaction to Geweke', Journal of Econometrics, 100.

Lewellen, J., 1999, 'The Time-Series Relations among Expected Returns, Risk and Book-to-Market', Journal of Financial Economics 54.

Lim, T., 2001, 'Rationality and Analysts' Forecast Bias', Journal of Finance, 56.

Lintner, J., 1965, 'The Valuation of Risk Assets and the Selection of Risky Investment in Stock Portfolios and Capital Budgets', Review of Economics and Statistics 47.

Lo, A.W. and A.C. MacKinlay, 1990, 'An Econometric Analysis of Nonsynchronous Trading', Journal of Econometrics 45.

Lo, A.W. and A.C. MacKinlay, 1990, 'Data-Snooping Biases in Tests of Financial Asset Pricing Models', Review of Financial Studies, 3.

MacKinlay, A.C., 1985, 'An Analysis of Multivariate Financial Tests, Ph.D. Dissertation, Graduate School of Business, University of Chicago.

MacKinlay, A.C., 1987, 'On Multivariate Tests of the CAPM', Journal of Financial Economics, 18.

MacKinlay, A.C. and L. Pástor, 2000, 'Asset Pricing Models: Implications for Expected Returns and Portfolio Selection', Review of Financial Studies, 13.

Markowitz, H., 1952, 'Portfolio Selection', Journal of Finance, 7, pages 77-91.

Mehra , R. and E. Prescott, 1985, 'The Equity premium : A Puzzle, Journal of Monetary Economics', 15.

Merton, R.C., 1973, 'An Intertemporal Capital Asset Pricing Model', Econometrica 41.

Michaely, R. and K. Womack, 1999, 'Conflict of Interest and the Credibility of Underwriter Analyst Recommendations', Review of Financial Studies, 12.

Moskowitz, T. and M. Grinblatt, 1999, 'Tax Loss Selling and Return Autocorrelation: New Evidence', Working paper, University of Chicago.

Odean, T., 1998a, 'Volume, Volatility, Price and Profit When all Traders are above Average', Journal of Finance, 53.

Odean, T., 1998b, 'Are Investors Reluctant to Realize their Losses?', Journal of Finance, 53.

Ohlson, J.A., 1995, 'Earnings, Book Values and Dividends in Equity Valuation', Contemporary Accounting Research, 11 (2).

Olsen, R.A., 1996, 'Implications of Herding Behavior for Earnings Estimation, Risk Assessment, and stock returns', Financial Analyst Journal, 52.

Parkinson, M., 1980, 'The Extreme Value Method for Estimating the Variance of the Rate of Return', Journal of Business, 53.

Penman, S. H., 1996, 'The Articulation of Price-earnings Ratios and Market-to-book ratios and the evaluation of growth', Journal of Accounting Research, 34 (2).

Penman, S. H., 1998, 'Combining Earnings and Book Value in Equity Valuation', Contemporary Accounting Research, 15 (3).

Penman, S. H. and T. Sougiannis, 1998, 'A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation', Contemporary Accounting Research, 15.

Roll, R., 1977, 'A Critique of the Asset Pricing Theory's Tests: part I: on Past and Potential Testability of the Theory', Journal of Financial Economics 4.

Roll, R., 1978, 'Ambiguity When Performance is Measured by the Securities Market Line', Journal of Finance 33.

Ross, S.A., 1976, 'The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory 13.

Rouwenhorst, G., 1998, 'International Momentum Strategies', Journal of Finance, 53.

Shafer, G. and A. Tversky, 1988, 'Languages and Designs for Probability Judgment', in: D. E. Bell, H. Raiffa and A. Tversky (eds.), Decision Making, Descriptive, Normative, and Prescriptive interactions, Cambridge University Press.

Shiller, R.J., 1998, 'Human Behavior and the Efficiency of the Financial System', NBER working paper, 6375.

Stein, C., 1955, 'Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution', Proceedings of the third Berkeley Symposium on Probability and Statistics 1, Berkeley, University of California Press.

Strobaek, M., 1997, 'Making creative use of otherwise useless P/E ratios', SBC Prospects, 2.

Thompson, R., 1978, 'The Information Content of Discounts and Premiums on Closed-End Fund Shares', Journal of Financial Economics, 6.

Tversky, A. and D. Kahneman, 1988, 'Rational Choice and the Framing of Decisions', in: D. E. Bell, H. Raiffa and A. Tversky (eds.), 'Decision Making, Descriptive, Normative, and Prescriptive interactions', Cambridge University Press.