ESSAYS ON EMPIRICAL ASSET PRICING USING BAYESIAN METHODS

by

Alexandre Rubesam

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

CITY UNIVERSITY

Faculty of Finance, Cass Business School

January/2009

Supervisor: Dr. Soosung Hwang

Table of Contents

List of Tables

List of Illustrations

Acknowledgements

I am very grateful to Dr. Soosung Hwang for his invaluable guidance and support. I would also like to thank the academic and general staff at Sir John Cass Business School.

This research could not have been done without the financial support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a Brazilian governmental agency which sponsored this project. I am thankful to them and to the Brazilian taxpayers, who are ultimately responsible for that support.

I would like to thank my direct and extended family and specially my parents, who have supported my decision to pursue a PhD degree in many ways.

I would like to thank my friend Dr. Helder Parra Palaro, with whom I have shared the experience of higher education over the last 10 years.

Finally, I am especially grateful to Beatriz Singer, who has witnessed first-hand a good part of my PhD journey and has given me great support through her love.

Copyright waiver

Abstract

This thesis is composed of three essays related to empirical asset pricing. In the first essay of the thesis, we investigate recent rational explanations of the value premium using a regime-switching approach. Using data from the US stock market, we investigate the risk of value and growth in different market states and using alternative risk measures such as downside beta and higher moments. Our results provide little or no evidence that value is riskier than growth, and that evidence is specific to pre-1963 period (including the Great Depression). Within the post-1963 sample, there are periods when the value premium can be explained by the CAPM, whilst during other periods the premium is explained by the fact that the returns on value firms increase more than the returns on growth stocks in periods of strong market performance, whilst in downturns growth stocks suffer more than value, and these features are captured by different upside/downside betas or higher moments. These results are not consistent with a risk-based explanation of the value premium.

The second essay of the thesis contributes to the debate about the momentum premium. We investigate the robustness of the momentum premium in the US over the period from 1927 to 2006 using a model that allows multiple structural breaks. We find that the risk-adjusted momentum premium is significantly positive only during certain periods, notably from the 1940s to the mid-1960s and from the mid-1970s to the late 1990s, and we find evidence that momentum has disappeared since the late 1990s. Our results suggest that the momentum premium has been slowly eroded away since the early 1990s, in a process which was delayed by the occurrence of the high-technology stock bubble of the 1990s. In particular, we estimate that the bubble accounts for at least 60% of momentum profits during the period from 1995 to 1999.

In the final essay of this thesis, we study the question of which asset pricing factors should be included in linear factor asset pricing model. We develop a simple multivariate extension of a Bayesian variable selection procedure from the statistics literature to estimate posterior probabilities of asset pricing factors using many assets at once. Using a dataset of thousands of individual stocks in the US market, we calculate posterior probabilities of 12 factors which have been suggested in the literature. Our results indicate strong and robust evidence that a linear factor model should include the excess market return, the size and the liquidity factors, and only weak evidence that the idiosyncratic volatility and downside risk factors matter. We also apply our methodology to portfolios of stocks commonly used in the literature, and find that the famous Fama and French (1993, 1996) HML factor has high posterior probability only if portfolios formed on book-to-market ratio are used.

Chapter 1

Introduction

The ultimate task of asset pricing, as Cochrane (2001) puts it, is "to understand and measure the sources of aggregate or macroeconomic risk that drive asset prices". Advances in theory over the last 50 years - as well as the ever-increasing availability of large amounts of data and cheap computational power - have expanded our understanding of how asset prices behave enormously. However, one could argue that this task remains unfinished.

The development of the Capital Asset Pricing Model by Sharpe (1964), Lintner (1965) and Black (1972) provided an elegant theoretical solution to understand the returns on risky assets. In a Markowitz (1952) efficient market, the expected returns on risky assets are determined solely by their covariation with the market portfolio of wealth. In other words, expected returns are a linear function of betas. The simplicity and intuitive appeal of this model made it extremely popular amongst practitioners. However, subsequent empirical studies provided and continue to provide evidence that beta is unable to explain asset prices. The typical study of this kind looks for variables – most often firm characteristics - associated with a spread in average returns. Stocks are ranked according to their corresponding values of these variables and a hedge portfolio, which is long stocks in one end of the spectrum and short in the other, is formed. If the risk-adjusted return from this hedge portfolio (i.e. the CAPM alpha) is significantly positive, the researcher concludes that this is evidence against the CAPM. Three cases in particular have received a vast amount of attention from academic works, namely the so-called size, value and momentum effects, which refer to the inability of the CAPM to price portfolios formed on market capitalisation, the ratio of book to market equity and past returns, respectively.

The presence of these empirical irregularities is interpreted in a variety of ways. Whilst some researchers believe that they indicate the presence of market inefficiencies or biased investor behaviour and perception, others believe that they represent additional sources of risk which are not captured by the CAPM beta. Another possibility is that they are the result of data-snooping, i.e. by looking at thousands of possible firm characteristics, researchers have found some which are related to average returns by sheer chance. The debate about the value premium is illustrative of these different views. Fama and French (1993, 1996), for instance, argue very strongly that the value effect is related to a financial distress risk factor, but other studies (e.g. Lakonishok, Shleifer and Vishny (1994)) argue in favour of the behavioural story. Fama and French further propose a three-factor model including the market return and two additional factors related to the size and value effects. The idea that multiple risk factors are important to determine asset prices is incorporated in the theories of the Intertemporal CAPM of Merton (1973) and the Arbitrage Pricing Theory of Ross (1976). These theories, however, are silent as to which and how many these factors might be, and this is left as a largely empirical issue.

There are two closely related issues in empirical asset pricing. The first one focuses in understanding the sources of these empirical irregularities, particularly whether they are related to non-diversifiable risk. The second is to identify, amongstst many empirical possibilities, which should be used as risk factors in asset pricing models. The first two essays in this thesis are related to the first issue, specifically to the value and momentum anomalies. The third essay is related to the second issue; in particular, which factors should be included in a linear factor model to prices stock returns.

The first essay in this thesis, entitled *Is Value Really Riskier Than Growth?*, contributes to the debate surrounding the value premium. It is a commonly-held belief that growth options are riskier than assets in place, because growth options depend more on future and uncertain economic conditions. However, value firms (whose values come mostly from existing assets) earn higher average returns than growth firms (whose values comes mostly from growth options) but have lower betas in the post-1963 period. The value premium has drawn considerable attention from both academics and practitioners alike. Academics would like to definitively explain the source of the premium. If the value premium is, as Fama and French (1993) advocate, related to a real, aggregate, non-diversifiable source of risk, academics would like to understand exactly what this risk is. Practitioners, on the other hand, are concerned with whether, if genuine, this anomaly is going to persist in the future.

A recent rational theory to explain the value premium, developed by Zhang (2005), argues that value firms are riskier than growth firms, especially in bad states of the economy, because of costly reversibility (it is more costly for value firms, with more assets in place than growth firms, to scale down production) and countercyclical price of risk. On the other hand, in good times or expansions, value firms benefit from having more assets in place, whilst growth firms need to invest in order to take advantage of the favourable environment. Although this theory suggests a role for time-varying risk in explaining the premium, the empirical evidence in that respect is mixed.

We contribute to the debate about the value premium by investigating the risk of value, growth and value-minus-growth portfolios using a regime-switching approach. We do this in two different ways. First, we estimate the systematic risk of value and growth in different market states, inferred from a regime-switching model of the market return. This approach is appealing because it does not rely on subjective choices of conditioning variables, or a high degree of model parameterisation, whilst allowing us to test the implications of Zhang's (2005) theory. Second, we propose a regime-switching model which allows alternative risk measures such as downside beta or higher moments to be used at different time periods. This model could be considered a generalised equilibrium model when investors' risk aversion changes over time and asset returns do not follow the normal distribution but are governed by higher moments such as skewness and kurtosis. These asymmetric risk measures might be particularly relevant, considering that theory suggests that value and growth are expected to react differently to periods of good and bad economic conditions. We consider two alternatives to the CAPM: the Lower Partial Moment CAPM (LCAPM), originally developed by Bawa and Lindenberg (1977) and Harlow and Rao (1989), which is a general framework to consider the systematic risk of an asset when the return is below or above a certain threshold (from which the downside beta is derived), and the Higher Moment CAPM (HCAPM), introduced by Kraus and Litzenberger (1976). We use the version of the HCAPM in Hwang and Satchell (1999), which includes two measures of risk in addition to the assets' beta: coskewness and cokurtosis.

