Essays in Empirical Asset Pricing

Johan Parmler

AKADEMISK AVHANDLING

Som för avläggande av ekonomie doktorsexamen
vid Handelshögskolan i Stockholm framläggs för offentlig granskning
Torsdagen den 15 December 2005, kl. 10.15
i sal KAW Handelshögskolan
Sveavägen 65

Essays in Empirical Asset Pricing

STOCKHOLM SCHOOL
OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

EFI Mission

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in the management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics. At EFI, the researchers select their projects based on the need for theoretical or practical development of a research domain, on their methodological interests, and on the generality of a problem.

Research Organization

The research activities at the Institute are organized in 23 Research Centres.
Centre Directors are professors at the Stockholm School of Economics.

EFI Research Centre:	*Centre Director:*
Management and Organisation (A)	Sven-Erik Sjöstrand
Centre for Ethics and Economics (CEE)	Hans de Geer
Centre for Entrepreneurship and Business Creation (E)	Carin Holmquist
Public Management (F)	Nils Brunsson
Information Management (I)	Mats Lundeberg
Centre for People and Organization (PMO)	Andreas Werr (acting)
Centre for Innovation and Operations Management (T)	Christer Karlsson
Centre for Risk Research (CFR)	Lennart Sjöberg
Economic Psychology (P)	Guje Sevón
Centre for Consumer Marketing (CCM)	Magnus Söderlund
Centre for Information and Communication Research (CIC)	Per Andersson (acting)
Marketing, Distribution and Industrial Dynamics (D)	Björn Axelsson
Centre for Strategy and Competitiveness (CSC)	Örjan Sölvell
Centre for Business and Economic History (BEH)	Håkan Lindgren
Accounting and Managerial Finance (B)	Johnny Lind
Centre for Financial Analysis and Managerial Economics in Accounting (BFAC)	Kenth Skogsvik
Finance (FI)	Clas Bergström
Centre for Health Economics (CHE)	Bengt Jönsson
International Economics and Geography (IEG)	Mats Lundahl
Economics (S)	Lars Bergman
Economic Statistics (ES)	Anders Westlund
Law (RV)	Erik Nerep
Centre for Tax Law (SR)	Bertil Wiman

Chair of the Board: Professor Carin Holmquist
Director: Associate Professor Filip Wijkström

Address

EFI, Box 6501, SE-113 83 Stockholm, Sweden • Homepage: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Essays in Empirical Asset Pricing

Johan Parmler

STOCKHOLM SCHOOL
OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

Dissertation for the Degree of Doctor of Philosophy, Ph.D
Stockholm School of Economics

ISBN NR 91-7258-691-5

Keywords:
Asset Pricing Models, Model Selection, Bayesian Econometrics

Distributed by:
EFI, Stockholm School of Economics
Box 6501, S-113 83 Stockholm, Sweden.
`www.hhs.se/efi`

Printed by:
Elanders Gotab, Stockholm 2005

To my family

Contents

Preface

Five years have passed since I first entered the doors of the Stockholm School of Economics.

I am very grateful for the advice and kindness shown to me by a number of people. In particular, I would like to thank my advisor Sune Karlsson for his encouragement and for pushing me at the end of the thesis work. I wish to express my sincere gratitude to my co-author and friend Andrés González.

I would also like to thank the Stockholm School of Economics and the people at the Department of Economic Statistics for a pleasant environment and the necessary resources for developing my research. Special thanks also to my fellow PhD students for not falling asleep during my seminars and to Pernilla Watson for help with practical problems over the years.

Special thanks to Håkan Lyckeborg for interesting and fruitful discussions regarding teaching and other matters in the early mornings at the department.

Most of all, I thank my wife Catharina, who has given me her loving companionship and full support specially at times of falling enthusiasm. I know I have been lost in my own world sometimes.

Last but not least, I would like to thank Bankforskningsinstitutet, and the Economic Research Institute (EFI) at the Stockholm School of Economics for financial support.

Stockholm, November 2005
Johan Parmler

Part I

Summary of Thesis

Introduction and Summary

The Capital Asset Pricing Model (CAPM) is the most widely used model in asset pricing. Several authors have contributed to this model. Sharpe (1963, 1964) is considered to be the forerunner and Mossin (1966), Lintner (1965, 1969) and Black (1972) made their contributions a few years later.

This model evaluates the asset return in relation to the market return and the sensitivity of the security to the market. The CAPM is not a predictive equation. Rather, the CAPM implies that contemporaneous movements in expected asset returns are linked to contemporaneous changes in the market excess return.

The CAPM predicts that only the covariance of returns between a specific asset and the market portfolio influences the cross-section of asset returns across assets. No additional variables such as size of the firm or fundamentals like dividend-price ratio should influence the cross-section of expected returns. However, the evidence supporting the CAPM is mixed. The first significant failure of the model was documented in Banz (1981).

Alternatives to the CAPM in determining the expected rate of return on portfolios and stocks was introduced by Ross (1976) through the Arbitrage Pricing Theory and by Merton (1973) through the Intertemporal CAPM. In contrast to the CAPM, where there is only one factor, these models allow a large number of factors to affect the rate of return. The introduction of these more general models raised the following important question: how should the risk factors in a multifactor pricing model be specified? Since the multifactor model theory is not very explicit regarding the number or nature of the factors the selection of factors has, to a large extent, become an empirical issue.

There are two strands in the empirical literature on selecting appropriate factors in a multifactor asset pricing model. One focuses on unobservable or latent factors and the second on observable factors. The first approach uses statistical techniques like factor analysis and principal components to extract the source of common variation in the asset returns. Two important studies using this approach are Lehmann and Modest (1988) and Connor and Korajczyk (1988). This approach has the advantage that the model does not make any prior assumptions about the number and the nature of the factors. Instead, the drawback lies in the difficulty in interpreting the factors obtained since they are linear combinations of more fundamental underlying economic forces. The second approach makes use of observable factors justified theoretically on the ground that they capture economy-wide risks. This approach makes the interpretation of the model straightforward. Two studies using this more theoretical approach to factor identification are Fama and French (1993) and Chen, Roll, and Ross (1986).

Multifactor pricing models are utilized in many areas of practical concern. For example, multifactor models are used to quantify the impact of events on stock returns. Assessing the performance of different investments constitutes another example. However, before a multifactor model can be used, the factors need to be identified. The first three chapters in this thesis consider the problem of selecting factors in a multifactor pricing model.

In the first and the second chapters, we conduct an exhaustive evaluation of multifactor asset pricing models based on observable factors. From a large set of factors Bayesian techniques are used to rank all the possible models based on posterior model probabilities.

In contrast to the first two chapters, the third chapter take the approach of using latent factors. In this chapter we set up the determination of the number of factors as a model selection problem. Again, Bayesian techniques are used.

In the first three chapters, Bayesian techniques are used to rank a large set of competing models. A Bayesian approach offers several advantages. Especially, and in contrast to a classical approach, it gives a coherent framework for addressing model uncertainty and comparison of non-nested models is straightforward. With that said, it should be mentioned that Bayesian analysis has its difficulties. Firstly the researcher needs to assign prior beliefs regarding the different models and the model parameters. With many models this is usually a challenging task. Secondly there are computational issues involved. See Hoeting, Madigan, Raftery, and Volinsky (1999) and Fernández, Ley, and Steel (2001) for a review and references.

The market is efficient if the prices of assets reflect all available information. In particular, the market is said to be weak-form efficient if today's prices reflect information contained in past prices. Consequently, it should be impossible to earn risk adjusted abnormal returns by exploiting investment strategies based on past prices. However, Jegadeesh and Titman (1993) document that over a span of three to 12 months, past winners continue to outperform past losers by about 1% per month on average, thus showing that there is "momentum" in stock prices.

While the momentum effect has been well documented, the cause of momentum is an open issue. Some have argued that the results provide strong evidence of market inefficiency and others have argued that returns from momentum strategies are compensation for risk. Finally, some claim that the profit obtained from momentum strategies is the product of data-snooping. The effect of data-snooping is probably the hardest to address since empirical research is limited by data availability.

In the final chapter we investigate if a momentum strategy is superior to a benchmark model once the effects of data-snooping have been accounted for. The procedure used is known as the "Reality Check", which was devised by White (2000).

A detailed summary of the chapters follows.

Chapter 1. Choosing Factors in a Multifactor Asset Pricing Model: A Bayesian Approach [1]

In this paper we conduct an evaluation of multifactor asset pricing models based on observable factors. The factors used are based on theoretical considerations and previous empirical studies. The first set is stock- and bond-market factors and includes returns on a market portfolio of stocks and mimicking portfolios for the size, momentum, book-to-market, and factors related to the term-structure of interest rates. The second category brings together models where the factors are macroeconomic variables.

From the set of factors Bayesian techniques are used to rank all possible factor pricing models based on the posterior model probabilities. Two kinds of priors are used. The first one is referred to as a reference prior since the prior for the model parameters is relatively uninformative and the second prior is based on the ideas of Pastor and Stambaugh (1999, 2000) where we take into account the prior degree of confidence in an asset pricing model.

In the first set of potential factors we find strong evidence that a multifactor pricing model should include the market excess return, the size premium, and the value premium, which is consistent with the famous three factor model of

---

[1] This is a joint work with Sune Karlsson.

Fama and French (1992, 1993, 1996). The support for the three factor model of Fama and French is stronger when the test assets are portfolios formed on size and/or book-to-market and during the first subperiod 196307 - 198212. In the second set of factors we consider macroeconomic variables and the result is rather inconsistent over different investment universes. Typically, only one factor shows up with a high probability of inclusion. The growth rates in real per capita consumption, personal savings rate and yearly growth rate in industrial production are factors that show up most frequently

Chapter 2. Choosing Factors in a Multifactor Asset Pricing Model When Returns Are Nonnormal[2]

Most empirical work in the asset pricing literature starts with the assumption that returns are drawn from a multivariate normal distribution. However, there is evidence that stock returns do not follow a normal distribution (Fama (1965), Affleck-Graves and McDonald (1989), Richardson and Smith (1993)). Still, normality is a common working assumption in most of the empirical work in finance.

In this chapter we consider the problem of selecting observable factors in a multifactor asset pricing model when the assumption of normally distributed returns is relaxed. Instead, we assume that asset returns are multivariate Student-t distributed. This setup allows us to capture the fat tail property of asset returns.

From a set of factors we construct all possible linear pricing models and use Bayesian techniques to rank them based on their posterior model probabilities. The factors included are based on theoretical considerations and previous empirical studies. Data from both the US and Swedish stock markets is examined.

For the US data, using return-based factors, we find evidence that a multifactor pricing model should include the market excess return, size and value premium and the momentum factor. The results for the macroeconomic factors are mixed and depends heavily on the test assets. The results for the Swedish data show little support for the Fama-French three factor model, except for when portfolios are based on book-to-market.

Finally, we find strong evidence of deviation from normality, which makes our approach of modelling the data with a Student-t likelihood more appropriate.

---

[2] This is a joint work with Sune Karlsson.

Chapter 3. Choosing Factors in a Multifactor Asset Pricing Model When the Factors Are Unobserved

As an alternative to the CAPM, Ross (1976) developed the Arbitrage Pricing Theory (APT). However, to make the APT operational, one must specify the number of pervasive factors. Empirically, this has been handled in several ways. One approach, taken by Lehmann and Modest (1988), is to estimate and test the model using different number of factors and examine if the tests are sensitive to increasing the number of factors. A second approach, adopted in Connor and Korajczyk (1993) in an approximate factor model, is to test explicitly for the adequacy of a specific number of factors.

In this chapter, I set up the determination of the number of factors as a model selection problem. Furthermore, a Bayesian approach is used. Different kind of factor structures are considered. In particular, time series dependence is introduced in the strict and the approximate factor structure.

Using data from the US market, 4 to 6 pervasive factor were generally found. It seems like that when time series dependence is introduced, the number of factors decreases. Furthermore, the data speaks in favor of an approximate factor structure with time series dependence through a common AR(1) process across assets.

Chapter 4. Is Momentum due to Data-Snooping?[3]

In this chapter, we examine if a momentum strategy is superior to a benchmark model once the effects of data-snooping have been accounted for. Data snooping occurs when a given set of data is used more than once for inference or model selection. As argued by Lo and MacKinlay (1990), the data-snooping bias can be substantial in financial studies.

The procedure used is known as the "Reality Check" which was devised by White (2000). A problem associated with White's Reality Check is that the power of the test is sensitive to the inclusion of a poor model. This issue is addressed by Hansen (2004) who proposed a modified version of White's test. In our, paper we also implement Hansen's modification.

Many studies of momentum and weak market efficiency have been conducted on US data. In contrast to the US studies, the evidence on the Swedish stock market is limited. Therefore, this paper also examines the momentum

---

[3] This is a joint work with Andrés González.

effect on Swedish stock returns and portfolios formed on size, book-to-market and industries.

The result shows that data-snooping bias can be very substantial. In this study, neglecting the problem would lead to very different conclusions. For the US data there is strong evidence of a momentum effect and we reject the hypothesis of weak market efficiency. For the Swedish data the results indicates that momentum strategies based on individual stocks generate positive and significant profits. Interestingly, a very weak or non at all, momentum effect can be found when stocks are sorted by size, book-to-market and industry.

Bibliography

AFFLECK-GRAVES, J., AND B. MCDONALD (1989): "Nonnormalities and Tests of Asset Pricing Theories," *Journal of Finance*, 44, 889–908.

BANZ, R. (1981): "The Relationship between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3–18.

BLACK, S. (1972): "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business*, 45, 444–455.

CHEN, N., R. ROLL, AND S. ROSS (1986): "Economic Forces and the Stock Market," *Journal of Business*, 59, 383–403.

CONNOR, G., AND R. KORAJCZYK (1988): "Risk and Return in an Equilibrium APT: Application of a New Test Methodology," *Journal of Financial Economics*, 21, 255–289.

——— (1993): "A Test for the Number of Factors in an Approximate Factor Model," *Journal of Finance*, 48, 1263–1291.

FAMA, E. (1965): "The Behaviour of Stock Market Prices," *Journal of Business*, 38, 34–105.

FAMA, E., AND K. FRENCH (1993): "Common Risk Factors in the Returns of Stock and Bonds," *Journal of Financial Economics*, 33, 3–56.

FERNÁNDEZ, C., E. LEY, AND M. STEEL (2001): "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, 38, 381–427.

HANSEN, P. (2004): "A Test for Superior Predictive Ability," Discussion Paper 01-06, Brown Univ. Dept. of Economics.

HOETING, J., D. MADIGAN, A. RAFTERY, AND C. VOLINSKY (1999): "Bayesian model averaging: A tutorial (with Discussion)," *Statistical Science*, 14, 382–401. Correction: vol. 15, pp. 193–195.

JEGADEESH, N., AND S. TITMAN (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48, 65–91.

LEHMANN, B., AND D. MODEST (1988): "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics*, 21, 213–254.

LINTNER, J. (1965): "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economic and Statistics*, 47, 13–37.

——— (1969): "The Aggregation of Investor's Diverse Judgements and Preferences in Purely Competitive Security Markets," *Journal of Financial and Quantitative Analysis*, 4, 347–400.

LO, A., AND A. MACKINLAY (1990): "Data-snooping biases in tests of financial asset pricing models," *Review of Financial Studies*, 3, 431–467.

MERTON, R. (1973): "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–887.

MOSSIN, J. (1966): "Equilibrium in a Capital Asset Market," *Econometrica*, October, 768–783.

PASTOR, L., AND R. F. STAMBAUGH (1999): "Costs of Equity Capital and Model Mispricing," *Journal of Finance*, (54), 67–121.

——— (2000): "Comparing Asset Pricing Models: An Investment Perspective," *Journal of Financial Economics*, (56), 335–381.

RICHARDSON, M., AND T. SMITH (1993): "A Test of Multivariate Normality of Stock Returns," *Journal of Business*, 66, 295–321.

ROSS, S. A. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

SHARPE, W. (1963): "A Simplified Model for Portfolio Analysis," *Management Science*, January, 277–293.

——— (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425–442.

WHITE, H. (2000): "A Reality Check for Data Snooping," *Econometrica*, 68, 1097–1127.

Part II

The Chapters

Chapter 1

Choosing Factors in a Multifactor Asset Pricing Model: A Bayesian Approach

Acknowledgement: An earlier version of this paper was presented at the Workshop on Econometrics and Computational Economics, Helsinki, November 2002; the $(EC)^2$ conference, Bologna, December 2002; the Nordic Econometric Meeting, Bergen, May 2003; and the Econometric Society European Meeting, Stockholm, August 2003. We wish to thank participants for their comments, and any errors in the paper are ours alone.

1.1 Introduction

The capital asset pricing model, (CAPM), developed by Sharpe (1964), Lintner (1965) and Black (1972), predicts that the expected asset return is a linear function of the risk, where the risk is measured by the covariance between its return and that of a market portfolio. The empirical evidence on the CAPM is mixed. Black, Jensen, and Scholes (1972), Fama and MacBeth (1973) and Blume and Friend (1973) find support for CAPM whereas Basu (1977) and Banz (1981), Fama and French (1992, 1993), DeBondt and Thaler (1985) and Jegadeesh and Titman (1993) find evidence against the CAPM. The mixed evidence naturally leads to the consideration of multifactor asset pricing models.

Multifactor pricing models are introduced by Ross (1976) through the Arbitrage Pricing Theory and by Merton (1973) through the Intertemporal CAPM. The multifactor pricing model implies that the expected return on an asset is a linear function of factor risk premiums and their associated factor sensitivities. The underlying theory is, however, not very explicit on the exact nature of these factors. The selection of an appropriate set of factors is thus largely an empirical issue. There are two strands in the empirical literature on multifactor asset pricing models. One focusing on unobservable or latent factors and the second focusing on observable factors.

The first approach uses latent (unobservable) factors as a source of common variation. These common factors are themselves extracted from the asset returns by using statistical techniques like factor analysis and principal components. Connor and Korajczyk (1988), who use principal components, find evidence for one to six latent factors in the cross-section of stock returns. Lehmann and Modest (1988), who use factor analysis, find weak evidence in favor of a ten-factor model but they also argue that the tests have little power to distinguish between models with different numbers of factors. This approach has the advantage that the model does not make any prior assumptions about the number and the nature of the factors. Instead, the drawback with using this kind of models then comes from the difficulty in interpreting the factors obtained. Furthermore, these models are not able to explicitly associate the estimated factors with the underlying state of the economy.

The second approach suggests the use of observable factors. The factors are assumed to capture wide economic risk associated with asset returns. Unfortunately, as in many economic applications, the theory is not very explicit about the nature of these factors. Chen, Roll, and Ross (1986) find evidence of five priced macroeconomic factors. Fama and French (1992, 1993, 1996) use firm characteristics to form factor portfolios and this resulted in the well known three-factor model. In addition, Carhart (1997) finds evidence for a fourth momentum factor. Overall, there is thus a lack of consensus about the number and the identity of the factors.

In this paper we conduct an exhaustive evaluation of multifactor asset pricing models based on observable factors. From a set of K factors Bayesian techniques are used to rank the 2^K possible models based on the posterior model probabilities. Two kinds of priors are considered. The first one is referred to as a reference prior since the prior for the model parameters is relatively uninformative, which ensures that the posterior results are dominated by the data. The second prior is based on the ideas of Pastor and Stambaugh (1999, 2000) where we take into account the prior degree of confidence in an asset pricing model.

The factors used are based on theoretical considerations and previous empirical studies. The first set is stock- and bond-market factors and includes returns on a market portfolio of stocks and mimicking portfolios for the size, momentum, book-to-market, and factors related to the term-structure of interest rates. The second category brings together models where the factors are macroeconomic variables.

The rest of the chapter is organized as follows. In the next section a general multifactor pricing model is presented. Section 1.3 then describes the Bayesian model selection procedure. Sections 1.4 and 1.5 contain the data and empirical results respectively, and Section 1.6 contains a conclusion.

1.2 The Multifactor Asset Pricing Model

In general, a multifactor pricing model states that the returns on different assets are explained by a set of common factors in a linear model. For the excess return on N assets, $\mathbf{r}$, we have the general multifactor model

$$E(\mathbf{r}) = \beta_1\lambda_1 + \beta_2\lambda_2 \tag{1.1}$$

where $E(\mathbf{r})$ is the expected excess return, $\lambda_j, j = 1, 2$ are vectors of factor risk premia. The empirical counterpart is

$$\mathbf{r}_t = \mathbf{a} + \beta_1\mathbf{f}_{1t} + \beta_2\mathbf{f}_{2t} + \varepsilon_t \tag{1.2}$$

where $\mathbf{r}_t$ is a $N \times 1$ vector of excess returns in time t, $\mathbf{a}$ is a $N \times 1$ vector of intercepts, $\mathbf{f}_{1t}$ is a $K_1 \times 1$ vector of general factors and $\mathbf{f}_{2t}$ is a $K_2 \times 1$ vector of factors that are portfolio returns. The error term ε_t is a $N \times 1$ normal distributed random vector with $E[\varepsilon_t] = 0$ and $E[\varepsilon_t \varepsilon_t'] = \mathbf{\Sigma}$. The matrices β_1 and β_2 are factor sensitivities with dimension $N \times K_1$ and $N \times K_2$, respectively.

For convenience (1.2) is rewritten as a multivariate regression model

$$\mathbf{R} = \mathbf{XB} + \mathbf{E}, \tag{1.3}$$

where the rows of $\mathbf{R}$, $\mathbf{X}$ and $\mathbf{E}$ are given by $\mathbf{r}_t'$, $\begin{bmatrix} 1 & \mathbf{f}_{1t}' & \mathbf{f}_{2t}' \end{bmatrix}$ and ε_t'. Finally, $\mathbf{B}' = \begin{bmatrix} \mathbf{a} & \beta_1 & \beta_2 \end{bmatrix}$.

Methods for the estimation and the testing of model (1.1) within the classical framework have been provided by Shanken (1992) and Velu and Zhou (1999). The testing is usually done by imposing a restriction on $\mathbf{a}$ in (1.2) implied by (1.1). In the case where the factors are traded portfolios, i.e, $\mathbf{f}_{1t}$ is absent, this is very straightforward since (1.1) implies a zero intercept. However, before any estimation and testing can take place the factors have to be identified.

Generally, asset pricing theory offers little guidance when selecting the factors. Theory suggests that assets will have to pay high average returns if they do poorly in bad times, in which investors would particularly like their investments not to perform badly and are willing to sacrifice some expected return in order to ensure that it is the case. Consumption, or more correctly marginal utility, should provide the purest measure of bad times. Investors consume less when their incomes are low or if they think future returns will be bad. But, the empirical evidence that relates asset returns to consumption is weak.[1] Therefore, empirical asset pricing models examine more indirect measures of good or bad times, interest rates, returns on broadbased portfolios, and growth in consumption, production and other macroeconomic variables that measure the state of the economy. Furthermore, it is also reasonable to include variables that signal change in the future, such as term premiums, credit spreads, etc.

The set of possible factors we consider is based on previous studies. Fama and French (1992,1993,1996), (hereinafter FF), advocate a model with the market return, the return of small less big stocks (SMB) and the return of high less low book-to-market stocks (HML) as factors. Carhart (1997) finds support for a four-factor model with the three factors of Fama and French and one additional factor that captures the momentum anomaly. Several authors have used macroeconomic variables as factors. Jagannathan and Wang (1996)

---

[1] See Cochrane (2001), chapter 2 for more details.

and Reyfman (1997) use labour income. Chen, Roll, and Ross (1986), (hereinafter CCR), test whether innovations in several macroeconomic variables are risks that are rewarded in the stock market. The variables included are: the spread between the long and short-interest rate, expected and unexpected inflation, industrial production, the spread between high and low-grade bonds, market portfolio, aggregate consumption and oil price. Other empirical evidence suggests that yields and yield spreads in corporate and Treasury bond markets are important in asset pricing models.[2]

1.3 Bayesian Model Selection

The Bayesian approach to model selection offers several advantages. In particular, the Bayesian approach is conceptually the same, regardless of the number of models under consideration, and the interpretation of the Bayes factor and the posterior model probabilities are straightforward.

From a given set of K factors, we evaluate all 2^K different models by the extent to which they describe the data as given by the posterior model probabilities. That is, we consider all possible models of the form

$$M_i : \mathbf{R} = \mathbf{X}_i \mathbf{B}_i + \mathbf{E}, \ i = 1, ..., 2^K$$

where $\mathbf{X}_i$ is $T \times (q_i + 1)$, q_i is the number of factors included in the model, and the parameter matrix $\mathbf{B}_i$ is $(q_i + 1) \times N$.

Given the prior distribution,

$$\pi(\mathbf{B}_i, \mathbf{\Sigma} | M_i)$$

for the parameters in model i, the marginal likelihood is obtained as

$$m(\mathbf{R}|M_i) = \int L(\mathbf{R}|\mathbf{B}_i, \mathbf{\Sigma}, M_i) \pi(\mathbf{B}_i, \mathbf{\Sigma}|M_i) d\mathbf{B}_i d\mathbf{\Sigma} \qquad (1.4)$$

where $L(\mathbf{R}|\mathbf{B}_i, \mathbf{\Sigma}, M_i)$ is the likelihood for model M_i. The marginal likelihood measures how well the model (and the prior) fits the data. A model comparison can be conducted through the use of Bayes factors. The Bayes factor for M_i versus M_j is given by

$$B_{ij} = \frac{m(\mathbf{R}|M_i)}{m(\mathbf{R}|M_j)} = \frac{\int L(\mathbf{R}|\mathbf{B}_i, \mathbf{\Sigma}, M_i) \pi(\mathbf{B}_i, \mathbf{\Sigma}|M_i) d\mathbf{B}_i d\mathbf{\Sigma}}{\int L(\mathbf{R}|\mathbf{B}_j, \mathbf{\Sigma}, M_j) \pi(\mathbf{B}_j, \mathbf{\Sigma}|M_j) d\mathbf{B}_j d\mathbf{\Sigma}}$$

---

[2] Ferson and Harvey (1991, 1999), Schwert (1990), Kothari and Shanken (1997), Whitelaw (1997), Campbell and Shiller (1988), and Campbell (1987).

and measures how much our belief in M_i relative to M_j has changed after viewing the data. If prior probabilities $P(M_i)$, $i = 1, ..., 2^K$, for the models are available, the Bayes factor can be used to compute the posterior model probabilities

$$P(M_i|\mathbf{R}) = \frac{m(\mathbf{R}|M_i)P(M_i)}{\sum_{j=1}^{2^K} m(\mathbf{R}|M_j)P(M_j)} = \left[\sum_{j=1}^{2^K} \frac{P(M_j)}{P(M_i)} B_{ji}\right]^{-1}.$$

Finally we note that if $P(M_i) = 1/2^K$ the posterior model probabilities are given by the normalized marginal likelihoods

$$P(M_i|\mathbf{R}) = \frac{m(\mathbf{R}|M_i)}{\sum_{j=1}^{2^K} m(\mathbf{R}|M_j)} = \left[\sum_{j=1}^{2^K} B_{ji}\right]^{-1}. \qquad (1.5)$$

There are two main problems with the Bayesian model selection. Firstly, we have to select prior distributions for the parameters of each model. In general, these priors must be informative since improper noninformative priors yield indeterminate marginal likelihoods. Secondly, to obtain the Bayes factors and the posterior model probabilities we need to compute the integration in (1.4). The second problem is addressed by using conjugate priors, which yield a closed form expression for the marginal likelihood. In order to be able to specify reasonable priors on the parameters of a large number of models the priors must be "automatic" and depend on a small number of hyperparameters. Two variations on the prior structure are considered. One, which is largely uninformative, is based on the reference prior proposed by Fernández, Ley, and Steel (2001) for model selection in univariate regression models. The second prior structure borrows ideas from Pastor and Stambaugh (1999, 2000) and explicitly incorporates information from economic theory in the form of the investor's degree of confidence in asset pricing models. The second prior specification is used when only return-based factors are considered in the set of potential factors.

1.3.1 Reference Prior

In this case, we use the natural conjugate prior for the factor sensitivities, $\mathbf{B}$, and for the covariance matrix, $\mathbf{\Sigma}$, we follow Berger and Pericchi (1998) and specify a diffuse prior since $\mathbf{\Sigma}$ is common for all models and the indeterminate factors cancel in the Bayes factor. The prior for $\mathbf{B}_i$ given $\mathbf{\Sigma}$ is given by the matrix variate normal distribution[3]

$$\mathbf{B}_i|\mathbf{\Sigma}, M_i \sim MN_{(q_i+1)\times N}\left(\mathbf{B}_i|\bar{\mathbf{B}}_i, \mathbf{\Sigma}, \mathbf{Z}_i^{-1}\right)$$

and the improper prior for $\mathbf{\Sigma}$ is given by

$$\pi\left(\mathbf{\Sigma}\right) \propto |\mathbf{\Sigma}|^{-\frac{1}{2}(N+1)}.$$

Using the above prior settings, the marginal likelihood for model M_i can be derived analytically. Let $\hat{\mathbf{B}}_i$ be the OLS estimator of $\mathbf{B}_i$ and let $\mathbf{S}_i = (\mathbf{R} - \mathbf{X}_i\hat{\mathbf{B}}_i)'(\mathbf{R} - \mathbf{X}_i\hat{\mathbf{B}}_i)$. Then, the Bayes factor for model M_i versus M_j is

$$B_{ij} = \frac{|\mathbf{Z}_i|^{N/2}|\mathbf{A}_i|^{-N/2} C_{IW}\left(\mathbf{S}_i^*, T, N\right)}{|\mathbf{Z}_j|^{N/2}|\mathbf{A}_j|^{-N/2} C_{IW}\left(\mathbf{S}_j^*, T, N\right)}$$

where $\mathbf{S}_i^* = \mathbf{S}_i + (\bar{\mathbf{B}}_i - \hat{\mathbf{B}}_i)'\left[\mathbf{Z}_i^{-1} + (\mathbf{X}_i'\mathbf{X}_i)^{-1}\right]^{-1}(\bar{\mathbf{B}}_i - \hat{\mathbf{B}}_i)$, $\mathbf{A}_i = \mathbf{Z}_i + \mathbf{X}_i'\mathbf{X}_i$ and

$$C_{IW}\left(\mathbf{S}, v, q\right) = 2^{\frac{1}{2}vq}\pi^{\frac{1}{4}q(q-1)}\prod_{i=1}^{q}\Gamma\left(\frac{v+1-i}{2}\right)|\mathbf{S}|^{-\frac{1}{2}v}.$$

Choosing the prior hyperparameters can be difficult in the absence of prior information. Reflecting the lack of consensus in the finance literature about the identity of the factors the prior mean of $\mathbf{B}$ conditional on a specific model is $\bar{\mathbf{B}}_i = 0$ and for the prior covariance matrix we follow Fernández, Ley and Steel (2001), Hall, Hwang, and Satchell (2002) and Smith and Kohn (2000) and use the g-prior of Zellner (1986). Thus,

$$\mathbf{Z}_i = g\left(\mathbf{X}_i'\mathbf{X}_i\right)$$

where $g > 0$. The parameter g is chosen such that the prior variance is large relative to the OLS counterpart. Finally, the Bayes factor simplifies to

$$B_{ij} = \left(\frac{g}{1+g}\right)^{\frac{N}{2}(q_i-q_j)}\left(\frac{\left|\mathbf{S}_j + \hat{\mathbf{B}}_j'\frac{g}{g+1}\left(\mathbf{X}_j'\mathbf{X}_j\right)\hat{\mathbf{B}}_j)\right|}{\left|\mathbf{S}_i + \hat{\mathbf{B}}_i'\frac{g}{g+1}\left(\mathbf{X}_i'\mathbf{X}_i\right)\hat{\mathbf{B}}_i)\right|}\right)^{\frac{1}{2}T} \tag{1.6}$$

and we can easily calculate the posterior model probabilities given by (1.5).

---

[3] That is $E(vec\mathbf{B}_i) = vec(\bar{\mathbf{B}}_i)$ and $Cov(vec\mathbf{B}_i) = \mathbf{\Sigma}\otimes\mathbf{Z}_i^{-1}$, where $\otimes$ denotes the Kronecker product.

1.3.2 Informative Prior

The prior setup in the previous section is very convenient and commonly used in Bayesian model selection problems. However, the prior may not be very realistic. Firstly, the prior mean of $\bar{\mathbf{B}}_j$ conditional on any specific model is zero. This is quite unreasonable since this leads to a zero expected return on all assets. Secondly, the beta for the market excess return has to be close to one on average. Finally, in the absence of macroeconomic factors, the pricing model in (1.1) implies that the intercept or misspricing is zero. In this section, we present a more realistic prior where we take into account the degree of confidence in an asset pricing model. This is done by following the ideas of Pastor and Stambaugh (1999, 2000) .

More formally, in Equation (1.3), let $\mathbf{X} = [l_T \ \mathbf{F}]$ where l_T is a vector of ones, $\mathbf{F}$ contains excess returns or zero-investment portfolios and $\mathbf{B}' = [\mathbf{a} \ \beta_2]$. The prior for the factor sensitivities is

$$\mathbf{B}|\Sigma \sim MN_{N\times K}(\bar{\mathbf{B}}, \mathbf{\Sigma}, \mathbf{Z}^{-1})$$

where the prior means are equal to zero except for the market excess return where the prior mean is equal to one. The prior for $\mathbf{\Sigma}$ is given by the inverted Wishart distribution.

$$\mathbf{\Sigma} \sim iW(\mathbf{S}_0, v_0).$$

The hyperparameters for Σ are difficult to choose. We follow Kandel and Stambaugh (1996) and use statistics from the actual sample. The prior is made relatively uninformative by setting $v_0 = N+2$ and $\mathbf{S}_0 = s^2(v_0 - N - 1)\mathbf{I}_N$ where s^2 is the average of the diagonal elements of the sample covariance matrix, $\frac{1}{T}\sum_{t=1}^{T}(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_t - \bar{\mathbf{r}})'$. These choices ensure that the prior expectation of Σ exists with $E(\Sigma) = s^2\mathbf{I}_N$.

In this setup, we can incorporate a prior degree of confidence in an asset pricing model. For a given asset pricing model, asset pricing theory implies that the intercept is equal to zero. Hence, a natural choice for the prior mean for the intercept is zero. The prior confidence in the model implications that $\mathbf{a} = 0$ is then expressed through the prior variance for $\mathbf{a}$. Let Z^{-1} be

$$Z^{-1} = \begin{pmatrix} \frac{\sigma_\alpha^2}{s^2} & 0 \\ 0 & \frac{(g\mathbf{F}'\mathbf{F})^{-1}}{s^2} \end{pmatrix}.$$

Then the unconditional variance of each element of $\mathbf{a}$ is σ_α^2. The value of σ_α^2 represents a prior degree of belief that the pricing model holds. A dogmatic belief in asset pricing is then characterized with a very low value of σ_α^2. Pastor and Stambaugh (1999) introduced this measure of mispricing uncertainty.

Since we use an informative prior on the covariance matrix, $\mathbf{\Sigma}$, the prior is in the full natural conjugate framework. The marginal likelihood is given by the matricvariate Student-t density and the Bayes factor is equal to

$$\begin{aligned} B_{ij} = & \frac{|\mathbf{I}_T + \mathbf{X}_i \mathbf{Z}_i^{-1} \mathbf{X}_i'|^{-0.5N}}{|\mathbf{I}_T + \mathbf{X}_j \mathbf{Z}_j^{-1} \mathbf{X}_j'|^{-0.5N}} \\ & \times \frac{|\mathbf{S}_0 + (\mathbf{R} - \mathbf{X}_i \bar{\mathbf{B}}_i)'(\mathbf{I}_T + \mathbf{X}_i \mathbf{Z}_i^{-1} \mathbf{X}_i')^{-1}(\mathbf{R} - \mathbf{X}_i \bar{\mathbf{B}}_i)|^{-0.5(T+v_0)}}{|\mathbf{S}_0 + (\mathbf{R} - \mathbf{X}_j \bar{\mathbf{B}}_j)'(\mathbf{I}_T + \mathbf{X}_j \mathbf{Z}_j^{-1} \mathbf{X}_j')^{-1}(\mathbf{R} - \mathbf{X}_j \bar{\mathbf{B}}_j)|^{-0.5(T+v_0)}} \end{aligned} \qquad (1.7)$$

and the posterior model probability can easily be calculated by (1.5).