We investigate a number of value, growth and value-minus-growth portfolios from the US stock market. Our results do not support the risk-based explanation of the value premium. When we identify the market state through regimes extracted from the market return, there is little or no difference in the risk of value-minus-growth portfolios across market states, and this difference does not explain the value premium. The analysis with different risk measures suggests that, in the post-1963 period, there are periods when the value premium can be explained by the CAPM, and other periods when the returns on value stocks increase much more than those on growth stocks, which is captured by different upside/downside betas or higher moments. Our results also suggest that the value premium is likely to be high during periods of bad market performance because of the negative returns of the growth portfolio.

In the second essay of this thesis, entitled *Disappearance of Momentum*, we investigate whether momentum strategies have been consistently profitable over time. The momentum anomaly has been a major focus of research since the publication of an influential paper by Jegadeesh and Titman (1993), who documented that simple strategies that buy stocks that had high returns in the previous 3 to 12 months and short stocks that had low returns in the same period earn an abnormal return of approximately 1% per month over a holding period of up to 12 months. How and why such a profitable opportunity appeared to persist for such a long period of time is a perplexing question; in an efficient market, arbitrageurs are expected to quickly drive away these profits.

To investigate this issue, we examine the profitability of several momentum strategies formed with stocks from the US market over a long period from 1927 to 2006, using a model which allows multiple structural breaks. We find that the momentum premium is significantly positive only during certain periods of time. Particularly, the last structural break we find occurred around the year 2000, and since then the momentum premium becomes insignificant. Also, although there have been periods of insignificant momentum premium in the past, we find that the momentum premium in this recent period is not probable considering the past distribution of momentum premia, which indicates that the anomaly might have been eroded away.

We further try to answer the question of why it took so long for the momentum premium to disappear, once the anomaly was first reported in the early 1990s. We seek an answer from the extraordinary boom and bust in the hi-technology and telecommunication sectors in the late 1990s. During that period, we observed several years of extraordinary performance in hi-tech and telecom stocks, and these acted as winners in momentum strategies. In order to see this, we decompose momentum profits by sectors of the industry and find that, during the 1995-1999 period, hi-tech and telecom stocks were responsible for approximately 60% of the profits of the momentum strategy. When we remove the effect of these stocks, a declining pattern for the profitability of the momentum strategy emerges since the early 1990s. We interpret this pattern as a slow erosion of the profits of the momentum strategy by market participants. Consistent with this idea, during this period the number and size of hedge funds - arbitrageurs who can undertake long-short trading strategies such as the momentum strategy - have increased enormously. During the 1990-2006 period, the number of hedge funds has increased from around

500 to more than 9000, and the assets under management by these funds has risen from 50 billion to over 1.5 trillion US dollars[1].

Finally, in the last essay in this thesis, *Fishing with a Licence: an Empirical Search for Asset Pricing Factors*, we investigate which factors should be included in a linear factor model to explain stock returns. We use a Bayesian approach to calculate posterior probabilities of possible factors. Our methodology is based on a Bayesian variable selection procedure from the statistics literature called Stochastic Search Variable Selection (SSSV), introduced by George and McCulloch (1993). We extend their approach to a simple multivariate case with N assets, i.e. we are interested in calculating the posterior probabilities of factors to explain many assets simultaneously. Our approach has several advantages. First, our method focuses on obtaining the posterior probabilities of the more promising models directly, without the need to estimate the posterior probabilities of all possible models, which can be quite an overwhelming task considering the growing number of possible factors. Second, it allows us to use data on thousands of individual stocks, which is important considering the data-snooping biases inherent when factors created by sorting on variables are tested on portfolios related to these variables (Lo and MacKinlay (1990), Ferson, Sarkissian and Simin (1999) and Berk (2000)).

We test 12 factors that have been reported in the literature. These are: the excess market return, size, value, momentum, asset growth, idiosyncratic volatility, trading volume, long-term reversal, liquidity, coskewness, cokurtosis and downside risk. We apply our methodology to a large number of individual stocks as well as portfolios of stocks from the US market. Our results suggest that a linear factor model for stock returns should contain the excess market return, the size factor and

[1] HedgeFund Intelligence Press Release.

the liquidity factor. We find only weak evidence that the idiosyncratic volatility, the value/growth and the downside risk factors should be included. Also, our results with individual stocks and portfolios of stocks differ dramatically. The posterior probability of the Fama and French (1993) value/growth factor (HML) is high only when it is estimated with portfolios formed using the book-to-market ratio, but the size factors (SMB) has high posterior probability regardless of the assets used.

References

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

Berk, Jonathan B., 2000, Sorting out sorts, *The Journal of Finance* 55, 407-427.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *The Journal of Business* 45, 444-455.

Cochrane, John, 2001, *Asset pricing* (Princeton University Press, Princeton and Oxford).

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns of stocks and bonds, *The Journal of Financial Economics* 33, 357-384.

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

Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin, 1999, The alpha factor asset pricing model: A parable, *Journal of Financial Markets* 2, 49-68.

George, Edward I., and Robert E. McCulloch, 1993, Variable selection via gibbs sampling, *Journal of the American Statistical Association* 88, 881-889.

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

Hwang, S., and S. E. Satchell, 1999, Modelling emerging market risk premia using higher moments, *International Journal of Finance and Economics* 4.

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

Kraus, A., and R. H. Litzenberger, 1976, Skewness preference and the valuation of risky assets, *The Journal of Finance* 31, 1085-1100.

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

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

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

Markowitz, H., 1952, Portfolio selection, *The Journal of Finance* 7, 77-91.

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

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

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

Zhang, L., 2005, The value premium, *The Journal of Finance* 60, 67-103.

Chapter 2

Is Value Really Riskier than Growth?

2.1. Introduction

Many studies have tried to explain the value premium identified by Rosenberg, Reid and Lanstein (1985) and Fama and French (1992, 1993). To the proponents of conventional asset pricing theory, the value premium is a puzzle, since growth firms, whose values come mostly from growth options, are expected to depend more on future business cycles and thus be riskier than value firms, whose values are derived more from existing assets. Thus growth stocks are expected to have higher betas and higher returns than value stocks. Empirical evidence, however, does not support this argument; value stocks have higher returns than growth stocks but have lower betas. For example, the famous Fama-French's HML (high book-to-market portfolio returns minus low book-to-market portfolio returns) returns 0.45% a month (standard error 0.13) for the period from January 1963 to December 2006. Even after considering the systematic risk, the estimated alpha is 0.58% which is highly significant. Recently Chen, Petkova and Zhang (2006) show that there is not much evidence that the value premium has weakened in recent times, so the anomaly persists.

Fama and French (1993, 1995) argue that HML is a risk factor which represents financial distress of weak firms with low earnings, which tend to have high book-to-market ratios. On the other hand, Lakonishok, Shleifer and Vishny (1994) suggest investors' incorrect extrapolation of the past earnings growth of firms as the source of the value premium. Others try to explain the premium in the framework of the CAPM, with mixed results. For example, Jagannathan and Wang (1996) and Ang and Chen (2007) propose conditional CAPM models. Lewellen and Nagel (2006) and Petkova and Zhang (2005) also use conditional CAPM models to investigate the value premium, but their results are not as strong as those of Ang and Chen (2007). Campbell and Vuolteenaho (2004), on the other hand, decompose the beta of a stock into the 'good' beta that comes from news about the discount rate and the 'bad' beta from news about the future cash flows, and show that value stocks have higher proportion of 'bad' betas. Some of these studies, however, have been criticised by Daniel and Titman (2005), who show that their favourable results could be due to the low power of the tests used.

Recently, risk-based explanations for the value premium have been proposed which seek answers from the investment inflexibility of value firms upon economic conditions in the framework of real options models. Zhang (2005), for example, provides an explanation in the neoclassical framework with rational expectations and competitive equilibrium; value firms are less flexible and thus riskier than growth firms, especially in bad times, since value firms have more assets in place and thus it is more costly for these firms to scale down production than for growth firms which have fewer assets in place. On the other hand, in expansions value firms are favoured in detriment of growth firms, which need to invest to take advantage of the optimistic economic environment. Most of the empirical evidence supporting this argument is provided in the firm or industry level using firm characteristics. Xing and Zhang (2005) show that value firms in the manufacturing sector perform worse than growth firms in the negative business cycle and vice versa, using variables such as earnings growth, sales growth, investment growth, and investment rate, whilst Cooper, Gerard and Wu (2005) investigate the link between the rate of capacity and the degree of investment irreversibility and the book-to-market ratio.

One caveat with the empirical studies above is that they use subsets of stocks, whilst the value premium is calculated using a much larger number of stocks from the whole market. For example, the number of manufacturing firms Xing and Zhang (2005) use in their study is only 21% (37% in market capitalisation) of the firms publicly traded in the market. In addition, the firm characteristics used by these studies may reflect business cycles, but are not necessarily concurrent with the movements in financial markets because the lead and lag relationship between the firm characteristics and the dynamics in the stock market is not likely to be constant.

Some of the studies mentioned above attempt to model time-varying risk and the expected market risk premium directly. We do not follow this path for two reasons. First, as mentioned before, studies based on this approach fail to show conclusive evidence that time-varying risk explains the value premium. Second, the choice to use a conditional model and to estimate the expected market risk premium involves either a subjective decision about which conditioning variables to use, or a high degree of model parameterisation (as in Ang and Chen (2007)).

Instead, motivated by the literature on regime-switching models and business cycles, particularly the work of Hamilton and Lin (1996), we model market returns with a regime-switching process to identify different market states as regimes with different mean market returns and volatilities. This approach allows us to study the risk of the value-minus-growth strategy in different market conditions, whilst avoiding a high degree of parameterisation and subjective choices of conditioning variables. Specifically, we investigate the risk of value and growth by estimating the CAPM conditioned on the state of the market as inferred from the regime-switching model of the market return. If value is riskier than growth during bad states when the market is doing poorly, there should be an increase in the beta of HML in the corresponding market regime[2].

The approach above, however, may be criticised in two points. First, this analysis assumes that the risk of HML is related to the state of the economy as measured by the market regimes. Even if the risk of HML is not related to the different market regimes, there could still be increases (decreases) in this risk over different periods. Second, using beta as the measure of risk might not be adequate, since investors are not likely to have quadratic utility and asset returns do not follow the normal distribution. If the investors' utility function is better specified by power utility and higher moments such as skewness and kurtosis matter in asset return distribution, asymmetry or fat tails should be priced. These asymmetric risk measures might be particularly relevant, considering that value and growth are expected to react differently to periods of good and bad economic conditions. Therefore we consider the possibility that different risk measures other than the CAPM beta, such as the downside beta and higher moments, matter to explain portfolios' returns over time, without the explicit assumption that they must be related to the state of the market or economy.

[2] This approach is related to the one followed by Petkova and Zhang (2005, section 4), except that they regress HML returns on a dummy variable indicating whether the estimated expected market risk premium is lower than its sample average (good times) or higher (bad times).

Several asset pricing models have been proposed to explain asymmetry and fat tails. Amongst these, we choose two widely known equilibrium asset pricing models in addition to the CAPM: the lower partial moment CAPM (henceforth LCAPM) and the higher moment CAPM (henceforth HCAPM). The LCAPM, which was developed by Bawa and Lindenberg (1977) and Harlow and Rao (1989), includes asymmetric reactions to downside and upside markets separately. Chan (1988), De Bondt and Thaler (1987), and Petkova and Zhang (2005) use it to investigate the value premium but give us mixed results. The HCAPM introduced by Kraus and Litzenberger (1976) prices higher moments. A closely related study is Harvey and Siddique (2000) who model conditional skewness. Higher moments explain asset returns with asymmetry or fat tails, but are not necessarily the same as the upside and downside betas.

One major difference of the approach above from those of previous studies is that any of the three models − CAPM, LCAPM, and HCAPM − can explain asset returns in the regime switching framework we employ. Since these are based on equilibrium models, our approach is still within the rationality framework. Thus we seek explanations on the value premium in the conventional risk-return framework by concentrating on the empirical possibility of changing risk measures and its impact on asset pricing. When there are no differences in upside and downside betas or when higher moments are not priced, the model is equivalent to the conventional CAPM. Therefore the LCAPM or HCAPM are selected only when asymmetries or fat tails matter in asset pricing.

With appropriate risk measures chosen for different time periods, we investigate whether or not value firms are riskier than growth firms. As argued by Zhang (2005), if value firms are riskier than growth firms during troughs, the asymmetric models should show that the downside beta is higher than upside beta for value-minus-growth portfolios during bear markets or that the coefficients on higher moments should be significant during bear markets such that value firms become riskier.

Our results show that, when we identify the market state through a regime-switching model for the market return, there is little or no difference in the risk of value-minus-growth portfolios across market regimes, and this difference does not explain the value premium, except when the pre-1963 sample (including the Great Depression) is used. Moreover, when we investigate the value premium using different risk measures, we find that there are periods of time when the premium can be explained by the CAPM, whilst during other periods the premium is explained by the fact that the returns of value firms increase more than the returns on growth stocks in periods of strong market performance, whilst in downturns growth stocks suffer more than value stocks. These features are captured by the upside/downside betas in the LCAPM or by the coefficient on the square and cube of the market return in the HCAPM. Overall, our results are not consistent with a risk-based explanation of the value premium.

The rest of this work is organised as follows. In Section 2.2 we explain the methodology used. Section 2.3 contains the explanation of the data set and the empirical results including robustness tests. Section 2.4 concludes.

2.2. Regime-Switching Model

In this section we first explain different risk measures and why they could be used to model the value premium. We then introduce a regime switching model that allows these risk measures to be used over time, and the method used to estimate it.

2.2.1. Lower Partial Moment CAPM

Previous studies test either betas, consumption betas or conditional betas of value-minus-growth portfolios as the relevant measure of risk. However, the asymmetric reactions of value and growth firms to market conditions, proposed by Zhang (2005), might not be well explained by symmetric models such as the CAPM. An alternative model to test this theory would be the downside/upside CAPM, which allows different response to positive and negative market movements.

This upside/downside CAPM model (the Lower Partial Moment CAPM, LCAPM) initially developed by Bawa and Lindenberg (1977) and Harlow and Rao (1989), has been a popular method to investigate asymmetric reactions to market movements or if downside beta is priced.[3] This model suggests that investors react differently to the returns above or below a specified target return:

$$r_{pt} = \alpha_{LCAPM} + \beta^{-} r_{mt}^{-} + \beta^{+} r_{mt}^{+} + \varepsilon_t , \tag{2.1}$$

[3] See Kim and Zwalt (1979), Chen (1982), De Bondt and Thaler (1987), Chan (1988), Harlow and Rao (1989), Petkova and Zhang (2005), Post and Van Vliett (2005) and Ang, Chen and Xing (2006).

where r_{pt} is the excess return on portfolio *p*, $r_{mt} = R_{mt} - r_f$ is the excess market return, $r_{mt}^{+} = r_{mt} I\left(r_{mt} > r_{\text{target}}\right)$ and $r_{mt}^{-} = r_{mt} I\left(r_{mt} < r_{\text{target}}\right)$ are the positive and negative components of the excess market return relative to the target return r_{target}, *I(.)* is the indicator variable, and ε_t is an error term with mean zero and standard deviation σ_p.

Although the LCAPM has an appealing property for the explanation of asymmetry in asset returns, there are a few issues which remain unclear. First, this model cannot be directly estimated through OLS using equation (2.1), due to the issue discussed in Post and Van Vliett (2005): if the constant is included, the downside and upside betas calculated by OLS are not consistent with what the LCAPM suggests[4]. Consequently, throughout this study, we first estimate β^{-} and β^{+} by running a regression without the constant to obtain the estimates that are consistent with the LCAPM theory, and then calculate $\hat{\alpha}_{LCAPM}$ using these estimates as the mean and standard error of the residuals $\hat{\alpha}_{LCPAM} = r_{pt} - \hat{\beta}^{+} r_{mt}^{+} - \hat{\beta}^{-} r_{mt}^{-}$. Second, there is no agreement about how we should define the upside and downside markets that are inherently unobservable. Popular methods are setting the target return to zero or to the average market return. Petkova and Zhang (2005) obtain four market states (peak, expansion, recession, and trough) depending on the expected market risk premium calculated with four macroeconomic variables. Although this method is consistent with theory and thus appealing, the estimated market premium could be noisy, and the breakpoints between peak and expansion and between recession and

[4] Ignoring this issue, as most studies do, produces non-trivial distortions in upside and downside betas. For example, when we use the HML returns from 1963 to 2006 (data from Kenneth French's data library), the estimation of downside and upside betas (where the target return is zero) for the HML portfolio estimated *with* the constant term are -0.288 and -0.261 (not statistically different from each other), whilst those *without* the constant are -0.360 and -0.168 (different at less than 1% significance level). The conclusion in each case is quite different; in the former, there is no difference between upside and downside betas, whilst in the latter upside beta is larger than downside beta.

trough, i.e. 10% and -10%, are arbitrary. We will come back to this issue later in Section 2.3.3.