1.4 The Data

The data in this study contains monthly observations on US stock excess returns and a set of factors spanning from July 1963 through December 2003. The estimation and the testing of multifactor asset pricing models are typically done on portfolios of assets, rather than on individual assets. The reason for this is that the returns must be stationary, in the sense that they have approximately the same mean and covariance. Individual assets are usually very volatile, which makes it hard to obtain precise estimates. In this study, we use eight sets of portfolios[4]. The first set contains the six benchmark portfolios of Fama and French sorted on size[5] and book-to-market[6] (B/M). The second set contains the 25 Fama and French (1993) portfolios formed on size and B/M. The third set contains 10 industry portfolios. The last five sets contains 10 portfolios formed on size, book-to-market, cashflow, earnings and dividends respectively.[7]

Based on theoretical considerations and previous empirical studies, two sets of candidate factors are specified in our evaluation. The first set is stock- and bond-market factors and includes returns on a market portfolio of stocks and mimicking portfolios for the size, momentum, book-to-market, and factors related to the term-structure of interest rates. This will be referred to as return-based factors. The second set contains macroeconomic factors.

---

[4] The portfolios include all NYSE, AMEX, and NASDAQ stocks

[5] Market equity (size) is price times shares outstanding

[6] Book equity to market equity.

[7] We thank Kenneth R. French for providing the data at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

1. Market excess returns (MKT-RF), the difference between value-weighted return on all NYSE, AMEX, and NASDAQ stocks and the one-month Treasury bill rate, size premium (SMB), value premium (HML) and a momentum factor (UMD). The credit risk spread (RP) is the difference between yields of Moody's Baa and the yields of Moody's Aaa rated bonds. This is a state variable that measures changes in the risk of corporate bonds. Proxies for unexpected change in interest rates are the difference in the annualized yield of ten-year and one-year Treasuries (UTS(L)), and the difference between the one-year Treasuries and the three-month Treasury bill rate (UTS(S)).

2. The macroeconomic factors are monthly (MP) and yearly (YP) growth rate in industrial production, unanticipated inflation (UI), the change in expected inflation (DEI), growth rate in real per capita personal consumption (CG) and monthly change in oil price (OG). The inflation series was obtained by following the procedures in Fama and Gibbons (1984). In addition to these CRR factors we add the following, growth rate in real per capita disposable income (IC), growth rate in personal savings rate (PSR) and growth rate in unemployment rate (UNR).

Note that in some cases the intercept is treated as one of the factors. The set of potential factors is summarized in Table 1.1.

1.5 Empirical Results

In this section, we will try to identify the nature of the factors in a multifactor pricing model. First, we will examine the return-based and non-return-based factors separately and then we will merge the two. Equations (1.6) and (1.7) compute the Bayes factors and by allocating the prior model probabilities equally to all models (1.5) yields the posterior model probabilities. As a starting point we will use the reference prior outlined in Section 1.3.1. Hence, in the prior settings we only need to specify the parameter g, the amount of prior information relative to the information in the data. The results presented here are based on $g = 0.05$. That is, the prior information corresponds to 5% of the sample. The analysis will then be followed by the extended prior described in Section 1.3.2 and finally we will examine the sensitivity with respect to the sample period and the prior specification.

Table 1.1: The set of potential factors

Symbol	Variable
MKT-RF	Market excess return
SMB	Size premium
HML	Value premium
UMD	Momentum premium
RP	The credit risk spread
UTS(S)	Term spread (short)
UTS(L)	Term spread (long)
MP	Monthly growth rate in industrial production
YP	Yearly growth rate in industrial production
CG	Monthly growth rate in consumption
IC	Monthly growth rate in income
UI	Unanticipated inflation
DEI	Change in expected inflation
OG	Monthly growth rate in oil price
PSR	Monthly growth rate in private savings
UNR	Monthly growth rate in unemployment rate

1.5.1 Return-Based Factors

In the case of only return-based factors, the asset pricing theory implies that the intercept or misspricing is zero. By including the intercept in the set of factors we can evaluate the extent of misspricing by the posterior probability that the intercept should be included in the model. A zero or small posterior probability indicates that there is no misspricing and a large posterior probability provides evidence of misspricing. This results in 8 factors and $2^8 = 256$ models where 128 of them are potential pricing models, which is the number of models without intercept.

In Table 1.2a we report the posterior probability of inclusion for the factors and the different sets of portfolios. It is computed as the total sum of the posterior probabilities of all models in which the particular factor is included.

Focusing on what is common among the different portfolios, Table 1.2a shows that size premium, value premium and market excess return all have a high probability of inclusion. This indicates that each of these factors has a high probability of appearing in a pricing model. In addition, the momentum factor has a high probability of inclusion except when portfolios constructed by size and/or book-to-market are used. It is worth noting that risk factors related to the bond market do not seem to be very important, except for the 6 size-B/M portfolios where the long term spread has a probability of inclusion

Table 1.2a: Probability of Inclusion: Reference prior $g = 0.05$

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
INT	0.037	0.035	0.000	0.000	0.424	0.001	0.000	0.000
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.007	0.070	0.025	0.001	0.944	0.999	0.941	0.723
RP	0.049	0.000	0.000	0.000	0.053	0.000	0.000	0.000
UTS(S)	0.001	0.000	0.000	0.000	0.003	0.000	0.000	0.000
UTS(L)	0.924	0.253	0.001	0.000	0.000	0.000	0.000	0.000

B/M = book-to-market; INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

of 0.92. Except for the industry portfolios we find no evidence of misspricing.

One major advantage of the Bayesian approach is that model uncertainty is easily quantified. In Table 1.2b attention is paid to the three best models with the highest posterior model probabilities represented by combinations of zeros and ones, where one indicates that a specific factor is included in the model.

Starting with the 25 size-B/M portfolios as the investment universe, the best model has a posterior model probability of 0.65. The factor pricing model includes size and value premiums, and the market excess return. This is consistent with the three- factor model of Fama and French (1993). For the second and the third model the posterior model probabilities are 0.25 and 0.07 respectively. This indicates the importance of model uncertainty in asset pricing models. For the six portfolios also constructed by sorting stocks on size and book-to-market the result differs from the previous case in several ways. First, we note that the model with the highest probability contains the long term-spread in addition to the three FF factors. Secondly, the posterior model probability for the best model is much higher, namely 0.91. In columns three and four we use portfolios formed by book-to-market and size. The best models clearly dominate with a posterior model probability of 0.97 and 0.99 respectively, and yield strong support for the FF model.

In the set of results for portfolios formed on size and/or book-to-market, the dependent returns and the two explanatory returns, SMB and HML, are portfolios formed on the same firm attribute. Thus, it is possible that the

Table 1.2b: Three best models: Reference prior $g = 0.05$

Factor	6 Size-B/M			25 Size-B/M			B/M			Size		
INT	0	0	1	0	0	0	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	0	1	0	1	0	0	1	0
RP	0	1	0	0	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	1	0	0	0	1	0	0	0	1	0	0	1
Prob	0.906	0.041	0.033	0.644	0.251	0.068	0.973	0.025	0.001	0.999	0.000	0.000
Factor	Industry			Dividend			Earning			Cashflow		
INT	0	1	0	0	1	0	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	1	1	1	1	1	1	1	0	1	1	0	1
RP	0	0	1	0	0	1	0	0	0	0	0	1
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	1	0	0	0
Prob	0.511	0.379	0.051	0.999	0.000	0.000	0.941	0.059	0.000	0.722	0.277	0.000

B/M = book-to-market; INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

inclusion of these two factors is spurious. To investigate this we examine whether these factors explain returns on portfolios formed on other variables. Furthermore, by using different portfolios as the investment universe we alleviate data snooping to some extent. As noted by Lo and MacKinlay (1990) and Ericsson and Gonzalez (2003) the effect of data snooping can in financial studies be substantial.

The first part in the second panel in Table 1.2b shows the result when stocks are sorted by industry. The best model includes the three FF factors and momentum. However, the posterior model probability for the best model is only 0.51, indicating substantial model uncertainty. Most of the uncertainty is over the inclusion of the intercept. The data is rather informative when we use portfolios formed on dividends, earnings and cashflow. The best model in all three sets of portfolios contain the same factors. The model includes the FF factors and the momentum factor with a posterior model probability equal to 0.99, 0.94 and 0.72 respectively.

The FF three factor model has received a lot of attention over the last ten years. However, Tables 1.2a and 1.2b reveal some interesting issues. It seems like the support for the FF model is strongest when the investment universe contains portfolios sorted with respect to size and/or book-to-market. This highlights the data snooping problem and the importance of using different portfolios as test assets.

The results so far are based on the reference prior, as outlined in Section 1.3.1. This prior setup is convenient but, in some cases, not very realistic. In the final part of this section, we will consider a more realistic prior setup, as described in Section 1.3.2. The major difference from the reference prior is that the degree of confidence in an asset pricing model is taken into account. The prior confidence in the model implication that $\mathbf{a} = 0$ is expressed through the prior variance for $\mathbf{a}$, given by $\sigma_\alpha^2 \mathbf{I}_N$. Hence, a dogmatic belief in the asset pricing model is characterized by a very low value of σ_α. Tables 1.3a to 1.3c show the results when we use the informative prior with $\sigma_\alpha = \{0.01, 0.100, 1.000\}$. As σ_α increases the confidence in the pricing model declines.

The effect of increasing σ_α on the probability of inclusion is mixed, except for the three Fama and French factors where the probabilities are always large. The risk premium obtains a higher probability for larger σ_α while the other factors have a lower probability. One exception is the momentum factor for the cashflow portfolios where the probability increases with σ_α. More interestingly, our prior belief in asset pricing seems to have an affect on model uncertainty. When we have a very strong prior belief in asset pricing, the model uncertainty is low and when our confidence decreases, that is σ_α increases, the model

Table 1.3a: Probability of Inclusion: Informative prior $g = 0.05$.

$\sigma_\alpha = 0.01$

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.006	0.020	0.020	0.000	0.976	0.999	0.904	0.646
RP	0.047	0.000	0.000	0.000	0.076	0.000	0.000	0.000
UTS(S)	0.002	0.000	0.000	0.000	0.004	0.000	0.000	0.000
UTS(L)	0.956	0.125	0.001	0.000	0.000	0.000	0.000	0.000

$\sigma_\alpha = 0.100$

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.005	0.000	0.022	0.000	0.941	1.000	0.695	0.916
RP	0.057	0.000	0.003	0.001	0.005	0.001	0.001	0.015
UTS(S)	0.003	0.000	0.000	0.001	0.000	0.000	0.001	0.000
UTS(L)	0.523	0.000	0.000	0.000	0.000	0.000	0.000	0.000

$\sigma_\alpha = 1.000$

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.004	0.000	0.022	0.000	0.932	0.999	0.632	0.934
RP	0.338	0.006	0.219	0.025	0.278	0.035	0.047	0.758
UTS(S)	0.005	0.000	0.000	0.005	0.000	0.000	0.006	0.001
UTS(L)	0.224	0.000	0.000	0.000	0.000	0.000	0.000	0.000

B/M = book-to-market; INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

Table 1.3b: Three best models: Informative prior $g = 0.05$.

$\sigma_\alpha = 0.01$

Factor	6 Size-B/M			25 Size-B/M			B/M			Size		
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	0	0	1	0	0	1	0	1	0	0	1	0
RP	0	1	0	0	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	1	0	1	0	1	0	0	0	1	0	0	1
Prob	0.944	0.042	0.006	0.855	0.125	0.020	0.980	0.020	0.001	0.999	0.000	0.000
Factor	Industry			Dividend			Earning			Cashflow		
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	1	1	0	1	1	1	1	0	1	1	0	1
RP	0	1	0	0	1	0	0	0	0	0	0	1
UTS(S)	0	0	0	0	0	1	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	1	0	0	0
Prob	0.899	0.074	0.021	0.999	0.000	0.000	0.904	0.096	0.000	0.715	0.285	0.000

$\sigma_\alpha = 0.100$

Factor	6 Size-B/M			25 Size-B/M			B/M			Size		
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	1	0	0	1	0	0	0	0
RP	0	0	1	0	0	0	0	0	1	0	0	1
UTS(S)	0	0	0	0	0	0	0	0	0	0	1	0
UTS(L)	1	0	0	0	0	1	0	0	0	0	0	0
Prob	0.515	0.420	0.051	1.000	0.000	0.000	0.975	0.021	0.003	0.998	0.000	0.000
Factor	Industry			Dividend			Earning			Cashflow		
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	1	0	1	1	1	1	1	0	1	1	0	1
RP	0	0	1	0	1	0	0	0	1	0	0	1
UTS(S)	0	0	0	0	0	1	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	0	0	0	0
Prob	0.936	0.059	0.005	0.999	0.000	0.000	0.694	0.305	0.000	0.902	0.083	0.014

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

Table 1.3c: Three best models: Informative prior $g = 0.05$.

	$\sigma_\alpha = 1.000$											
Factor	6 Size-B/M			25 Size-B/M			B/M			Size		
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	0	1	0	0	1	0	0	0
RP	0	1	0	0	1	0	0	1	0	0	1	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	1
UTS(L)	0	0	1	0	0	0	0	0	0	0	0	0
Prob	0.459	0.310	0.196	0.994	0.006	0.000	0.762	0.215	0.018	0.969	0.025	0.005
Factor	Industry			Dividend			Earning			Cashflow		
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	1	1	0	1	1	1	1	0	1	1	1	0
RP	0	1	0	0	1	0	0	0	1	1	0	1
UTS(S)	0	0	0	0	0	1	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	0	0	0	0
Prob	0.676	0.256	0.046	0.965	0.035	0.000	0.601	0.346	0.028	0.706	0.228	0.052

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

uncertainty becomes larger. However, the best model does not change for different values of σ_α. The only case where the best model changes is when the 6 size and book-to-market portfolios are the investment universe and σ_α changes from 0.1 to 1. Again, the support for the FF model is highest for portfolios based on size and/or book-to-market.

It is important to note that, even if the choice of σ_α does not have a big impact on the factors selected, it has an impact on asset pricing. In Tables 1.4a and 1.4b we present the posterior mean and standard deviation of $\mathbf{a}$, the vector of intercepts for the model with the highest posterior probability. By looking at the posterior mean of $\mathbf{a}$ we note that the misspricing increases with its prior variance. This is something we can expect since when σ_α is low, more weight is allocated to the prior and the posterior mean of α shrinks towards its prior mean. Note that the posterior standard deviation also increases with the misspricing prior variance. The posterior means and standard deviations are lowest for the size sorted portfolios. For the other portfolios the misspricing is approximately the same in magnitude. In summary, the estimated intercepts show that the best models leave a large unexplained return and that the unexplained returns are bigger when the investor becomes more uncertain about establishing an asset pricing model with zero intercept. On the other hand, a 90% highest posterior density region would cover zero in most cases.

1.5.2 Non-Return-Based Factors

A major criticism of the Fama and French three factor model is the interpretation of the risk factors. In particular, it is not clear what kind of economic risks these are proxies for. This is a common problem for all asset pricing models based on fundamental factors. Regarding this issue it is useful and interesting to examine macroeconomic factors directly in an asset pricing context. The drawback is that the implication of a zero intercept does not hold any longer. Consequently, we always include the intercept in the model and select macroeconomic factors. Furthermore, the market factor is also always included and the selected models can be viewed as an extended CAPM with macroeconomic factors.

In Table 1.5a we report the posterior probability of inclusion for the macroeconomic factors and the different sets of portfolios. In this case, the result is rather inconsistent over different investment universes. Typically, only one factor obtains a high probability of inclusion. The growth rate in real per capita consumption, growth rate in personal savings rate and yearly growth rate in industrial production are factors that show up most frequently with a high posterior probability of inclusion. These factors are also the ones that are

Table 1.4a: Posterior means and standard deviations for the intercept in the model with the highest posterior model probability: Informative prior $g = 0.05$.

	B/M			Size		
Portfolio	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$
1	0.008 (0.018)	0.154 (0.075)	0.187 (0.083)	0.000 (0.022)	0.004 (0.096)	0.005 (0.106)
2	0.004 (0.017)	0.083 (0.073)	0.100 (0.081)	-0.001 (0.017)	-0.015 (0.075)	-0.018 (0.083)
3	0.001 (0.017)	0.014 (0.076)	0.017 (0.083)	-0.001 (0.016)	-0.011 (0.069)	-0.014 (0.076)
4	-0.002 (0.019)	-0.045 (0.083)	-0.054 (0.091)	0.000 (0.016)	-0.008 (0.070)	-0.010 (0.077)
5	-0.001 (0.019)	-0.026 (0.080)	-0.031 (0.088)	0.002 (0.016)	0.042 (0.068)	0.051 (0.075)
6	0.001 (0.017)	0.027 (0.075)	0.032 (0.082)	-0.002 (0.016)	-0.033 (0.069)	-0.040 (0.076)
7	0.002 (0.017)	0.029 (0.075)	0.035 (0.082)	0.002 (0.016)	0.043 (0.067)	0.052 (0.074)
8	-0.001 (0.015)	-0.016 (0.066)	-0.019 (0.073)	0.001 (0.015)	0.020 (0.067)	0.025 (0.074)
9	-0.002 (0.017)	-0.028 (0.074)	-0.034 (0.082)	0.000 (0.014)	0.009 (0.060)	0.011 (0.067)
10	-0.003 (0.022)	-0.050 (0.096)	-0.060 (0.105)	0.004 (0.011)	0.071 (0.046)	0.086 (0.051)
	Industry			Dividend		
Portfolio	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$
1	0.006 (0.026)	0.125 (0.113)	0.153 (0.125)	0.006 (0.022)	0.122 (0.099)	0.150 (0.109)
2	0.002 (0.030)	0.034 (0.134)	0.042 (0.149)	0.002 (0.020)	0.039 (0.087)	0.047 (0.096)
3	-0.003 (0.019)	-0.067 (0.084)	-0.082 (0.093)	0.008 (0.021)	0.148 (0.092)	0.182 (0.102)
4	0.001 (0.037)	0.029 (0.163)	0.035 (0.180)	-0.002 (0.020)	-0.034 (0.090)	-0.042 (0.100)
5	0.017 (0.032)	0.327 (0.139)	0.401 (0.154)	-0.005 (0.022)	-0.107 (0.096)	-0.131 (0.106)
6	0.009 (0.034)	0.169 (0.151)	0.207 (0.167)	0.002 (0.020)	0.031 (0.086)	0.038 (0.096)
7	0.007 (0.028)	0.127 (0.124)	0.156 (0.138)	0.002 (0.018)	0.033 (0.082)	0.040 (0.090)
8	0.018 (0.032)	0.355 (0.140)	0.435 (0.155)	0.007 (0.019)	0.138 (0.084)	0.169 (0.093)
9	-0.007 (0.029)	-0.144 (0.129)	-0.177 (0.143)	0.005 (0.020)	0.107 (0.088)	0.131 (0.097)
10	-0.003 (0.020)	-0.060 (0.088)	-0.073 (0.098)	0.002 (0.024)	0.037 (0.107)	0.046 (0.118)
	Earning			Cashflow		
Portfolio	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$
1	0.007 (0.021)	0.142 (0.094)	0.175 (0.104)	0.007 (0.019)	0.143 (0.085)	0.175 (0.094)
2	0.002 (0.018)	0.030 (0.078)	0.037 (0.087)	0.004 (0.018)	0.088 (0.080)	0.108 (0.088)
3	0.004 (0.020)	0.069 (0.088)	0.084 (0.098)	0.004 (0.018)	0.076 (0.080)	0.093 (0.089)
4	0.003 (0.018)	0.049 (0.078)	0.060 (0.086)	0.002 (0.018)	0.048 (0.081)	0.059 (0.090)
5	-0.001 (0.019)	-0.015 (0.083)	-0.019 (0.092)	0.005 (0.019)	0.095 (0.082)	0.117 (0.091)
6	0.002 (0.019)	0.032 (0.083)	0.039 (0.092)	0.000 (0.019)	0.003 (0.086)	0.003 (0.095)
7	0.006 (0.018)	0.122 (0.080)	0.150 (0.089)	-0.003 (0.019)	-0.057 (0.086)	-0.070 (0.095)
8	0.004 (0.019)	0.081 (0.084)	0.099 (0.093)	-0.002 (0.019)	-0.033 (0.086)	-0.040 (0.095)
9	0.004 (0.021)	0.082 (0.092)	0.101 (0.102)	0.006 (0.019)	0.119 (0.083)	0.146 (0.092)
10	0.002 (0.022)	0.035 (0.098)	0.043 (0.109)	0.001 (0.021)	0.016 (0.094)	0.020 (0.104)

Table 1.4b: Posterior means and standard deviations for the intercept in the model with the highest posterior model probability: Informative prior $g = 0.05$.

		6 Size-B/M		
Size	BM	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$
Small	Low	-0.007 (0.018)	-0.130 (0.078)	-0.158 (0.086)
	2	0.005 (0.014)	0.087 (0.062)	0.106 (0.068)
	High	0.006 (0.013)	0.112 (0.058)	0.136 (0.063)
Big	Low	0.007 (0.013)	0.133 (0.056)	0.162 (0.062)
	2	-0.001 (0.015)	-0.024 (0.064)	-0.029 (0.071)
	High	-0.004 (0.014)	-0.073 (0.061)	-0.089 (0.067)
		25 Size-B/M		
Size	BM	$\sigma_\alpha = 0.01$	$\sigma_\alpha = 0.1$	$\sigma_\alpha = 1$
Small	Low	-0.018 (0.029)	-0.341 (0.123)	-0.415 (0.135)
	2	0.003 (0.022)	0.050 (0.097)	0.061 (0.106)
	3	0.003 (0.019)	0.063 (0.080)	0.076 (0.089)
	4	0.012 (0.018)	0.220 (0.078)	0.268 (0.085)
	High	0.009 (0.019)	0.176 (0.081)	0.214 (0.090)
2	Low	-0.008 (0.023)	-0.144 (0.098)	-0.175 (0.108)
	2	-0.004 (0.020)	-0.071 (0.084)	-0.086 (0.093)
	3	0.005 (0.018)	0.098 (0.077)	0.119 (0.085)
	4	0.005 (0.017)	0.097 (0.073)	0.118 (0.081)
	High	0.002 (0.018)	0.037 (0.078)	0.045 (0.086)
3	Low	-0.002 (0.021)	-0.040 (0.090)	-0.049 (0.100)
	2	0.001 (0.020)	0.018 (0.088)	0.021 (0.097)
	3	-0.002 (0.019)	-0.033 (0.081)	-0.041 (0.090)
	4	0.003 (0.018)	0.049 (0.077)	0.059 (0.085)
	High	0.003 (0.020)	0.054 (0.088)	0.066 (0.097)
4	Low	0.007 (0.020)	0.123 (0.086)	0.150 (0.095)
	2	-0.006 (0.020)	-0.111 (0.088)	-0.135 (0.097)
	3	0.000 (0.020)	0.007 (0.084)	0.009 (0.093)
	4	0.004 (0.018)	0.074 (0.079)	0.090 (0.087)
	High	-0.001 (0.023)	-0.022 (0.098)	-0.027 (0.108)
Big	Low	0.010 (0.016)	0.187 (0.069)	0.227 (0.076)
	2	0.001 (0.017)	0.011 (0.075)	0.013 (0.082)
	3	0.000 (0.019)	0.002 (0.082)	0.002 (0.090)
	4	-0.003 (0.017)	-0.056 (0.074)	-0.068 (0.081)
	High	-0.009 (0.023)	-0.163 (0.101)	-0.199 (0.111)

Table 1.5a: Probability of Inclusion, Macro factors: Reference prior $g = 0.05$

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
MP	0.000	0.000	0.071	0.000	0.000	0.000	0.000	0.056
DEI	0.002	0.000	0.004	0.000	0.014	0.000	0.000	0.000
UI	0.006	0.000	0.018	0.002	0.002	0.061	0.001	0.001
CG	0.998	0.066	0.031	0.977	0.000	0.000	0.000	0.001
IC	0.001	0.000	0.006	0.000	0.000	0.000	0.000	0.000
OG	0.003	0.000	0.196	0.010	0.359	0.937	0.000	0.002
PSR	0.013	0.933	0.147	0.008	0.119	0.001	0.999	0.001
UNR	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.894
YP	0.008	0.000	0.521	0.001	0.899	0.000	0.000	0.042

B/M = book-to-market; MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production. 1 indicates inclusion and 0 indicates exclusion.

included in the best model for the different test portfolios, as shown in Table 1.5b where the three best models for the different portfolios are presented. Note that only one factors is included in the model with the highest posterior model probability. Furthermore, the factor in the best model is generally not included in the second best model. Exceptions are the 6 size-B/M, industry and earnings portfolios.

1.5.3 All Factors

In this section, we merge the return-based factors and the macroeconomic factors into one large set of 17 factors. The intercept is one of the factors. We are aware that by using both return- and non-return based factors, any simple interpretation of the intercept is lost. Furthermore, we will compare models based on return-based factors and models based on macroeconomic factors. This is done by comparing the marginal likelihoods through the Bayes factor for the best model based on return-based factors and the best model based on macroeconomic factors.

The posterior probability of inclusion and the best models are presented in Tables 1.6a and 1.6b. The return-based factors generally obtain a higher probability of inclusion than the macroeconomic factors. This is also highlighted in Table 1.6b where the best models almost only contain return-based factors. The best asset pricing model for the four sets of portfolios formed

Table 1.5b: Three best models, Macro Factors: Reference prior $g = 0.05$

Factor	6 Size-B/M			25 Size-B/M			B/M			Size		
MP	0	0	0	0	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	0	0	0	0	0	0	0
CG	1	1	1	0	1	0	0	0	0	1	0	0
IC	0	0	0	0	0	0	0	0	0	0	0	0
OG	0	0	0	0	0	0	0	1	0	0	1	0
PSR	0	1	0	1	0	0	0	0	1	0	0	1
UNR	0	0	0	0	0	0	0	0	0	0	0	0
YP	0	0	1	0	0	1	1	0	0	0	0	0
Prob	0.964	0.012	0.933	0.933	0.066	0.001	0.521	0.196	0.976	0.783	0.010	0.008
Factor	Industry			Dividend			Earning			Cashflow		
MP	0	0	0	0	0	0	0	0	0	0	1	0
DEI	0	0	0	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	1	0	0	1	1	0	0	0
CG	0	0	0	0	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0	0	0	0	0
OG	0	1	1	1	0	0	0	0	0	0	0	0
PSR	0	0	0	0	0	1	1	1	0	0	0	0
UNR	0	0	0	0	0	0	0	0	0	1	0	0
YP	1	1	0	0	0	0	0	0	0	0	0	1
Prob	0.557	0.242	0.066	0.937	0.061	0.001	0.998	0.001	0.001	0.894	0.057	0.042

B/M = book-to-market; MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production.

Table 1.6a: Probability of Inclusion, All Factors: Reference prior $g = 0.05$

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
INT	0.036	0.035	0.000	0.000	0.193	0.000	0.000	0.000
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.007	0.070	0.025	0.001	0.918	0.999	0.936	0.709
RP	0.049	0.000	0.000	0.000	0.051	0.000	0.000	0.000
UTS(S)	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000
UTS(L)	0.924	0.253	0.001	0.000	0.000	0.000	0.000	0.000
MP	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001
DEI	0.001	0.000	0.000	0.000	0.014	0.000	0.000	0.000
UI	0.003	0.000	0.000	0.000	0.002	0.014	0.001	0.000
CG	0.006	0.000	0.001	0.000	0.000	0.000	0.000	0.000
IC	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
OG	0.001	0.000	0.000	0.001	0.216	0.653	0.000	0.000
PSR	0.003	0.000	0.002	0.001	0.280	0.000	0.776	0.000
UNR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.081
YP	0.001	0.000	0.003	0.000	0.935	0.007	0.000	0.001

B/M = book-to-market; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long. MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production. 1 indicates inclusion and 0 indicates exclusion.

on size and/or book-to-market and the cashflow portfolios does not contain any macroeconomic factors. In the case where stocks are sorted by industry, dividend and earning, the best model includes one macroeconomic factor in addition to the FF factor and the momentum factor.

Table 1.7 presents the comparison between return-based factor pricing models and macroeconomic factor pricing models. The marginal likelihoods are expressed in log format. The natural log of the Bayes factor is the difference between the log marginal likelihood for the best model using return-based factors and the log marginal likelihood for the best model based on macroeconomic factors. The results provide clear evidence in favor of a pricing model based on return-based factors. All log Bayes factors are positive and very large. The lowest Bayes factor is obtained for the industry portfolios where the difference between the log marginal likelihoods is 253. However,

Table 1.6b: Three best models, All Factors: Reference prior $g = 0.05$

Factor	6 Size-B/M			25 Size-B/M			B/M			Size		
INT	0	0	1	0	0	0	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	0	1	0	1	0	0	0	0
RP	0	1	0	0	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	1	0	0	0	1	0	0	0	0	0	0	0
MP	0	0	0	0	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	0	0	0	0	0	0	0
CG	0	0	0	0	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0	0	0	0	0
OG	0	0	0	0	0	0	0	0	0	0	1	0
PSR	0	0	0	0	0	0	0	0	0	0	0	1
UNR	0	0	0	0	0	0	0	0	0	0	0	0
YP	0	0	0	0	0	0	0	0	1	0	0	0
Prob	0.887	0.041	0.032	0.644	0.251	0.068	0.962	0.025	0.003	0.997	0.010	0.000
Factor	Industry			Dividend			Earning			Cashflow		
INT	0	0	0	0	0	0	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1
UMD	1	1	1	1	1	1	1	1	0	1	0	1
RP	0	0	0	0	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	0	0	0	0
MP	0	0	0	0	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	1	0	0	0	0	0	0
CG	0	0	0	0	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0	0	0	0	0
OG	0	0	1	1	0	0	0	0	0	0	0	0
PSR	0	1	0	0	0	0	1	0	1	0	0	0
UNR	0	0	0	0	0	0	0	0	0	0	0	1
YP	1	1	1	0	0	0	0	0	0	0	0	0
Prob	0.338	0.137	0.083	0.647	0.331	0.013	0.724	0.211	0.050	0.662	0.254	0.045

B/M = book-to-market; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long. MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production. 1 indicates inclusion and 0 indicates exclusion.

Table 1.7: Return-Based Factors vs Macroeconomic Factors: Reference prior $g = 0.05$

Portfolio	Macroeconomic Factors	Return-Based Factors	$\log(BF)$
6 Size-B/M	-5829.95	-4390.45	1439.50
25 Size-B/M	-23879.40	-22579.60	1299.80
B/M	-9659.72	-9169.24	490.48
Size	-7950.46	-7317.35	633.11
Industry	-11930.60	-11722.30	208.30
Dividend	-10123.20	-9870.11	253.09
Earning	-9948.64	-9628.36	320.28
Cashflow	-9863.08	-9653.52	209.56

it is important to remember that the return-based factors can be viewed as factor-mimicking portfolios. As argued by Cochrane (2001), a model with factor-mimicking portfolios will almost always outperform a model with real economic factors. Hence, the result in Table 1.7 is something we can expect.

1.5.4 Sensitivity Analysis

The exact results obtained are dependent on a number of choices such as the composition of the portfolios, the sample used and the prior specification. The preceding section gave some results on the sensitivity to portfolio composition. In this section we address the latter two issues.

First, we consider two subsamples, 196307 - 198212 and 198301 - 200312. The results for the return-based factors are displayed in Tables 1.8a and 1.8b and the results for the macroeconomic factors are presented in Tables 1.9a and 1.9b.

For the return-based factors we note several differences between the subsamples. Firstly, the model probabilities for the best model are generally higher for the first subsample except for the book-to-market and cashflow portfolios. Secondly, the probability of inclusion is higher for more factors during the second periods. The momentum factor obtains a substantially larger probability of inclusion in the later period. Finally, the evidence in favor of the FF three-factor model seems to be strongest over the first period. Using the informative prior does not change the results and they are therefore not reported.

When we consider the set of macroeconomic factors we note that the difference between the two periods is substantial. The selected factors for the two time periods are very different. In none of the portfolios are the selected

Table 1.8a: Probability of Inclusion, Return-Based Factors: Reference prior $g = 0.05$

Sample Period: 196307-198212

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
INT	0.001	0.000	0.000	0.000	0.005	0.000	0.000	0.000
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.058	0.002	0.027	0.024	0.999	0.999	0.000	0.598
RP	0.003	0.000	0.000	0.000	0.013	0.000	0.000	0.001
UTS(S)	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000
UTS(L)	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Sample Period: 198301-200312

Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
INT	0.362	0.700	0.000	0.000	0.001	0.000	0.000	0.000
MKT-RF	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
SMB	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
HML	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
UMD	0.941	0.999	0.004	0.286	0.821	0.826	0.988	0.637
RP	0.323	0.079	0.000	0.000	0.000	0.000	0.000	0.000
UTS(S)	0.088	0.004	0.000	0.000	0.000	0.000	0.000	0.000
UTS(L)	0.185	0.157	0.000	0.000	0.000	0.000	0.000	0.000

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

Table 1.8b: The Best Model, Return-Based Factors: Reference prior $g = 0.05$

	6 Size-B/M		25 Size-B/M		B/M		Size	
Factor	P1	P2	P1	P2	P1	P2	P1	P2
INT	0	1	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1
UMD	0	1	0	1	0	0	0	0
RP	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	1	0	0	0	0
Prob	0.935	0.322	0.998	0.7	0.973	0.995	0.976	0.714
	Industry		Dividend		Earning		Cashflow	
Factor	P1	P2	P1	P2	P1	P2	P1	P2
INT	0	0	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1
SMB	1	1	0	1	1	1	1	1
HML	1	1	1	1	1	1	1	1
UMD	1	1	1	1	0	1	1	1
RP	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0
Prob	0.981	0.820	0.820	0.826	0.999	0.988	0.597	0.637

P1 is 196307-198212 and P2 is 198301-200312. INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

Table 1.9a: Probability of Inclusion, Macroeconomic Factors: Reference prior $g = 0.05$

Sample Period: 196307-198212								
Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
MP	0.006	0.000	0.009	0.095	0.004	0.046	0.018	0.000
DEI	0.011	0.000	0.001	0.010	0.656	0.054	0.016	0.000
UI	0.043	0.000	0.001	0.020	0.003	0.338	0.179	0.000
CG	0.695	0.000	0.001	0.548	0.000	0.188	0.055	0.000
IC	0.031	0.000	0.003	0.003	0.000	0.046	0.025	0.000
OG	0.015	0.000	0.169	0.002	0.000	0.063	0.029	0.139
PSR	0.194	0.000	0.002	0.138	0.007	0.108	0.151	0.005
UNR	0.006	0.009	0.002	0.128	0.047	0.078	0.164	0.010
YP	0.004	0.990	0.807	0.053	0.392	0.074	0.360	0.881
Sample Period: 198301-200312								
Factor	6 Size-B/M	25 Size-B/M	B/M	Size	Industry	Dividend	Earning	Cashflow
MP	0.003	0.000	0.010	0.001	0.000	0.000	0.000	0.055
DEI	0.020	0.000	0.006	0.014	0.000	0.000	0.000	0.004
UI	0.004	0.000	0.002	0.024	0.000	0.222	0.001	0.187
CG	0.376	0.081	0.010	0.251	0.003	0.002	0.002	0.362
IC	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.001
OG	0.281	0.000	0.907	0.021	0.990	0.772	0.000	0.370
PSR	0.304	0.918	0.062	0.049	0.067	0.001	0.994	0.007
UNR	0.011	0.000	0.000	0.001	0.000	0.002	0.000	0.012
YP	0.071	0.000	0.000	0.638	0.008	0.000	0.000	0.002

P1 is 196307-198212 and P2 is 198301-200312. B/M = book-to-market; MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production.

Table 1.9b: The Best Model, Macroeconomic Factors: Reference prior $g = 0.05$

	6 Size-B/M		25 Size-B/M		B/M		Size	
Factor	P1	P2	P1	P2	P1	P2	P1	P2
MP	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	0	0	0
CG	1	1	0	0	0	0	1	0
IC	0	0	0	0	0	0	0	0
OG	0	0	0	0	0	1	0	0
PSR	0	0	0	1	0	0	0	0
UNR	0	0	0	0	0	0	0	0
YP	0	0	1	0	1	0	0	1
Prob	0.69	0.33	0.99	0.91	0.808	0.91	0.548	0.63
	Industry		Dividend		Earning		Cashflow	
Factor	P1	P2	P1	P2	P1	P2	P1	P2
MP	0	0	0	0	0	0	0	0
DEI	1	0	0	0	0	0	0	0
UI	0	0	1	0	0	0	0	0
CG	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0
OG	0	1	0	1	0	0	0	1
PSR	0	0	0	0	0	1	0	0
UNR	0	0	0	0	0	0	0	0
YP	0	0	0	0	1	00	1	0
Prob	0.546	0.92	0.33	0.77	0.36	0.99	0.84	0.36

P1 is 196307-198212 and P2 is 198301-200312. B/M = book-to-market; MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production.

factors found to be important in the first period, also included in the set of factors for the second period. One exception is the 6-size-B/M portfolios where the growth rate in real per capita consumption is selected in both periods. Hence, the selected factors seem to be very sensitive to the sample period under investigation.