2.2.2. Higher Moment CAPM

Alternative risk measures could be obtained by generalising the assumption of the CAPM, i.e. the unrealistic mean-variance assumption based on normality or quadratic utility function. Assume that asset returns are non-normal *and* investors' utility function is not quadratic, which we believe is more realistic than normality or a quadratic utility function. The utility function can be linearised using the Taylor series expansion assuming compact returns (see Ingersoll (1987)). In equilibrium the expanded utility function is equivalent to the CAPM only when moments higher than the second moment are trivial. In general, for investors whose utility function is linear risk tolerant, asset prices need to be modelled by the CAPM with additional higher moment terms. The requirement of the additional terms depends on the probability density function of asset returns, which changes over time. Although both the HCAPM and LCAPM can model asymmetry in asset returns[5], the HCAPM models returns as a non-linear function of the market return, whilst in the LCAPM asset returns are priced linearly with market returns conditioning on up- and down-markets. Therefore the two models are not necessarily the same and in particular, the HCAPM can be used to price kurtosis in addition to skewness in asset returns.

[5] Bawa's (1978) study suggests that asset pricing with downside risk is consistent with a utility function that satisfies Arrow's properties of positive but decreasing marginal utility of wealth and decreasing absolute risk aversion.

There are several different approaches to include higher moments (see for example Kraus and Litzenberger (1976), Friend and Westerfield (1980), Sears and Wei (1985), Barone-Adesi (1985), Harvey and Siddique (2000)). As explained in Kraus and Litzenberger (1976) and Hwang and Satchell (1999), we use the following cubic market model which is consistent with the four moment (coskewness and cokurtosis) CAPM. The specification is:

$$r_{pt} = \alpha_{HCAPM} + \beta_1 r_{mt} + \beta_2 \left(R_{mt} - E\left(R_{mt}\right)\right)^2 + \beta_3 \left(R_{mt} - E\left(R_{mt}\right)\right)^3 + \varepsilon_t \,. \qquad (2.2)$$

2.2.3. Regime-Switching Model with Alternative Risk Measures

We consider three equilibrium based models, i.e. the traditional CAPM, the LCAPM that allows different responses to up- and down-market states and the HCAPM that models skewness and fat tails in addition to the traditional beta. Empirically, most of the previous studies test one of these models against the other models for certain sample periods and then conclude which one explains assets' returns better than the others. For example, beta appears to be priced before 1968 (Fama and MacBeth (1973)), but not from 1963 to 1990 (Fama and French (1992)). Lim (1989) reports that skewness is priced in some sub-periods. These empirical results indicate that asset returns are priced with different risk measures for different time periods.

In order to explain these empirical results, let us assume a market where asset returns are not normally distributed and investors' utility functions have the desirable properties of positive but decreasing marginal utility of wealth and decreasing absolute risk aversion as in Arrow (1971). Moreover assume that both the probability density function of asset returns and the investors' degree of risk aversion change over time (Campbell and Cochrane (1999)). Then, in this generalised framework the choice of risk measures does not lie in selecting one that dominates the others for the entire sample period, but in finding out which one is selected over time. When asset returns are normally distributed for a specific period of time, for example, the CAPM explains asset returns. However, when asset returns become non-normal, i.e. skewed or fat-tailed, and investors become far more risk averse, assets are priced not only by mean and variance but also by higher moments or downside risk. Therefore, different risk measures may be required for different periods of time to explain asset returns.

In order to model asset returns with different risk measures over different periods of time, we assume that there are *N* regimes defined by S_t, a random Markov regime variable that for each time *t* assigns a value in $\{1,\ldots,N\}$. When a dummy variable S_{jt}, *j*=1, 2, …, *N*, is defined for each regime, that is, $S_{jt}=1$ when $S_t = j$ and $S_{jt}=0$ otherwise, our model is given by

$$r_{pt} = \alpha_p + \sum_{j=1}^{N} S_{jt} m_{jt} + \varepsilon_{pt} , \tag{2.3}$$

where m_{jt} is a fully-specified relationship between the portfolio return r_{pt}, and the set of factors in regime *j*, and $\varepsilon_{pt} \sim \left(0, \sigma_{pt}^2\right)$, where $\sigma_{pt}^2 = \sum_{j=1}^{N} \sigma_{p,j}^2 S_{jt}$.

For m_{jt} the following three models are used. First, for the CAPM we have

$$m_{1t} = \beta r_{mt} . \tag{2.4}$$

Following the discussion in subsections 2.2.1 and 2.2.2, we allow the LCAPM and the HCAPM to be selected by defining

$$m_{2t} = \beta^- r_{mt}^- + \beta^+ r_{mt}^+ \tag{2.5}$$

and

$$m_{3t} = \beta_1 r_{mt} + \beta_2 \left(R_{mt} - E\left(R_{mt}\right)\right)^2 + \beta_3 \left(R_{mt} - E\left(R_{mt}\right)\right)^3 . \tag{2.6}$$

When the data generating processes follow equations (2.4), (2.5) and (2.6), they are *equivalent* to the CAPM, LCAPM, and HCAPM, respectively. Therefore, our regime switching model can be presented as:

$$\begin{aligned} r_{pt} = \alpha_p + S_{1t}\left[\beta r_{mt}\right] + S_{2t}\left[\beta^- r_{mt}^- + \beta^+ r_{mt}^+\right] \\ + S_{3t}\left[\beta_1 r_{mt} + \beta_2 \left(R_{mt} - E\left(R_{mt}\right)\right)^2 + \beta_3 \left(R_{mt} - E\left(R_{mt}\right)\right)^3\right] + \varepsilon_{pt} \end{aligned} \tag{2.7}$$

Note that the transition probability matrix will describe how likely it is to migrate from one regime, say the CAPM, to another, say, the HCAPM. Since we specify a first-order Markov chain, the only information that matters to predict the regime at time *t+1* is the regime at time *t*.

This regime-dependent risk measure model is quite flexible since asset returns can be modelled with time variation in both asset returns' distribution and investors' risk aversion. Over different periods of time, any one of the models can dominate the other two, or there may be no dominant model. The estimates of the parameters and probabilities of regimes could provide answers to the questions of whether and when value firms are riskier than growth firms; e.g. by comparing β, β^-, and β^+ of value and growth portfolios. If $\beta^- > \beta^+$ in the LCAPM during bear markets, the portfolio is riskier in downside markets. On the other hand, when

$\beta_2 < 0$ $(\beta_2 > 0)$ in the HCAPM, the portfolio is expected to show lower returns (higher returns) and be riskier (less risky) than other portfolios that just follow the CAPM. If $\beta_3 > 0$ $(\beta_3 < 0)$ and market returns are positively skewed, the portfolio is expected to show higher (lower) returns than the symmetric CAPM and the portfolio is less risky (riskier).

2.2.4. Estimation Method

We estimate the regime-switching model (2.7) via a Bayesian Markov Chain Monte Carlo (MCMC) Gibbs-sampling approach. As Kim and Nelson (1999) point out, using the Gibbs sampler to estimate unobserved variables as well as parameters allows us to draw from the relevant distributions simultaneously. Moreover, the model has conditioning features that make it simple to implement the Gibbs sampler. Another reason to use an MCMC method is that it provides posterior distributions from which we draw the parameter estimates and conduct significance tests directly. Finally, by averaging the generated values of the regime dummy variables we obtain estimates of the smoothed probabilities of regime selection over time, which are useful to study the implications of our model.

The Gibbs sampling estimation of model (2.7) consists of two steps. Let $\boldsymbol{\theta} = \left(\alpha, \beta, \beta^-, \beta^+, \beta_1, \beta_2, \beta_3, \sigma_1, \sigma_2 \sigma_3, P\right)$ denote the vector of parameters in the model. In the first step, conditional on $\boldsymbol{\theta}$, we sample from the distribution of $\tilde{\mathbf{S}}_T = \left(S_1 \ldots S_T\right)$ using the multi-move algorithm of Carter and Kohn (1994). Conditional on $\tilde{\mathbf{S}}_T$, the model reduces to a regression model with known structural breaks, so in the second step each parameter in $\boldsymbol{\theta}$ is sampled in turn, conditioned on these structural breaks. We use standard conjugate Gaussian distributions for the regression coefficients and the inverted gamma distribution for the variance, as in Zellner (1971). Given the issue discussed in Section 2.2.1, we take special care to draw β^{-} and β^{+} from the correct distributions (i.e. these parameters are drawn from the distribution which disregards the constant term, so that correct downside and upside betas are estimated). The transition probabilities are estimated using conjugate beta priors. We allow for a large number of burn-in iterations to guarantee convergence. All results are obtained with 10,000 iterations after 30,000 burn-in iterations.

2.3. Empirical Results

This section is organised as follows. The data are described in subsection 2.3.1. In subsection 2.3.2 we analyse the value premium with OLS estimates of the CAPm, LCAPM and HCAPM. In subsection 2.3.3, we report the results of our investigation of the risk of value and growth portfolios according to the regime of the market as inferred from a regime-switching model. In subsection 2.3.4, we consider the results obtained with the regime-switching model with alternative risk measures. Robustness tests are discussed in subsection 2.3.5.

2.3.1. Data

There are many different ways of constructing value/growth portfolios. Book-to-market, earnings and other firm characteristics have been used with different breakpoints in the literature. We focus on several widely used value, growth and value-minus-growth portfolios.

The main results in this paper are reported for two sets of portfolios. The first set of portfolios are Fama and French's (1993) H (high book-to-market or value portfolio), L (low book-to-market or growth portfolio), and HML (value-minus-growth) portfolios, which are constructed using a two-by-three sort on size and book-to-market. HML has been used as a standard measure of the value premium since Fama and French (1993), and has been used in many studies, including Petkova and Zhang (2005) and Fama and French (2006a). The second set of portfolios considers size, since the value premium is supposedly stronger amongst smaller firms (Loughran (1997) and Fama and French (2006a)). From a five-by-five sort on size and book-to-market the small-value (Hs) and small-growth (Ls) portfolios can be calculated, and the small value premium HMLs is the small-value portfolio (Hs) minus the small-growth portfolio (Ls). To check the robustness of our results, we also use portfolios based on a decile sort on earnings-to-price ratio[6]. The excess market return is the CRSP value-weighted portfolio return minus the one-month Treasury bill rate. All the data are obtained from Kenneth French's data library.[7]

[6] Fama and French (2006a) show that the value premium becomes less dependent on size if earnings-to-price ratio is used instead of book-to-market ratio

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

We consider different subsamples within the period from July 1926 to December 2006: the full sample (1926-2006), the post-depression sample (1935-2006), and the post-1963 sample (1963-2006). We focus more on the post-1963 sample because the value premium is more difficult to explain during this period. The value premium in the period from 1926 to 1963 can be explained using the simple CAPM or the conditional CAPM; see Ang and Chen (2007) and Fama and French (2006a). We consider the post-depression sample because the Great Depression was a remarkably unique event which could alter the results significantly. In our robustness checks we also consider the pre-1963 sample.

The value premium can be measured by average returns or unconditional alphas of value-minus-growth portfolios. Table 2-1 reports several descriptive statistics for our portfolios. Consistent with previous studies, i.e. Fama and French (1992, 1993, 2006) and Davis, Fama and French (2000), the value premium exists and is stronger for small stocks. The value premia from HML and HMLs in the full sample (Panel A) are, in terms of average returns (CAPM alphas), 0.42% and 0.51% (0.32% and 0.48%), respectively. When we exclude the Great Depression (Panel B), the value premia from the HML and HMLs portfolios increase to 0.48% and 0.56% (0.51% and 0.66%) in terms of average returns (unconditional alphas). Finally, in the more recent period from 1963 to 2006 (Panel C), the average returns (unconditional alphas) of the HML and HMLs portfolios still remain high, at 0.45% and 0.62% (0.58% and 0.80%), respectively.

Other statistics suggest that all value portfolios have fatter tails (higher kurtosis) than the growth portfolios and the market in all sample periods, whilst growth portfolios tend to have less extreme returns than the market. However, the value portfolios have larger skewness than the growth portfolios and thus the value-minus-growth portfolios have positive skewness.

2.3.2. Preliminary Results with Unconditional Models

In this subsection we investigate the value premium with OLS estimation of the CAPM, LCAPM and HCAPM. Panels A, B and C of Table 2-2 report the results using the full, post-depression and post-1963 periods, respectively. From the estimates of the CAPM, there is very little evidence that value is riskier than growth; indeed, the betas of HML and HMLs are positive only when the sample includes the Great Depression (see Panel A), whilst in the post depression and post-1963 samples betas are negative and mostly significant (Panels B and C).

Accounting for different responses to up- and down-markets reduces alpha significantly. The LCAPM alphas are much smaller than the CAPM alphas in all three samples. For instance, in the full sample the LCAPM explains the value premium ($\hat{\alpha}_{LCAPM} = -0.15\%$, t-stat = -1.39), whilst the CAPM does not ($\hat{\alpha}_{CAPM} = 0.32\%$, t-stat = 2.69). This results has been obtained by Chan (1988), De Bondt and Thaler (1987) and was replicated in Petkova and Zhang (2005). Value stocks seem to benefit from a larger increase in returns in up-markets and a smaller decrease in returns in down-markets (i.e. $\beta^{+} > \beta^{-}$), whilst growth stocks behave in the opposite way ($\beta^{-} > \beta^{+}$). However, value is not riskier than growth in any of the samples; in the full samples the downside betas of HML and HMLs are not significantly different from zero, whilst in the post-depression and post-1963 samples the downside betas are negative and statistically significant. Specifically, in the post-1963 sample, the downside betas of HML and HMLs (-0.36 and -0.50) are roughly double the upside betas (-0.17 and -0.22), so during this period we expect the value-minus-growth strategies to have high and positive returns when the market return is negative, but negative returns when the market return is positive.

Higher moments seem to be significant mostly when the sample includes the Great Depression. In the full sample, the coefficients of HML on the square and cube of the market return are positive and significant. However, they do not represent risk; the positive coefficient on the square market return (b_2) suggests that the return on HML increases further when the market return is positive, but decreases less when market returns are negative. Also, since the market has positive skewness in the full sample (Panel A, Table 2-1), HML increases even further with the positive b_3. The HCAPM seems to reduce the size and significance of the premium; for instance over the full sample, the CAPM alpha is 0.32% (t-stat = 2.69) but the HCAPM alpha is lower at 0.25% and borderline significant (t-stat = 2.00). The higher moments are not significant in the post-depression samples, but are significant for the growth portfolios in the post-1963 sample, increasing their risk relative to the CAPM.

This exploratory analysis indicates that asymmetries play a large role in the relationship between value, growth and value-minus-growth portfolios and the market. These results are not consistent with Zhang's (2005) implication that value firms are in advantage in favourable economic situations (in up-market) since the upside betas of growth firms are higher than those of value firms, and overall do not elicit evidence that value might be riskier than growth.

2.3.3. Is Value Riskier than Growth? Reinvestigation with Market Regimes

In this subsection we reinvestigate if value is riskier than growth in different market states. As pointed out earlier, we first show that market states derived from the estimated expected market risk premium may be too noisy. For example, Petkova and Zhang (2005) obtain market states from estimates of the expected market risk premium obtained with macroeconomic variables. Although they argue this is a better measure of the market return than ex-post returns, it is not clear whether the lagged four macroeconomic variables could provide unbiased and noise-free estimates of the expected market risk premium.

To investigate this, as in Petkova and Zhang (2005), we run Center for Research in Security Prices (CRSP) value-weighted market returns in excess of one month Treasury Bill rate (r_{mt}) on the following four lagged macroeconomic variables: the one month Treasury Bill (TBill), credit spread (CS) (the difference between Moody's AAA and BAA rated corporate bonds), term spread (TS) (the difference between the US 10 year and the 1 year treasury bond rates), and dividend yield (DY) (the CRSP value-weighted dividend yields).[8] For the period between January 1963 and December 2004 (504 monthly observations) we have

$$r_{m,t+1} = \underset{(0.060)}{0.0004} - \underset{(1.432)}{2.211} TBill_t + \underset{(2.053)}{1.461} CS_t + \underset{(0.178)}{0.042} TS_t - \underset{(0.072)}{0.084} DY_t + \hat{\varepsilon}_{m,t+1} \qquad (2.8)$$

[8] These monthly data are from the database of the Federal Reserve Bank of St. Louis and CRSP.

where the numbers in brackets are the Newey-West heteroskedasticity consistent *t* statistics.[9] None of the lagged macroeconomic variables are significant at the 1% level, and only one lagged variable, CS, is significant at 5% level.[10] Moreover, the value of R-square is only 1.85%. In other words, less than 2% of *ex post* market returns reflects *ex ante* market returns if the regression is a proper way to estimate the expected market returns. These results suggest that even though we admit that individual beliefs are not homogeneous (Ross (1978)), the difference between *ex post* and *ex ante* returns appears to be too large to justify using the simple regression to approximate the expected market returns. In addition, by including variables (such as DY, TS, and TBill) that do not appear to explain next month excess market returns, the estimated market risk premium is likely to have large measurement errors. Although the four macroeconomic variables are economically motivated and thus are widely used in the literature (see e.g. the discussion in Ferson, Sarkissian and Simin (2003)) , we could not conclude that the estimated expected market return from the regression is a good proxy for the expected market risk premium. This is supported by Cooper and Gubellini (2008), who show that the results from conditional models (including the one used by Petkova and Zhang (2005)) are extremely sensitive to the conditioning variables used.