Addressing the issue of prior sensitivity we first consider the choice of g, measuring the tightness or information content of the prior. Letting g take the values $\{1/T, 1/K^2, 0.05, 0.5\}$, we find some sensitivity to g. Tables 1.10a and 1.10b show the results for the return-based factor and Tables 1.10c and 1.10d show the results for the macroeconomic factors.

In general, as g increases, the prior is made more informative, the probability of inclusion increases and becomes substantial for several factors. This holds for both return-based and macroeconomic factors. One additional factor is usually selected when the prior is very informative, corresponding to $g = 0.5$. Furthermore, the model uncertainty increases with the value of g as shown in Tables 1.10b and 1.10d. The result that a more informative prior increases the probability of inclusion and the number of factors may seem counterintuitive. However, when the prior becomes more informative the posterior for the factor sensitivities will shrink to zero. Since this seems to be true for many of the factors when looking at the data, more factors will be found to be important. Overall, it seems like the selected factors are fairly insensitive to different values for g.

Finally, we consider the prior for the innovation variance, Σ, in the reference prior setup. Specifying a proper inverse Wishart prior as in Section 1.3.2 for the variances instead of the improper Jeffrey's prior leads to a well defined marginal likelihood and might thus be preferable. The results are, however, not affected in any substantial way by this change in the prior specification and are therefore not reported. The results are available from the authors on request.

Table 1.10a: Probability of Inclusion, Return-Based Factors: Reference prior with different g

Factor	6 Size-B/M				25 Size-B/M				B/M				Size			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
INT	0.01	0.03	0.04	0.22	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.04
MKT-RF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
SMB	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
HML	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
UMD	0.00	0.00	0.00	0.26	0.00	0.00	0.07	0.99	0.00	0.00	0.02	0.76	0.00	0.00	0.00	0.43
RP	0.00	0.03	0.04	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.03
UTS(S)	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.06
UTS(L)	0.33	0.93	0.92	0.89	0.00	0.00	0.25	0.98	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.14
Factor	Industry				Dividend				Earning				Cashflow			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
INT	0.00	0.00	0.42	0.79	0.00	0.00	0.00	0.31	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.11
MKT-RF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
SMB	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
HML	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
UMD	0.93	0.99	0.94	0.96	0.87	0.12	0.99	0.99	0.00	0.12	0.94	0.99	0.00	0.01	0.72	0.99
RP	0.00	0.00	0.05	0.62	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.31
UTS(S)	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.07
UTS(L)	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.03

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

Table 1.10b: The best model, Return-Based Factors: Reference prior with different g

Factor	6 Size-B/M				25 Size-B/M				B/M				Size			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
RP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0
Prob	0.657	0.929	0.906	0.293	0.999	0.999	0.644	0.984	0.999	0.999	0.973	0.618	0.999	0.999	0.999	0.415

Factor	Industry				Dividend				Earning				Cashflow			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
INT	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
MKT-RF	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
UMD	1	1	0	1	1	1	1	1	0	0	1	1	0	0	1	0
RP	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Prob	0.933	0.973	0.511	0.400	0.871	0.999	0.999	0.433	0.999	0.878	0.941	0.741	0.999	0.982	0.902	0.559

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long.

Table 1.10c: Probability of Inclusion, Macroeconomic Factors: Reference prior with different g

Factor	6 Size-B/M				25 Size-B/M				B/M				Size			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
MP	0.000	0.000	0.000	0.082	0.000	0.000	0.000	0.000	0.021	0.042	0.071	0.154	0.000	0.053	0.000	0.081
DEI	0.000	0.000	0.002	0.202	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.042	0.000	0.016	0.000	0.084
UI	0.000	0.000	0.006	0.290	0.000	0.000	0.000	0.013	0.020	0.021	0.018	0.079	0.002	0.792	0.002	0.162
CG	0.998	0.998	0.998	0.993	0.058	0.059	0.066	0.837	0.010	0.019	0.031	0.205	0.958	0.024	0.977	0.945
IC	0.000	0.000	0.001	0.245	0.000	0.000	0.000	0.568	0.002	0.003	0.006	0.087	0.000	0.007	0.000	0.020
OG	0.000	0.000	0.003	0.178	0.000	0.000	0.000	0.197	0.602	0.385	0.196	0.342	0.030	0.024	0.010	0.408
PSR	0.000	0.001	0.013	0.503	0.941	0.939	0.933	0.972	0.140	0.159	0.147	0.374	0.007	0.060	0.008	0.179
UNR	0.000	0.000	0.000	0.124	0.000	0.000	0.000	0.019	0.000	0.001	0.002	0.040	0.000	0.010	0.000	0.082
YP	0.000	0.000	0.008	0.340	0.000	0.000	0.000	0.903	0.200	0.363	0.521	0.586	0.000	0.009	0.001	0.258
Factor	Industry				Dividend				Earning				Cashflow			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K$	0.05	0.5
MP	0.000	0.000	0.000	0.400	0.000	0.000	0.000	0.121	0.000	0.000	0.000	0.024	0.040	0.046	0.056	0.253
DEI	0.002	0.003	0.014	0.902	0.000	0.000	0.000	0.062	0.000	0.000	0.000	0.012	0.000	0.000	0.000	0.060
UI	0.002	0.001	0.002	0.160	0.016	0.026	0.061	0.387	0.000	0.000	0.001	0.540	0.002	0.001	0.001	0.169
CG	0.000	0.000	0.000	0.078	0.000	0.000	0.000	0.115	0.000	0.000	0.000	0.026	0.000	0.000	0.001	0.144
IC	0.000	0.000	0.000	0.451	0.000	0.000	0.000	0.058	0.000	0.000	0.000	0.007	0.000	0.000	0.000	0.085
OG	0.071	0.084	0.359	0.989	0.983	0.973	0.937	0.802	0.000	0.000	0.000	0.038	0.009	0.004	0.002	0.199
PSR	0.035	0.031	0.119	0.982	0.000	0.000	0.001	0.275	0.999	0.999	0.999	0.994	0.000	0.001	0.001	0.207
UNR	0.000	0.000	0.000	0.152	0.000	0.000	0.000	0.046	0.000	0.000	0.000	0.075	0.913	0.909	0.894	0.831
YP	0.887	0.880	0.899	0.997	0.000	0.000	0.000	0.080	0.000	0.000	0.000	0.038	0.033	0.036	0.042	0.281

B/M = book-to-market; MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production.

Table 1.10d: The best model, Macroeconomic Factors: Reference prior with different g

Factor	6 Size-B/M				25 Size-B/M				B/M				Size			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
MP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CG	1	1	1	1	0	0	0	1	0	0	0	0	1	1	1	1
IC	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
OG	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
PSR	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
UNR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
YP	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0
Prob	0.998	0.998	0.964	0.090	0.941	0.939	0.933	0.358	0.603	0.385	0.681	0.521	0.178	0.978	0.976	0.212
Factor	Industry				Dividend				Earning				Cashflow			
g	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5	$1/T$	$1/K^2$	0.05	0.5
MP	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
DEI	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
CG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OG	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
PSR	0	0	0	1	0	0	0	0	1	1	1	1	0	0	0	0
UNR	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
YP	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Prob	0.888	0.880	0.577	0.174	0.983	0.973	0.937	0.269	0.999	0.998	0.998	0.425	0.913	0.909	0.894	0.178

B/M = book-to-market; MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production.

1.6 Summary and Conclusions

In this paper we use Bayesian techniques to select the factors in a general multifactor asset pricing model. From a given set of K factors we evaluate and rank all 2^K different pricing models by their posterior model probabilities. Two sets of factors are considered; the first set includes returns on a market portfolio of stocks and mimicking portfolios for the size, momentum, book-to-market, and factors related to the term-structure of interest rates and the second set of factors contains macroeconomic variables. The resulting pricing models are evaluated using eight different sets of portfolios.

In the first set of potential factors we find strong evidence that a general multifactor pricing model should include the market excess return, the size premium, and the value premium. The evidence in favor of the momentum factor is more sensitive to the sample used and to the selection of the test asset. Risk factors related to the bond market do not seem to be very important. It seems like the support for the three factor model of Fama and French (1992, 1993, 1996) is stronger when the test assets are portfolios formed on size and/or book-to-market. Furthermore, the evidence in favor of the three factor models is stronger during the first subperiod 196307 - 198212 and the evidence for the additional momentum factor can be traced to the later subperiod. Introducing a prior where we take into account the prior degree of confidence in an asset pricing model does not affect the selection of factors in any substantial way.

The interpretation of the momentum and the three factors of Fama and French as risk factors have caused a large debate in the finance literature. In the second set of factors we therefore consider macroeconomic variables. The model uncertainty is substantial and the factors selected depends on the test assets and the sample period. In general, only one factor is selected and we find some support that the growth rate in real per capita consumption, growth rate in personal savings rate and yearly growth rate in industrial production are important factors.

The identified factors are consistent with what others have found. However, we believe this study adds some interesting aspects concerning the evaluation of asset pricing models. Most importantly, by using a Bayesian approach we can easily address model uncertainty, which we found to be substantial.

Bibliography

BANZ, R. (1981): "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3–18.

BASU, S. (1977): "The Investment Performance of Common Stock in Relation to their Price to Earnings Ratios: A Test of the Efficient Market Hypothesis," *Journal of Finance*, 32, 663–682.

BERGER, J. O., AND L. R. PERICCHI (1998): "Objective Bayesian Methods for Model Selection," Discussion paper.

BLACK, J., M. JENSEN, AND M. SCHOLES (1972): "The Capital Asset Pricing Model: Some Empirical Tests," in *Studies in the Theory of Capital Markets*, ed. by M. Jensen, New York. Praeger.

BLACK, S. (1972): "Capital Market Equilibrium with Restricted Borrowing," *Journal of Business*, 45, 444–455.

BLUME, O., AND L. FRIEND (1973): "A New Look at the Capital Asset Pricing Model," *Journal of Finance*, 28, 19–33.

CAMPBELL, J. (1987): "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373–399.

CAMPBELL, J., AND R. SHILLER (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.

CARHART, M. M. (1997): "Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

CHEN, N., R. ROLL, AND S. ROSS (1986): "Economic Forces and the Stock Market," *Journal of Business*, 59, 383–403.

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*, 21, 255–289.

DEBONDT, W., AND R. THALER (1985): "Does the Stock Market Overreact?," *Journal of Finance*, 40, 739–805.

ERICSSON, J., AND A. GONZALEZ (2003): "Is Momentum Due to Data-Snooping?," Working Paper Series in Economics and Finance 536, Department of Economic Statistics Stockholm School of Economics.

FAMA, E., AND K. FRENCH (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

——— (1993): "Common Risk Factors in the Returns of Stock and Bonds," *Journal of Financial Economics*, 33, 3–56.

——— (1996): "Multifactor Explanations for Asset Pricing Anomalies," *Journal of Finance*, 51, 55–94.

FAMA, E., AND M. GIBBONS (1984): "A Comparison of Inflation Forecasts," *Journal of Financial Economics*, 13, 327–384.

FAMA, E., AND J. MACBETH (1973): "Risky Return and Equilibrium: Empirical Tests," *Journal of Financial Political Economy*, 71, 607–636.

FERNÁNDEZ, C., E. LEY, AND M. STEEL (2001): "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, 38, 381–427.

HALL, A. D., S. HWANG, AND E. S. SATCHELL (2002): "Using Bayesian Variable Selection Methods to Choose Style Factors in Global Stock Return Models," *Journal of Banking and Finance*, 26, 2301–2325.

JAGANNATHAN, R., AND Z. WANG (1996): "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3–53.

JEGADEESH, N., AND S. TITMAN (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 51, 65–91.

KANDEL, S., AND R. STAMBAUGH (1996): "On the Predictability of Stock Returns: An Asset Allocation Perspective," *Journal of Finance*, 51, 385424.

KOTHARI, S., AND J. SHANKEN (1997): "Book-to-Market, Dividend Yield and Expected Market Returns: A Time Series Analysis," *Journal of Financial Economics*, 44, 169–203.

LEHMANN, B., AND D. MODEST (1988): "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics*, 21, 213–254.

LINTNER, J. (1965): "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economic and Statistics*, 47, 13–37.

LO, A., AND A. C. MACKINLAY (1990): "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431–468.

MERTON, R. (1973): "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–887.

PASTOR, L., AND R. F. STAMBAUGH (1999): "Costs of equity capital and model mispricing," *Journal of Finance*, (54), 67–121.

——— (2000): "Comparing asset pricing models: An investment perspective," *Journal of Financial Economics*, (56), 335–381.

REYFMAN, A. (1997): "Labor Market Risk and Expected Asset Returns," Ph.D. thesis, University of Chicago.

ROSS, S. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

SCHWERT, G. (1990): "Stock Returns and Real Activity: A Century of Evidence," *Journal of Finance*, 45, 1237–1257.

SHANKEN, J. (1992): "On the Estimation of Beta-Pricing Models," *Review of Financial Studies*, 5, 425–442.

SHARPE, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425–442.

SMITH, M., AND R. KOHN (2000): "Nonparametric Seemingly Unrelated Regression," *Journal of Econometrics*, 98, 257–281.

VELU, R., AND G. ZHOU (1999): "Testing Multi-Beta Asset Pricing Models," *Journal of Empirical Finance*, 6, 219–241.

WHITELAW, R. (1997): "Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns," *Journal of Finance*, 41-2, 515–541.

ZELLNER, A. (1986): "On Assessing Prior Distributions and Bayesian Regression Analysis with G-Prior Distributions," in *Bayesian Inference and Decision Techniques - Essays in Honor of Bruno de Finetti*, ed. by P. K. Goel, and A. Zellner, pp. 223–243, Amsterdam. North Holland.

Chapter 2

Choosing Factors in a Multifactor Asset Pricing Model when Returns are Nonnormal

Acknowledgement: An earlier version of this paper was presented at the Nordic Econometric Meeting, Helsinki, May 2005 and the Forecasting Financial Markets, Marseilles, June 2005. We wish to thank participants for their comments, and any errors in the paper are ours alone.

2.1 Introduction

Ever since the work by Fama (1965) there is evidence that stock returns do not follow a normal distribution. Similar results have been presented by Affleck-Graves and McDonald (1989) and Richardson and Smith (1993). However, normality is still the working assumption in much of the empirical work in finance. The main reason for this is probably due to tractability with respect to estimation and for making statistical inference. Hence, many important findings in empirical finance are based on the normality assumption.

Most empirical work in the asset pricing literature starts with the assumption that returns are drawn from a multivariate normal distribution. There are however a few exceptions. Groenewold and Fraser (2002) use Australian data and examine whether standard tests of asset pricing models are sensitive to deviations from the assumptions that returns are identically, independently and normally distributed. They find that the test outcomes are generally robust. Tu and Zhou (2004) incorporate the uncertainty about the data generating process into the portfolio analysis using a Bayesian approach. The result shows that accounting for fat tails leads to nontrivial changes in both parameter estimates and optimal portfolio weights but that the normality assumption works well in evaluating portfolio performance for a mean-variance investor. In a capital asset pricing model in an international setting Harvey and Zhou (1993) adjust the multivariate test of efficiency to account for alternative distributional specifications: multivariate t and multivariate mixture normal. Although the p-values are generally lower, the basic inference is unchanged.

Recent papers that have addressed the normality assumption in asset pricing models have focused on the estimation and the testing of pricing models. However, none of the recent papers have examined the problem of selecting relevant factors in an asset pricing model when normality is relaxed. For example, Tu and Zhou (2004) assume that the investor has knowledge of a set of probability distributions for the returns that are possible candidates for the true data generating process but the factors in the asset pricing model are assumed to be known.

In this paper we consider the problem of selecting observable factors in a multifactor asset pricing model when the assumption of normally distributed returns is relaxed. More precisely, we assume that asset returns are multivariate Student-t distributed. Even if the Student-t distribution only adds one more parameter, this setup allows us to capture the well known fat tail property of asset returns. Furthermore, multivariate Student-t is a return distribution for which mean-variance analysis is consistent with expected utility maximization, making the choice theoretically appealing. From a set of K factors, Bayesian techniques are used to rank the 2^K possible models based on the posterior model probabilities. The factors used are based on theoretical considerations and previous empirical studies. The first set consists of stock- and bond-market factors and includes returns on a market portfolio of stocks and mimicking portfolios for the size, momentum, book-to-market, and term-structure factors in returns. The second category brings together models where the factors are macroeconomic variables.

The rest of the paper is organized as follows, in the next section we present the model. Section 2.3 introduces the prior and the posterior, and in Section 2.4 the Bayesian model selection procedure is described. Sections 2.5 and 2.6 contain the data and empirical results respectively and in Section 2.7 a conclusion is given.

2.2 The Model

In general, a multifactor pricing model states that the returns on different assets are explained by a set of common factors in a linear model. For the excess return on N assets, $\mathbf{r}$, we have the following model

$$E(\mathbf{r}) = \beta_1\varphi_1 + \beta_2\varphi_2 \tag{2.1}$$

where $E(\mathbf{r})$ is the expected excess return, $\varphi_j, j = 1, 2$ are vectors of factor risk premia and

$$\mathbf{r}_t = \mathbf{a} + \beta_1\mathbf{f}_{1t} + \beta_2\mathbf{f}_{2t} + \varepsilon_t \tag{2.2}$$

where $\mathbf{r}_t$ is a $N \times 1$ vector of excess returns at time t, $\mathbf{a}$ is a $N \times 1$ vector of intercepts, $\mathbf{f}_{1t}$ is a $K_1 \times 1$ vector of general factors and $\mathbf{f}_{2t}$ is a $K_2 \times 1$ vector of factors that are portfolio returns and ε_t is a $N \times 1$ random error vector. The matrices β_1 and β_2 are factor sensitivities with dimension $N \times K_1$ and $N \times K_2$, respectively.

In many applications of the normal linear asset pricing model there is evidence that the probability of an unusually large or small value of the outcome $\mathbf{r}_t$ is substantially greater than indicated by a Gaussian distribution. This is a well documented phenomenon in the case of financial asset returns. Therefore, we assume that the errors in (2.2) are multivariate Student-t distributed with the density function

$$p(\varepsilon_t|\nu, \Sigma) = \frac{\Gamma\left(\frac{\nu+N}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\pi\nu)^{N/2}|\Sigma|^{1/2}}\left[1 + \frac{1}{\nu}\varepsilon_t'\Sigma^{-1}\varepsilon_t\right]^{-(\nu+N)/2}.$$

Since the Student-likelihood is difficult to work with, we note that by data augmentation we can express the model in (2.2) with multivariate-t errors in a more convenient form (Geweke (1993))

$$\mathbf{r}_t|\mathbf{B},\Sigma,\lambda_t \sim N\left(\mathbf{B}'\mathbf{f}_t,\Sigma/\lambda_t\right)$$

where λ_t is Gamma distributed with unitary mean[1], $\mathbf{B} = \begin{bmatrix} \mathbf{a}' & \beta_1' & \beta_2' \end{bmatrix}'$ and $\mathbf{f}_t = \begin{bmatrix} 1 & \mathbf{f}_{1t}' & \mathbf{f}_{2t}' \end{bmatrix}'$. By stacking $\mathbf{r}_t$, $\mathbf{f}_t$ and ε_t row-wise, (2.2) can be written as

$$\mathbf{R} = \mathbf{fB} + \mathbf{E}$$

where $E|\mathbf{\Psi} \sim MN_{T\times N}(0,\Sigma,\mathbf{\Psi}^{-1})$, $\mathbf{\Psi} = diag\left\{\lambda_1,...,\lambda_T\right\}$ and MN denotes the matrix variate normal distribution. The conditional likelihood is then given by

$$\begin{aligned} L(\mathbf{R}|\mathbf{B},\Sigma,\lambda) &= \left(\frac{1}{2\pi}\right)^{TN/2} |\Sigma|^{-T/2} |\mathbf{\Psi}|^{N/2} \\ &\times \exp\left\{-\frac{1}{2} tr\left[\Sigma^{-1}\left(\mathbf{R}-\mathbf{fB}\right)'\mathbf{\Psi}\left(\mathbf{R}-\mathbf{fB}\right)\right]\right\} \\ &= \left(\frac{1}{2\pi}\right)^{TN/2} |\Sigma|^{-T/2} |\mathbf{\Psi}|^{N/2} \exp\left\{-\frac{1}{2}\left(\mathbf{y}-\mathbf{Z}\beta\right)'\mathbf{\Phi}^{-1}\left(\mathbf{y}-\mathbf{Z}\beta\right)\right\} \end{aligned}$$

where $\mathbf{y} = vec(\mathbf{R})$, $\mathbf{Z} = \mathbf{I}_N \otimes \mathbf{f}$, $\beta = vec(\mathbf{B})$, $\mathbf{\Phi} = \mathbf{\Sigma} \otimes \mathbf{\Psi}^{-1}$ and $\otimes$ denotes the Kronecker product.

2.3 The Prior and the Posterior

In principle, we can choose any prior for the parameters since we need to approximate the posterior numerically or analytically anyway. However, these priors must be informative since improper noninformative priors yield indeterminate marginal likelihoods and Bayes factors. A natural choice is the following

$$\begin{aligned} &\mathbf{B} \sim MN_{K\times N}\left(\mathbf{B}_0, I, M_0^{-1}\right) \\ &\Sigma \sim iW(\mathbf{S}_0, v_0) \\ &v \sim Ga(1,\theta) \end{aligned}$$

---

[1] That is, $\lambda_t \sim Ga(v/2, 2/v)$ for $t = 1,...,T$.

where *iW* and *Ga* denote the inverse Wishart and the Gamma distribution respectively.

Choosing the prior hyperparameters can be difficult in the absence of substantive prior information. Reflecting the lack of consensus in the finance literature about the identity of the factors the prior mean of **B** conditional on specific model is zero and for the prior covariance matrix we follow the ideas of Fernández, Ley, and Steel (2001), Hall, Hwang, and Satchell (2002) and Smith and Kohn (2000) and use the g-prior of Zellner (1986). Thus,

$$M_0 = g\mathbf{f}'\mathbf{f}$$

where $g > 0$. The parameter g is chosen such that the prior is made relatively uninformative. Note that the prior for **B** with the g-prior is equivalent to $\beta \sim N_{KN}(\beta_0, (\mathbf{Z}'\mathbf{Z})^{-1}/g)$, where $\beta_0 = vec(\mathbf{B_0})$ and $\mathbf{Z} = \mathbf{I}_N \otimes \mathbf{f}$.

The g-prior is particularly suitable for a variable selection exercise. Let M_J be the model with all potential regressors included and for model j partition $\beta = (\beta_j, \beta_{-j})$ and $Z = (Z_j, Z_{-j})$ conformably where Z_j are the variables included in model j. It is then easy to show that conditioning on $\beta_{-j} = \beta_{0,-j}$ in the prior for the full model yields the prior for β_j in the subset model. That is, $\pi(\beta_j|\beta_{-j} = \beta_{0,-j}, M_J) = \pi(\beta_j|M_j)$. Recalling that $\beta_0 = 0$ it is clear that this provides a consistent set of prior distributions.

The hyperparameters for Σ are more difficult to choose. We follow Kandel and Stambaugh (1996) and use statistics from the actual sample, which are given by

$$\begin{aligned}\bar{\mathbf{r}} =& \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t \\ \hat{\mathbf{V}} =& \frac{1}{T}\sum_{t=1}^{T}(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_t - \bar{\mathbf{r}})'.\end{aligned}$$

Then the hyperparameters in the inverted Wishart distribution are specified as

$$\begin{aligned}\mathbf{S}_0 =& s^2\mathbf{I}_N \\ v_0 =& N + 2.\end{aligned}$$

where where s^2 is the average of the diagonal elements of the sample covariance matrix $\hat{\mathbf{V}}$ and $E(\Sigma) = s^2\mathbf{I}_N$.

The gamma prior for the degrees of freedom is specified with $\theta = 25$, which allocates substantial probability to both fat tailed error distributions with $v_0 < 10$ and approximate normal error distributions with $v_0 > 40$.

The posterior is proportional to the likelihood times the priors,

$$\pi(\beta,\Sigma,\lambda,v|\mathbf{R}) \propto L(\mathbf{R}|\mathbf{B},\Sigma,\lambda)\pi(\beta)\pi(\Sigma)\pi(\lambda)\pi(v)$$

where $\lambda = \{\lambda_1, ..., \lambda_T\}$. It is clear that the joint posterior density does not have the form of any known density and therefore, cannot be used in a simple way for posterior inference. However, some of the conditionals of the posterior are simple. It can be verified that the full conditionals are given by the following distributions.

1. The conditional posterior distribution of β is multivariate normal,

   $$\beta|\Sigma,\lambda,v,\mathbf{R} \sim N_{KN}\left(\beta_1, M_1^{-1}\right). \tag{2.3}$$

   where $M_1 = \tilde{M}_0 + \mathbf{Z}'\mathbf{\Phi}^{-1}\mathbf{Z}$, $\beta_1 = M_1^{-1}\left(\tilde{M}_0\beta_0 + \mathbf{Z}\mathbf{\Phi}^{-1}\mathbf{Z}\widehat{\beta}\right)$, $\widehat{\beta} = (\mathbf{Z}'\mathbf{\Phi}^{-1}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{\Phi}^{-1}\mathbf{y}$ and $\tilde{M}_0 = \mathbf{I}_N \otimes M_0$.

2. The conditional posterior distribution of Σ is an inverse Wishart,

   $$\Sigma|\beta,\lambda,v,\mathbf{R} \sim iW\left(\mathbf{E}'\mathbf{\Psi}\mathbf{E} + \mathbf{S}_0, T + v_0\right). \tag{2.4}$$

3. The conditional density of λ_t, $t = 1, ..., T$ is Gamma

   $$\lambda_t|\beta,\Sigma,v,\mathbf{R} \sim Ga\left(\frac{N+v}{2}, \frac{2}{\varepsilon_t'\Sigma^{-1}\varepsilon_t + v}\right). \tag{2.5}$$

4. The conditional for v does not take the form of a known density, but the kernel of the conditional posterior density for v is

   $$\pi(v|\beta,\Sigma,\lambda,\mathbf{R} \propto \left(\frac{v}{2}\right)^{\frac{Tv}{2}} \Gamma\left(\frac{v}{2}\right)^{-T} \exp\{-v\eta\} \tag{2.6}$$

   where $\eta = \left[\frac{1}{2}\sum_{t=1}^{T}\lambda_t - \ln\lambda_t\right] + \frac{1}{\theta}$. To generate draws from (2.6) a Metropolis-Hastings algorithm can be used.

The posterior simulator is a Metropolis-within-Gibbs algorithm with draws of β and Σ taken from (2.3) and (2.4) respectively and draws from λ are taken using (2.5). The degrees of freedom, v, is updated in a Metropolis-Hastings step with a normal random walk proposal. Candidate draws of v, which are less or equal to zero, have the acceptance probability set to zero.

2.4 Bayesian Model Selection

Consider the problem of comparing a collection of models $\{M_l, l = 1, ..., L\}$ that reflect competing hypotheses about the data, y. The basis for hypothesis testing and model selection in the Bayesian framework is the marginal likelihood, which measures how well the model (and the prior) fits the data. Given the prior distribution for the parameters, $p(\theta|M_l)$, the marginal likelihood is

$$m(y|M_l) = \int L(y|\theta, M_l)p(\theta|M_l)d\theta$$

where $L(y|\theta, M_l)$ is the likelihood. Model comparison can be conducted through the use of Bayes factors. The Bayes factor for M_i versus M_j is given by

$$B_{ij} = \frac{m(y|M_i)}{m(y|M_j)} = \frac{\int L(y|\theta, M_i)p(\theta|M_i)d\theta}{\int L(y|\theta, M_j)p(\theta|M_j)d\theta}.$$

and measures how much our belief in M_i relative to M_j has changed after viewing the data. If prior probabilities $P(M_l)$, $l = 1, ..., L$, of the models are available, the Bayes factor can be used to compute the posterior model probabilities

$$P(M_i|\mathbf{y}) = \frac{m(y|M_i)P(M_i)}{\sum_{l=1}^{L} m(y|M_l)P(M_l)} = \left[\sum_{j=1}^{L} \frac{P(M_j)}{P(M_i)} B_{ji}\right]^{-1}.$$

The marginal likelihood can only be calculated analytically in special cases. In other cases, numerical or asymptotic methods are needed. One possibility is to approximate $m(y|M_l)$ by Laplace's method (Kass, Tierney, and Kadane (1988); Tierney, Kass, and Kadane (1989)). In the case where the models contain a relative small number of parameters, the Laplace approach can provide an excellent approximation. In our case, we have a rather large number of parameters and numerical optimization is needed to obtain the posterior mode $\bar{\theta}$ of $\ln L(y|\theta, M_l) + \ln p(\theta|M_l)$ and its inverse Hessian, which can be difficult even if the first and second derivatives are used in the optimization routine. Instead we will rely on the Markov Chain Monte Carlo (MCMC) methods. There are at least two methods that are straightforward to implement. The first method for marginal likelihood estimation is outlined in Chib (1995) and extended in Chib and Jeliazkov (2001). The second approach is the Savage-Dickey density ratio proposed by Dickey (1971) or the generalized Savage-Dickey density ratio proposed by Verdinelli and Wasserman (1995). Since Chib's method needs several MCMC runs for each competing model where the Savage-Dickey density ratio only needs one, we will use the latter method, which is presented below.

2.4.1 The Savage-Dickey Density Ratio

The Savage-Dickey density ratio is a simple method for calculating the Bayes factor for nested models. Suppose for instance, that the unrestricted model, M_2, has the parameters $\theta = (\omega, \psi)$ with the likelihood $L(y|\theta, M_2)$ and prior $\pi(\theta|M_2)$. The restricted model, M_1, has the restriction $\omega = \omega_0$ with $L(y|\psi, M_1)$ and $\pi(\psi|M_1)$ as the corresponding likelihood and prior. Restrictions of the form $R\omega = r$ are a simple extension. Dickey (1971) showed that if the priors in the two models satisfy

$$\pi(\psi|\omega = \omega_0, M_2) = \pi(\psi|M_1) \tag{2.7}$$

then the Bayes factor comparing M_1 to M_2 has the following form

$$B_{12} = \frac{\pi(\omega = \omega_0|y, M_2)}{\pi(\omega = \omega_0|M_2)} \tag{2.8}$$

where $\pi(\omega = \omega_0|y, M_2)$ and $\pi(\omega = \omega_0|M_2)$ are the marginal posterior and prior, i.e.

$$\pi(\omega|y, M_2) = \int \pi(\omega, \psi|y, M_2)d\psi.$$

Note that the condition given by (2.7) on the prior is sensible in practice. For example, in most cases, it is reasonable to use the same prior for parameters which are common in competing models. For such a prior, the condition in (2.7) is fulfilled. However, (2.7) is a much weaker condition since the prior for ψ in the restricted and the unrestricted model must be the same only at $\omega = \omega_0$. In the case where the condition on the prior is not satisfied the generalized Savage-Dickey density ratio proposed by Verdinelli and Wasserman (1995) can be used.

In our setup ω corresponds to β_{-j} and ψ corresponds to $(\beta_j, \Sigma, \lambda, v)$. The different models we want to compare contain a different number of factors, which can be imposed by restricting the elements corresponding to β_{-j} to zero in the model containing all factors. That is, the restricted model imposes $R\beta = \mathbf{0}$ to M_2 where R is a $NK \times Nl$ matrix of zeros and ones with l number of coefficients restricted to be equal to zero. For example, if the restricted models exclude the first factor then R is given by

$$R = \mathbf{I}_N \otimes e$$

where e' is a $k \times 1$ vector with the first element equal to one.

The denominator of (2.8), $\pi(R\beta = \mathbf{0}|M_2)$ is easily calculated since the marginal prior for β is normal. The numerator is more difficult to evaluate. However, using the output from the sampler $\pi(R\beta = \mathbf{0}|y, M_2)$ can be estimated. Assume that we have generated a sample $\left\{\beta^{(i)}, \Sigma^{(i)}, \lambda^{(i)}, v^{(i)}\right\}_{s=1}^{S}$ from the posterior. Since $\pi(\beta|y,\Sigma, \lambda, v, M_2)$ is in a closed form we estimate the marginal posterior by averaging the full conditionals posterior over the MCMC output, $\pi(R\beta = \mathbf{0}|y, M_2)$ by

$$\hat{\pi}(R\beta = \mathbf{0}|y, M_2) = \frac{1}{S}\sum_{s=1}^{S} \pi(R\beta = \mathbf{0}|y,\Sigma^{(s)}, \lambda^{(s)}, v^{(s)}, M_2).$$

2.5 The Data

Much of the empirical work on asset pricing has been conducted on US data. In this paper data from both the US and Sweden will be considered.

2.5.1 US Data

The US data in this study contains monthly observations on stock excess returns and a set of factors spanning the period July 1963 to December 2003. Asset pricing models are generally evaluated using portfolio returns and this paper is no exception. Returns on portfolios, market return, size premium, value premium and momentum were kindly provided by Kenneth French.[2]

Test Assets

The test asset consists of ten size-sorted portfolios, ten book-to-market sorted portfolios and ten industry portfolios. Furthermore, we consider ten portfolios sorted by cashflow, dividend and earnings respectively. The portfolios include all NYSE, AMEX, and NASDAQ stocks. Tables A.1a and A.1b in Appendix A contain summary statistics for the different portfolios. The results shows a widespread departure from normality in the returns.

Factors

The factors can be divided into two groups. The first set is stock- and bond-market factors and includes returns on a market portfolio of stocks and mim-

---

[2] A description of the data obtained from Kenneth French can be found at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

icking portfolios for the size, momentum, book-to-market and term-structure factors in returns. The second set contains macroeconomic factors.

The stock-market factors included are the market excess return (MKT-RF), size premium (SMB), value premium (HML) and a momentum factor (UMD). The first three factors correspond to the three factor model of Fama and French (1993). Adding the momentum factor corresponds to the model of Carhart (1997).

The bond-market factors consist of the Credit risk spread (RP), the difference between the yields of Moody's Baa and the yields of Moddy's Aaa rated bonds. This is a state variable that measures changes in the risk of corporate bonds. Proxies for unexpected change in interest rates are the difference in the annualized yield of ten-year and one-year Treasuries (UTS(L)), and the difference between the one-year Treasuries and the Federal Funds rate (UTS(S)).

The macroeconomic factors are monthly (MP) and yearly (YP) growth rate in industrial production, unanticipated inflation (UI), the change in expected inflation[3] (DEI), growth rate in real per capita personal consumption (CG) and the monthly change in the oil price (OG). In addition to these factors the following were added: growth rate in real per capita disposable income (IC), growth rate in the personal savings rate (PSR) and growth rate in the unemployment rate (UNR).

2.5.2 Swedish Data

The data covers the period January 1979 to June 2003 and consist of all stocks listed on the Stockholm Stock Exchange. The data is collected from the database "Trust". Information on accounting data is collected from the firm's annual statements and the data for macroeconomic variables are from the database Ecowin and Reuters. Due to data availability the test assets and the set of factors differ from the US data.

Test Assets

The test assets are portfolios formed on book value to market value, size, cashflow and dividends. Size is measured by the market value, price per share times shares outstanding, and the book value is the total value of stockholders equity. The book value, dividend and cashflow used to form a portfolio in June of year t are from the fiscal year ending in calender year $t-1$. The market value

---

[3] The inflation variables are constructed by the procedure in Fama and Gibbons (1984).

Table 2.1: The set of potential factors: US data

Symbol	Variable
MKT-RF	Market excess returns
SMB	Size premium
HML	Value premium
UMD	Momentum premium
RP	The credit risk spread
UTS(S)	Term spread (short)
UTS(L)	Term spread (long)
MP	Monthly growth rate in industrial production
YP	Yearly growth rate in industrial production
CG	Monthly growth rate in consumption
IC	Monthly growth rate in income
UI	Unanticipated inflation
DEI	Change in expected inflation
OG	Monthly growth rate in oil price
PSR	Monthly growth rate in private savings
UNR	Monthly growth rate in unemployment rate

used to form size portfolios in June of year t is the market value at the end of June of year t. The summary statistics in Tables A.1c and A.1d in Appendix A show that deviation from normality in the returns is very substantial.