[9] Further studies on the properties of these four macroeconomic variables suggest that the augmented Dick-Fuller test fails to reject that TBill and DY have unit root. Thus as shown by Ferson, Sarkissian and Simin (2003), the regression results may be further undermined by the high level of autocorrelation in these independent variables.

[10] We also estimated the regression equation using 5-year sub-sample periods, and obtained similar results.

2.3.3.1. Regime Switching Models for the Market Return

We consider a different approach to identify the market states, which is motivated by the regime-switching literature and the modelling of business cycles. Following Hamilton (1989) and Schwert (1989), Turner, Startz and Nelson (1989) and Hamilton and Susmel (1994) show that there are distinct regimes in the S&P 500 index in terms of mean and volatility. Moreover, Hamilton and Lin (1996) investigate the relationship between stock returns and real output growth in industrial production, and conclude that economic recessions are the primary drivers of fluctuations in market volatility. More recently, Perez-Quiros and Timmermann (2000) use regime-switching models to study the risk of firms with different sizes in expansions and recessions. These studies suggest that regime switching models are an effective tool to identify regimes which are linked to economic conditions.

We model the data-generating process of the market return as a three-state first-order Markov process with switching mean and variance[11]. Our model for the market return can be described in the following way. Let S_t be a regime variable assuming the values of 1, 2 or 3 according to the appropriate regime, and $S_{jt} = I\left(S_t = j\right)$ be dummy variables which assume the value of 1 when the market is in regime $j, j = 1,2,3$, where *I*(.) is the indicator variable. The model is then given by

$$\begin{aligned} r_{mt} &= \mu_1 S_{1t} + \mu_2 S_{2t} + \mu_3 S_{3t} + \sigma_{m,t}\varepsilon_t \\ \sigma^2_{m,t} &= \sigma^2_{m,1} S_{1t} + \sigma^2_{m,2} S_{2t} = \sigma^2_{m,3} S_{3t} \end{aligned}, \qquad (2.9)$$

[11] We also considered a model with two regimes, but found that three regimes better specify the market states than two regimes. Namely, for all three periods a model with three-regimes captures the features of the data (such as heteroskedasticity, skewness and excess kurtosis) well, in the sense that a Jarcque-Bera test of the standardised residuals does not reject normality.

where μ_j and $\sigma_{m,j}$ are the expected excess market return and volatility, respectively, in regime *j*. Following Hamilton (1989), we allow the regime variable S_t to be governed by a first-order Markov chain with a transition probability matrix $P = \{p_{ij}\}$, where $p_{ij} = P(S_t = j \mid S_{t-1} = i)$ is the probability that regime *i* at time *t*-1 is followed by regime *j* at time *t*. We interpret each regime according to estimates of the mean and volatility of the market.

2.3.3.2. Description of Market Regimes

We estimate the regime-switching model for the three sample periods using the Bayesian MCMC Gibbs-sampling estimation, which is standard from the regime-switching literature (see Kim and Nelson (1999) for example). From the MCMC estimates of the dummy variable S_t, we estimate the smoothed probabilities of the market being in each regime in each month. Figure 2-1 to Figure 2-3 plot the smoothed probabilities of each regime for the full, post-Depression and post-1963 samples, respectively. Table 2-3 reports descriptive statistics for the market return, and also the value, growth and value-minus-growth portfolios in each of these regimes. We label each regime according to the characteristics of the market return. We name Regime 3 as "Bear market", since the mean return is negative and the volatility is high. Likewise, Regime 2 is labelled "Bull market", since the mean return is positive and the volatility is low. Finally, Regime 1 is labelled "Transition", when the market is neither bullish nor bearish.

For example, for the full sample, the average monthly market return in the bull and bear markets are 1.26% and -059%, respectively, whilst the volatilities are 2.95% and 11.19% (see the column labelled "Market" in Table 2-3). Figure 2-1 shows that the bear market regime is selected mostly during the Great Depression and during the period from the middle of 1937 through the middle of 1940, which includes the first two years of the Second World War, and then during short periods such as the 1973 oil crisis and the Black Monday month of October 1987. When we exclude the Great Depression period (Figure 2-2), we find that the bear market regime now includes many other periods.

In the post-1963 sample (Figure 2-3), the bear market regime captures periods such as the two oil crises, the Mexican moratorium of 1982, the Black Monday of 1987, the Russian crisis of 1998, the burst of the internet bubble and the period following the terrorist attacks of 2001. The bull regime includes most of the economic expansion of the US economy in the early 1960s, and also most of the early 1990s. The overall picture we obtain with these results is that the return on the aggregate stock market can be modelled as a mixture of longer expansion or bull market periods with high returns and low volatility (the mean market return and volatility in the bull regime are 1.57% and 3.53%), and infrequent and shorter periods of contraction during which the stock market does very poorly and has high volatility (the mean market return and volatility in the bear regime are -1.54% and 6.34%), which agrees with the modelling of regimes and the business cycle using macroeconomic data as in Hamilton (1989).

It is difficult to infer a relationship between the state of the market and the average value-minus-growth return. From Table 2-3, when the whole sample is used, the average HML return seems to increase with the market return: it is highest

(0.53% in terms of average return) when the market is in the bull market regime and lowest (0.19%) when the market is in the bear regime (the difference in the median HML across regimes is statistically significant at 1%). However, in the more recent post-1963 period, we find that, as expected from the estimates of the LCAPM for this period (Panel C, Table 2-2), the average HML return is higher in the bear market regime, at 1.21% a month) and actually negative in the bull market regime (these differences are also statistically significant). Therefore, over the last 40 years, growth stocks do marginally better than value stocks during bull markets, but do much worse than value stock in the bear market regime. This is contrary to what one expects from the theory of Zhang (2005)[12].

2.3.3.3. The Value Premium in Different Market Regimes

We now proceed to investigate the risk of the value, growth and value-minus-growth portfolios in different market states. We estimate the CAPM conditioned on the state of the market (transition, bull or bear). The purpose of this subsection is to investigate the asymmetric reactions of value and growth firms to market conditions suggested by Zhang (2005).

Panels A, B and C of Table 2-4 report the estimation results for the H, L and HML portfolios using the full, post-depression and post-1963 periods, respectively[13]. The risk of the value portfolio increases during bear market only if the sample includes the period of the Great Depression. When we exclude the Great Depression

[12] However, we notice that the post-1963 period has relatively few recessions and also contains the unprecedented boom in the stock market in the late 1990s (which was driven by growth stocks), so we should be cautious regarding this interpretation.

[13] The results with the Hs, Ls and HMLs portfolios are largely similar, so we do not present or comment on them.

(Panel B), value is not riskier than growth in any of the regimes, whilst alpha is positive and statistically significant in the bull and transition regimes. In the post-1963 sample (Panel C), value is significantly less risky than growth in the bull and bear regimes, whilst in the transition regime beta is not statistically different from zero. Also, as expected, the value premium is very strong in the bear market regime (the bear-market alpha is 0.82%, t-statistic 2.56), but not very strong in the bull regime (the bull-market alpha is only 0.31% and not statistically significant).

The period since 1963 is particularly troublesome for a risk-based explanation of the value premium. Not only does value have lower betas than growth, whilst having higher average returns; growth is riskier than - or at least as risky as - value in all market regimes. Other studies such as Ang and Chen (2007) and Petkova and Zhang (2005) provide little or no evidence that value is riskier than growth for this period. Ang and Chen's estimate of time-varying betas for the period are mostly negative and they argue that the (unconditionally) high value premium during the period can be explained by small-sample properties of the unconditional alpha when betas are time-varying. Petkova and Zhang's (2005) evidence for this period goes in opposite directions from a risk-based explanation of the value premium: value betas covary negatively with the market risk premium, whilst growth betas have no significant covariation with the market risk premium. When the results in panels A and C are compared, we conclude that it is only when the Great Depression period is included that these conditional models explain the value premium and that value is riskier than growth. Since the Great Depression was an extreme recession with unique characteristics (as pointed out by Bernanke (1983)), for our main results in the next section we focus on the post-1963 period.

2.3.4. Regime-Switching Risk Measures

The results so far have revealed no conclusive empirical evidence supporting Zhang's (2005) theory or Petkova and Zhang's (2006) result that value is riskier than growth in bad times. In this subsection we investigate whether or not our regime switching model with different risk measures can capture the asymmetric risk pattern of HML over time. We combine the three models (CAPM, LCAPM and HCAPM) in the regime-switching model in Equation (2.7), so each model has the possibility of being selected in each month. If there are periods when the relationship between the portfolios and the market is symmetric, then the CAPM should be selected. Asymmetries can be modelled by two alternative models: either in dichotomous up and down markets (LCAPM) or in a continuous framework (HCAPM).