Factors

Many of the factors for Sweden are the same as for the US data. However, some specific factors have been added that should be important for Sweden as a small open economy. More specifically, the movement in the exchange rate is added which is a proxy for the relative competitive strength of the Swedish economy. Furthermore, we take the opportunity to calculate factor mimicking portfolios for all non-return based factors since we have access to all individual stocks. First, we present the potential factors and in the next section we describe how the mimicking portfolios are constructed.

Firstly, there is the market excess return, which is the difference between the value-weighted return on all stocks and the three month Treasury bill rate. Fundamental factors are represented by two commonly used firm characteristics. The first one is the ratio of book value to market value (BM) and the second one is size (SIZE). The technical factor, (UMD), is the past stock return beginning seven months before the formation period and ending one month before it. This factor is supposed to capture the momentum anomaly. The factor related to the term structure is the difference between the five-year Treasuries and the three month Treasury bill (SLOPE).

The macroeconomic factors included are the monthly growth rates in industrial production (MP), consumption (CON), disposable income (INC) and unemployment rate (UNR). The consumption and income data was disaggregated from quarterly to monthly frequency by the method of Boot, Feibes, and Lisman (1967). The change in the USD/SEK exchange rate is a proxy for the overall variation in the currency market. Finally, we include the monthly change in expected inflation (DEI) and unanticipated inflation (UI) as potential factors. We fit an ARMA model to the monthly change in the CPI and the forecast from the model serves as expected inflation and the unanticipated is the forecast error.

Constructing Factor Mimicking Portfolios

We generally follow Chan, Karceski, and Lakonishok (1998) in constructing the portfolios. A factor mimicking portfolio is constructed by taking a long position in the portfolio with high loading and a short position in the portfolio with low loading on the factor.

For the accounting-based factors, book-to-market (BM), and size (ME), the procedure is as follows. At the end of June of each year t, stocks are sorted by a particular attribute (BM or ME) and allocated to a portfolio based on their ranks. Five portfolios are formed so stocks with the lowest and highest value of the attribute are assigned to portfolio 1 and 5 respectively. Equally weighted returns are then calculated from July to the following June. The mimicking portfolio return for the factor is then calculated each month as the difference between the return on the highest-ranked and the lowest-ranked portfolio. The BM ratio in the formation period of year t is the book value for the fiscal year ending in the calendar year $t-1$ divided by the market value at the end of December of year $t-1$. The ME in the formation period of year t is the market value at the end of December of year $t-1$. The momentum factor is constructed in a similar way except that we reform the portfolios every six months.

The macroeconomic factors are not expressed in returns and since a model with factor-mimicking portfolios will almost always outperform a model with real economic factors it is useful to construct factor mimicking portfolios for these factors as well. In this case, the relevant attribute is a stocks loading on the factor. To estimate the loading for each firm we regress the excess returns

Table 2.2: The set of potential factors: Swedish data

Symbol	Variable
RM-RF	Market excess returns
SIZE	Size premium
BM	Book-to-market ratio
UMD	Momentum premium
SLOPE	Slope of the yield curve
MP	Monthly growth rate in industrial production
CON	Monthly growth rate in consumption
INC	Monthly growth rate in income
UI	Unanticipated inflation
DEI	Change in expected inflation
USD	Change in exchange rate SEK/USD

of the stocks on the factor using the most recent past 24 months of data before the portfolio formation period. The regression coefficient is then the attribute on which stocks are ranked and assigned to portfolios. The procedure to form portfolios is then the same as for the accounting based factors.

2.6 Empirical Results

In the prior settings we still need to specify the parameter g. The results presented here are based on $g = 0.05$. The variance of the proposal density in the MCMC is calibrated until a value is found which yields reasonable acceptance probability. In our case, the acceptance probabilities are around 0.45. Then a final long run of 60000 replications, with 30000 burn-in replications discarded, is taken.

First, we will examine the US data where the return-based and non-return-based factors are analyzed separately and, thereafter, we will analysis the Swedish data.

2.6.1 US Data

In the case of only return-based factors, the asset pricing theory implies that the intercept or misspricing is zero. Including the intercept in the set of potential factors leads to a simple test of this aspect of the pricing model. This results in 8 factors and $2^8 = 256$ models where 128 of them are potential factor pricing models which is the number of models without intercept.

One major advantage of the Bayesian approach is that model uncertainty is easily quantified. In Table 2.3 we present the three best models with the highest posterior model probabilities, represented by combinations of zeros and ones, where one indicates that a specific factor is included in the model.

Starting with the size and book-to-market portfolios the best model has a posterior model probability of 0.90 and 0.66 respectively. The factor model includes the size and value premiums, and the market excess return. This is consistent with the three factor model of Fama and French (1993). The best model for the book-to-market portfolios also includes a term spread as one additional factor. Note that the best model for the size portfolios has a higher posterior model probability than the best model for the book-to-market portfolios. Hence, the model uncertainty is higher for the book-to-market sorted portfolios. Furthermore, it seems like the uncertainty is over the inclusion of the intercept.

Next we consider the case when the investment universe consists of industry portfolios. The best model has a posterior model probability of 0.82 and all factors are included except the two term spreads. The major difference from the results when stocks are sorted by size and book-to-market portfolios concerns the intercept. In the industry portfolios, the intercept is included in all top three models. Hence, the theoretical property of an asset pricing model is not fulfilled for the selected model when the industry portfolios are used as the investment universe. In addition, we found support for the momentum factor, which is included in the two best models.

Finally, we turn to the results when portfolios are formed on cashflow, dividends and earnings. The best model in all three sets of portfolios contains the three-factor model of Fama and French and the momentum factor. The posterior model probabilities are 0.72, 0.67 and 0.98 respectively.

So far, we have only considered return-based factors. A major criticism of these type of factors, such as the size, value and momentum factor, is their interpretation. It is not clear what kind of economic risk these are proxies for. Therefore, it is interesting to investigate macroeconomic factors directly in an asset pricing context. The drawback is that the implication of a zero intercept does not hold any more. Consequently, we always include the intercept in the model when selecting macroeconomic factors. Furthermore, the market excess return is also always included.

Table 2.4 reports the result for the macroeconomic factors. The results indicate that the model uncertainty is generally higher than for the return-based factors and the selected factors differ widely for the test assets under investigation. Typically only one factor shows up in the best models. The

Table 2.3: Three best models: Return-based factors US Data

Factors	Size			Book-to-Market			Industry		
INT	0	0	1	0	1	0	1	1	1
MKT-RF	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	0	0	1	1	0
RP	0	0	0	0	0	0	1	1	1
UTS(S)	0	1	0	0	0	0	0	1	0
UTS(L)	0	0	0	1	0	0	0	0	0
Prob	0.898	0.058	0.029	0.661	0.113	0.088	0.819	0.059	0.046

Factors	Cashflow			Dividend			Earning		
INT	0	0	0	0	1	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1
UMD	1	1	0	1	1	1	1	1	1
RP	0	1	0	0	0	1	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	1
UTS(L)	0	0	0	0	0	0	0	1	0
Prob	0.723	0.118	0.106	0.670	0.261	0.050	0.979	0.010	0.007

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long. 1 indicates inclusion and 0 indicates exclusion.

Table 2.4: Three best models:Macroeconomic factors US Data

Factors	Size			Book-to-Market			Industry		
MP	0	0	0	0	0	0	1	1	0
DEI	0	0	0	0	1	0	1	1	1
UI	0	0	1	0	0	0	0	0	0
CG	1	0	0	0	0	1	0	0	0
IC	0	0	0	0	0	0	0	1	1
OG	0	0	0	0	0	0	1	1	1
PSR	0	1	0	0	0	0	1	1	1
UNR	0	0	0	0	0	0	0	0	0
YP	0	0	0	1	0	0	1	1	1
Prob	0.839	0.087	0.042	0.91	0.017	0.01	0.169	0.166	0.135

Factors	Cashflow			Dividend			Earning		
MP	0	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	1	0	0	0
UI	0	0	0	1	0	1	0	1	0
CG	0	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0	0
OG	0	0	0	0	1	0	0	0	0
PSR	0	0	0	0	0	0	1	1	1
UNR	1	0	1	0	0	0	0	0	1
YP	0	1	1	0	0	0	0	0	0
Prob	0.442	0.366	0.068	0.59	0.295	0.029	0.675	0.287	0.009

MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production. 1 indicates inclusion and 0 indicates exclusion.

best model for the size portfolios only contains the growth rate in real per capita consumption and the best model for the book-to-market portfolios only contains the yearly growth rate in industrial production. The posterior model probability for the best model is 0.83 and 0.91 for the size and book-to-market portfolios respectively. When the industry portfolios are considered, more factors are included. However, the model uncertainty is very substantial since the posterior probability is evenly spread among the top three models. The results are also very mixed when we consider portfolios formed on cashflow, dividends and earnings.

Overall, the selection of macroeconomic factors depend heavily on the choice of test assets. All six test portfolios generate different asset pricing models.

Until now we have focused on the selection of factors. Another important issue is related to normality. Table 2.5 contains posterior results for the degrees of freedom, v, for the full model. The posterior mean when return-based factors are considered lies between 6.4-9.0 and the standard deviations are around 0.6-1.0. Hence, the result indicates a substantial deviation from normality. The posterior mean for the degrees of freedom when macroeconomic factors are selected is higher. The mean is around 14 for all portfolios except for the dividend portfolios where the mean is 8.1. The corresponding standard deviations lie between 0.92 and 1.35. Even if the posterior means are higher, we still find deviation from normality.

In Figures A.1 and A.2, in the Appendix, the estimated posterior density for v is shown. The Figures indicate that $\pi(v|\mathbf{R})$ has a shape which is slightly skewed and it confirms that all of the posterior probability is allocated to small values for the degrees of freedom parameter.

An important practical issue involves the assessment of the convergence of the sampling process used to estimate parameters. The property of the Markov chain for the degrees of freedom parameter is shown in Table 2.5 and Figure A.1 and A.2.[4] Estimates of the numerical standard error and relative numerical efficiency (RNE) using the spectral estimator are presented in the two last columns in Table 2.5. The RNEs for the return-based factors indicate that we need almost 10 times as many draws from the sampler as when sampling directly from the posterior. The high RNEs can be explained by the sample autocorrelation functions displayed in Figure A.1. The autocorrelation dies out first after 25 lags which leads to a reduction in efficiency. The first graph in the Figures displays Geweke diagnostics. Convergence implies that

---

[4] The diagnostics for the other parameters indicate speedy convergence and are not reported but are available on request.

Table 2.5: Posterior simulation results for the degrees of freedom US Data

Return-based factors

			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	7.3321	0.7428	0.0099	10.5948
Book-to-Market	9.0411	1.0072	0.0145	12.4305
Industry	6.4244	0.613	0.0077	9.5812
Cashflow	6.7753	0.6691	0.0087	10.1933
Dividend	6.9665	0.6805	0.0089	10.2521
Earning	7.0921	0.6986	0.0091	10.2651

Macroeconomic factors

			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	14.5431	1.3386	0.0068	1.5443
Book-to-Market	13.1017	1.2355	0.0067	1.7784
Industry	14.0983	1.1403	0.0067	2.0763
Cashflow	14.3115	1.2632	0.0069	1.7787
Dividend	8.0792	0.917	0.0047	1.5435
Earning	14.8865	1.3459	0.0073	1.7736

the calculated statistics should be within the two lines. A few more significant test statistics that we would like, especially for the cashflow portfolios, but the overall judgment is that the chains seem to have converged.

The Markov chain property for the degrees of freedom when macroeconomic factors are considered is different from the return-based factors. Firstly, the RNEs are much lower. All of them are around 1.54-2.08. Secondly, the autocorrelation dies out much faster. Finally, the Geweke diagnostics indicate that the chains have converged.

Sensitivity Analysis

As in any empirical study, the results can be sensitive to the assumptions and choices we make. The preceding section gave some results on the sensitivity to portfolio composition. In this section we address the sensitivity with respect to the sample period by considering two subsamples, 196307 - 198212 and 198301 - 200312. Furthermore, the robustness regarding the prior configuration is investigated.

The result of the model selection result is displayed in Tables 2.6a and 2.6b. The results for the return-based factors in Table 2.6a indicate that the model uncertainty is higher for the second period. However, the factors selected in the first period are also included in the best model during the second period. Hence, the results for the return-based factors are rather insensitive to the selected time period under investigation. It seems to be the other way around for the macroeconomic factors. Table 2.6b shows that different factors are selected for the two time periods. Again, usually only one factor is found to be important, except for the industry portfolios.

The posterior results for the degrees of freedom, v, for the two subperiods are reported in Table 2.6c. In general, the posterior mean for the degrees of freedom is almost twice as high in the first period, compared with the second period. Hence, the deviation from normality is more substantial during the later period, 198301-200312. Estimates of the numerical standard error and relative numerical efficiency (RNE) using the spectral estimator are also presented in the two last columns. The RNEs for the return-based factors are again higher than for the macroeconomic factors. Figures of the estimated posterior densities for v and the convergence diagnostics are not reported but are available from the authors on request.

Addressing the issue of prior sensitivity we first consider the choice of g, measuring the tightness or information content of the prior for the factor sensitivities. Letting g take the values $\{1/T, 1/K^2, 0.05\}$, we find that the

Table 2.6a: Three best models: Return-based factors US Data

Factors	Size		Book-to-Market		Industry	
	P1	P2	P1	P2	P1	P2
INT	0	0	0	1	1	1
MKT-RF	1	1	1	1	1	1
SMB	1	1	1	1	1	1
HML	1	1	1	1	1	1
UMD	0	0	1	0	1	1
RP	0	0	0	0	1	0
UTS(S)	0	1	0	0	0	1
UTS(L)	0	0	0	0	0	0
Prob	0.862	0.552	0.67	0.589	0.812	0.376

Factors	Cashflow		Dividend		Earning	
	P1	P2	P1	P2	P1	P2
INT	1	0	0	0	0	0
MKT-RF	1	1	1	1	1	1
SMB	1	1	1	1	1	1
HML	1	1	1	1	1	1
UMD	1	1	1	1	0	1
RP	1	0	0	0	0	0
UTS(S)	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0
Prob	0.813	0.766	0.837	0.744	0.6	0.981

P1 is 196307-198212 and P2 is 198301-200312. INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long. 1 indicates inclusion and 0 indicates exclusion.

Table 2.6b: Three best models:Macroeconomic factors US Data

Factors	Size		Book-to-Market		Industry	
	P1	P2	P1	P2	P1	P2
MP	0	0	0	0	1	0
DEI	0	0	0	0	1	0
UI	0	0	0	0	0	0
CG	0	1	0	0	0	0
IC	0	0	0	0	0	0
OG	0	0	0	1	1	1
PSR	0	0	0	0	1	1
UNR	1	0	0	0	0	0
YP	0	0	1	0	1	1
Prob	0.356	0.536	0.971	0.56	0.103	0.516

Factors	Cashflow		Dividend		Earning	
	P1	P2	P1	P2	P1	P2
MP	0	0	0	0	0	0
DEI	0	0	1	0	0	0
UI	0	0	0	0	0	0
CG	0	0	0	0	0	0
IC	0	1	0	0	0	0
OG	1	0	0	1	0	0
PSR	0	0	0	0	0	1
UNR	0	0	0	0	0	0
YP	1	0	0	0	1	0
Prob	0.574	0.236	0.368	0.905	0.728	0.928

P1 is 196307-198212 and P2 is 198301-200312. MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production. 1 indicates inclusion and 0 indicates exclusion.

Table 2.6c: Posterior simulation results for the degrees of freedom US Data

196307-198212

Return-based factors

			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	14.055	3.865	0.016	9.787
Book-to-Market	9.942	1.520	0.006	8.893
Industry	11.182	2.116	0.008	9.201
Cashflow	8.551	1.374	0.011	8.809
Dividend	8.567	1.288	0.005	8.673
Earning	7.706	1.163	0.004	8.407

Macroeconomic factors

			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	14.543	1.339	0.007	1.544
Book-to-Market	13.102	1.236	0.007	1.778
Industry	14.098	1.140	0.007	2.076
Cashflow	14.312	1.263	0.007	1.779
Dividend	8.079	0.917	0.005	1.544
Earning	14.887	1.346	0.007	1.774

198301-200312

Return-based factors

			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	6.067	0.815	0.002	7.649
Book-to-Market	9.322	1.475	0.006	8.865
Industry	6.169	0.786	0.005	7.424
Cashflow	6.025	0.819	0.002	7.665
Dividend	6.265	0.852	0.005	7.735
Earning	6.471	0.822	0.007	5.715

Macroeconomic factors

			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	7.448	1.302	0.003	3.953
Book-to-Market	6.358	1.357	0.004	4.267
Industry	6.735	1.253	0.003	3.954
Cashflow	7.048	1.359	0.004	4.076
Dividend	6.002	1.227	0.003	3.581
Earning	7.818	1.504	0.003	4.925

selection of factors is very robust. The results are presented in Tables 2.7a 2.7b. We also investigate the prior sensitivity for the degrees of freedom. It might be of interest to consider a noninformative prior for v. However, a flat prior on v on $(0, \infty)$ yields a posterior that is not integrable. As an alternative, Bauwens and Lubrano (1998) make use of the following proper, but rather uninformative prior

$$\pi(v) \propto (1 + v^2)^{-1}$$

if $v > 1$. This is the half-right side of a Cauchy centred at 0. Implementing this prior into the Markov Chain does not change the results in any substantial way. More specifically, the posterior mean and standard deviation for the degrees of freedom are about the same with the above prior compared with the exponential prior. Detailed results for the half-Cauchy prior is available from the authors on request.

2.6.2 Swedish Data

As explained earlier, the set of potential factors for the Swedish data is all expressed in returns by factor mimicking. Hence, the stock- and bond-market factors and the macroeconomic factors do not have to be treated separately. Furthermore, the intercept is included in the set of factors since asset pricing theory implies that misspricing is zero. This results in 12 factors.

In Table 2.8 the three best models with the highest posterior model probabilities are presented. Focusing on what is common among the different portfolios, Table 2.8 shows that the market excess return and the momentum factor seem to be important. In the book-to-market portfolios and the cashflow portfolios we find support for the size factor. Hence, we found no support for the three factor model of Fama and French. Taking a closer look at the posterior probabilities we note that the data is very informative about one single model. The posterior model probability for the best model is over 0.90 in all test assets except for the book-to-market portfolios where the probability for the best model is equal to 0.45. Furthermore, we note that the intercept is not included in any of the top models.

In Table 2.9 and in Figure A.3 in the Appendix, we summarize the posterior results for the degrees of freedom parameter, v, for the model containing all factors. The posterior means are low in all cases, which indicates substantial deviation from normality. The posterior mean is between 2.8 and 3.7. Furthermore, the standard deviations are all low. The estimated densities show that all of the posterior probability is allocated to small values for the

Table 2.7a: The best model using different values for g: Return-based factors US Data

	Size			Book-to-Market			Industry		
g	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05
INT	0	0	0	0	0	0	1	1	1
MKT-RF	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1
UMD	0	0	0	0	0	0	1	1	1
RP	0	0	0	0	0	0	1	1	1
UTS(S)	0	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	1	1	1	0	0	0
Prob	0.893	0.931	0.899	0.385	0.75	0.666	0.663	0.854	0.819

	Cashflow			Dividend			Earning		
g	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05
INT	1	1	0	1	0	0	0	0	0
MKT-RF	1	1	1	1	1	1	1	1	1
SMB	1	1	1	1	1	1	1	1	1
HML	1	1	1	1	1	1	1	1	1
UMD	1	1	1	1	1	1	1	1	1
RP	1	1	0	0	0	0	0	0	0
UTS(S)	0	0	0	0	0	0	0	0	0
UTS(L)	0	0	0	0	0	0	0	0	0
Prob	0.929	0.862	0.73	0.491	0.543	0.666	0.96	0.982	0.981

INT = intercept; MKT-RF = excess return on the market; SMB = size premium; HML = value premium; UMD = momentum factor; RP = risk premium; UTS(S) = term spread short; UTS(L) = term spread long. 1 indicates inclusion and 0 indicates exclusion.

Table 2.7b: The best model using different values for g:Macroeconomic factors US Data

	Size			Book-to-Market			Industry		
g	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05
MP	0	0	0	0	0	0	1	1	1
DEI	0	0	0	0	0	0	1	1	1
UI	0	0	0	0	0	0	1	0	0
CG	1	1	1	0	0	0	0	0	0
IC	0	0	0	0	0	0	1	1	0
OG	0	0	0	0	0	0	1	1	1
PSR	0	0	0	0	0	0	1	1	1
UNR	0	0	0	0	0	0	0	0	0
YP	0	0	0	1	1	1	1	1	1
Prob	0.975	0.576	0.887	0.496	0.67	0.92	0.264	0.304	0.204

	Cashflow			Dividend			Earning		
g	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05
MP	0	0	0	0	0	0	0	0	0
DEI	0	0	0	0	0	0	0	0	0
UI	0	0	0	0	0	0	0	0	0
CG	0	0	0	0	0	0	0	0	0
IC	0	0	0	0	0	0	0	0	0
OG	0	0	0	1	1	1	0	0	0
PSR	0	0	0	1	0	0	1	1	1
UNR	0	0	1	0	0	0	0	0	0
YP	1	1	0	0	0	0	0	0	0
Prob	0.442	0.573	0.475	0.151	0.305	0.501	0.839	0.86	0.956

MP = monthly growth rate in industrial production; DEI = change in expected inflation; UI = unanticipated inflation; CG = growth rate in real per capita consumption; IC = growth rate in real per capita disposable income; OG = growth rate in oil prices; PSR = growth rate in personal savings rate; UNR = growth rate in unemployment rate; YP = yearly growth rate in industrial production. 1 indicates inclusion and 0 indicates exclusion.

Table 2.8: Three best models Swedish Data

Factors	Size			Book-to-Market		
INT	0	0	0	0	0	0
RM-RF	1	1	1	1	1	1
SIZE	0	0	0	1	0	1
BM	0	0	0	1	1	1
UMD	0	1	0	1	1	1
MP	0	0	0	0	0	0
CON	0	0	0	0	0	0
INC	0	0	0	0	1	1
USD	0	0	0	0	0	0
DEI	0	0	1	0	0	0
UI	0	0	0	0	0	0
SLOPE	1	0	0	0	0	0
Prob	0.921	0.05	0.016	0.447	0.258	0.102

Factors	Cashflow			Dividend		
INT	0	0	0	0	0	0
RM-RF	1	1	1	1	1	1
SIZE	1	1	1	0	0	0
BM	0	0	0	0	0	0
UMD	1	1	1	1	1	1
MP	0	0	0	0	0	0
CON	0	1	0	1	0	1
INC	1	0	0	0	1	0
USD	0	0	0	0	0	0
DEI	0	0	0	0	0	0
UI	0	0	1	0	0	0
SLOPE	0	0	0	0	0	1
Prob	0.999	0.001	0	0.967	0.033	0

INT = intercept; RM-RF = Market excess return; SIZE = size premium; BM = book-to-market ratio; UMD = momentum premium; UI = unanticipated inflation; DEI = change in expected inflation; USD = change in USD/SEK exchange rate; SLOPE = slope of the yield curve; MP = monthly growth rate in industrial production; CON = growth rate in personal consumption; INC = growth rate in disposable income. 1 indicates inclusion and 0 indicates exclusion.

Table 2.9: Posterior simulation results for the degrees of freedom Swedish Data

Return-based factors				
			Spectral	
Portfolio	Mean	St.dev	NSE	RNE
Size	3.6988	0.4132	0.0047	13.6807
Book-to-Market	2.7673	0.2770	0.0032	7.8188
Cashflow	3.3016	0.3488	0.0045	9.9769
Dividend	3.0173	0.3142	0.0032	8.4750

degrees of freedom parameter. The degrees of freedom parameters are lower for the Swedish data compared with the US data. One possible explanation is diversification. The US portfolios contain many more stocks than the Swedish portfolios.

Table 2.9 and Figure A.3 also show the property of the Markov chain for the degrees of freedom parameter.[5] The numerical standard errors are quite low. Due to the autocorrelation, displayed in the graphs, the RNEs are high. They indicate that we need almost 8 to 13 times as many draws from the samples as when sampling directly from the posterior. However, Geweke diagnostics indicate that the chains seem to have converged.

Next, the robustness of the result with respect to the prior specification is examined.

Sensitivity Analysis

In this section we address the sensitivity with respect to the prior configuration. Due to the small sample period we will not consider any subsamples.

Table 2.10 shows the included factors in the best models when g is equal to $\{1/T, 1/K^2, 0.05\}$ respectively. We find some sensitivity to g. This is especially the case for the book-to-market portfolios. The best model for the book-to-market portfolios includes the market excess return and consumption when $g = 1/T$. When g is equal to $1/K^2$ and 0.05 four factors are selected, the market excess return, size, book-to-market and the momentum factor. As in the case of the US data, we also examined the prior sensitivity for the degrees of

---

[5] The diagnostics for the other parameters indicates convergence and are not reported but are available on request.

freedom by using the prior proposed by Bauwens and Lubrano (1998). Again, the results do not change in any significant way and the results are available from the authors on request.

Table 2.10: The best model using different values for g Swedish Data

	Size			Book-to-Market		
g	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05
INT	0	0	0	0	0	0
RM-RF	1	1	1	1	1	1
SIZE	0	0	0	0	1	1
BM	0	0	0	0	1	1
UMD	1	1	0	0	1	1
MP	0	0	0	0	0	0
CON	0	0	0	1	0	0
INC	0	0	0	0	0	0
USD	0	0	0	0	0	0
DEI	0	0	0	0	0	0
UI	0	0	0	0	0	0
SLOPE	0	0	1	0	0	0
Prob	0.675	0.708	0.921	0.877	0.373	0.447

	Cashflow			Dividend		
g	$1/T$	$1/K^2$	0.05	$1/T$	$1/K^2$	0.05
INT	0	0	0	0	0	0
RM-RF	1	1	1	1	1	1
SIZE	1	1	1	0	0	0
BM	0	0	0	0	0	0
UMD	1	1	1	1	1	1
MP	0	0	0	0	0	0
CON	0	0	0	1	1	1
INC	1	1	1	0	0	0
USD	0	0	0	0	0	0
DEI	0	0	0	0	0	0
UI	0	0	0	0	0	0
SLOPE	0	0	0	0	0	0
Prob	0.986	0.999	0.999	0.969	0.972	0.967

INT = intercept; RM-RF = Market excess return; SIZE = size premium; BM = book-to-market ratio; UMD = momentum premium; UI = unanticipated inflation; DEI = change in expected inflation; USD = change in USD/SEK exchange rate; SLOPE = slope of the yield curve; MP = monthly growth rate in industrial production; CON = growth rate in personal consumption; INC = growth rate in disposable income. 1 indicates inclusion and 0 indicates exclusion.

2.7 Conclusions

In this chapter, Bayesian techniques are used to select the factors in a multifactor asset pricing model when the assumption of normally distributed returns is relaxed. More precisely, we assume that asset returns are multivariate t-distributed. This setup allows us to capture the well known fat tail property of asset returns. Interest rates, premiums, returns on broadbased portfolios and macroeconomic variables are included in the set of factors considered. Furthermore, we examine data from the US and Swedish stock markets.

For the US data, using return-based factors, we find evidence that a general multifactor pricing model should include the market excess return, size and value premium and the momentum factor. It is however problematic that the intercept is included when the industry portfolios are the investment universe. Asset pricing theory implies a zero intercept and, hence, the intercept should not be selected. The results for the macroeconomic factors are mixed. The factor selection depend heavily on the test assets. The model uncertainty is substantial, at least for the industry sorted portfolios.

The results for the Swedish data show little support for the Fama-French three factor model except for when portfolios are based on book-to-market. However, the model uncertainty is also higher than for the other investment universes. The results are mixed and the model uncertainty is, in some cases, substantial. The important factors are the market excess return and the factor related to the momentum anomaly. Furthermore, none of the best models include the intercept which indicates that we have found factors that do price the assets under investigation.

The estimated densities for the degrees of freedom parameter have a shape, which is slightly skewed and indicates that all of the posterior probability is allocated to small values for the degrees of freedom parameter. Hence, we find a strong indication of deviation from normality, which makes our approach to modelling the data with a Student-t likelihood more appropriate.

Bibliography

AFFLECK-GRAVES, J., AND B. MCDONALD (1989): "Nonnormalities and Tests of Asset Pricing Theories," *Journal of Finance*, 44, 889–908.

BAUWENS, L., AND M. LUBRANO (1998): "Bayesian Inference on GARCH Models using the Gibbs Sampler," *Econometrics Journal*, 1, 23–46.

BOOT, J., W. FEIBES, AND J. LISMAN (1967): "Further Methods of Derivation of Quarterly Figures from Annual Data," *Applied Statistics*, 16, 65.

CARHART, M. M. (1997): "Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

CHAN, L., J. KARCESKI, AND J. LAKONISHOK (1998): "The Risk and Return from Factors," *Journal of Financial and Quantitativ Analysis*, 33, 159–188.

CHIB, S. (1995): "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, 1313–1321.

CHIB, S., AND I. JELIAZKOV (2001): "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270–281.

DICKEY, J. (1971): "The Weighted Likelihood Ratio, Linear Hypotheses on Normal Location Parameters," *Annals of Statistics*, 42, 204–224.

FAMA, E. (1965): "The Behaviour of Stock Market Prices," *Journal of Business*, 38, 34–105.

FAMA, E., AND K. FRENCH (1993): "Common Risk Factors in the Returns of Stock and Bonds," *Journal of Financial Economics*, 33, 3–56.

FAMA, E., AND M. GIBBONS (1984): "A Comparison of Inflation Forecasts," *Journal of Financial Economics*, 13, 327–384.

FERNÁNDEZ, C., E. LEY, AND M. STEEL (2001): "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, 38, 381–427.

GEWEKE, J. (1993): "Bayesian Treatment of the Independent Student-t Linear Model," *Journal of Applied Econometrics*, 8, 19–40.

GROENEWOLD, N., AND P. FRASER (2002): "Violation of the IID-Normal Assumption: Effects on Tests of Asset Pricing Models using Australian Data," *International Review of Financial Analysis*, 11, 491–510.

HALL, A. D., S. HWANG, AND E. S. SATCHELL (2002): "Using Bayesian Variable Selection Methods to Choose Style Factors in Global Stock Return Models," *Journal of Banking and Finance*, 26, 2301–2325.

HARVEY, C. R., AND G. ZHOU (1993): "International Asset Pricing with Alternative Distributional Specification," *Journal of Empirical Finance*, 1, 107–131.

KANDEL, S., AND R. STAMBAUGH (1996): "On the Predictability of Stock Returns: An Asset Allocation Perspective," *Journal of Finance*, 51, 385424.

KASS, R. E., L. TIERNEY, AND J. B. KADANE (1988): "Asymptotics in Bayesian Computations," in *Bayesian Statistics 3*, ed. by J. M. Bernardo, M. H. DeGroot, and D. V. Lindley, Valencia. University Press.

RICHARDSON, M., AND T. SMITH (1993): "A Test of Multivariate Normality of Stock Returns," *Journal of Business*, 66, 295–321.

SMITH, M., AND R. KOHN (2000): "Nonparametric Seemingly Unrelated Regression," *Journal of Econometrics*, 98, 257–281.

TIERNEY, L., R. E. KASS, AND J. B. KADANE (1989): "Approximate Marginal Densities for Nonlinear Functions," *Biometrika,*, 76, 425–433.

TU, J., AND G. ZHOU (2004): "Data-Generating Process Uncertainty: What difference Does it Make in Portfolio Decisions?," *Journal of Financial Economics*, 72, 385–421.

VERDINELLI, I., AND A. WASSERMAN (1995): "Computing Bayes Factors by Using a Generalization of the Savage-Dickey Density Ratio," *Journal of the American Statistical Association*, 90, 614–618.

ZELLNER, A. (1986): "On Assesing Prior Distributions and Bayesian Regression Analysis with G-Prior Distributions," in *Bayesian Inference and*

Decision Techniques - Essayes in Honor of Bruno de Finetti, ed. by P. K. Goel, and A. Zellner, pp. 223–243, Amsterdam. North Holland.

Appendix A

Tables and Figures

Table A.1a: Descriptive statistics: US test assets

Size

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	0.8021	6.493	-0.1082	5.4987	125.0951
2	0.7456	6.365	-0.2325	5.603	139.1364
3	0.7362	6.0603	-0.4365	5.2044	111.8377
4	0.7025	5.8611	-0.5278	5.2923	126.8194
5	0.729	5.6115	-0.5429	5.3703	135.3918
6	0.6059	5.3037	-0.5452	5.0774	109.5744
7	0.6541	5.1902	-0.4207	5.3119	120.4514
8	0.6101	5.0672	-0.3816	4.6543	65.8639
9	0.5216	4.6178	-0.3378	4.5541	56.9227
Large	0.4117	4.3334	-0.2968	4.6786	62.8531

Book-to-market

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	0.356	5.3138	-0.1867	4.2934	35.7654
2	0.4991	4.8426	-0.4075	4.779	76.0479
3	0.5006	4.7956	-0.5277	5.4732	144.0394
4	0.4986	4.7499	-0.3668	5.1047	98.7515
5	0.5191	4.4306	-0.4093	5.9709	189.2795
6	0.6237	4.4199	-0.4207	5.5558	144.1708
7	0.7026	4.3709	0.0322	4.9575	76.071
8	0.7209	4.3502	-0.0764	5.226	98.8848
9	0.7661	4.71	-0.1614	4.9854	80.2803
Large	0.9056	5.4473	0.0045	6.3081	218.1534

Industry

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
NonDurables	0.6366	4.5454	-0.2744	5.0098	86.1882
Durabless	0.527	5.6042	-0.0978	4.5534	48.4663
Oil	0.435	4.8461	-0.4154	5.7421	163.544
Chemicals	0.5566	5.2206	0.0871	4.6327	53.3377
Manufacturing	0.5454	6.7692	-0.1483	4.1127	26.0868
Telecom	0.3871	4.9952	-0.1017	4.8611	69.4781
Utilities	0.5947	5.4504	-0.2305	5.4907	127.637
Shops	0.6885	5.1238	0.0728	5.4345	118.2521
Money	0.3207	4.139	0.1481	4.0193	22.13
Other	0.5637	5.1614	-0.3518	4.4789	53.1542

JP is the Jarque-Bera test for normality and is χ_2^2 distributed with a critical value of 5.9915.

Table A.1b: Descriptive statistics: US test assets

Cashflow

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	0.3853	5.7095	-0.3081	4.6142	59.1774
2	0.4271	4.8222	-0.0558	4.723	59.0194
3	0.492	4.6045	-0.332	4.7761	71.3519
4	0.4875	4.7176	-0.3754	5.0996	98.8341
5	0.6098	4.5712	-0.5439	5.6643	165.0645
6	0.5374	4.4791	-0.514	4.9992	100.5442
7	0.593	4.4345	-0.3289	5.5119	134.1828
8	0.614	4.4057	0.0115	5.2357	99.2873
9	0.8408	4.4775	0.2413	5.7365	153.7281
Large	0.8273	5.0856	-0.256	5.4899	128.5567

Dividend

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	0.5406	5.7364	-0.3838	4.7244	70.7227
2	0.474	5.1286	-0.4145	4.6908	70.4072
3	0.5572	4.9849	-0.2049	5.0095	83.487
4	0.4996	4.8011	-0.3253	4.8233	74.3842
5	0.4013	4.6726	-0.1468	5.5372	129.7611
6	0.5358	4.5252	-0.3489	5.0942	96.8411
7	0.5623	4.4176	-0.388	4.8458	79.6297
8	0.6663	4.3288	-0.1914	5.0709	88.0529
9	0.6448	4.1167	0.0239	4.3517	36.0718
Large	0.5822	4.0292	0.7392	7.6139	469.2517

Earnings

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	0.3522	5.868	-0.1921	4.3479	38.7939
2	0.3649	4.8278	-0.2745	4.8597	74.6034
3	0.5008	4.6815	-0.2443	4.9058	76.8082
4	0.5011	4.4323	-0.3199	5.2217	106.2684
5	0.4423	4.6218	-0.3459	5.001	89.0596
6	0.5986	4.4107	-0.3268	5.3162	115.1956
7	0.7717	4.4372	-0.2234	5.1238	93.547
8	0.7713	4.4663	-0.0653	5.0539	84.0469
9	0.7755	4.7658	-0.0245	5.4608	120.4455
Large	0.9331	5.3647	-0.1094	5.9315	172.1043

JP is the Jarque-Bera test for normality and is χ_2^2 distributed with a critical value of 5.9915.