We estimate the regime-switching model (2.7) for the H, L, HML, Hs, Ls and HMLs portfolios for the post-1963 sample. We focus on this sample for a few reasons. First, as stated before, the value premium is more difficult to explain during this period. Second, there is a structural break in the early 1960s (see for instance Section 3.1 of Petkova and Zhang (2005), and Figure 1 of Fama French (2006)), and thus using a long time series without considering the breaks could be misleading. Our earlier results also confirm that including the Great Depression period could give us wrong inferences about the value premium for the last 40 years. Finally, our regime switching model allows different classes of risks, but not time-varying risks within a regime.

The results are displayed in Table 2-5, in which we report the posterior means and standard deviations of the parameters from the MCMC iterations. The smoothed probabilities in Figure 2-5 can be interpreted as the probability that, at each month, the CAPM, LCAPM and HCAPM are selected.

It is important to note that the regime switching risk measures explain the value premium in terms of alpha. The posterior distributions of the alphas of HML and HMLs suggest that HML and HMLs can be explained by the model at the 1% significance level. In the next subsection, we show that the value premium is explained by the higher upside betas of the value-minus-growth portfolios, relative to their downside betas, and the positive coefficient on the squared and cubed market returns. This result suggests that it is not increased downside risk during bearish markets which drives the value premium. In the following subsections, we examine the risk of the value portfolio and the value premium in more detail.

2.3.4.1. Is Value Riskier than Growth?

From the estimates in Table 2-5, there is no evidence that value is riskier than growth in any of the regimes. First, the CAPM beta is negative (and significant) for both portfolios: for HML (HMLs) the average posterior beta is -0.16 (-0.28), with a standard deviation of 0.07 (0.05). Second, in the LCAPM regime, the downside betas of both portfolios are also negative and significant. The downside beta of HML (HMLs) is -0.69 (-0.87), with a standard deviation of 0.16 (0.14). Finally, in the HCAPM regime, beta is not significantly different from zero for either of the value-minus-growth portfolios, but the coefficients on the square and cube of the market return are positive and statistically significant[14]. A positive coefficient on the second

[14] In the sense that their estimated posterior distributions do not include the value zero.

moment of excess market return makes HML concave to market movements, increasing returns whilst decreasing risk.

These estimates also suggest that the average HML (HMLs) return might be higher in the LCAPM and HCAPM regimes. In the LCAPM regime, this is expected because even though both the downside and the upside betas are negative, the downside beta is larger than the upside beta, so the return on HML will increase more when the market return is negative than it will decrease when the market return is positive[15]. In the HCAPM regime, in addition to the increase due to the positive coefficient on the squared market return, the positive coefficient on the third moment of excess market returns increases the returns of HML even further because the market has positive skewness in this regime (not reported).

Table 2-6 displays the average HML and HMLs returns in each of the three regimes. The t-statistics show whether the average return within a regime is significantly different from that of the whole sample. The average value-minus-growth returns in the CAPM regime are significantly lower than those of the whole period: they are only 0.01% and -0.10% for HML and HMLs respectively. As expected, in the LCAPM and HCAPM regimes the average returns are much higher and in some cases significant. We examine whether these differences are significant with a non-parametric median test for robustness, since we do not know the distribution of market returns within a regime. The hypothesis that the median return is the same across regimes is rejected at the 1% significance level for both HML and HMLs. On the other hand, the difference in market returns across regimes is not significant.

[15] The probability that the upside beta is higher than the downside beta is around 90% (99%) for the HML (HMLs) portfolio. This probability is calculated by counting the number of cases in the MCMC iterations that upside beta is higher than downside beta and is more robust compared to methods that assume a specific probability density function.

Again, these results are at odds with a risk-based explanation of the value premium. In the first place, they show no sign that value is riskier than growth. Secondly, even though we found that the selection of the different risk measures is linked to differences in the average returns of the value-minus-growth portfolios, the regimes do not seem to be linked to the condition of the market in terms of average market returns. According to Zhang (2005), value becomes riskier than growth in bad times and thus the value premium is expected to be linked to market conditions. Chen, Petkova and Zhang (2006), for instance, argue that the value premium peaks in recessions in the post-1963 sample. However, from the 77 months of recession in our sample (according to the NBER), 40 months occur in the LCAPM regime, and thus our results suggest that the value premium is higher during recessions not because value becomes riskier than growth, but because of the difference in upside and downside betas between value and growth stocks. Indeed, the average excess value (H) and growth (L) returns in the LCAPM regime are 0.66% and -0.45% respectively, which shows that a large part of the positive HML return comes from the negative returns of growth stocks. On the other hand, in the HCAPM regime the average HML return increases due to the higher moments associated with the value portfolio, and the average excess returns on the H and L portfolios during this regime are 1.38% and 0.36%. Therefore in this regime the high value-minus-growth returns come from the outperformance of value relative to growth, especially when the market return is high.

2.3.4.2. Selection of Regimes

In this subsection we study the selection of each regime through time. The transition probabilities tell us how likely it is to remain in each model (i.e. the CAPM, LCAPM or HCAPM) or to move to another one. Also, we analyse the smoothed probabilities that each regime has been selected at each month.

The regimes are persistent for all portfolios except the L portfolio. Figure 2-5 shows that for the L portfolio, the only persistent regime is the LCAPM regime. Since the higher moment coefficients are not significant, and the betas in the CAPM and HCAPM regimes are very close, the L portfolio could be well described by a model with two regimes (CAPM and LCAPM). So there are periods when growth stocks behave similarly in up and down markets, and other periods when downside risk is increased. For the H portfolio, all regimes are persistent, as indicated by the transition probabilities.

The HML and HMLs portfolios tend to exhibit similar (but weaker) patterns to the H and Hs portfolios, respectively (see Figure 2-5). It should be noticed that, in the case of portfolio Hs, the HCAPM regime is quite persistent, even though the higher moments are not significant. The estimate of beta in this regime, though, is much smaller than in the CAPM regime, so we can attribute the persistence to the difference in beta.

The smoothed probabilities of the regimes for HML and HMLs seem to be quite similar, which indicates that both strategies behave similarly regarding the selection of the regimes. We know that the average value-minus-growth return is close to zero (for HML) or negative (for HMLs) when the CAPM regime is selected, that is, when the relationship between the return on the value-minus-growth strategies and the return on the market is symmetrical. This regime is selected over short one or two-year long periods, with the exception of two longer periods, one in the late 1970s and another from 1985 to 1992. Overall, this is the most prevalent regime: it is selected in 241 (249) months when the model is estimated with the HML (HMLs) portfolio, which corresponds to almost half of the whole sample. Therefore, the value premium is close to zero during half of the post-1963 sample.

The LCAPM regime is selected during some turbulent periods, such as the oil crisis of 1973 and a four-year period following the burst of the internet bubble in 2000. The average excess return on the H and L portfolios during this regime are 0.66% and -0.45%, respectively, and thus much of the HML return during this period comes from the negative returns of growth firms.

The HCAPM is the least persistent regime for both HML and HMLs; the probability of remaining in this regime (p_{33}) is 0.84 and 0.82 for the HML and HMLs portfolios respectively. This regime is selected in 131 (120) months for the HML (HMLs) portfolio, which corresponds to around 25% (22%) of the whole sample. It is selected over short periods, usually less than a year, except for the period from the middle of 2003 until early 2005.

These results are quite interesting because although the value premium is high in both the LCAPM and HCAPM regimes, the reasons are quite different. In the former regime, the value premium comes from the negative returns of growth firms, especially when the market is doing poorly, whilst in the latter it stems from the outperformance of value relative to growth when the market is doing well.

2.3.5. Robustness Checks

So far we have reported results obtained using two sets of portfolios, both of which are based on sorting procedures using market equity and the book-to-market ratio. We have also focused more on the more recent sample period from 1963 onwards. We address the first issue by replicating our results using value, growth and value-minus-growth portfolios based on a decile sort on Earnings/Price ratio. We define the highest decile to be the value portfolio (H_EP) and the lowest decile to be the growth portfolio (L_EP), and the value-minus-growth portfolio (HML_EP) is the H_EP portfolio minus the L_EP portfolio. To address the second issue, we replicate our results for the pre-1963 sample period, using the H, L and HML portfolios.

2.3.5.1. Portfolios Formed on Earning/Price Ratio

Panel A of Table 2-7 reports descriptive statistics for the value, growth and value-minus growth portfolios based on a decile sort on Earnings/Price. The value premium defined by Earnings/Price ratio is higher than Fama and French's HML; the average HML_EP return over the 1963-2006 period is 0.60%, compared to 0.45% of HML. The CAPM alpha of HML_EP is also higher at 0.69%, compared to 0.58% of HML. The results obtained with the CAPM, LCAPM and HCAPM are similar to the ones obtained with the size and book-to-market-sorted portfolios, so we do not report them. Next, we estimate the CAPM in the different market regimes. The results are reported on Table 2-8. Similarly to our previous results in Table 2-4, value (H_EP) is not riskier than growth (L_EP) in any of the market regimes, and alpha is large in the bear and transition regimes (although it is not statistically significant in the bear regime).