Table A.1c: Descriptive statistics: Swedish test assets

Size

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	2.0888	8.4467	0.456	5.7887	90.39021275
2	1.8907	12.8862	0.9647	7.957	297.0915503
3	2.0594	8.9464	0.4351	5.0382	51.57082644
4	1.8843	8.6901	0.2009	4.216	17.22104202
5	1.2436	8.8469	0.4141	6.5295	138.0044976
6	1.2701	7.687	0.2053	4.1196	14.93201346
7	1.5485	9.4242	0.4345	9.2343	416.0274036
8	1.077	8.3751	0.7366	5.166	72.04967952
9	1.3107	7.7446	0.2218	4.1004	14.78044176
Large	1.6815	7.1161	0.1867	4.6901	31.45658849

Book-to-market

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	1.9085	4.268	0.8912	6.8979	192.8910288
2	0.8017	11.891	0.6347	7.8523	264.1400123
3	1.5189	8.7566	0.5298	5.6516	85.61421456
4	0.7918	8.3465	0.8385	6.7398	176.3835469
5	1.3122	10.8046	0.8531	5.8	112.8867436
6	2.0131	8.9531	0.9324	6.5808	171.1458806
7	1.7535	7.5837	0.9851	6.2079	148.8092597
8	2.1183	10.8308	0.8734	5.8478	117.1933883
9	1.5105	13.9326	0.8734	6.0947	132.5990225
Large	4.0467	7.2982	0.6808	4.9875	60.94316351

JP is the Jarque-Bera test for normality and is χ^2_2 distributed with a critical value of 5.9915.

Table A.1d: Descriptive statistics: Swedish test assets

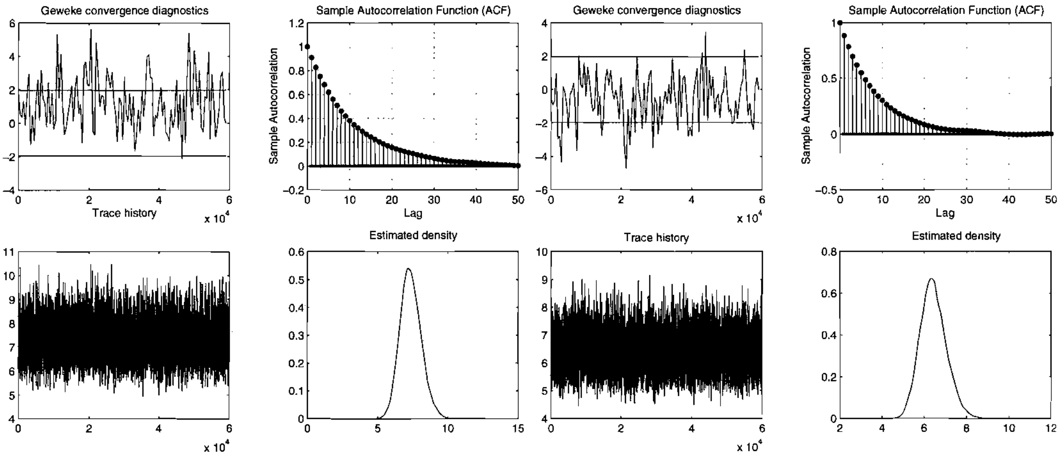
Cashflow

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	2.164	9.8434	0.3456	8	123.9823
2	1.9922	8.2567	0.4284	4.728	39.4848
3	1.7433	7.3163	0.6932	6.113	124.3605
4	1.3971	7.4346	0.4152	4.899	45.6334
5	1.7661	6.7845	0.3874	5.5029	73.1633
6	1.5901	7.402	0.7549	7.1033	205.1844
7	1.8426	7.4836	0.5479	5.0371	57.0143
8	1.5583	7.7181	0.1979	4.5557	27.1611
9	1.9518	10.9093	0.4939	5.9834	128.3423
Large	1.1719	8.3821	0.3917	4.7536	39.1288

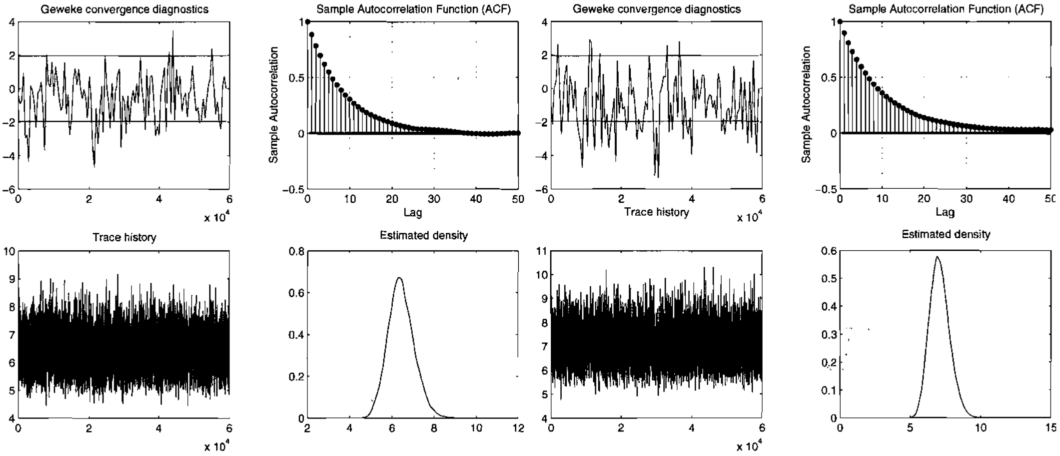
Dividend

Portfolio	Average Return	Std.dev	Skewness	Kurtosis	JB
Small	1.1123	8.3997	-0.0212	5.0287	43.6122
2	1.5036	7.2492	0.4054	6.0803	108.43
3	1.3432	6.3779	0.2912	4.7119	34.5997
4	1.4757	6.6624	0.2238	4.7516	34.5575
5	1.9125	6.1123	-0.0223	5.4528	63.9917
6	1.4737	6.3229	0.0872	4.4577	22.6371
7	1.9008	6.4272	-0.1558	4.1941	15.9184
8	1.6963	6.8235	0.0747	5.1783	50.5793
9	2.2472	10.5355	-0.3454	6.3454	34.3423
Large	2.4721	10.6271	0.2342	4.3468	54.6286

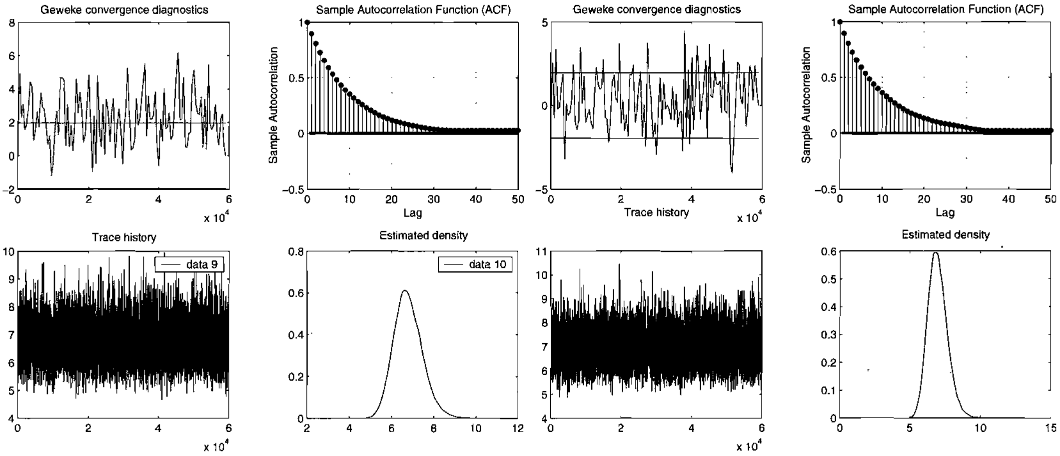
JP is the Jarque-Bera test for normality and is χ_2^2 distributed with a critical value of 5.9915.



(a) Size (b) Book-to-market

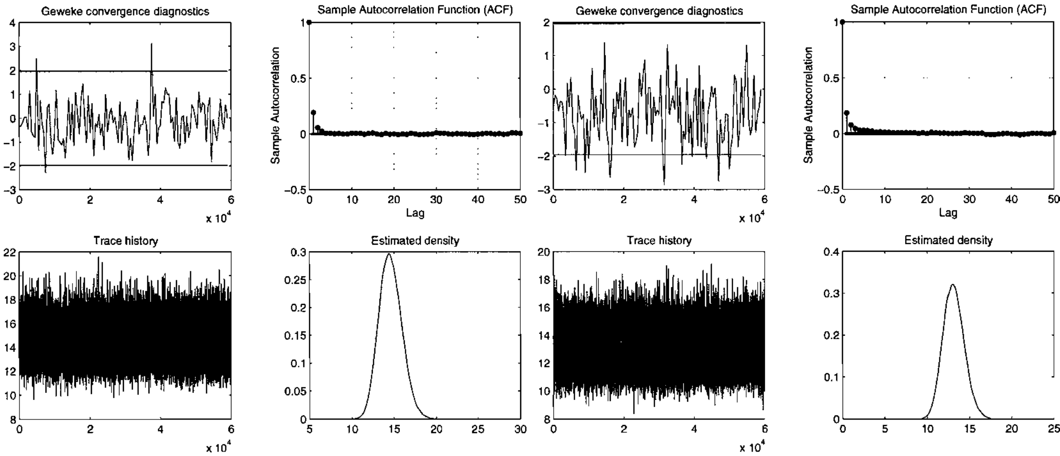


(c) Industry (d) Earning

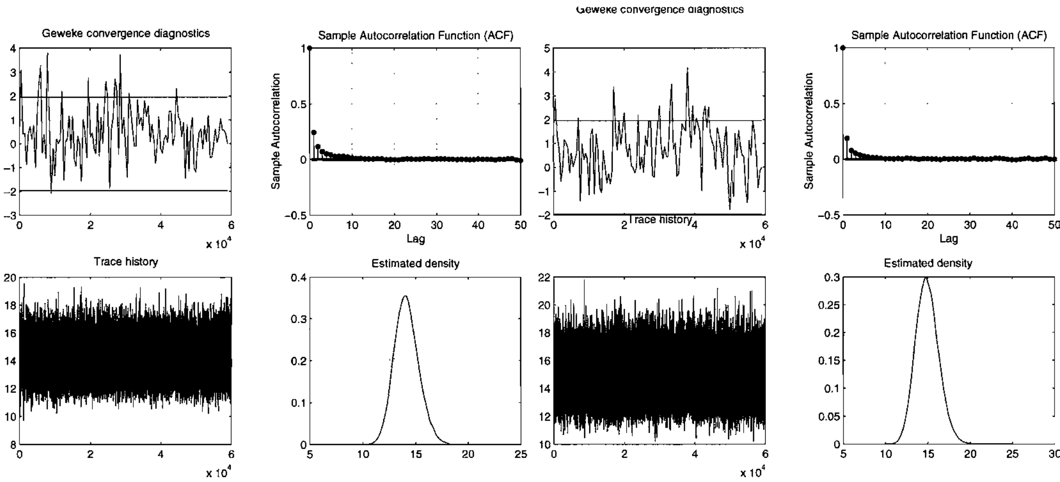


(e) Cashflow (f) Dividend

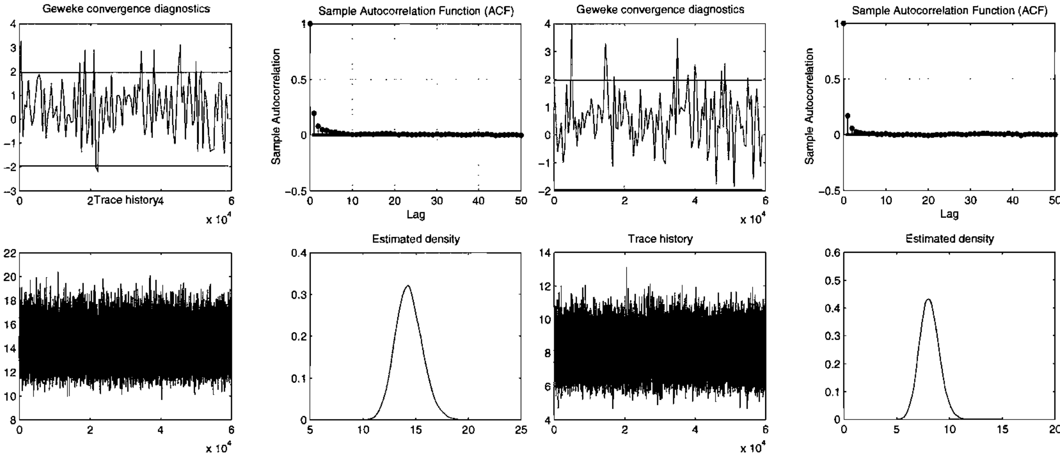
Figure A.1: Markov chain properties for the degrees of freedom v: US Data, Return-based factors.



(a) Size (b) Book-to-market

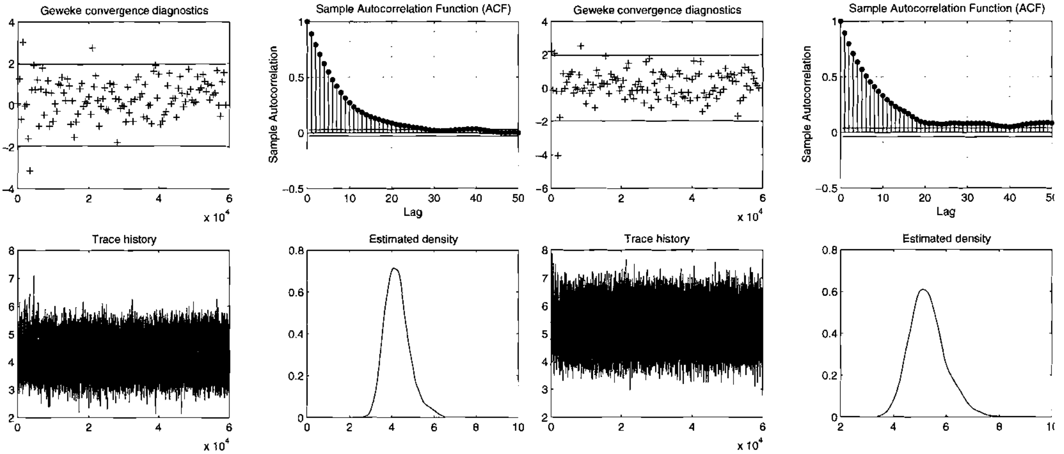


(c) Industry (d) Earning

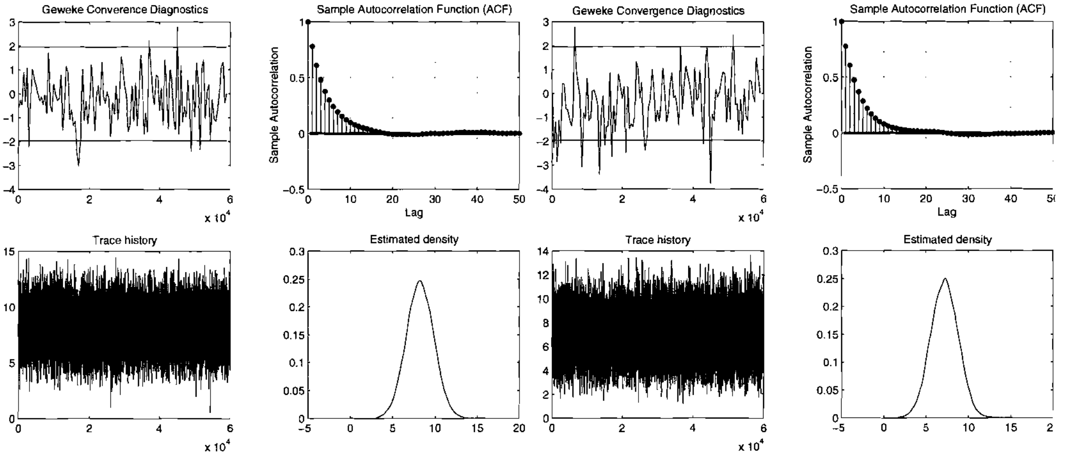


(e) Cashflow (f) Dividend

Figure A.2: Markov chain properties for the degrees of freedom v: US Data, Macroeconomic factors.



(a) Size (b) Book-to-market



(c) Cashflow (d) Dividend

Figure A.3: Markov chain properties for the degrees of freedom v: Swedish Data

Chapter 3

Choosing Factors in a Multifactor Asset Pricing Model when Factors are Unobserved

3.1 Introduction

The Arbitrage Pricing Theory (APT) was introduced by Ross (1976) as an alternative to the Capital Asset Pricing Model (CAPM). In contrast to the CAPM, the APT allows for multiple risk factors and the model implies that the expected return on an asset is a linear function of factor risk premiums and their associated factor sensitivities.

The estimation and testing of multifactor pricing models assumes that the identity of the factors is known. However, economic theory is not very explicit on the number and nature of these factors. The selection of the number of factors and an appropriate set of factors are therefore an empirical issue. Two approaches are common in the literature. One focusing on unobservable or latent factors and the second focusing on observable factors. In this paper, attention is paid to the first approach where the factors are unobserved, which is the framework for the APT model.

In the APT framework, the systematic, unobservable factors are usually extracted by using statistical techniques like factor analysis and principal components. The original derivation of the APT model in Ross (1976) is based on a strict factor structure where the idiosyncratic returns are uncorrelated across assets. Chamberlain and Rothschild (1983) and Ingersoll (1984) generalize the results of the APT in the case of an approximate factor structure. In an approximate factor structure, the idiosyncratic returns are allowed to be correlated across assets, at least to some extent. The approximate factor structure is more general and may therefore be more attractive than the strict version. Furthermore, the approximate version is more realistic. For example, we can expect that a few firms or industries might have specific components of their return which are not pervasive sources of uncertainty for the whole economy.

The problem of determining the number of factors in the APT framework has been analyzed in both the classical and the Bayesian framework. Bai and Ng (2002) present several model selection criteria which hold under weak serial and cross-section dependence . In an application to the US market they found support for two pervasive factors. Connor and Korajczyk (1988) who use principal components find evidence for one to six latent factors in the cross-section of stock returns and Lehmann and Modest (1988) who use factor analysis find weak evidence in favor of a ten-factor model. Geweke and Zhou (1996) provide a Bayesian approach for analyzing the pricing error in the APT. They find that there is little improvement in reducing the pricing errors by including more factors than one. Inference on the number of factors itself has received relatively little attention in the Bayesian literature. Recently, Lopes and West (2004) explore several Bayesian techniques regarding the uncertainty about the number of latent factors in a strict factor model. They found that the Bayesian approach is just as successful as classical selection criteria.

In the strict and the approximate factor structure the idiosyncratic returns and the factors are assumed to be serially independent. Hence, so are the returns. Nevertheless, empirical results indicate that returns are autocorrelated to some extent. For example, Geweke and Zhou (1996) argue that the autocorrelation does not seem to be severe and therefore, adopt the working assumption that the returns are independent and identically distributed. In contrast to the standard working assumption, two factor structures are introduced in this chapter where the returns are allowed to be serially correlated.

I set up the determination of the number of factors as a model selection problem. A Bayesian approach is used and the selection of the number of factors is based on posterior model probabilities. In using a Bayesian approach I also have the opportunity to quantify the uncertainty about the number of factors. Furthermore, I relax the assumption of a strict factor structure and in addition, by letting the error term follow an AR(1) process, I also allow for some time series dependence.

The rest of the chapter is organized as follows. In the next section we present the model. Section 3.3 describes the prior and the algorithm for making posterior inference and Section 3.4 describes the Bayesian model selection procedure. Section 3.5 contains the empirical results and Section 3.6 provides a conclusion.

3.2 The Model

Let $\mathbf{r}_t$ denote a $N \times 1$ vector of excess returns on N assets in period t. The APT assumes that the return generating process is

$$\mathbf{r}_t = \mu + \Lambda \mathbf{f}_t + \varepsilon_t \tag{3.1}$$

where μ is a $N \times 1$ vector of constants, f_t is a $m \times 1$ vector of pervasive factors, Λ is a $N \times m$ matrix of factor betas or loadings and ε_t is a $N \times 1$ vector of idiosyncratic returns. In what follows, it will be convenient to work with the matrix form of the model,

$$\mathbf{R} = e \otimes \mu + \mathbf{F}\Lambda' + \mathbf{E} \tag{3.2}$$

where e is the N-vector of ones and the rows of $\mathbf{R}$, $\mathbf{F}$ and $\mathbf{E}$ are given by $\mathbf{r}_t'$, $\mathbf{f}_t'$ and ε_t'. The vec version of (3.2) will also be useful

$$\begin{aligned} r &= vec(e \otimes \mu) + Z\lambda + \varepsilon \\ &= vec(e \otimes \mu) + \Upsilon \mathbf{f} + \varepsilon \end{aligned} \tag{3.3}$$

where $r = vec(\mathbf{R})$, $Z = (\mathbf{I}_N \otimes \mathbf{F})$, $\lambda = vec(\Lambda')$, $\Upsilon = (\Lambda \otimes \mathbf{I}_T)$, $\mathbf{f} = vec(\mathbf{F})$, and finally, $\varepsilon = vec(\mathbf{E})$.

The standard assumptions on the factor model are

$$E(\mathbf{f}_t) = 0 \quad E(\mathbf{f}_t\mathbf{f}_t') = \mathbf{I}_m \quad E(\varepsilon_t|\mathbf{f}_t) = 0 \quad E(\varepsilon_t\varepsilon_t'|\mathbf{f}_t) = \Sigma \tag{3.4}$$

and that the idiosyncratic component is normally distributed. A strict factor structure imposes the condition that the covariance matrix of idiosyncratic returns Σ is a diagonal matrix. Ross (1976) assumes a strict factor structure in his original development of the APT.

Chamberlain and Rothschild (1983) and Ingersoll (1984) generalize the results of the APT to the case of an approximate factor structure. In an approximate factor structure the idiosyncratic components of the returns may be correlated across assets and hence, the idiosyncratic covariance matrix is not restricted to be diagonal.

In this paper, I will consider four specifications regarding the idiosyncratic component.

1. The first specification follows from the strict factor structure where Σ is assumed to be diagonal.
2. The second follows from the approximate factor model where the assumption of a diagonal covariance matrix is relaxed.
3. In the third specification, time series dependence is introduced in the approximate factor model by letting the error term follow an AR(1) process
4. In the final specification we return to the strict factor model where the time series dependence can differ across assets.

For simplicity, but without losing any generality, we assume that the returns have been demeaned.

3.3 The Prior and the Posterior

In general, we need to be informative in the prior setting since improper non-informative priors yield indeterminate marginal likelihoods. In the next subsections the prior distributions are discussed and the posterior distributions are derived.

3.3.1 Strict Factor Structure

Assuming a strict factor structure it follows that the distribution of the idiosyncratic returns is given by,

$$\mathbf{E} \sim MN_{TN}(0, \Sigma, \mathbf{I}_N)$$

where MN denotes the matrix variate normal distribution and $\Sigma = diag\{\sigma_{11}^2, \ldots, \sigma_{NN}^2\}$. The likelihood is given by

$$L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, m) \propto |\Sigma|^{-T/2} \exp\left\{-\frac{1}{2} tr \Sigma^{-1} (\mathbf{R} - \mathbf{F}\Lambda')'(\mathbf{R} - \mathbf{F}\Lambda')\right\}.$$

Diffuse but proper priors are used to represent the uncertainty about the parameters. To complete the model specification we make use of the following priors

$$\begin{aligned} \mathbf{f}_t|m &\sim N_m(0, \mathbf{I}_m) \\ \Lambda|m &\sim MN_{Nm}(\Lambda_0, \mathbf{H}_0, \mathbf{I}_N) \\ \sigma_{ii}^2 &\sim IG_2(s_0, v_0) \end{aligned}$$

where Λ_0, $\mathbf{H}_0$, s_0 and v_0 are hyperparameters to be assessed and IG_2 denotes the inverse Gamma-2 distribution given in Bauwens, Lubrano, and Richard (2000).

The assessment of the hyperparameters is usually a difficult task and this case is no exception. Following the recommendation in Lopes and West (2004) we take Λ_0 to be equal to zero and $\mathbf{H}_0 = h_0 \mathbf{I}_m$ with h_0 a large value. For each of the elements in Σ we assume a common inverse Gamma-2 prior. The degrees of freedom parameter v_0 is taken to be a low value to produce a diffuse but still proper prior and s_0 is chosen in such a way that the prior mean is equal to one.

Using Bayes rule, the posterior distribution for the unknown parameters is obtained by combining the likelihood and the priors

$$P(\mathbf{F}, \Lambda, \Sigma|\mathbf{R}, m) \propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, m)\pi(\mathbf{F}|m)\pi(\Lambda|m)\pi(\Sigma).$$

The marginal posteriors cannot be derived analytically but the conditional posterior distributions can be obtained and Gibbs sampling can be used for posterior inference.

The conditional posterior of the factor score at time t is

$$\begin{aligned}P(\mathbf{f}_t|\Lambda,\Sigma,\mathbf{R},m) &\propto L(\mathbf{r}_t|\mathbf{f}_t,\Lambda,\Sigma,m)\pi(\mathbf{f}_t|m)\\ &\propto \exp\left\{-\frac{1}{2}[(\mathbf{r}_t-\Lambda\mathbf{f}_t)'\Sigma^{-1}(\mathbf{r}_t-\Lambda\mathbf{f}_t)+\mathbf{f}_t'\mathbf{f}_t]\right\}\\ &\propto \exp\left\{-\frac{1}{2}(\mathbf{f}_t-\tilde{\mathbf{f}}_t)'(\Lambda'\Sigma^{-1}\Lambda+\mathbf{I}_m)(\mathbf{f}_t-\tilde{\mathbf{f}}_t)\right\}\end{aligned} \tag{3.5}$$

where

$$\tilde{\mathbf{f}}_t=(\Lambda'\Sigma^{-1}\Lambda+\mathbf{I}_m)^{-1}\Lambda'\Sigma^{-1}\mathbf{r}_t.$$

That is, the conditional posterior distribution of the factor scores is normal.

In deriving the conditional posterior distribution for the factor loadings it is convenient to use the vec form of the model given by (3.3). In addition, let $\lambda_0=vec(\Lambda_0')$. The conditional posterior of the factor loadings is

$$\begin{aligned}P(\Lambda|\mathbf{F},\Sigma,\mathbf{R},m) &\propto L(\mathbf{R}|\mathbf{F},\Lambda,\Sigma,m)\pi(\Lambda|m)\\ &\propto \exp\left\{-\frac{1}{2}[(r-Z\lambda)'(\Sigma\otimes\mathbf{I}_T)^{-1}(r-Z\lambda)]\right\}\\ &\times \exp\left\{-\frac{1}{2}[(\lambda-\lambda_0)'(\mathbf{I}_N\otimes\mathbf{H}_0)^{-1}(\lambda-\lambda_0)]\right\}\\ &\propto \exp\left\{-\frac{1}{2}(\lambda-\tilde{\lambda})'(Z'(\Sigma\otimes\mathbf{I}_T)^{-1}Z+(\mathbf{I}_N\otimes\mathbf{H}_0)^{-1})(\lambda-\tilde{\lambda})\right\}\end{aligned} \tag{3.6}$$

where

$$\tilde{\lambda}=(\Sigma^{-1}\otimes\mathbf{F}'\mathbf{F}+\mathbf{I}_N\otimes\mathbf{H}_0^{-1})^{-1}((\Sigma^{-1}\otimes\mathbf{F}')r+(\mathbf{I}_N\otimes\mathbf{H}_0^{-1})\lambda_0).$$

Using the fact that the covariance matrix is diagonal I obtain

$$\begin{aligned}P(\lambda^{(i)}|\mathbf{F},\Sigma,\mathbf{R},m) &\propto L(\mathbf{R}|\mathbf{F},\Lambda,\Sigma,m)\pi(\lambda^{(i)}|m)\\ &\propto \exp\left\{-\frac{1}{2}(\lambda^i-\tilde{\lambda}^{(i)})'(\mathbf{F}'\mathbf{F}/\sigma_{ii}^2+\mathbf{H}_0)^{-1}(\lambda^i-\tilde{\lambda}^{(i)})\right\}\end{aligned} \tag{3.7}$$

where $\lambda^{(i)}$ is column (i) in Λ',

$$\tilde{\lambda}^{(i)} = (\mathbf{F}'\mathbf{F}/\sigma_{ii}^2 + \mathbf{H}_0^{-1})^{-1}(\mathbf{F}'r^{(i)}/\sigma_{ii}^2 + \mathbf{H}_0^{-1})\lambda_0^{(i)}),$$

,$r^{(i)}$ is column (i) in $\mathbf{R}$ and finally, $\lambda_0^{(i)}$ is column (i) in Λ_0'. The conditional posterior distribution of the factor loadings is normal.

The conditional posterior of the diagonal element σ_{ii}^2 in Σ, the covariance matrix, is

$$\begin{aligned} P(\sigma_{ii}^2|\mathbf{F}, \Lambda, \mathbf{R}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, m)\pi(\Sigma) \\ &\propto |\Sigma|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}tr\Sigma^{-1}[\mathbf{E}'\mathbf{E}]\right\} \\ &\times \left\{\frac{1}{\sigma_{ii}^2}\right\}^{\frac{v_0+2}{2}} \exp\left\{\frac{1}{2\sigma_{ii}^2}s_0\right\} \\ &\propto \left\{\frac{1}{\sigma_{ii}^2}\right\}^{\frac{T+v_0+2}{2}} \exp\left\{-\frac{1}{2\sigma_{ii}^2}(\varepsilon_i'\varepsilon_i + s_0)\right\} \end{aligned}$$

where ε_i is column i in $\mathbf{E} = [\varepsilon_1, ..., \varepsilon_N]$. That is, the conditional posterior distribution of the element σ_{ii}^2 of Σ is inverse Gamma-2.

Given the number of factors m, a Gibbs algorithm can be implemented as follows for posterior estimation and inference.

1. draw $\breve{\mathbf{F}}_{j+1}$ by sampling from $P(\mathbf{f}_t|\breve{\Lambda}_j, \breve{\Sigma}_j, \mathbf{R}, m)$, $t = 1, \ldots, T$
2. draw $\breve{\Lambda}_{j+1}$ by sampling from $P(\Lambda|\breve{\mathbf{F}}_{j+1}, \breve{\Sigma}_j, \mathbf{R}, m)$
3. draw $\breve{\Sigma}_{j+1}$ by sampling from $P(\sigma_{ii}^2|\breve{\mathbf{F}}_{j+1}, \breve{\Lambda}_{j+1}, \mathbf{R}, m)$, $i = 1, \ldots, N$.

Cycling through these steps a large number of times will generate a sample from the posterior.

3.3.2 Approximate Factor Structure

In this section, the assumption of a diagonal covariance matrix is relaxed and the distribution of the idiosyncratic term is now given by

$$\mathbf{E} \sim MN_{TN}(0, \Sigma, \mathbf{I}_N)$$

where Σ is not restricted to be diagonal. Still, the prior will be based on traditional beliefs of a strict factor structure containing pervasive and idiosyncratic components.

The prior setup is the same as for the strict factor structure, except that the prior for the covariance matrix is an Inverse Wishart density

$$\Sigma \sim IW(\mathbf{S}_0, v_0)$$

where $\mathbf{S}_0$ and v are hyperparameters to be assessed. The degrees of freedom parameter v_0 is set to a low value to make the prior diffuse but proper, and $\mathbf{S}_0$ is chosen to be a diagonal matrix in such a way that the prior mean is equal to the identity matrix.

Combining the likelihood and the priors yields non-standard marginal posterior distributions but the full conditionals can be obtained. It turns out that the conditional posterior distribution for the factor scores and the conditional posterior distribution for the factor loadings follow from the strict factor structure in (3.5) and (3.6). The conditional posterior of Σ is, however, different and is given by

$$\begin{aligned} P(\Sigma|\mathbf{F}, \Lambda, \mathbf{R}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, m)\pi(\Sigma) \\ &\propto |\Sigma|^{-\frac{T+v_0+N+1}{2}} \exp\left\{-\frac{1}{2} tr\Sigma^{-1}[\mathbf{E}'\mathbf{E} + \mathbf{S}_0]\right\} \end{aligned}$$

which takes the form of an Inverse Wishart density.

Given the number of factors m a Gibbs algorithm can again be implemented as follows:

1. draw $\breve{\mathbf{F}}_{j+1}$ by sampling from $P(\mathbf{f}_t|\breve{\Lambda}_j, \breve{\Sigma}_j, \mathbf{R}, m)$, $t = 1, \ldots, T$
2. draw $\breve{\Lambda}_{j+1}$ by sampling from $P(\Lambda|\breve{\mathbf{F}}_{j+1}, \breve{\Sigma}_j, \mathbf{R}, m)$
3. draw $\breve{\Sigma}_{j+1}$ by sampling from $P(\Sigma|\breve{\mathbf{F}}_{j+1}, \breve{\Lambda}_{j+1}, \mathbf{R}, m)$.

3.3.3 Approximate Factor Structure with Time Series and Cross-Sectional Dependence

This section introduces the possibility of time dependence in the errors. More specifically, the distribution of the idiosyncratic component is given by

$$\mathbf{E} \sim MN_{TN}(0, \Sigma, \mathbf{\Phi})$$

where Σ is not restricted to be diagonal and $\mathbf{\Phi}$ is a $T \times T$ matrix given by

$$\mathbf{\Phi} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{pmatrix}$$

where $|\rho| < 1$. Hence, ε_t follows the process

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where u_t is multivariate normal, $N(0, \Sigma)$. This yields the following likelihood

$$L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, \mathbf{\Phi}, m) \propto |\Sigma|^{-T/2}|\mathbf{\Phi}|^{-N/2} \times \exp\left\{-\frac{1}{2}tr\Sigma^{-1}(\mathbf{R} - \mathbf{F}\Lambda')'\mathbf{\Phi}^{-1}(\mathbf{R} - \mathbf{F}\Lambda')\right\}.$$

Except for ρ, the prior follows the same setup as for the approximate factor structure and combining the likelihood and the prior yields non-analytical marginal posterior distributions. However, conditioning on $\mathbf{\Phi}$, the posterior conditionals are very similar to the ones in the previous section.

The conditional posterior of the factor scores is

$$\begin{aligned} P(f|\Lambda, \Sigma, \mathbf{\Phi}, \mathbf{R}, m) &\propto L(\mathbf{R}|F, \Lambda, \Sigma, \mathbf{\Phi}, m)\pi(F|m) \\ &\propto \exp\left\{-\frac{1}{2}[(r - \Upsilon\mathbf{f})'(\Sigma \otimes \mathbf{\Phi})^{-1}(r - \Upsilon\mathbf{f}) + \mathbf{f}'\mathbf{f}]\right\} \\ &\propto \exp\left\{-\frac{1}{2}(\mathbf{f} - \tilde{\mathbf{f}})[\Upsilon'(\Sigma \otimes \mathbf{\Phi})^{-1}\Upsilon + \mathbf{I}_{Tm}](\mathbf{f} - \tilde{\mathbf{f}})\right\} \end{aligned}$$

where

$$\tilde{\mathbf{f}} = (\Lambda'\Sigma^{-1}\Lambda \otimes \mathbf{\Phi}^{-1} + \mathbf{I}_{Tm})^{-1}(\Lambda'\Sigma^{-1} \otimes \mathbf{\Phi}^{-1})r.$$

The conditional posterior distribution of the factor scores is normal.

The conditional posterior of the factor loadings is derived in a similar way as for the strict factor structure by replacing $\mathbf{I}_T$ with $\mathbf{\Phi}$ in (3.6). This yields

$$\begin{aligned} P(\Lambda|\mathbf{F}, \Sigma, \mathbf{\Phi}, \mathbf{R}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, \mathbf{\Phi}, m)\pi(\Lambda|m) \\ &\propto \exp\left\{-\frac{1}{2}(\lambda - \tilde{\lambda})'(Z'(\Sigma \otimes \mathbf{\Phi})^{-1}Z \right. \\ &\left. + (\mathbf{I}_N \otimes \mathbf{H}_0)^{-1})(\lambda' - \tilde{\lambda}')\right\} \end{aligned}$$

where

$$\tilde{\lambda} = (\Sigma^{-1} \otimes \mathbf{F}'\mathbf{\Phi}^{-1}\mathbf{F} + \mathbf{I}_N \otimes \mathbf{H}_0^{-1})^{-1}((\Sigma^{-1} \otimes \mathbf{F}'\mathbf{\Phi}^{-1})r + (\mathbf{I}_N \otimes \mathbf{H}_0^{-1})\lambda_0)$$

leading to a normal conditional posterior.

The conditional posterior of the covariance matrix again takes the form of an Inverse Wishart density,

$$\begin{aligned} P(\Sigma|\mathbf{F}, \Lambda, \mathbf{R}, \mathbf{\Phi}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, \mathbf{\Phi}, m)\pi(\Sigma) \\ &\propto |\Sigma|^{-\frac{T+v_0+N+1}{2}} \exp\left\{-\frac{1}{2}tr\Sigma^{-1}[\mathbf{E}'\mathbf{\Phi}^{-1}\mathbf{E} + \mathbf{S}_0]\right\}. \end{aligned}$$

Next we turn to the prior and the conditional posterior for $\mathbf{\Phi}$, which depends on the single parameter ρ. A common prior for ρ in the Bayesian literature is the truncated normal density

$$\pi(\rho) \propto f_N(\rho|\rho_0, \sigma_\rho^2)1(\rho \in \Omega)$$

where f_N is the density of the normal distribution and $1(\rho \in \Omega)$ is the indicator function, which equals 1 for the stationary region and zero otherwise. The resulting conditional posterior is

$$\begin{aligned} P(\rho|\mathbf{F}, \Sigma, \Lambda, \mathbf{R}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Sigma, \mathbf{\Phi}, m)\pi(\rho) \\ &\propto \pi(\rho)|\mathbf{\Phi}|^{-(N)/2} \exp\left\{-\frac{1}{2}tr\mathbf{\Phi}^{-1}\mathbf{E}\Sigma^{-1}\mathbf{E}'\right\} \end{aligned}$$

The following results concerning the determinant and the inverse of $\mathbf{\Phi}$ are useful

$$|\mathbf{\Phi}| = (1 - \rho^2)^{-1}$$

$$\begin{aligned} \mathbf{\Phi}^{-1} &= \begin{pmatrix} 1 & -\rho & & & 0 \\ -\rho & (1+\rho^2) & -\rho & & \\ & \ddots & \ddots & \ddots & \\ & & & (1+\rho^2) & -\rho \\ 0 & & & -\rho & 1 \end{pmatrix} \\ &= \mathbf{I}_T - \rho M_1 + \rho^2 M_2 \end{aligned}$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are

$$\mathbf{M}_1 = \begin{pmatrix} 0 & 1 & & & 0 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & & & 1 & 0 \end{pmatrix}$$

and

$$\mathbf{M}_2 = \begin{pmatrix} 0 & 0 & & & 0 \\ 0 & 1 & & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & 0 \\ 0 & & & 0 & 0 \end{pmatrix}.$$

The conditional posterior can then be written as

$$\begin{aligned} P(\rho|\mathbf{F}, \Sigma, \Lambda, \mathbf{R}, m) &\propto (1-\rho^2)^{N/2} \\ &\quad \exp\left\{ -\frac{1}{2\sigma_\rho^2}(\rho-\rho_0)^2 \right. \\ &\quad \left. -\frac{1}{2} tr(I_T - \rho\mathbf{M}_1 + \rho^2\mathbf{M}_2)\mathbf{E}\Sigma^{-1}\mathbf{E}' \right\} 1(\rho \in \Omega) \\ &\propto (1-\rho^2)^{N/2} \exp\left\{ -\frac{1}{2\sigma_\rho^2}(\rho-\rho_0)^2 - \frac{c_2}{2}\left(\rho - \frac{c_1}{2c_2}\right)^2 \right\} \\ &\propto (1-\rho^2)^{N/2} \exp\left\{ -\frac{\sigma_\rho^2 c_2 + 1}{2\sigma_\rho^2}(\rho - \tilde{\rho})^2 \right\} \end{aligned} \tag{3.8}$$

where $c_1 = tr(\Sigma^{-1}\mathbf{E}'\mathbf{M}_1\mathbf{E})$, $c_1 = tr(\Sigma^{-1}\mathbf{E}'\mathbf{M}_2\mathbf{E})$ and $\tilde{\rho} = \frac{\sigma_\rho^2}{\sigma_\rho^2 c_2 + 1}\left(\frac{\rho_0}{\sigma_\rho^2} + \frac{c_1}{2}\right)$. The conditional posterior for ρ does not take the form of a known distribution. However, the exp{.} term in the last row in (3.8) is recognized as the kernel of a univariate normal density. Using this fact and making draws from a normal proposal in a Metropolis-Hastings step is straightforward.

Given the number of factors, m, a Metropolis-within-Gibbs algorithm can be implemented as follows:

1. draw $\breve{\mathbf{F}}_{j+1}$ by sampling from $P(\mathbf{F}|\breve{\Lambda}_j, \breve{\Omega}_j, \breve{\rho}_j, \mathbf{R}, m)$
2. draw $\breve{\Lambda}_{j+1}$ by sampling from $P(\Lambda|\breve{\mathbf{F}}_{j+1}, \breve{\Omega}_j, \breve{\rho}_j, \mathbf{R}, m)$
3. draw $\breve{\Sigma}_{j+1}$ by sampling from $P(\Sigma|\breve{\mathbf{F}}_{j+1}, \breve{\Lambda}_{j+1}, \breve{\rho}_j, \mathbf{R}, m)$
4. draw $\breve{\rho}_{j+1}$ with a Metropolis-Hastings step from $P(\rho|\breve{\mathbf{F}}_{j+1}, \breve{\Lambda}_{j+1}, \breve{\Sigma}_{j+1}, \mathbf{R}, m)$

3.3.4 Strict Factor Structure, Allowing for Different Time Series Dependence Across Assets

In this section, we return to the strict factor structure but we will make use of the time dependence from the previous section. However, the time dependence is not necessarily the same across assets. Consider the model in vec form

$$\begin{aligned} r &= (\mathbf{I}_N \otimes \mathbf{F})vec(\Lambda') + \varepsilon \\ r &= Z\lambda + \varepsilon \end{aligned} \tag{3.9}$$

where $\varepsilon \sim N(0, \Omega)$ and Ω is a $TN \times TN$ covariance matrix given by

$$\Omega = \begin{pmatrix} \sigma_{11}^2 \mathbf{\Phi}_1 & 0 & \dots & 0 \\ 0 & \sigma_{22}^2 \mathbf{\Phi}_2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_{NN}^2 \mathbf{\Phi}_N \end{pmatrix}$$

and the T by T matrices $\mathbf{\Phi}_i, i = 1, \dots, N$ are given by

$$\mathbf{\Phi}_i = \frac{1}{1-\rho_i^2} \begin{pmatrix} 1 & \rho_i & \rho_i^2 & \dots & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & \dots & \rho_i^{T-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \dots & 1 \end{pmatrix}$$

where $|\rho_i| < 1$.

The prior structure is the same as in the strict factor structure and the prior for $\rho_i, i = 1, \dots, N$ is the truncated normal density

$$\pi(\rho_i) \propto f_N(\rho_i | \rho_0, \sigma_\rho^2) 1(\rho \in \Omega).$$

The conditional posterior for the factor score and the factor loadings follows from the previous section by replacing $(\Sigma \otimes \mathbf{\Phi})$ with Ω. This yields

$$\begin{aligned} P(f|\Lambda, \Sigma, \Omega, \mathbf{R}, m) &\propto L(\mathbf{R}|F, \Lambda, \Omega, m)\pi(F|m) \\ &\propto \exp\left\{ -\frac{1}{2}(\mathbf{f} - \tilde{\mathbf{f}})[(\Lambda \otimes \mathbf{I}_T)'\Omega^{-1}(\Lambda \otimes \mathbf{I}_T) + \mathbf{I}_{Tm}](\mathbf{f} - \tilde{\mathbf{f}}) \right\} \end{aligned}$$

where

$$\tilde{\mathbf{f}} = ((\Lambda \otimes \mathbf{I}_T)'\Omega^{-1}(\Lambda \otimes \mathbf{I}_T) + \mathbf{I}_{Tm})^{-1}(\Lambda \otimes \mathbf{I}_T)'\Omega^{-1} r$$

and

$$\begin{aligned} P(\Lambda|\mathbf{F}, \Omega, \mathbf{R}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Omega, m)\pi(\Lambda|m) \\ &\propto \exp\left\{ -\frac{1}{2}(\lambda - \tilde{\lambda})'(Z'\Omega^{-1}Z + (\mathbf{I}_N \otimes \mathbf{H}_0)^{-1})(\lambda - \tilde{\lambda}) \right\} \end{aligned} \tag{3.10}$$

where

$$\tilde{\lambda} = (Z'\Omega^{-1}Z + \mathbf{I}_N \otimes \mathbf{H}_0^{-1})^{-1}(Z'\Omega^{-1}r + (\mathbf{I}_N \otimes \mathbf{H}_0^{-1})\lambda_0).$$

That is, the factor scores and factor loadings have normal conditional posterior distributions.

The conditional posterior of σ_{ii}^2 is

$$\begin{aligned} P(\sigma_{ii}^2|\mathbf{F}, \Lambda, \mathbf{\Phi}_i, \mathbf{R}, m) &\propto L(\mathbf{R}|\mathbf{F}, \Lambda, \Omega, m)\pi(\sigma_{ii}^2) \\ &\propto |\Omega|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}tr\Omega^{-1}[\varepsilon\varepsilon']\right\} \\ &\times \left\{\frac{1}{\sigma_{ii}^2}\right\}^{\frac{v_0+2}{2}} \exp\left\{-\frac{1}{2\sigma_{ii}^2}s_0\right\} \\ &\propto \left\{\frac{1}{\sigma_{ii}^2}\right\}^{\frac{T+v_0+2}{2}} \exp\left\{-\frac{1}{2\sigma_{ii}^2}(\varepsilon_i'\mathbf{\Phi}_i^{-1}\varepsilon_i + s_0)\right\} \end{aligned} \tag{3.11}$$

where ε_i column i in $\mathbf{E} = [\varepsilon_1, ..., \varepsilon_N]$. This is the form of an inverse Gamma-2.

Finally, we derive the conditionals for $\rho_i, i = 1, ..., N$. Note that the likelihood viewed as a function of ρ_i can be written as

$$L(\mathbf{R}|\mathbf{F}, \Lambda, \Omega, m) \propto |\mathbf{\Phi}_i|^{-1/2} \exp\left\{-\frac{1}{2\sigma_{ii}^2}\varepsilon_i'\mathbf{\Phi}_i^{-1}\varepsilon_i\right\}.$$

Combining the likelihood with the prior and using the special structure of $\mathbf{\Phi}_i$ yields the following conditional posterior density

$$\begin{aligned} P(\rho_i|\mathbf{F}, \Sigma, \Lambda, \mathbf{R}, m) \propto &(1 - \rho_i^2)^{1/2} \\ &\exp\left\{-\frac{1}{2\sigma_\rho^2}(\rho_i - \rho_0)^2\right. \\ &\left.- \frac{1}{2\sigma_{ii}^2}tr(I_T - \rho_i\mathbf{M}_1 + \rho_i^2\mathbf{M}_2)\varepsilon_i\varepsilon_i'\right\} 1(\rho \in \Omega). \end{aligned}$$

As in the previous section, this can be written as a factor times the kernel of a normal density. A Metropolis-Hastings step can then be used in a similar way as in the previous section to generate draws from the conditional posterior distribution of ρ_i.

Given the number of factors, m, a Metropolis-within-Gibbs algorithm is implemented as follows:

1. draw $\breve{\mathbf{F}}_{j+1}$ by sampling from $P(\mathbf{F}|\breve{\Lambda}_j, \breve{\Omega}_j, \mathbf{R}, m)$

2. draw $\breve{\Lambda}_{j+1}$ by sampling from $P(\Lambda|\breve{\mathbf{F}}_{j+1}, \breve{\Omega}_j, \mathbf{R}, m)$
3. draw $\breve{\sigma}^2_{ii,j+1}$ by sampling from $P(\sigma^2_{ii}|\breve{\mathbf{F}}_{j+1}, \breve{\Lambda}_{j+1}, \breve{\mathbf{\Phi}}_j, \mathbf{R}, m)$, $i = 1, \dots, N$.
4. draw $\breve{\rho}_{i,j+1}$ with a Metropolis-Hasting step from $P(\rho_i|\breve{\mathbf{F}}_{j+1}, \breve{\Lambda}_{j+1}, \breve{\sigma}_{ii,j+1}, \mathbf{R}, m)$, $i = 1, \dots, N$.

3.3.5 Identification

Nothing has been said about the well-known identification problem in the factor model. That is, that the model is invariant under transformation of the form $\Lambda^* = \Lambda P'$ and $f_t^* = P f_t$, where P is any orthogonal matrix. As in Geweke and Zhou (1996) and Lopes and West (2004), identification is obtained by restricting Λ^m, the upper $m \times m$ submatrix of Λ, to be lower triangular with positive elements on the diagonal.

Under the identification condition, all parameters, except for Λ, have the same posterior distributions as before. Imposing the restriction on Λ^m implies that the free elements of Λ are multivariate normal conditional on the zero elements. In the cases where Σ is diagonal it turns out that the rows of Λ are independent and conditioning on the zero elements in the same row is sufficient. Using that $\mathbf{H}_0 = h_0 \mathbf{I}_m$, the full conditional for the free elements in the first m rows is given by

$$P(\lambda^{i*}|\lambda^{i+}, \mathbf{F}, \Sigma, \mathbf{R}, m) \propto \exp\left\{-\frac{1}{2}(\lambda^{i*} - \tilde{\lambda}^{(i*)})'(\mathbf{F}_i'\mathbf{F}_i/\sigma^2_{ii} + \mathbf{H}_0^{-1})(\lambda^{i*} - \tilde{\lambda}^{(i*)})\right\}$$

where $\lambda^{(i*)} = (\lambda_{i,1}, ..., \lambda_{i,i})'$ $i = 1, ..., m$,

$$\tilde{\lambda}^{(i*)} = (\mathbf{F}_i'\mathbf{F}_i/\sigma^2_{ii} + \mathbf{I}h_0^{-1})^{-1}(\mathbf{F}_i' r^{(i)}/\sigma^2_{ii} + \mathbf{I}h_0^{-1})\lambda_0^{(i*)}),$$

$\mathbf{F}_i$ is column (i) in $\mathbf{R}$, λ^{i+} contains the element in row i in Λ equal to zero, and finally, $\lambda_0^{(i*)}$ is formed in a similar way as $\lambda^{(i*)}$. The posterior distribution for the factor loadings in the other cases where Σ is diagonal can be derived in a similar way.

In an approximate factor structure it is slightly more complicated. The reason is that the posterior covariance matrix for the conditional distribution for the factor loadings does not take the form of a block diagonal matrix. Instead, to generate the factor loadings from the distribution in (3.6) under the identification conditions we have to sample from the conditional distribution of $\lambda^{(i*)}$, conditioning not only on the elements restricted to be zero in row i in Λ, but also on the previous $i - 1$ rows. Since the conditional posterior distribution of Λ is normal the draws are straightforward to generate.

3.4 Bayesian Model Selection

Let us assume that there are L competing models $\{M_l, l = 1, ..., L\}$ under consideration. For model M_l, the posterior distribution takes the form

$$\pi(\theta_l|D, M_l) \propto L(\theta_l|D, M_l)\pi(\theta_l|M_l) \tag{3.12}$$

where $L(\theta_l|D, M_l)$ is the likelihood, D denotes the data, and $\pi(\theta_l|M_l)$ is the prior distribution. Then the marginal likelihood is given by

$$m(D|m_l) = \int L(\theta_l|D, M_l)\pi(\theta_l|M_l)d\theta_l. \tag{3.13}$$

To compare different models, we calculate the marginal likelihood $m(D|m_l)$ for $l = 1, ..., L$ and choose the model, which yields the largest marginal likelihood. Alternatively, a model comparison can be conducted through the use of Bayes factors. The Bayes factor for M_i versus M_j is given by

$$B_{ij} = \frac{m(D|M_i)}{m(D|M_j)} = \frac{\int L(D|\theta_i, M_i)p(\theta_i|M_i)d\theta_i}{\int L(D|\theta_j, M_j)p(\theta_j|M_j)d\theta_j}.$$

and measures how much our belief in M_i relative M_j has changed after viewing the data. If prior probabilities $P(M_l)$, $l = 1, ..., L$, of the models are available, the Bayes factor can be used to compute the posterior model probabilities

$$P(M_i|\mathbf{D}) = \frac{m(D|M_i)P(M_i)}{\sum_{l=1}^{L} m(D|M_l)P(M_l)} = \left[\sum_{j=1}^{L} \frac{P(M_j)}{P(M_i)} B_{ji}\right]^{-1}.$$

The marginal likelihood $m(D|m_l)$ in (3.13) is the inverse normalizing constant of the posterior distribution $\pi(\theta_k|D, M_l)$ in (3.12). Using (3.12) and (3.13), Chib (1995) obtains the following identity

$$m(D|M_l) = \frac{L(\theta_l|D, M_l)\pi(\theta_l|M_l)}{\pi(\theta_l|D, M_l)}. \tag{3.14}$$

Let θ_l^* denote the posterior mean or the posterior mode with respect to the posterior distribution $\pi(\theta_l|D, M_l)$.[1] Since the identity in (3.14) holds for any θ_l we have

$$m(D|M_l) = \frac{L(\theta_l^*|D, M_l)\pi(\theta_l^*|M_l)}{\pi(\theta_l^*|D, M_l)}. \tag{3.15}$$

---

[1] In principal, any θ_l^* can be chosen but the posterior mean or the posterior mode is often used to ensure numerical stability.

Computationally, it is more efficient to compute $\ln[m(D|M_l)]$ instead of directly computing $m(D|M_l)$,

$$\ln[m(D|M_l)] = ln[L(\theta_l^*|D, M_l)] + \ln[\pi(\theta_l^*|M_l)] - \ln[\pi(\theta_l^*|D, M_l)]. \tag{3.16}$$

It is usually straightforward to compute $L(\theta_l^*|D, M_l)$ and $\pi(\theta_l^*|M_l)$ whereas the difficulty in computing $\ln[m(D|M_l)]$ is $\ln[\pi(\theta_l^*|D, M_l)]$. There are several ways to calculate this posterior density ordinate. However, Chib (1995) and Chib and Jeliazkov (2001) present a method where $\ln[\pi(\theta_l^*|D, M_l)]$ is estimated using the output from a Gibbs sampler and a Metropolis-Hastings sampler, respectively.

In our case the posterior inference, conditioning on the number of factors, m, is made through Gibbs sampling and a Metropolis-within-Gibbs sampler, which makes Chib's method very suitable for estimating the marginal likelihood. In the first two cases, where the Gibbs sampler is used for making draws from the posterior, the posterior ordinate is $\pi(\Lambda^*, \Sigma^*|m, \mathbf{R}) = \pi(\Lambda^*|m, \mathbf{R})\pi(\Sigma^*|m, \Lambda, \mathbf{R})$. The first term in the right hand side is estimated by using draws from the full Gibbs sampler and averaging over the full conditionals of $\mathbf{F}$ and Σ. The second term is estimated by running an additional reduced Gibbs sampler, with Λ fixed at its posterior mean.

In case 3 and 4 where I allow for some time series dependence, the Metropolis-within-Gibbs is used for making draws from the posterior. The posterior ordinate is

$$\pi(\Lambda^*, \Sigma^*, \rho^*|m, \mathbf{R}) = \pi(\rho^*|m, \mathbf{R})\pi(\Lambda^*|m, \rho, \mathbf{R})\pi(\Sigma^*|m, \rho, \Lambda, \mathbf{R}).$$

Since the normalizing constant of $\pi(\rho|m, \mathbf{R})$ is not known it is estimated by using the output from the MCMC sampler as outlined in Chib and Jeliazkov (2001). The other two posterior ordinates are estimated in a similar way as in case 1 and 2. However, note that by placing ρ first in the decomposition of the posterior ordinate the reduced runs does not involve any Metropolis Hastings steps. Finally, in case 4, where the time series dependence can differ across assets, the ρ_i, $i = 1, \ldots, N$ are sampled in one block.

3.5 Empirical Results

The data in this study contains monthly observations on US stock excess returns from July 1963 through May 2005. The test asset consists of 17 and 30 industry portfolios and the 25 Fama and French (1993) portfolios formed on size and book-to-market. The portfolios include all NYSE, AMEX, and

NASDAQ stocks.[2] Each asset has been standardized with respect to its sample mean and standard deviation. This does not alter the factor structure analysis.

Before any analysis can be done the hyperparameters must be assessed. More specifically, hyperparameters in the inverse Gamma-2 prior are chosen to be $v_0 = 3$ and $s_0 = 1$ and in the inverse Wishart prior $v_0 = N + 2$ and $\mathbf{S}_0 = \mathbf{I}_n$. This implies that the prior mean of σ^2 and Σ are equal to 1 and $\mathbf{I}_n$ respectively. The factor loadings are assumed to be normal independent distributed with mean zero and variance $h_0 = 1$. The prior distribution for the ρ and ρ_i $i = 1, \ldots, N$ is given by the normal distribution with zero mean and variance 0.09.[3] Finally, all the models are equally probable before we have seen any data. For the MCMC algorithm we take 10 000 iterations as burn-in and save 30 000 replications for inference.

Tables 3.1a to 3.1c summarize the results for the three portfolios. We also address the sensitivity with respect to the sample period by considering two subsamples, 196307 - 198312 and 198401 - 200412.

Table 3.1a shows the results for the 17 industry portfolios. The number of factors varies both between the subsamples and the underlying assumptions regarding the idiosyncratic returns. In general, the number of factors is higher in the strict factor structure where 5 to 6 pervasive factors are found. Introducing time series dependence seems to reduce the number of factors. In both case 3 and 4, 4 to 5 common factors are found. The difference between the the two subsample is most substantial for Case 4, where 4 factor is found in the first period while only 2 factors is found in the second.

The results for the 30 industry portfolios are displayed in Table 3.1b. Generally six to seven common factors are found, which is one more than for the 17 industry portfolios. The results for Case 4, the strict factor model, where the time series dependence can differ across assets, are very mixed. In the first sample period, 5 factors are found whereas seven and four factors are found in the second period and the whole sample respectively.

Finally, Table 3.1c shows the results for the 25 size and book-to-market portfolios. The difference between the sample periods and the four cases is larger than for the industry portfolios. Firstly, the number of factors is generally higher in the second period. Secondly, there is substantial difference in the number of factors between the four cases. Again, introducing time series dependence seems to reduce the number of factors.

---

[2] We thank Kenneth R. French for providing the data at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[3] The results seems to be insensitive to the chosen values.

Table 3.1a: Posterior Model Probabilities: 17 Industry portfolios

	196307 - 200412			
Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00
4	0.00	0.00	0.90	0.66
5	0.40	0.00	0.00	0.32
6	0.00	1.00	0.00	0.01
7	0.39	0.00	0.00	0.00
8	0.13	0.00	0.09	0.01
9	0.08	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00
	196307 - 198312			
Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00
4	0.00	0.00	0.30	1.00
5	0.00	0.55	0.68	0.00
6	0.85	0.42	0.01	0.00
7	0.05	0.01	0.01	0.00
8	0.06	0.01	0.00	0.00
9	0.00	0.00	0.00	0.00
10	0.04	0.01	0.00	0.00
	198401 - 200412			
Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	1.00
3	0.00	0.00	0.00	0.00
4	0.00	0.00	0.34	0.00
5	0.74	0.93	0.66	0.00
6	0.00	0.06	0.00	0.00
7	0.00	0.00	0.00	0.00
8	0.24	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00

Case 1: The first specification follows from the strict factor structure where Σ is assumed to be diagonal. Case 2: The second follows from the approximate factor model where the assumption of a diagonal covariance matrix is relaxed. Case 3: In the third specification, time series dependence is introduced to the approximate factor model. Case 4: In the final specification we return to the strict factor model where the time series dependence can differ across assets.

Table 3.1b: Posterior Model Probabilities: 30 Industry portfolios

196307 - 200412

Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.01	0.00	0.00	0.01
4	0.00	0.00	0.00	0.68
5	0.00	0.00	1.00	0.00
6	0.01	0.64	0.00	0.02
7	0.97	0.31	0.00	0.05
8	0.00	0.04	0.00	0.02
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.23

196307 - 198312

Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.01
4	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.97
6	1.00	0.54	0.97	0.00
7	0.00	0.44	0.01	0.00
8	0.00	0.01	0.01	0.01
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00

198401 - 200412

Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.00	0.00	0.01	0.00
4	0.00	0.00	0.99	0.00
5	0.00	0.01	0.00	0.00
6	0.00	0.97	0.00	0.00
7	0.70	0.02	0.00	0.97
8	0.00	0.00	0.00	0.03
9	0.21	0.00	0.00	0.00
10	0.08	0.00	0.00	0.00

Case 1: The first specification follows from the strict factor structure where Σ is assumed to be diagonal. Case 2: The second follows from the approximate factor model where the assumption of a diagonal covariance matrix is relaxed. Case 3: In the third specification, time series dependence is introduced to the approximate factor model. Case 4: In the final specification we return to the strict factor model where the time series dependence can differ across assets.

Table 3.1c: Posterior Model Probabilities: 25 Size and Book-to-Market portfolios

	196307 - 200412			
Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.00	0.00	0.99	0.00
4	0.00	0.81	0.00	1.00
5	0.15	0.18	0.01	0.00
6	0.38	0.00	0.00	0.00
7	0.42	0.00	0.00	0.00
8	0.05	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00
	196307 - 198312			
Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	1.00
3	0.00	0.00	0.00	0.00
4	0.00	0.00	0.94	0.00
5	0.00	0.00	0.06	0.00
6	0.00	1.00	0.00	0.00
7	0.42	0.00	0.00	0.00
8	0.58	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00
	198401 - 200412			
Factors (m)	Case 1	Case 2	Case 3	Case 4
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	1.00
6	0.98	0.03	1.00	0.00
7	0.01	0.97	0.00	0.00
8	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00

Case 1: The first specification follows from the strict factor structure where Σ is assumed to be diagonal. Case 2: The second follows from the approximate factor model where the assumption of a diagonal covariance matrix is relaxed. Case 3: In the third specification, time series dependence is introduced to the approximate factor model. Case 4: In the final specification we return to the strict factor model where the time series dependence can differ across assets.

Table 3.2: Posterior simulation results for ρ, Case 3

17 Industry portolios

Time period	Mean	St.dev	Spectral NSE	RNE
196307-200412	0.0169	0.0178	0.0002	2.1723
196307-198312	0.0384	0.0188	0.0002	2.7037
198401-200412	-0.0100	0.0182	0.0002	2.5499

30 Industry portfolios

Time period	Mean	St.dev	Spectral NSE	RNE
196307-200412	-0.0052	0.0097	0.0001	2.8851
196307-198312	0.0108	0.0142	0.0001	2.6574
198401-200412	-0.0321	0.0137	0.0001	2.7488

25 size and book-to-market portfolios

Time period	Mean	St.dev	Spectral NSE	RNE
196307-200412	-0.0083	0.0100	0.0001	1.6481
196307-198312	-0.0295	0.0148	0.0001	1.9498
198401-200412	0.0015	0.0154	0.0001	2.2081

One major advantage of the Bayesian approach is that model uncertainty is easily quantified. For example, in the strict factor structure in Table 3.1a, the best model contains 5 factors with a posterior probability of 0.4 whereas the second best model contains 7 pervasive factors with a posterior probability of 0.39. Hence, the model uncertainty is very substantial. In contrast, the posterior model probability for the best model in case 2 and 3 is 1.0 and 0.9 respectively, which indicates that the data is informative about the number of factors. The model uncertainty seems to be higher for the industry portfolios than for the 25 size and book-to-market portfolios except for the strict factor structure where the uncertainty about the number of factors is more substantial for all portfolios.

In Case 3, time series dependence is introduced in the approximate factor model by letting the error term follow an AR(1) process. Table 3.2 contains posterior results for the parameter ρ. The time series dependence seems to be more substantial during the first subsample. Note that the posterior means are sometimes negative and sometimes positive. For the 17 industry portfolios the posterior mean of the AR(1) coefficient is negative in the later subperiod whereas they are negative for the whole sample and the second period when the 30 industry portfolios is examined. Finally, the posterior means are negative in the first and the whole sample period when the 25 size and book-to-market portfolios is considered. On the other hand, a 90% highest posterior density region would cover zero in most cases.

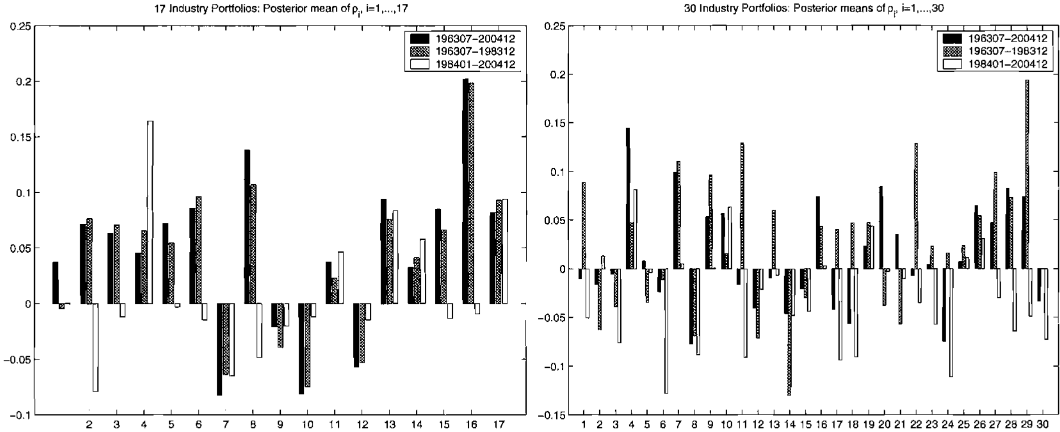
The property of the Markov chains for ρ is also shown in Table 3.2 and Figures A.1 to A.3.[4] The two last columns in Table 3.2 contains estimates of the numerical standard error and the relative numerical efficiency (RNE) using the spectral estimator. The RNEs for the industry portfolios indicate that we need 2-3 times as many draws from the sampler as when sampling directly from the posterior. The higher RNEs for the industry portfolios can be explained by the sample autocorrelation functions displayed in Figure A.1 and A.2. The autocorrelation is low but slowly decaying. The first graph in the figures displays Geweke diagnostics. Convergence implies that the calculated statistics should be within the two lines. The overall assessment is that the chains seem to have converged.

In the final specification, , Case 4, a strict factor model is assumed where the time series dependence can differ across assets. Figure 3.1 displays the posterior means of ρ_i, $i = 1, ..., N$. The posterior means are very different across assets. Especially for the 30 industry portfolios. However, the posterior standard deviations, not reported but available from the author on request, indicates that a 90% highest posterior density region would again cover zero in most cases.

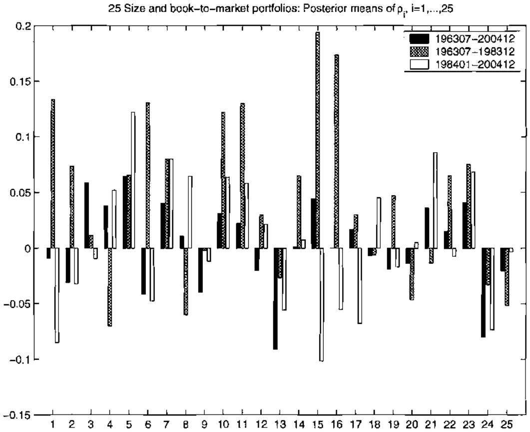
In a Bayesian approach it is easy to make comparisons between the different assumptions regarding the idiosyncratic term. Table 3.3 presents the comparison between the four cases. The marginal likelihoods are expressed in log format. The natural log of the Bayes factor is the difference between the log marginal likelihood for the best model in Case 2, 3 and 4 respectively and the log marginal likelihood for the best model assuming a strict factor structure, Case 1. The results provide clear evidence in favor of an approximate factor structure with time series dependence through a common AR(1) process across assets since the log Bayes factors are highest for case 3. Exceptions are found in the 30 industry portfolios during the second subperiod and in the 25 size and book-to-market portfolios in the first subperiod where the data favors the strict factor structure.

---

[4] Due to the large number of parameters the diagnostics for the other parameters are not reported but are available on request.



(a) 17 Industry Portfolios (b) 30 Industry Portfolios



(c) 25 Size and book-to-market Portfolios

Figure 3.1: Posterior means of ρ_i, $i = 1, ..., N$, Case 4.

Table 3.3: Bayes factors: Case 2, 3 and 4 vs Case 1

17 Industry portolios

Time period	Case 2	Case 3	Case 4
196307-200412	74.30	1081.70	6.40
196307-198312	109.10	464.57	203.27
198401-200412	288.17	438.17	29.60

30 Industry portfolios

Time period	Case 2	Case 3	Case 4
196307-200412	129.60	911.50	23.00
196307-198312	69.60	1143.12	1087.67
198401-200412	396.50	316.10	96.00

25 size and book-to-market portfolios

Time period	Case 2	Case 3	Case 4
196307-200412	659.10	1514.40	206.90
196307-198312	375.70	311.32	104.83
198401-200412	176.30	274.30	200.40

The table shows the natural log of the Bayes factor of Case 2 to 4 versus Case 1. Case 1: The first specification follows from the strict factor structure where Σ is assumed to be diagonal. Case 2: The second follows from the approximate factor model where the assumption of a diagonal covariance matrix is relaxed. Case 3: In the third specification, time series dependence is introduced to the approximate factor model. Case 4: In the final specification we return to the strict factor model where the time series dependence can differ across assets.

3.6 Summary and Conclusions

In this chapter a Bayesian framework is presented for examining the number of factors in a multifactor pricing when the factors are unobserved. The determination of the number of factors is viewed as a model selection problem. Furthermore, the assumption of a strict factor structure is relaxed and by letting the error term follow an AR(1) process some time series dependence has been allowed for.

Using industry portfolios and portfolios formed on size and book-to-market 4 to 6 pervasive factor were generally found. It seems like that when time series dependence is introduced, the number of factors decreases. The data points to an approximate factor structure with time series dependence through a common AR(1) process across assets.

The number of factors found is consistent with what other has found. Connor and Korajczyk (1993) argue for a 3 to 6 factor model using monthly returns NYSE and AMEX stocks over the period 1967 to 1991. In contrast to other studies, this study does not only select the number of factor, but also address model uncertainty. In particular, the model uncertainty is quite substantial when a strict factor structure is assumed.

Bibliography

BAI, J., AND S. NG (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221.

BAUWENS, L., M. LUBRANO, AND J. RICHARD (2000): *Bayesian Inference in Dynamic Econometric Models.* Oxford University Press.

CHAMBERLAIN, G., AND M. ROTHSCHILD (1983): "Arbitrage, Factor Structure, and Mean Variance Analysis on Large Asset Markets," *Econometrica*, 51, 1281–1304.

CHIB, S. (1995): "Marginal Likelihood from the Gibbs output," *Journal of the American Statistical Association*, 90, 1313–1321.

CHIB, S., AND I. JELIAZKOV (2001): "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270–281.

CONNOR, G., AND R. KORAJCZYK (1988): "Risk and Return in an Equilibrium APT: Application of a New Test Methodology," *Journal of Financial Economics*, 21, 255–289.

——— (1993): "A Test for the Number of Factors in an Approximate Factor Model," *Journal of Finance*, 48, 1263–1291.

FAMA, E., AND K. FRENCH (1993): "Common Risk Factors in the Returns of Stock and Bonds," *Journal of Financial Economics*, 33, 3–56.

GEWEKE, J., AND G. ZHOU (1996): "Measuing the Pricing Error of the Arbitrage Pricing Theory," *The Review of Financial Studies*, 2, 557–587.

INGERSOLL, J. E. (1984): "Some results in the Theory of Arbitrage Pricing," *Journal of Finance*, 39, 1021–1039.

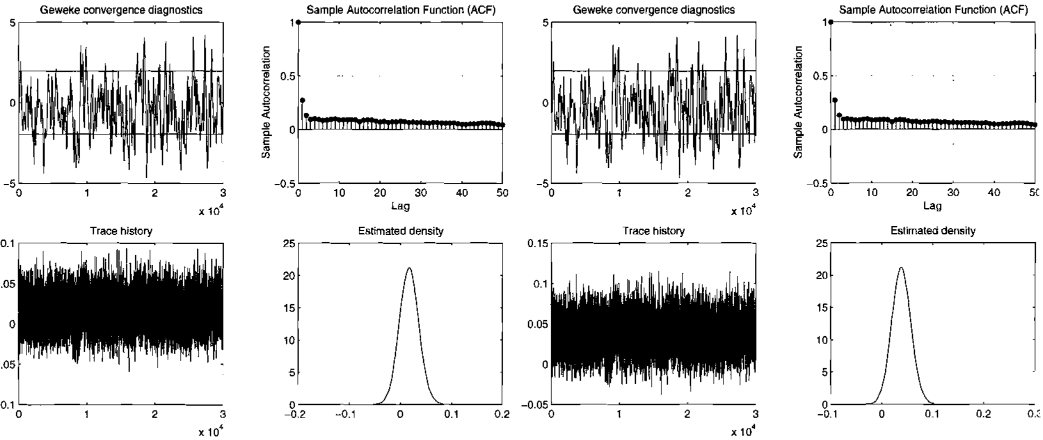
LEHMANN, B., AND D. MODEST (1988): "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics*, 21, 213–254.

LOPES, H, F., AND M. WEST (2004): "Bayesian Model Assessment in Factor Analysis," *Statistics Sinica*, 14, 41–67.

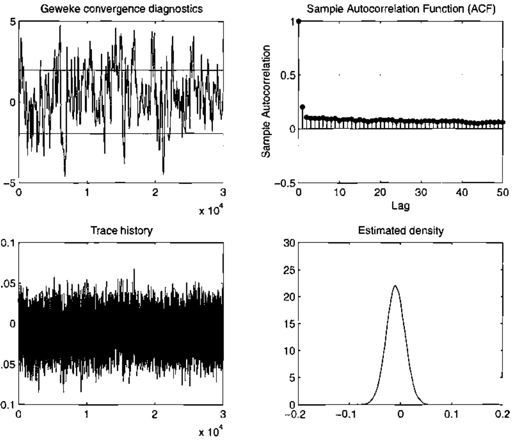
ROSS, S. A. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

Appendix A

Figures

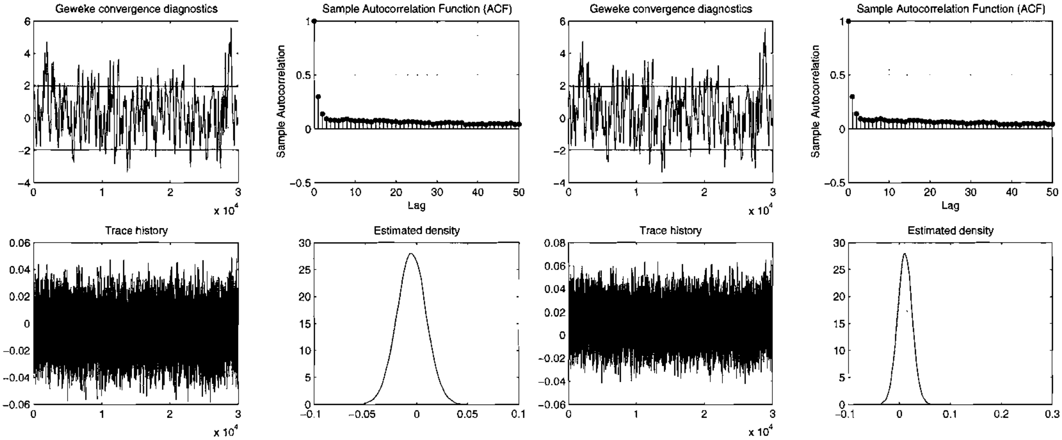


(a) 196307-200412 (b) 196307-198312

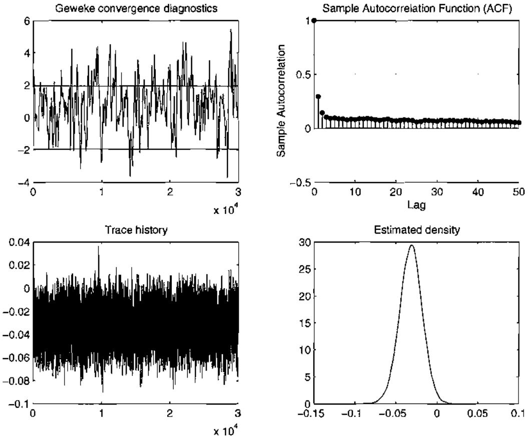


(c) 198401-200412

Figure A.1: 17 Industry portfolios: Markov chain properties for ρ in the approximate factor structure with time series and cross-sectional dependence (Case 3).

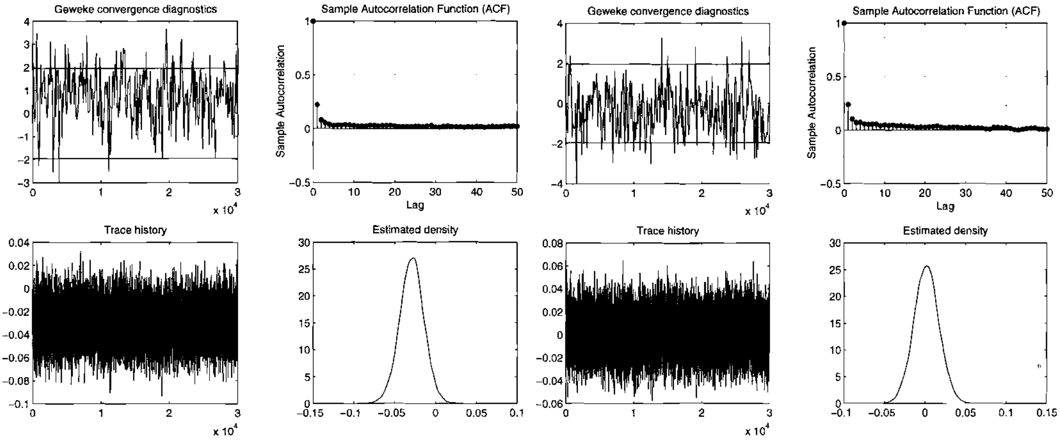


(a) 196307-200412 (b) 196307-198312

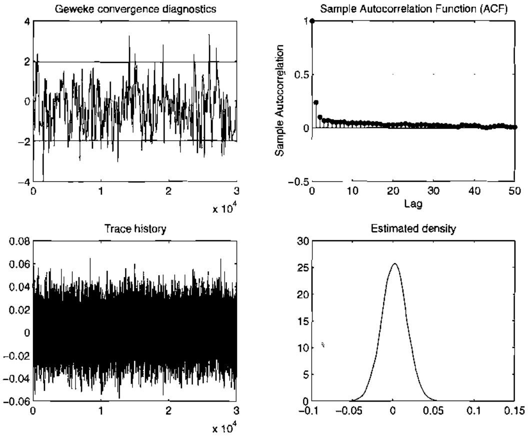


(c) 198401-200412

Figure A.2: 30 Industry portfolios: Markov chain properties for ρ in the approximate factor structure with time series and cross-sectional dependence (Case 3).



(a) 196307-200412 (b) 196307-198312



(c) 198401-200412

Figure A.3: 25 size and book-to-market portfolios: Markov chain properties for ρ in the approximate factor structure with time series and cross-sectional dependence (Case 3).

Chapter 4

Is Momentum Due to Data-Snooping?

Acknowledgement: We would like to thank Sune Karlsson and the seminar participants at the Department of Finance at the Stockholm School of Economics, the Workshop in Econometrics and Computational Economics in Helsinki 2003, and the Forecasting Financial Markets Conference in Paris 2004, for their valuable comments and suggestions. Additionally, we would like to thank Peter Schotman for suggesting the idea for the paper.

4.1 Introduction

In efficient markets, asset prices are assumed to fully reflect all available information in the market. Consequently, it should be impossible to earn risk adjusted abnormal returns by exploiting investment strategies based on past information. However, several empirical studies suggest that past returns are powerful predictors of future stock returns. For example, Jegadeesh and Titman (1993) document that over a span of three to 12 months, past winners continue to outperform past losers by about 1% per month on average, showing that there is "momentum" in stock prices. They state that one explanation could be that investors, who follow a momentum strategy temporarily, move prices away from their long run values. Conrad and Kaul (1998) find similar results as Jegadeesh and Titman (1993). These papers are all based on US stocks. Similar evidence in favor of the momentum effect has been documented for the European market by Rouwenhorst (1998) who shows that abnormal returns to momentum strategies could be found in twelve European markets. Evidence on momentum has been found for emerging markets by Rouwenhorst (1999) and van der Hart, Slagter, and van Dijk (2001). Recently, several authors have tried to link momentum to other factors than firm specific ones. Moskowitz and Grinblatt (1999) ague that momentum can be traced to industry factors and Lewellen (2002) extends the analysis to size and book-to-market sorted portfolios.

While the momentum effect has been well documented, the cause of momentum is still an open issue. Some have argued that the results provide strong evidence of market inefficiency and others have argued that returns from momentum strategies are compensation for risk. Finally, some claim that the profit obtained from momentum strategies is the product of data-snooping. The effect of data-snooping is probably the hardest to address since empirical research is limited by data availability.

An important issue when evaluating a large set of trading rules is data-snooping. As argued by Lo and MacKinlay (1990), the data-snooping bias can be substantial in financial studies. Data-snooping occurs when the same set of data is used more than once for inference or model selection. The problem of data-snooping has been mentioned in several papers. Savin (1984) and Lakonishok and Smidt (1984) have remarked that the actual size of a t-test that follows a search for the largest possible t-statistics can be very different from its nominal size. To address the question of data mining in momentum strategies Jegadeesh and Titman (2001) test for momentum using an extended data set. They find that momentum strategies continue to be profitable at about the same magnitude as in the earlier period.

The main purpose of our paper is to extend and enrich the existing research on momentum strategies by applying a procedure that permits us to ascertain whether momentum is a product of data-snooping or a result of market inefficiency. Hence, we investigate if a momentum strategy is superior to a benchmark model once the effects of data-snooping have been accounted for. This procedure is known as the "Reality Check" which was devised by White (2000) and utilized in Timmermann, White, and Sullivan (1998) to evaluate simple technical trading rules. A problem associated with White's Reality Check is that the power of the test is sensitive to the inclusion of a poor model. This issue is addressed by Hansen (2004) who proposed a modified version of White's test. In our paper we also implement Hansen's modification.

As noted above, many studies of momentum and weak market efficiency have been conducted on US data. Since the momentum effect is well documented on US stock returns we will follow the recent trends and instead consider portfolios formed on industries, size, book-to-market and size/book-to-market. In contrast to the US studies, the evidence on the Swedish stock market is limited.[1] Therefore, this paper also examines the momentum effect on Swedish stock returns and portfolios formed on size, book-to-market and industries. To our knowledge, this is the first time the momentum effect is investigated using all stocks listed on the Stockholm Stock Exchange.

This chapter is set out as follows. In Section 4.2, the data is presented and Section 4.3 describes the momentum strategies that we follow in formulating trading rules. In Section 4.4, White's Reality Check and the Hansen's modification are explained. In Section 4.5, our results are discussed and, in Section 4.6, a conclusion is given.

4.2 Data

4.2.1 US Data

The US data consists of all NYSE, AMEX, and NASDAQ stocks in the "CRSP" database. The analysis considers the period July 1963 to December 2004. Furthermore, the tests are performed on portfolios formed on industry, size, book-to-market and size/book-to-market. Returns on portfolios

---

[1] The only study we are aware of is Rouwenhorst (1998). However, Rouwenhorst (1998) uses a limited set of stocks from the Stockholm Stock Exchange.

were kindly provided by Kenneth French.[2]

Table A.1a in the Appendix reports summary statistics for 17 industry portfolios, 10 size and 10 book-to-market portfolios and the 25 size/book-to-market portfolios. The average returns for the industry portfolios range from 0.699% to 1.153% resulting in a annualized spread of 5.448%. An F-test of equal returns across industries is not rejected, suggesting there is little cross-sectional variation in the industry sample means. The average returns for the book-to-market and the size portfolios range from 0.815% to 1.299% and 0.867% to 1.180%, which result in annualized spreads of 5.808% and 3.756%, respectively. Again, the null hypothesis of equal returns are not rejected.

The cross-sectional variation is larger in the size/book-to-market portfolios. The annualized spread is 11.208%. However, the F-test of equal mean returns is not rejected.

4.2.2 Swedish Data

The Swedish data consists of monthly observations on Swedish stock returns over the period January 1979 to December 2003. The data is collected from the "Trust" database and the sample includes all stocks listed on the Stockholm Stock Exchange. The returns are corrected for dividends and capital changes such as splits. However, a stock must have been traded for at least a period of 24 months to be eligible. To avoid double counting in cases where companies listed both voting and limited voting shares, the one that is most actively traded is taken. The long time period covered and the large proportion of listed companies and the explicit inclusion of delisted companies should make the results robust to possible sample bias.

The portfolios are formed on book value to market value (BM), size (ME) and industries. Size is measured by the market value, price per share times shares outstanding, and the book value is the total value of stockholder's equity. The book value used to form portfolios in June of year t is from the fiscal year ending in the calender year $t-1$ divided by the market value at the end of December of $t-1$. The market value used to form size portfolios in June of year t is the market value at the end of June of year t. Ten portfolios are formed so stocks with the lowest and highest 10 percent value of the attribute (BM or ME) are assigned to the portfolios Low and High respectively. Equally weighted returns are then calculated from July to the following June.[3] The

---

[2] A description of the data obtained from Kenneth French can be found at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[3] Using value-weighted portfolios yields the same results

industry portfolios are value-weighted and cover the period 1977-1997 and include all shares from the so-called "A1-listan" excluding banks and financial firms.[4]

Table A.2 in the Appendix shows the descriptive statistics for the Swedish data. Starting with the size portfolios, we can see the common size effect, that is, the Low size portfolio generates a higher return than the High size portfolio. The annualized spread between the High and Low portfolio is 13.140 percent. Performing an F-test results in a p-value equal to 0.72 and the null of equal average returns is not rejected. The returns for the book-to-market and the industry portfolios are lower. The annualized spreads are 10.176 percent and 10.476 percent respectively. The null of equal average returns is once again not rejected.

4.3 Momentum Strategies

As in most of the literature on momentum our research is based on the methodology of the original work of Jegadeesh and Titman (1993). This consists of identifying "Winner" and "Loser" portfolios according to their past performance.

To form the momentum strategies we have to specify the length of the period over which we rank the assets k, the proportion of winners and losers to include q, and the length of the period for holding the selected assets l. Letting $k = k_i$, $i = 1, \ldots, K$, $l = l_i$, $i = 1, \ldots, L$ and $q = q_i$, $i = 1, \ldots, Q$ yields KQL different strategies. Let R_{it} be the net return for asset i in period t, then the return of each momentum strategy is calculated as follows.

1. Calculate the geometrically compounded return for asset i as

$$R_{it}(k) = \prod_{j=1}^{k}(1 + R_{i,t-j}) - 1, \qquad t > k.$$

2. To construct the portfolios of winners and losers the assets according to the returns calculated in Step 1 are sorted. The winner portfolio is then formed by giving equal weight to the q percent assets with the highest geometrically compounded return. Similarly, the loser portfolio is formed by giving equal weight to the q percent assets with the lowest

---

[4] Data on industrial portfolios were kindly provided by B, Asgharian. More details about the industrial portfolios can be found in Asgharian and Hansson (2003).

geometrically compounded return. The momentum strategy generates a zero-cost portfolio that buys the winners and sells the losers. This position is held for the next l months.

3. Every month a new momentum portfolio is formed and the oldest momentum portfolio is retired. After an initial ranking period of length K+L there are L different momentum portfolios. The return of the strategy is the average return from all L portfolios that are held simultaneously during period t.

4. Working through the T periods there is a time series of momentum returns for each of the KLQ momentum strategies.

In the case where the assets are individual stocks, firms must have k months of past returns. Hence, if a firm is delisted during a ranking period, it is excluded. No restriction is placed on survival going forward. If a stock is delisted during the holding period, the liquidation proceeds are assumed to be reinvested in the remaining stocks within the same portfolio until the end of the period.

4.4 The Reality Check

Imagine a large set of momentum trading rules. Our interest is to compare the return of each strategy to that of the benchmark return. A null hypothesis is formulated, where the momentum strategy with the largest return is not any better than the benchmark return. If the null hypothesis is rejected, there is at least one strategy that produces a significantly higher return than the benchmark. Since the momentum portfolios are by construction zero-cost portfolios the appropriate benchmark return is zero.

Formally, the testing proceeds as follows. Assume that there are returns from N momentum strategies over T periods. Let $\bar{\mathbf{d}} = [\bar{d}_1, \ldots, \bar{d}_N]'$ be an $N \times 1$ vector of momentum average returns

$$\bar{\mathbf{d}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{d}_t.$$

A hypothesis about a $N \times 1$ vector of moments $\mu = [E(d_1), \ldots, E(d_N)]'$ is tested. An appropriate null hypothesis is that the strategy with the highest average return is no better than the benchmark. Hence, the null hypothesis is

$$\mathrm{H}_0 : \max_{n=1,\ldots,N}\{E(d_n) \leq 0\}.$$

The alternative hypothesis is that the best strategy is superior to the benchmark and if the null hypothesis is rejected, there must be at least one strategy for which $E(d_n)$ is positive.

White (2000) proposed the following test statistic for testing the null hypothesis

$$V_{max}^{RC} = \max_{n=1,\ldots,N}\{\sqrt{T}(\bar{d}_n)\}. \qquad (4.1)$$

Since the asymptotic distribution of V_{max}^{RC} under the null hypothesis is non-standard, White (2000) shows that this null hypothesis can be evaluated by applying the stationary bootstrap of Politis and Romano (1994).[5]

The stationary bootstrap of Politis and Romano (1994) is based on pseudo time-series of the original data, i.e, the momentum returns. Let ξ_b, $b = 1, \ldots, B$ be $T \times 1$ vectors of indexes constructed by combining blocks of $\{1, ..., T\}$ with a random length. The block length follows a geometric distribution with parameter $\rho \in (0, 1]$ and an expected block length equal to $1/\rho$. The bootstrap samples are given by $\mathbf{d}_{b,t}^* = \mathbf{d}_{\xi_{b,t}}$, $t = 1, \ldots, T$, which leads to the sample averages $\bar{\mathbf{d}}_b^*$, $b = 1, \ldots, B$.

The bootstrap sample averages are used to construct the statistic

$$V_{max,b}^{RC*} = \max_{n=1,\ldots,N}\{\sqrt{T}(\bar{d}_{n,b}^* - \bar{d}_n)\}, \qquad b = 1, \ldots, B \qquad (4.2)$$

and White's Reality Check p-value is obtained as the fraction of times that $V_{max,b}^{RC*}$ is larger than V_{max}^{RC} for $b = 1, \ldots, B$. By using the maximum values over all N strategies, the estimated p-value is corrected for the effect of data-snooping.

The inclusion of $\bar{d}_n$ in (4.2) guarantees that the statistic satisfies the null hypothesis for all $n = 1, \ldots, N$. However, this makes the null hypothesis the least favorable to the alternative and a very conservative test.[6] Consequently, White's test may lack power. Hansen (2004) considers several adjustment to White's test and proposes a test that is supposed to be more powerful and

---

[5] The bootstrap samples can also be generated by the block bootstrap of Künsch (1989). The stationary bootstrap is used for two reasons. First, it generates stationary samples and second it does not require the determination of an optimal block-length. Instead of using the bootstrap, it is possible to use Monte Carlo simulation. In this case, one needs to estimate Ω consistently, the covariance matrix of the returns. Then one samples returns from $N(0, \widehat{\Omega})$ and the desired p-value can be obtained from the distribution of the extremes of $N(0, \widehat{\Omega})$. In our case, the number of strategies exceeds the number of observations so the only feasible implementation is the bootstrap. In addition, drawing from the empirical distribution has the advantage that it does not rely on restrictive distributional assumptions.

[6] Note that all negative values of $E(d_n)$ also conform with the null hypothesis

less sensitive to poor and irrelevant alternatives. Based on Hansen's recommendations, the statistic in (4.1) is modified as

$$V_{max}^{SPA} = \max\left[\max_{n=1,\ldots,N} \frac{\sqrt{T}(\bar{d}_n)}{\hat{\omega}_n}, 0\right] \tag{4.3}$$

where $\hat{\omega}_n^2$ is a consistent estimator of $\omega_n^2 = \mathrm{Var}(\sqrt{T}\bar{d}_n)$. The estimates of ω_n^2, $n = 1, \ldots, N$, can be obtained from the bootstrap samples or from the bootstrap population. Since the former usually requires a large number of bootstrap samples Hansen's recommendation has been followed and the estimates based on the bootstrap population have been used, which is given by

$$\hat{\omega}_n^2 = \hat{\gamma}_{0,n} + 2\sum_{t=1}^{T-1} \kappa(T,t)\hat{\gamma}_{t,n}$$

where

$$\hat{\gamma}_{t,n} = \frac{1}{T}\sum_{j=1}^{T-i}(d_{n,j} - \bar{d}_n)(d_{n,j+i} - \bar{d}_n), \qquad i = 0, \ldots, T-1$$

are the empirical covariances and the kernel under the stationary bootstrap is given by

$$\kappa(T,t) = \frac{T-i}{T}(1-\rho)^i + \frac{i}{T}(1-\rho)^{T-i}.$$

Furthermore, Equation (4.2) is modified as

$$V_{max,b}^{SPA*} = \max\left[\max_{n=1,\ldots,N} \sqrt{T}\left(\frac{\bar{d}_{n,b}^* - g(\bar{d}_n)}{\hat{\omega}_n}\right), 0\right], \qquad b = 1, \ldots, B$$

where different functions $g(\cdot)$ will produce different bootstrap distributions that are compatible with the null hypothesis. Hansen suggests three functions that generate a consistent estimate of the p-value as well as an upper and a lower bound. If $g(\bar{d}_n) = \max\{\bar{d}_n, 0\}$, the null hypothesis is the more favorable to the alternative and the corresponding p-value will be a lower bound for the true value. If $g(\bar{d}_n) = \bar{d}_n$ as in Equation (4.2) a standardized version of White's test is obtained with a corresponding p-value, which can be viewed as an upper bound for the true value. Finally, Hansen recommends the use of $g(\bar{d}_n) = \bar{d}_n \cdot 1_{\{\bar{d}_n \geq -\sqrt{(\hat{\omega}_n^2/T)2\log\log T}\}}$ which leads to a consistent estimate of the p-value.

In our empirical section, four reality check p-values are reported: the White Reality Check p-value, and the lower, the consistent and the upper bound p-value of Hansen.

4.5 Empirical Results

This section evaluates the profitability of momentum investment strategies described in the previous sections. In Section 4.5.1 we investigate the US data and in section 4.5.2 the data from the Swedish stock market is considered.

4.5.1 US Data

Tables 4.1a and 4.1b present the average returns on winner, loser and momentum portfolios between June 1965 and December 2004. Two years are lost due to the initial ranking period. The ranking period is 6 months and 12 months and the holding period ranges from 3 to 12 months. The proportion of winners/losers is 20 percent. In the 17 industry portfolios this corresponds to 3. The portfolios in Table 4.1a are formed at the end of the ranking period. Because the bid-ask bounce can attenuate the continuation effect the momentum returns are also calculated when the portfolio formation is delayed relative to the ranking by one month. They are reported in Table 4.1b.

Table 4.1a reports the average returns for the loser and winner portfolios and the momentum profits using different sets of portfolios. All momentum profits are positive. The most successful zero-cost strategy selects assets based on their returns over the past 12 months and then holds the portfolio for 3 months. Applying this strategy yields a profit equal to 0.62% for the size and book-to-market portfolios and around 0.52% considering portfolios formed on size and industry. The difference between losers and winners is smaller for the book-to-market portfolios. Furthermore, the magnitude of the momentum profits is consistent with Moskowitz and Grinblatt (1999) and Lewellen (2002) except that they find a reversal in the industry portfolios for 9 or more months after the formation. Finally, it seems that the momentum portfolios based on a 6-month ranking are more profitable than the momentum portfolios based on a 12-month ranking for longer holding periods. An exception is the size portfolios where the 12-month ranking always outperforms the shorter ranking period.

Table 4.1b shows the average returns when the portfolio formation is delayed relative to the ranking by one month. The average returns show the same pattern as in the previous table. In general, the momentum returns are

Table 4.1a: US Data: Returns of Relative Strength Portfolios

Momentum portfolios are formed immediately after the ranking.

Industry Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	0.91	0.94	0.95	0.89	0.85	0.81	0.80	0.81	0.82	0.83
6	Winner	1.12	1.16	1.17	1.20	1.23	1.24	1.22	1.12	1.19	1.17
	Momentum	0.21	0.22	0.22	0.31	0.33	0.42	0.42	0.39	0.37	0.34
	Loser	0.74	0.76	0.77	0.77	0.76	0.78	0.79	0.82	0.83	0.85
12	Winner	1.26	1.23	1.21	1.19	1.18	1.16	1.15	1.14	1.13	1.11
	Momentum	0.52	0.47	0.43	0.42	0.42	0.38	0.36	0.32	0.29	0.26

Size Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	1.02	1.01	1.02	1.0	0.98	0.96	0.96	0.95	0.95	0.96
6	Winner	1.20	1.20	1.21	1.23	1.25	1.26	1.26	1.26	1.26	1.26
	Momentum	0.18	0.19	0.20	0.23	0.28	0.30	0.31	0.31	0.30	0.30
	Loser	0.83	0.84	0.85	0.85	0.84	0.84	0.84	0.84	0.85	0.86
12	Winner	1.34	1.34	1.33	1.32	1.33	1.34	1.33	1.33	1.32	1.31
	Momentum	0.52	0.50	0.48	0.47	0.48	0.50	0.50	0.49	0.47	0.45

Book-to-Market Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	0.97	0.95	0.94	0.95	0.95	0.95	0.96	0.96	0.97	0.99
6	Winner	1.18	1.18	1.17	1.17	1.18	1.18	1.18	1.17	1.17	1.16
	Momentum	0.21	0.23	0.23	0.23	0.23	0.23	0.22	0.21	0.20	0.18
	Loser	0.94	0.98	1.01	1.03	1.05	1.07	1.08	1.09	1.10	1.10
12	Winner	1.24	1.23	1.22	1.20	1.19	1.18	1.19	1.18	1.17	1.16
	Momentum	0.30	0.25	0.21	0.17	0.13	0.12	0.11	0.09	0.08	0.07

Size and Book-to-Market Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	0.94	0.94	0.95	0.96	0.96	0.95	0.95	0.95	0.96	0.96
6	Winner	1.38	1.37	1.36	1.35	1.36	1.37	1.37	1.36	1.34	1.33
	Momentum	0.44	0.42	0.40	0.39	0.41	0.43	0.42	0.40	0.38	0.36
	Loser	0.83	0.86	0.88	0.89	0.90	0.90	0.91	0.92	0.93	0.94
12	Winner	1.45	1.42	1.39	1.37	1.36	1.35	1.34	1.32	1.30	1.28
	Momentum	0.62	0.56	0.51	0.47	0.46	0.45	0.43	0.40	0.37	0.34

higher when there is no lag between the ranking and the formation period. This is especially the case for short holding periods. The difference decreases when the holding period increases.

Next, we investigate whether momentum profits differ significantly from the benchmark return when taking into account the effect of data-snooping.

Is Momentum Due to Data-Snooping?

The Reality Check is applied to a universe of momentum strategies. For a given proportion of winner and loser portfolios q, we let the ranking k range from 3 to 12 months and the holding period l range from 3 to 12 months. The proportion of winner and loser portfolios is 10%, 20% and 30%. This results in 300 momentum portfolios. Since the industry contains 17 portfolios we cannot take exactly 10%, 20% and 30% winners and losers. The same problem exists for the 25 size-book/market portfolios. Instead, we take 2, 3 and 5, and 3, 5 and 8 loser/winner portfolios in industry and size-book/market. We also consider two sub-samples, 1965:08-1983:12 and 1984:01-2004:12.

In Table 4.2a we report the result for testing the null hypothesis that the best momentum strategy does not outperform the benchmark, which is the zero return. The momentum portfolios are formed immediately after the ranking. The table reports the Reality Check p-values and the corresponding nominal p-values (in brackets). The nominal p-value is the result of applying the bootstrap to the best trading rule only. Hence, by using the nominal p-value the effects of data-snooping are ignored and the difference between the two p-values yields the magnitude of data-snooping bias. The number of bootstrap samples is 10 000 and we set $\rho = 0.5$, which implies an expected block length of 2 observations.[7]

Firstly, we note that the data-snooping bias is very substantial. Neglecting the data-snooping effect means that we almost always reject the null hypothesis at the 5% significance level, while the evidence is mixed when taking account of data-snooping. Secondly, we note that the Reality Check p-values are very similar, especially for the three versions of the Reality Check provided by Hansen (2004).

Starting with the whole sample period, the momentum effect is significant for the industry, size, and the size and book-to-market portfolios. The p-values of Hansen are all lower than 1% while the p-values based on White

---

[7] We examined the robustness of the results with respect to ρ. It appears that our results are insensitive to the choice of ρ.

Table 4.1b: US Data: Returns of Relative Strength Portfolios

Momentum portfolios are formed one month after the ranking has take place.

Industry Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	0.97	0.97	0.90	0.85	0.81	0.80	0.80	0.81	0.83	0.85
6	Winner	1.17	1.1	1.21	1.25	1.25	1.23	1.21	1.19	1.17	1.14
	Momentum	0.20	0.20	0.31	0.39	0.44	0.44	0.40	0.38	0.34	0.29
	Loser	0.78	0.79	0.79	0.78	0.79	0.81	0.83	0.85	0.86	0.87
12	Winner	1.21	1.19	1.17	1.16	1.14	1.13	1.12	1.11	1.10	1.07
	Momentum	0.43	0.40	0.38	0.39	0.35	0.33	0.30	0.26	0.23	0.20

Size Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	1.03	1.03	1.00	0.98	0.96	0.95	0.95	0.95	0.95	0.96
6	Winner	1.19	1.21	1.23	1.26	1.27	1.27	1.26	1.26	1.26	1.26
	Momentum	0.17	0.19	0.23	0.28	0.31	0.31	0.32	0.31	0.31	0.29
	Loser	0.86	0.87	0.86	0.85	0.85	0.84	0.85	0.86	0.87	0.87
12	Winner	1.32	1.31	1.31	1.32	1.33	1.33	1.32	1.31	1.30	1.30
	Momentum	0.46	0.44	0.44	0.46	0.48	0.48	0.47	0.46	0.44	0.43

Book-to-Market Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	0.95	0.94	0.94	0.95	0.95	0.95	0.96	0.97	0.99	1.01
6	Winner	1.16	1.15	1.16	1.17	1.18	1.17	1.17	1.17	1.16	1.14
	Momentum	0.22	0.22	0.21	0.22	0.22	0.22	0.20	0.19	0.17	0.14
	Loser	1.01	1.04	1.06	1.08	1.09	1.10	1.11	1.13	1.13	1.12
12	Winner	1.23	1.21	1.20	1.18	1.18	1.18	1.17	1.16	1.16	1.15
	Momentum	0.22	0.17	0.1400	0.10	0.09	0.08	0.06	0.05	0.04	0.03

Size and Book-to-Market Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	0.98	0.98	0.98	0.97	0.969	0.96	0.97	0.97	0.97	0.99
6	Winner	1.34	1.34	1.33	1.35	1.37	1.36	1.35	1.33	1.32	1.30
	Momentum	0.37	0.36	0.35	0.38	0.41	0.40	0.3803	0.36	0.34	0.31
	Loser	0.89	0.91	0.92	0.93	0.93	0.93	0.94	0.95	0.96	0.97
12	Winner	1.38	1.36	1.34	1.33	1.33	1.31	1.23	1.28	1.26	1.25
	Momentum	0.49	0.45	0.41	0.41	0.40	0.38	0.36	0.33	0.31	0.28

Table 4.2a: US Data: Reality Check

Momentum portfolios are formed immediately after the ranking.

Industry Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
196506-198312	0.011 (0.002)	0.013 (0.002)	0.013 (0.002)	0.021 (0.008)
198401-200412	0.190 (0.051)	0.213 (0.053)	0.213 (0.053)	0.308 (0.094)
196506-200412	0.009 (0.002)	0.009 (0.002)	0.009 (0.002)	0.038 (0.008)

Size Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
196506-198312	0.014 (0.001)	0.014 (0.001)	0.014 (0.0005)	0.013 (0.002)
198401-200412	0.016 (0.000)	0.016 (0.000)	0.016 (0.0003)	0.022 (0.004)
196506-200412	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)

Book-to-Market Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
196506-198312	0.158 (0.020)	0.168 (0.021)	0.168 (0.0210	0.190 (0.071)
198401-200412	0.061 (0.001)	0.063 (0.001)	0.063 (0.001)	0.044 (0.009)
196506-200412	0.020 (0.001)	0.020 (0.001)	0.020 (0.001)	0.029 (0.007)

Size and Book-to-Market Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
196506-198312	0.014 (0.000)	0.014 (0.000)	0.014 (0.000)	0.005 (0.002)
198401-200412	0.041 (0.003)	0.045 (0.003)	0.045 (0.003)	0.024 (0.008)
196506-200412	0.002 (0.000)	0.003 (0.000)	0.003 (0.000)	0.001 (0.000)

Hansen_l, Hansen_c and Hansen_u corresponds to the lower, consistent and upper p-value of Hansen (2004).

are higher but still rather low. For the book-to-market portfolios the Reality Check p-values are between 2% and 3%.

Considering the two sub samples yields different results. Starting with the period 1965:08-1983:12 similar results are obtained as for the whole sample period. The p-values for the industry, size and size and book-to-market are all lower than 2%. During the second period 1984:01-2004:12 we have the opposite. The Reality Check p-values are between 19% and 30% for the industry portfolios and around 5% for the size and book-to-market portfolios. The only case where the p-values are low for both periods is when portfolios are formed on size. Overall, the results in Table 4.2a indicate that the profitability of momentum strategies is due to the high profitability over the first half of the sample period.

Table 4.2b shows the results when the Reality Check is applied to momentum returns when the portfolios are formed one month after the ranking has take place. The difference between Table 4.2a is that the p-values are a little bit higher. However, the main results are the same.

It is well known that riskier investments generally yield higher returns than investments that are free of risk. Hence, the result that that returns on winner portfolios dominate returns on loser portfolios may be because the securities in the winner portfolio are riskier. However, using risk-adjusted returns does not alter the results.[8] Instead, risk-adjusted returns of momentum strategies are significant and frequently larger than the raw returns of momentum strategies. Therefore, the results are not reported but available from the authors on request.

4.5.2 Swedish Data

Tables 4.3a and 4.3b present the average returns on winner, loser and momentum portfolios. Two years are lost due to the initial ranking period. The ranking period is 6 months and 12 months and the holding period ranges from 3 to 12 months. The proportion of winners/losers is 20 percent. The sample period for the individual stocks is January 1981 to December 2003 and the sample period for the size portfolios is June 1981 to December 2003. The sample for the book-to-market and industry portfolios starts in June 1982 and ends in December 2003 and December 1997 respectively.

Starting with the individual stocks in Table 4.3a we first note that the loser portfolio generates negative average returns. The winner returns have about

---

[8] We adjust for risk by using the CAPM benchmark and the Fama-French three factor model benchmark.

Table 4.2b: US Data: Reality Check

Momentum portfolios are formed one month after the ranking take place.

Industry Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
196506-198312	0.025 (0.002)	0.025 (0.02203)	0.025 (0.022)	0.030 (0.010)
198401-200412	0.180 (0.041)	0.215 (0.045)	0.215 (0.045)	0.255 (0.078)
196506-200412	0.018 (0.001)	0.010 (0.001)	0.019 (0.001)	0.019 (0.004)

Size Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
196506-198312	0.015 (0.000)	0.015 (0.000)	0.015 (0.000)	0.012 (0.003)
198401-200412	0.019 (0.000)	0.019 (0.000)	0.019 (0.000)	0.018 (0.001)
196506-200412	0.001 (0.000)	0.001 (0.000)	0.001 (0.0001)	0.001 (0.000)

Book-to-Market Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
196506-198312	0.161 (0.021)	0.181 (0.025)	0.181 (0.025)	0.191 (0.078)
198401-200412	0.133 (0.009)	0.142 (0.010)	0.142 (0.001)	0.151 (0.041)
196506-200412	0.036 (0.002)	0.037 (0.002)	0.037 (0.002)	0.678 (0.017)

Size and Book-to-Market Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
196506-198312	0.039 (0.001)	0.039 (0.001)	0.039 (0.001)	0.015 (0.005)
198401-200412	0.055 (0.004)	0.064 (0.005)	0.064 (0.005)	0.047 (0.014)
196506-200412	0.011 (0.000)	0.011 (0.000)	0.011 (0.000)	0.004 (0.002)

Hansen$_l$, Hansen$_c$ and Hansen$_u$ corresponds to the lower, consistent and upper p-value of Hansen (2004).

Table 4.3a: Swedish Data: Returns of Relative Strength Portfolios

Momentum portfolios are formed immediately after ranking.

Individual Stocks

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	-1.06	-1.04	-1.01	-0.97	-0.96	-0.95	-0.92	-0.89	-0.84	-0.80
6	Winner	1.22	1.16	1.10	1.04	1.00	0.96	0.92	0.88	0.83	0.79
	Momentum	2.28	2.20	2.10	2.01	1.96	1.91	1.84	1.77	1.67	1.59
	Loser	-1.12	-1.11	-1.06	-1.04	-1.00	-0.96	-0.91	-0.87	-0.82	-0.77
12	Winner	1.20	1.12	1.04	0.97	0.91	0.85	0.79	0.74	0.71	0.69
	Momentum	2.32	2.23	2.10	2.01	1.91	1.81	1.70	1.61	1.53	1.46

Size Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	1.38	1.43	1.44	1.45	1.48	1.49	1.50	1.52	1.52	1.52
6	Winner	1.93	1.94	1.89	1.86	1.87	1.86	1.88	1.88	1.87	1.86
	Momentum	0.55	0.51	0.45	0.41	0.39	0.37	0.38	0.36	0.35	0.34
	Loser	1.42	1.43	1.46	1.49	1.52	1.52	1.56	1.59	1.60	1.60
12	Winner	2.03	2.02	2.00	1.98	1.96	1.94	1.93	1.90	1.88	1.86
	Momentum	0.61	0.59	0.56	0.49	0.44	0.42	0.36	0.31	0.28	0.26

Book-to-Market Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	1.44	1.47	1.46	1.46	1.48	1.49	1.49	1.50	1.56	1.49
6	Winner	1.65	1.68	1.64	1.65	1.68	1.69	1.72	1.71	1.72	1.72
	Momentum	0.21	0.21	0.18	0.19	0.20	0.20	0.224	0.21	0.16	0.23
	Loser	1.39	1.42	1.44	1.45	1.46	1.46	1.47	1.48	1.47	1.47
12	Winner	1.78	1.78	1.76	1.76	1.75	1.76	1.76	1.75	1.74	1.75
	Momentum	0.39	0.36	0.32	0.31	0.29	0.30	0.29	0.27	0.27	0.28

Industry Portfolios

		Month after formation									
Rank		3	4	5	6	7	8	9	10	11	12
	Loser	1.22	1.20	1.171	1.133	1.123	1.133	1.118	1.098	1.093	1.101
6	Winner	1.84	1.83	1.791	1.763	1.757	1.774	1.761	1.745	1.727	1.693
	Momentum	0.62	0.63	0.620	0.631	0.634	0.641	0.642	0.647	0.633	0.593
	Loser	1.08	1.04	1.04	1.07	1.12	1.12	1.14	1.17	1.18	1.20
12	Winner	1.86	1.85	1.83	1.78	1.77	1.73	1.70	1.66	1.63	1.60
	Momentum	0.79	0.81	0.79	0.71	0.65	0.61	0.56	0.49	0.45	0.40

the same magnitude but they are all positive. The profit for a 12 month ranking and a 3 month holding is 2.318% and 2.283% when 6 months is used for the ranking. Turning to the portfolios in Table 4.3a we note that the profits from the momentum strategies are positive but they are lower than the momentum profits for the individual stocks. Sorting stock by industries generates the highest profit among the different portfolios. When we consider portfolios formed on size we note that the loser returns increase and the winner returns decrease for longer holding periods. The momentum returns are, however, still positive 12 months after the formation for both the 6- and 12-month ranking. Momentum returns based on a 12 month ranking decrease when the number of months after the formation increases. The strategies based on a 6 month ranking show a similar pattern except for the book-to-market and industry portfolios where the momentum returns do not change much for the different holding periods. Finally, the momentum returns based on a 6-month ranking are more profitable than the momentum portfolios based on a 12-month ranking for longer holding periods. An exception is the book-to-market portfolios.

Table 4.3b shows the return when there is a one-month lag between the formation period and the ranking period. The momentum profits are about the same magnitude as in Table 4.3a. Next, we examine whether the profits are significant using the Reality Check.

Is Momentum Due to Data-Snooping?

The Reality Check is applied to a universe of momentum strategies in a similar way as for the US data. The ranking and holding period range from 3 to 12 months, respectively and the proportion of winner and loser stocks is 10%, 20% and 30%. Two sub-samples are also considered. The test results are displayed in Tables 4.4a and 4.4b. The tables reports the Reality Check p-values and the corresponding nominal p-values (in brackets). The number of bootstrap samples is 10 000 and the expected block length is 2 observations.

Table 4.4a reports the estimated p-values when momentum portfolios are formed immediately after the ranking has taken place. The very low p-values for the individual stocks indicate that at least one of the momentum strategies yields a higher return than the benchmark. Furthermore, the null hypothesis of no momentum effect is rejected in the two sub-samples. Hence, there is strong evidence in favor of the momentum effect in individual stocks for the Swedish market. Even if the nominal p-values differ from the corrected ones, the same conclusion is drawn from both sets of p-values. Next, we examine whether the momentum anomaly is present when stocks are sorted into different portfolios.

Interestingly, the momentum effect for the different portfolios is not that

Table 4.3b: Swedish Data: Returns of Relative Strength Portfolios

Momentum portfolios are formed one month after the ranking .

Individual Stocks

Month after formation

Rank		3	4	5	6	7	8	9	10	11	12
6	Loser	-1.11	-1.05	-1.00	-0.98	-0.97	-0.93	-0.90	-0.84	-0.79	-0.75
	Winner	1.15	1.08	1.02	0.97	0.93	0.89	0.85	0.80	0.76	0.72
	Momentum	2.26	2.12	2.02	1.95	1.90	1.81	1.74	1.64	1.55	1.47
12	Loser	-1.18	-1.11	-1.07	-1.02	-0.98	-0.92	-0.87	-0.81	-0.76	-0.72
	Winner	1.09	1.00	0.93	0.86	0.80	0.74	0.69	0.66	0.64	0.63
	Momentum	2.27	2.11	2.00	1.88	1.77	1.66	1.56	1.47	1.40	1.35

Size Portfolios

Month after formation

Rank		3	4	5	6	7	8	9	10	11	12
6	Loser	1.45	1.45	1.46	1.49	1.50	1.52	1.53	1.53	1.53	1.55
	Winner	1.91	1.86	1.85	1.84	1.84	1.86	1.86	1.86	1.85	1.83
	Momentum	0.47	0.41	0.39	0.36	0.34	0.34	0.33	0.33	0.32	0.29
12	Loser	1.42	1.46	1.500	1.53	1.53	1.58	1.60	1.62	1.62	1.63
	Winner	2.01	2.00	1.98	1.95	1.93	1.91	1.88	1.86	1.85	1.82
	Momentum	0.59	0.54	0.48	0.42	0.40	0.34	0.28	0.25	0.23	0.19

Book-to-Market Portfolios

Month after formation

Rank		3	4	5	6	7	8	9	10	11	12
6	Loser	1.45	1.44	1.45	1.47	1.48	1.49	1.50	1.50	1.49	1.49
	Winner	1.69	1.63	1.66	1.69	1.70	1.73	1.72	1.73	1.72	1.71
	Momentum	0.24	0.19	0.21	0.22	0.22	0.24	0.20	0.23	0.23	0.22
12	Loser	1.42	1.44	1.46	1.46	1.47	1.48	1.49	1.46	1.48	1.47
	Winner	1.81	1.79	1.78	1.77	1.78	1.77	1.76	1.75	1.76	1.76
	Momentum	0.39	0.35	0.33	0.30	0.31	0.29	0.27	0.28	0.28	0.30

Industry Portfolios

Month after formation

Rank		3	4	5	6	7	8	9	10	11	12
6	Loser	1.14	1.12	1.09	1.08	1.10	1.09	1.07	1.066	1.076	1.11
	Winner	1.80	1.76	1.73	1.73	1.75	1.74	1.73	1.708	1.673	1.64
	Momentum	0.66	0.64	0.65	0.65	0.65	0.66	0.66	0.642	0.597	0.53
12	Loser	0.98	0.99	1.04	1.10	1.11	1.13	1.16	1.18	1.200	1.23
	Winner	1.85	1.83	1.76	1.75	1.71	1.68	1.64	1.60	1.58	1.56
	Momentum	0.87	0.84	0.73	0.66	0.60	0.56	0.49	0.42	0.38	0.33

Table 4.4a: Swedish Data:Reality Check

Momentum portfolios are formed immediately after the ranking.

Individual Stocks

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
198101-199212	0.011 (0.004)	0.011 (0.004)	0.011 (0.004)	0.009 (0.007)
199301-200312	0.011 (0.001)	0.012 (0.001)	0.012 (0.001)	0.004 (0.002)
198101-200312	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.000 (0.000)

Size Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
198106-199212	0.065 (0.011)	0.079 (0.011)	0.079 (0.011)	0.075 (0.026)
199301-200312	0.109 (0.017)	0.111 (0.017)	0.111 (0.017)	0.099 (0.034)
198106-200312	0.024 (0.005)	0.033 (0.005)	0.033 (0.005)	0.014 (0.006)

Book-to-Market Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
198205-199212	0.156 (0.004)	0.156 (0.004)	0.156 (0.004)	0.131 (0.024)
199301-200312	0.290 (0.069)	0.311 (0.079)	0.311 (0.079)	0.264 (0.117)
198206-200312	0.151 (0.024)	0.170 (0.024)	0.171 (0.024)	0.141 (0.054)

Industry Portfolios

Time Period	Hansen$_l$	Hansen$_c$	Hansen$_u$	White
198206-198912	0.039 (0.013)	0.040 (0.013)	0.040 (0.013)	0.117 (0.038)
199001-199712	0.308 (0.083)	0.308 (0.083)	0.308 (0.083)	0.230 (0.108)
198206-199712	0.068 (0.008)	0.088 (0.008)	0.088 (0.008)	0.101 (0.017)

Hansen$_l$, Hansen$_c$ and Hansen$_u$ corresponds to the lower, consistent and upper p-value of Hansen (2004).

strong as for the individual stocks. The estimated p-values for the book-to-market portfolios are all greater than 13% which indicates that no momentum strategy generates a higher return than the benchmark. The momentum strategies when portfolios are formed on size and industry seem to be more profitable during the first sub-sample, June 1981 to December 1992 and June 1982 to December 1989. However, as for the individual stocks, the evidence is not that strong. The p-values for the size portfolios are around 7% and for the industry portfolios around 4% for Hansen's versions and 13% for White's version.

Overall, Table 4.4a reveals several interesting issues. Firstly, the momentum effect seems to be very strong for the individual stocks, but weak or even non existent when stocks are sorted into portfolios. Secondly, and perhaps more importantly, the data-snooping bias is very substantial. Neglecting the effect of data-snooping means that we almost always reject the null hypothesis of no momentum effect for the size, book-to-market and industry portfolios.

Table 4.4b shows the results when the Reality Check is applied to momentum returns when the portfolios are formed one month after the ranking has taken place. The results for the individual stocks are very similar or even more stronger in favor of a momentum effect. The main results regarding the portfolios remain the same.

Momentum profits of risk-adjusted returns are very similar to the profits obtained on raw returns and the adjustment only serve to strengthen the previous results. Therefore, the results are not reported but available from the authors on request.

An important issue when evaluating different trading strategies is transaction costs. This is especially the case when considering individual stocks. However, incorporating the transaction cost, which will reduce the profits, is not straightforward. For example, institutional traders can often secure trade discounts relative to individual retail investors. Furthermore, stocks with smaller market capitalization are more likely to be traded at a wider bid-ask spread, compared to firms with larger market capitalization.

Table 4.4b: Swedish Data:Reality Check

Momentum portfolios are formed one month after the ranking take place.

Individual Stocks

Time Period	Hansen_l	Hansen_c	Hansen_u	White
198101-199212	0.007 (0.002)	0.007 (0.002)	0.007 (0.002)	0.006 (0.005)
199301-200312	0.005 (0.001)	0.007 (0.002)	0.007 (0.002)	0.005 (0.003)
198101-200312	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)

Size Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
198106-199212	0.051 (0.009)	0.051 (0.009)	0.051 (0.009)	0.041 (0.019)
199301-200312	0.109 (0.015)	0.119 (0.016)	0.119 (0.016)	0.179 (0.036)
198106-200312	0.025 (0.001)	0.025 (0.001)	0.025 (0.001)	0.016 (0.004)

Book-to-Market Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
198205-199212	0.156 (0.009)	0.196 (0.016)	0.196 (0.016)	0.108 (0.043)
199301-200312	0.324 (0.054)	0.325 (0.054)	0.325 (0.054)	0.180 (0.083)
198206-200312	0.151 (0.008)	0.151 (0.008)	0.151 (0.008)	0.071 (0.028)

Industry Portfolios

Time Period	Hansen_l	Hansen_c	Hansen_u	White
198206-198912	0.043 (0.023)	0.046 (0.023)	0.046 (0.023)	0.102 (0.048)
199001-199712	0.298 (0.076)	0.308 (0.083)	0.308 (0.083)	0.228 (0.098)
198206-199712	0.074 (0.004)	0.078 (0.004)	0.078 (0.004)	0.111 (0.028)

Hansen_l, Hansen_c and Hansen_u corresponds to the lower, consistent and upper p-value of Hansen (2004).

4.6 Summary and Conclusion

This paper explores the profitability of momentum strategies. Two data sets are considered. The first set of data consists of all NYSE, AMEX, and NASDAQ stocks on the CRSP database. The analysis considers the period from July 1963 to December 2004 and the tests are performed on portfolios formed on industry, size, book-to-market and double sorted on size and book-to-market. The second set of data consist of all stocks listed on the Stockholm Stock Exchange over the period January 1979 to December 2003 and the tests are performed on individual stocks and on portfolios formed on size, book-to-market and industry.

The departure from earlier studies lies in the way we test for profitability. To avoid the serious problem of data-snooping we apply the procedure provided by White (2000) and the modified version provided by Hansen (2004). Hence, we examine whether a momentum strategy is superior to a benchmark model once the effects of data-snooping have been accounted for.

For the US data there is strong evidence of a momentum effect where an investor takes a long position in the winner portfolio and a short position in the loser portfolio. Hence, we reject the hypothesis of weak market efficiency. By splitting the sample into two parts, 1965:08 to 1983:12 and 1984:01 to 2003:12, we find that the momentum strategy was profitable during the first period and not during the second. The overall significance is thus driven by events in the earlier part of the sample and it appears that the market has become more efficient.

To our knowledge, this paper is the first one that examine the momentum effect using all stocks listed on the Stockholm Stock Exchange. The long time period covered and the explicit inclusion of delisted companies should make the results free from possible sample bias. The results indicates that momentum strategies based on individual stocks generate positive and significant profits. The same result is obtained when two sub-samples are considered. Hence, we reject the hypothesis of weak market efficiency. Interestingly, a very weak or no momentum effect can be found when stocks are sorted by size, book-to-market and industry.

Finally, and perhaps most importantly, our results show that data snooping bias can be very substantial. In this study, neglecting the problem would lead to very different conclusions.

Bibliography

CONRAD, J., AND G. KAUL (1998): "An Anatomy of Trading Strategies," *Review of Financial Studies*, 11, 489–519.

HANSEN, P. (2004): "A Test for Superior Predictive Ability," Discussion Paper 01-06, Brown Univ. Dept. of Economics.

JEGADEESH, N., AND S. TITMAN (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.," *Journal of Finance*, 48, 65–91.

JEGADEESH, N., AND S. TITMAN (2001): "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance*, 56, 699–720.

KÜNSCH, H. R. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 12171241.

LAKONISHOK, J., AND S. SMIDT (1984): "Turn-of the Year Behavior," *Journal of Financial Economics*, 13, 435–456.

LEWELLEN, J. (2002): "Momentum and Autocorrelation in Stock Returns," *The Review of Financial Studies*, 15, 533–563.

LO, A., AND A. MACKINLAY (1990): "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431–467.

MOSKOWITZ, T., AND M. GRINBLATT (1999): "Do Industries Explain Momentum?," *Journal of Finance*, 54, 1249–1290.

POLITIS, D., AND J. ROMANO (1994): "The Stationary Bootstrap.," *Journal of the American Statistical Association*, 89, 1303–1313.

ROUWENHORST, G. K. (1998): "International Momentum Strategies," *Journal of Finance*, 53, 267–284.

——— (1999): "Local Return Factors and Turnover in Emerging Stock Markets," *Journal of Finance*, 54, 1439–1464.

SAVIN, N. E. (1984): "Multiple Hypothesis Testing," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. Intriligator, pp. 93–124. North Holland, Amsterdam.

TIMMERMANN, A., H. WHITE, AND R. SULLIVAN (1998): "Data-Snooping, Technical Trading, Rule Performance and the Bootstrap," FMG Discussion Papers dp303, Financial Markets Group, available at http://ideas.repec.org/p/fmg/fmgdps/dp303.html.

VAN DER HART, J., E. SLAGTER, AND D. VAN DIJK (2001): "Stock Selection Strategies in Emerging Markets," *Tinbergen Institute Discussion Papers*, No 01-009/4.

WHITE, H. (2000): "A Reality Check for Data Snooping," *Econometrica*, 68, 1097–1127.

Appendix A

Tables

Table A.1a: Descriptive statistics: US Data

Industry

Portfolio	Average Return	Std.dev
Food	1.135	4.531
Mining and Minerals	0.884	6.362
Oil	1.034	5.246
Textiles	0.955	6.119
Consumer Durables	1.028	5.757
Chemicals	0.887	5.308
Drugs, Soap, Tobacco	1.153	4.817
Construction	0.973	5.869
Steel	0.699	6.369
Fabricated Products	0.898	5.423
Machinery Equipment	0.946	6.534
Automobiles	0.871	6.031
Transportation	0.991	5.902
Utilities	0.773	4.131
Retail Stores	1.062	5.601
Finance	1.045	5.108
Other	0.866	5.034

Book-to-Market

Portfolio	Average return	Std.dev
Low	0.8152	5.3385
2	0.9340	4.8461
3	0.9458	4.8024
4	0.9305	4.7247
5	0.9552	4.4304
6	1.0728	4.4267
7	1.1496	4.3656
8	1.190	4.3586
9	1.2441	4.6578
High	1.2996	5.4573

Size

Portfolio	Average return	Std.dev
Small	1.1802	6.4667
2	1.1310	6.3371
3	1.1529	6.0494
4	1.1161	5.8716
5	1.1466	5.5599
6	1.0408	5.3194
7	1.0690	5.1775
8	1.0418	5.0764
9	0.9728	4.6413
Large	0.8675	4.3398

Table A.1b: Descriptive statistics: US Data

25 Size-book-to-market			
Size	B/M	Average Return	Std.dev
Small	Low	0.632	8.276
	2	1.200	7.087
	3	1.280	6.110
	4	1.491	5.675
	High	1.566	5.944
2	Low	0.801	7.561
	2	1.063	6.114
	3	1.326	5.405
	4	1.389	5.173
	High	1.422	5.759
3	Low	0.830	6.904
	2	1.143	5.503
	3	1.155	4.974
	4	1.288	4.724
	High	1.433	5.377
4	Low	0.955	6.158
	2	0.931	5.196
	3	1.131	4.892
	4	1.255	4.681
	High	1.346	5.414
Large	Low	0.883	4.877
	2	0.911	4.599
	3	0.951	4.375
	4	1.056	4.308
	High	1.019	4.799

Table A.2: Descriptive statistics: Swedish Data

Size

Portfolio	Average Return	Std.Dev
Low	2.510	9.552
2	1.694	7.578
3	1.845	7.522
4	1.499	6.971
5	1.507	6.822
6	1.386	6.861
7	1.378	6.672
8	1.388	6.506
9	1.410	6.184
High	1.415	6.305

Book-to-Market

Portfolio	Average Return	Std.Dev
Low	1.181	6.380
2	1.503	6.223
3	1.272	6.929
4	1.387	6.197
5	1.407	6.512
6	1.484	6.455
7	1.410	6.551
8	1.872	7.738
9	1.386	6.988
High	2.029	8.499

Industry

Portfolio	Average Return	Std.Dev
Low	1.009	9.725
2	1.567	7.896
3	1.882	8.538
4	1.669	7.218
5	1.532	7.714
6	1.335	9.161
7	1.322	8.934
8	1.734	7.319
9	1.205	5.832
High	1.706	8.337

EFI, The Economic Research Institute

Reports since 2000
A complete publication list can be found at www.hhs.se/efi

Published in the language indicated by the title.

2005

Books

Andersson, Per, Susanne Hertz and Susanne Sweet (eds). *Perspectives on market networks – boundaries and new connections.*
Charpentier, Claes. *IT inom omsorgen. Förväntade effekter av införande av IT-system för utförarna inom äldre- och handikappomsorgen.*
Lind, Johnny och Göran Nilsson (redaktörer). *Ekonomistyrningens metoder, sammanhang och utveckling – En vänbok till Lars A Samuelson.*
Samuelson, Lars A. *Organizational governance and control – a summary of research in the Swedish society.*

Dissertations

Andersson, Martin. *Making a Difference – Project Result Improvement in Organizations.*
Arvidsson, Per. *Styrning med belöningssystem – Två fallstudier om effekter av belöningssystem som styrmedel.*
Bems, Rudolfs. *Essays in International Macroeconomics.*
Berg-Suurwee, Ulrika. *Nya trender, nya nämnder – effekter av en stadsdelsnämndsreform inom kultur och fritid.*
Björkman, Hans. *Learning from members – Tools for strategic positioning and service innovation in trade unions.*
Bodnaruk, Andriy. *Essays in Empirical Corporate Finance and Portfolio Choice.*
Clapham, Eric. *Essays in Real Estate Finance.*
Dareblom, Jeanette. *Prat, politik och praktik – Om individers möten med strukturer i en kommunal satsning på kvinnors företagande.*
Hjelström, Anja, *Understanding International Accounting Standard Setting – A Case Study of IAS 12, Accounting for Deferred Tax.*
Hortlund, Per. *Studies on Swedish Banking 1870-2001.*
Lindahl, Therese. *Strategic and Environmental Uncertainty in Social Dilemmas.*
Linnarsson, Håkan. *Alliance for Innovation. A structural perspective on new business development in cooperative ventures.*
Madestam, Andreas. *Developing Credit Markets*
Nilsson, Roland. *The Market Impact of Short-Sale Constraints.*
Nordfält, Jens. *Is consumer decision-making out of control? Non-conscious influences on the consumer decision-making process for fast moving consumer goods.*
Nordin, Fredrik. *Externalising Services – Walking a Tightrope between Industrial and Service Logics.*
Simbanegavi, Witness. *Price Discrimination, Advertising and Competition.*
Thodenius, Björn. *Användning av ledningsinformationssystem: en longitudinell studie av svenska storföretag.*
Tolis, Christofer. *Framing the Business – Business Modelling for Business Development.*
Östberg, Per. *Corporate Disclosure and Investor Recognition.*

2004

Books

Ahrne, Göran och Nils Brunsson (red). *Regelexplosionen.*

Lind, Johnny. *Strategi och ekonomistyrning. En studie av sambanden mellan koncernstrategi, affärsstrategi och ekonomistyrning.*

Lind, Johnny och Walter Schuster (red). *Redovisningens teori, praktik och pedagogik. En vänbok till Lars Östman.*

Sevón, Guje och Lennart Sjöberg (red). *Emotioner och värderingar i näringslivet.* EFIs Årsbok 2004.

Wijkström, Filip and Stefan Einarsson. *Foundations in Sweden – Their scope, roles and visions.*

Dissertations

Anderson, Anders. *Essays in Behavioral Finance.*

Balsvik, Gudrun. *Information Technology Users: Studies of Self-Efficacy and Creativity among Swedish Newspaper Journalists.*

Blix, Magnus. *Essays in Mathematical Finance – Modelling the Futures Price.*

González Gómez, Andrés. *Nonlinear dynamics and smooth transition models.*

Grönqvist, Erik. *Selection and Moral Hazard in Health Insurance: Taking Contract Theory to the Data.*

Ivarsson Westerberg, Anders. *Papperspolisen – varför ökar administrationen i moderna organisationer.*

Jutterström, Mats. *Att påverka beslut – företag i EUs regelsättande.*

Jönsson, Kristian. *Macroeconomic Aspects of Capital Flows to Small Open Economies in Transition.*

Larsson, Pär. *Förändringens villkor. En studie av organisatoriskt lärande och förändring inom skolan.*

Lagerwall, Björn. *Empirical Studies of Portfolio Choice and Asset Prices.*

Malmsten, Hans. *Properties and Evaluation of Volatility Models.*

Marshall, Cassandra. *Dating for Innovation. Creating and Recognizing Opportunities through Collaborative Interorganizational Relationships in Fluid Environments.*

Mattsson, Susanna. *På gränsen mellan ordning och oordning – tingens betydelse vid marknadsombildningar. En studie av svenska postväsendets ombildning under 1990-talet.*

Nilsson, Charlotte. *Studies in Environmental Economics: Numerical Analysis of Greenhouse Gas Policies.*

Nilsson, Hans. *Medborgaren i styrsystemet – beskrivning av VAD och HUR i styrning av kommunal verksamhet.*

Nystedt, Jens. *Competition, Regulation and Integration in International Financial Markets.*

Pajuste, Anete. *Corporate Governance and Controlling Shareholders.*

Richtnér, Anders. *Balancing Knowledge Creation. Organizational Slack and Knowledge Creation in Product Development.*

Salabasis, Mickael. *Bayesian Time Series and Panel Models – Unit Roots, Dynamics and Random Effects.*

Sandberg, Rickard. *Testing the Unit Root Hypothesis in Nonlinear Time Series and Panel Models.*

Skallsjö, Sven. *Essays on Term Structure and Monetary Policy.*

Strikholm, Birgit. *Essays on Nonlinear Time Series Modelling and Hypothesis Testing.*

Söderström, John. *Från Produkt till Tjänst. Utveckling av affärs- och miljöstrategier i produktorienterade företag.*

Talia, Krim. *The Scandinavian Currency Union, 1873–1924 – Studies in Monetary Integration and Disintegration.*

2003

Books

Lundahl, Mats (editor). *Globalization and Its Enemies.* EFIs Årsbok 2003.

Sundgren, Bo, Pär Mårtensson, Magnus Mähring and Kristina Nilsson (editors). *Exploring Patterns in Information Management. Concepts and Perspectives for Understanding IT-Related Change.*

Dissertations

Andersson, Henrik. *Valuation and Hedging of Long-Term Asset-Linked Contracts.*

Bergman, Malin. *Essays on Human Capital and Wage Formation.*

Damsgaard, Niclas. *Deregulation and Regulation of Electricity Markets.*

Eklund, Bruno. *Four Contributions to Statistical Inference in Econometrics.*

Hakkala, Katariina. *Essays on Restructuring and Production Decisions in Multi-Plant Firms.*

Holgersson, Charlotte. *Rekrytering av företagsledare. En studie i homosocialitet.*

Ivaschenko, Iryna. *Essays on Corporate Risk, U.S. Business Cycles, International Spillovers of Stock Returns, and Dual Listing.*

Lange, Fredrik. *Brand Choice in Goal-derived Categories – What are the Determinants?*

Le Coq, Chloé. *Quantity Choices and Market Power in Electricity Market.*

Magnusson, Peter R. *Customer-Oriented Product Development – Experiments Involving Users in Service Innovation.*

Meisiek, Stefan. *Beyond the Emotional Work Event Social Sharing of Emotion in Organizations.*

Mårtensson, Anders. *Managing Mission-Critical IT in the Financial Industry.*

Nilsson, Göran. *Processorientering och styrning – Regler, mål eller värderingar?*

Sandberg, Robert. *Corporate Consulting for Customer Solutions Bridging Diverging Business Logics.*

Sturluson, Jon Thor. *Topics in the Industrial Organization of Electricity Markets.*

Tillberg, Ulrika. *Ledarskap och samarbete – En jämförande fallstudie i tre skolor.*

Waldenström, Daniel. *Essays in Historical Finance.*

Wallén, Ulrika. *Effektivitet i grundskolan i anslutning till en stadsdelsnämndsreform.*

Ögren, Anders. *Empirical Studies in Money, Credit and Banking – The Swedish Credit Market in Transition under the Silver and the Gold Standards, 1834–1913.*

2002

Books

Schuster, Walter. *Företagets Valutarisk – En studie av horisontella och vertikala styrprocesser.*

Sjöstrand, Sven-Erik och Pernilla Petrelius. *Rekrytering av koncernstyrelsen – Nomineringsförfaranden och styrelsesammansättning med fokus på kvinnors ställning och möjligheter.* EFI/SNS Förlag

Löwstedt, Jan och Bengt Stymne (red). *Scener ur ett företag – Organiseringsteori för kunskapssamhället.* EFIs Årsbok 2002.EFI/Studentlitteratur.

Dissertations

Barinaga, Ester. *Levelling Vagueness – A Study of Cultural Diversity in an International Project Group.*

Berglund, Johan. *De otillräckliga – En studie av personalspecialisternas kamp för erkännande och status.*
Bolander, Pernilla. *Anställningsbilder och rekryteringsbeslut.*
Damjanovic, Tatiana. *Essays in Public Finance.*
Ekman, Mattias. *Studies in Health Economics – Modelling and Data Analysis of Costs and Survival.*
Heyman, Fredrik. *Empirical Studies on Wages, Firm Performance and Job Turnover.*
Kallifatides, Markus. *Modern företagsledning och omoderna företagsledare.*
Kaplan, Michael. *Acquisition of Electronic Commerce Capability – The Cases of Compaq and Dell in Sweden.*
Mähring, Magnus. *IT Project Governance.*
Nilsson, Mattias. *Essays in Empirical Corporate Finance and Governance.*
Schenkel, Andrew. *Communities of Practice or Communities of Discipline – Managing Deviations at the Øresund Bridge.*
Skogsvik, Stina. *Redovisningsmått, värderelevans och informationseffektivitet.*
Sundén, David. *The Dynamics of Pension Reform.*
Ternström, Ingela. *The Management of Common-Pool Resources – Theoretical Essays and Empirical Evidence.*
Tullberg, Jan. *Reciprocitet – Etiska normer och praktiskt samarbete.*
Westling, Gunnar. *Balancing Innovation and Control – The Role of Face-to-face Meetings in Complex Product Development Projects.*
Viklund, Mattias. *Risk Policy – Trust, Risk Perception, and Attitudes.*
Vlachos, Jonas. *Risk Matters – Studies in Finance, Trade and Politics.*

2001

Books

Sjöstrand, Sven-Erik. *Center for Management and Organization 50 (1951–2001).*
Charpentier, Claes. *Uppföljning av kultur- och fritidsförvaltningen efter stadsdelsnämnds-reformen.*
Hedlund, Andreas. *Konsumentens erfarenhet – och dess inverkan på livsmedelsinköp på Internet.*
Hvenmark, Johan. *Varför slocknar elden? Om utbrändhet bland chefer i ideella organisationer.*
Jönsson, Bengt (red). *Läkemedel – Kostnad eller resurs för sjukvården?* EFIs Årsbok 2001.
Lange, Fredrik och Richard Wahlund. *Category Management – När konsumenten är manager.*
Ljunggren, Ulrika. *Nyckeltal i grundskolan i Stockholms stad före och efter stadsdelsnämnds-reformen.*
Thorén, Bertil. *Stadsdelsnämndsreformen och det ekonomiska styrsystemet – Om budget-avvikelser.*

Dissertations

Adolfson, Malin. *Monetary Policy and Exchange Rates – Breakthrough of Pass-Through.*
Andersson, Patric. *Expertise in Credit Granting: Studies on Judgment and Decision-Making behavior.*
Björklund, Christina. *Work Motivation – Studies of its Determinants and Outcomes.*
Dahlén, Micael. *Marketing on the Web – Empirical Studies of Advertising and Promotion Effectiveness.*
Eckerlund, Ingemar. *Essays on the Economics of Medical Practice Variations.*

Ekelund, Mats. *Competition and Innovation in the Swedish Pharmaceutical Market.*
Engström, Stefan. *Success Factors in Asset Management.*
Ericsson, Daniel. *Kreativitetsmysteriet – Ledtrådar till arbetslivets kreativisering och skrivandets metafysik.*
Eriksson, Rickard. *Price Responses to Changes in Costs and Demand.*
Frisell, Lars. *Information and Politics.*
Giordani, Paulo. *Essays in Monetary Economics and Applied Econometrics.*
Gustavsson, Patrik. *Essays on Trade, Growth and Applied Econometrics.*
Hill, Martin. *Essays on Environmental Policy Analysis: Computable General Equilibrium Approaches Applied to Sweden.*
Hägglund, Peter B. *Företaget som investeringsobjekt – Hur placerare och analytiker arbetar med att ta fram ett investeringsobjekt.*
Höök, Pia. *Stridspiloter i vida kjolar, om ledarutveckling och jämställdhet.*
Johansson, Christer. *Styrning för samordning.*
Josephson, Jens. *Evolution and Learning in Games.*
Kjellberg, Hans. *Organising Distribution – Hakonbolaget and the efforts to rationalise food distribution, 1940–1960.*
Liljenberg, Anders. *Customer-geared competition – A socio-Austrian explanation of Tertius Gaudens.*
Lindkvist, Björn. *Kunskapsöverföring mellan produktutvecklingsprojekt.*
Löf, Mårten. *On Seasonality and Cointegration.*
Martensen, Kaj. *Essays on Entry Externalities and Market Segmentation.*
Matros, Alexandr. *Stochastic Stability and Equilibrium Selection in Games.*
Mårtensson, Pär. *Management Processes – An Information Perspective on Managerial Work.*
Nilsson, Arvid. *Market Transparency.*
Norberg, Peter. *Finansmarknadens amoralitet och det kalvinska kyrkorummet – En studie i ekonomisk mentalitet och etik.*
Persson, Björn. *Essays on Altruism and Health Care Markets.*
Rech, Gianluigi. *Modelling and Forecasting Economic Time Series with Single Hidden-layer Feedforward Autoregressive Artificial Neural Networks.*
Skoglund, Jimmy. *Essays on Random Effects Models and GARCH.*
Strand, Niklas. *Empirical Studies of Pricing.*

2000

Books

Bengtsson, Lars, Johnny Lind och Lars A. Samuelson (red). *Styrning av team och processer – Teoretiska perspektiv och fallstudier.*
Berg-Suurwee, Ulrika. *Styrning före och efter stadsdelsnämndsreform inom kultur och fritid – Resultat från intervjuer och enkät.*
Brodin, Bengt, Leif Lundkvist, Sven-Erik Sjöstrand och Lars Östman. *Koncernchefen och ägarna.*
Charpentier, Claes och Lars A. Samuelson. *Effekter av en sjukvårdsreform – En analys av Stockholmsmodellen.*
Emling, Emil. *Svenskt familjeföretagande.*
Ericson, Mona. *Strategi, kalkyl, känsla.*
Gunnarsson, Jonas, Richard Wahlund och Helena Flink. *Finansiella strategier i förändring: segment och beteenden bland svenska hushåll.*
Ljunggren, Ulrika. *Styrning av grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen – Resultat från intervjuer och enkät.*

Schwarz, Brita och Susanne Weinberg. *Serviceproduktion och kostnader – Att söka orsaker till kommunala skillnader.*

Söderlund, Magnus (red). *I huvudet på kunden.* EFIs Årsbok 2000. EFI/Liber Förlag.

Dissertations

Bergkvist, Lars. *Advertising Effectiveness Measurement: Intermediate Constructs and Measures.*

Bornefalk, Anders. *Essays on Social Conflict and Reform.*

Edman, Jan. *Information Use and Decision Making in Groups – A Study of an Experimental Oligopoly Market with the Use of a Business Game.*

Hellman, Niclas. *Investor Behaviour – An Empirical Study of How Large Swedish Institutional Investors Make Equity Investment Decisions.*

Hyll, Magnus. *Essays on the Term Structure of Interest Rates.*

Håkansson, Per. *Beyond Private Label – The Strategic View on Distributor Own Brands.*

Karlsson Stider, Annelie. *Familjen och firman.*

Ludvigsen, Johanna. *The International Networking between European Logistical Operators.*

Nittmar, Henrik. *Produktutveckling i samarbete – Strukturförändring vid införande av nya informationssystem.*

Robertsson, Göran. *International Portfolio Choice and Trading Behavior.*

Stenström, Emma. *Konstiga företag.*

Sweet, Susanne. *Industrial Change Towards Environmental Sustainability – The Case of Replacing Chloroflouorocarbons.*

Tamm Hallström, Kristina. *Kampen för auktoritet – standardiseringsorganisationer i arbete.*