Finally, we estimate the regime-switching model with different risk measures for the H_EP, L_EP and HML_EP portfolios. The results are reported in Panel A of Table 2-9 and are quite similar to those obtained before. In Panel A of Table 2-10, we repeat our analysis of the average value premium and market return in each regime, and the results are also similar to those obtained with the HML portfolio: the value premium is highest in the LCAPM regime, at 1.29%, and nearly zero in the CAPM regime at 0.08%. It is also very high in the HCAPM regime, at 0.93%. As with the results for the HML portfolio, the difference in the median HML_EP return across regimes is statistically significant.

2.3.5.2. The Pre-1963 Period

In this subsection we repeat our analyses for the pre-1963 sample. This period is quite interesting because it contains the Great Depression, but also the remarkable post-war expansion of the US economy. When we estimate the regime-switching model for the market return over this period, we find that one of the regimes captures almost exclusively the Great Depression (see Figure 2-4), whilst the other two regimes have characteristics of bull and bear markets. Therefore we name the regimes 'Great Depression', 'Bull' and 'Bear'. The average market return in the

Great Depression, Bull and Bear market regimes are -0.42%, 1.95% and -1.13% respectively, and the market volatilities are 12.99%, 2.93% and 5.85% (not reported).

The descriptive statistics over the pre-1963 period are reported in Panel B of Table 2-7, which shows that there is no value premium over the period: the average return of the HML strategy is 0.37%, and the CAPM alpha is only 0.06%, and not statistically significant. The estimates of the CAPM in the different market regimes are reported in Panel B of Table 2-8. The table shows that value is significantly riskier than growth in all three regimes, and the CAPM explains the returns on HML in all three regimes. The risk of HML is highest in the Great Depression (beta is 0.41 and statistically significant), whilst it is similar in the bull and bear regimes.

The estimate of the regime-switching model with alternative risk measures (Panel B, Table 2-9) shows that value is riskier than growth in the CAPM and LCAPM regimes (the beta of HML in the CAPM regime is 0.20, and the downside beta in the LCAPM regime is 0.45, both statistically significant). Also, the upside beta of HML is positive and larger than its downside beta, suggesting that value firms are favoured by increases in the market return, relative to growth firms. In the HCAPM regime, none of the coefficients is statistically significant. Unlike our previous results, the average HML premium in the pre-1963 period does not seem to be related to these regimes (see Panel B of Table 2-10).

Summarising our robustness checks, the results using the alternative portfolios formed on Earnings/Price largely confirm our evidence for the post-1963 sample. The results with the pre-1963 sample are supportive of the risk base explanation of the value premium: during this period, value is riskier than growth, the value premium is explained by the CAPM and the risk of HML is highest in bad states of the market. We conclude that the evidence supporting the theory, e.g. Petkova and Zhang (2005), is mainly driven by the inclusion of the pre-1963 period in the sample.

2.4. Conclusion

This work contributes to the debate about the value premium in two ways. First, we show that the empirical conclusion of Petkova and Zhang (2005) that value is riskier than growth in bad times is driven by the pre-1963 period and the method they use to estimate the market risk premium. When we use a different method (a regime-switching model) to identify the market state, there is little or no evidence that value is riskier than growth, and thus the value premium is not explained by higher risk of value firms in bad times.

Second, we propose a regime-switching model which allows three different risk measures to be selected over time, relaxing the CAPM statement which is derived under restrictive assumptions such as normality of returns and quadratic utility. When both the probability density function of asset returns and investors' risk aversion change over time, the simple mean-variance analysis could misspecify asset returns. Periods when returns are normally distributed can be modelled by the CAPM, but periods when returns become heavy-tailed and/or skewed or when downside and upside betas differ may be explained by higher moments in the HCAPM or dichotomous downside/upside betas in the LCAPM. We find that the value premium in the post-1963 sample can be empirically explained with this generalised model, but value is not riskier than growth in any of the regimes.

Overall, our results are not consistent with the risk-based explanation of the value premium proposed by Zhang (2005). We find that value stocks' returns increase more in up-markets than they decrease in down-markets, which is reflected by the higher upside beta of value relative to its downside beta (or the positive coefficients on the square and cube of the market return), whilst growth stocks behave in the opposite way (their returns decrease more in down markets than they increase in up markets). However, overall the upside betas of growth portfolio are still larger than those of the value portfolio, which is the opposite of what is expected from Zhang's theory.

These results make a risk-based explanation of the premium less likely and also allow us to view the anomaly from a different perspective. As in Zhang (2005), during bull periods investors would gladly pay more to hold value firms but would not be willing to pay much for growth firms. Unless these asymmetric risk patterns can be linked to some kind of fundamental risk which is not captured by beta, our results point to behavioural explanations of the value premium, such as the one proposed by Lakonishok, Shleifer and Vishny (1994). In this case, the question is why the anomaly persists, many years after having been documented.

Table 2-1 Descriptive statistics of value, growth and value-minus-growth portfolios

This table reports descriptive statistics for several value, growth and value-minus-growth portfolios. H and L are the value and growth portfolios obtained with a 2 by 3 sort on size and book-to-market ratio. Hs and Ls are the small-value and small-growth portfolios obtained with a five-by-five sort on size and book-to-market, and HML (HMLs) is a strategy which is long the H (Hs) portfolio and short the L (Ls) portfolio. The t-statistic are adjusted using the Newey and West (1987) method with 3 lags.

Panel A. January 1927 - December 2006

	H	L	HML	Hs	Ls	HMLs	Market
Monthly return (%)	1.084	0.668	0.415	1.231	0.722	0.510	0.650
Monthly volatility (%)	7.652	6.386	3.592	8.356	7.854	3.668	5.437
Skewness	2.052	0.459	1.896	2.292	1.051	0.970	0.220
Kurtosis	23.919	10.317	18.928	25.528	13.558	12.096	10.911
Alpha (CAPM) (%)	0.258	-0.065	0.323	0.364	-0.114	0.478	-
t-stat	2.449	-1.128	2.690	2.888	-0.996	3.813	-

Panel B. January 1935 - December 2006

	H	L	HML	Hs	Ls	HMLs	Market
Monthly return (%)	1.154	0.675	0.479	1.284	0.721	0.563	0.696
Monthly volatility (%)	5.742	5.511	2.939	6.385	6.794	3.174	4.527
Skewness	0.185	-0.413	0.756	0.302	-0.239	0.334	-0.536
Kurtosis	11.021	5.985	8.968	11.438	5.982	6.874	6.292
Alpha (CAPM) (%)	0.371	-0.138	0.510	0.454	-0.205	0.659	-
t-stat	3.869	-2.530	4.564	3.839	-1.867	5.470	-

Panel C. January 1963 - December 2006

	H	L	HML	Hs	Ls	HMLs	Market
Monthly return (%)	0.894	0.441	0.453	1.075	0.456	0.619	0.477
Monthly volatility (%)	4.633	5.558	2.901	5.329	6.961	3.325	4.380
Skewness	-0.318	-0.458	0.011	-0.331	-0.328	-0.135	-0.510
Kurtosis	6.859	4.706	5.552	7.291	4.842	5.832	5.135
Alpha (CAPM) (%)	0.448	-0.135	0.583	0.593	-0.202	0.795	-
t-stat	4.072	-1.754	4.415	4.030	-1.312	5.358	-

Table 2-2 OLS regressions of CAPM, LCAPM (Lower Partial Moment CAPM) and HCAPM (Higher-moment CAPM)

The table reports OLS estimation of the following three models:

$$r_{pt} = \alpha_{CAPM} + \beta r_{mt} + \varepsilon_{pt}$$
$$r_{pt} = \alpha_{LCAPM} + \beta^{-} r_{mt}^{-} + \beta^{+} r_{mt}^{+} + \varepsilon_{pt} \quad ,$$
$$r_{pt} = \alpha_{HCAPM} + \beta_1 r_{mt} + \beta_2 \left(R_{mt} - E\left(R_{mt}\right)\right)^2 + \beta_3 \left(R_{mt} - E\left(R_{mt}\right)\right)^3$$

where r_{pt} denotes the excess return on either a value, growth or value-minus-growth portfolio, r_{mt} denotes the excess market return, $r_{mt}^{-} = r_{mt} I(r_{mt} < 0)$ and $r_{mt}^{+} = r_{mt} I(r_{mt} > 0)$ are the negative and positive components of the market return and R_{mt} is the raw market return. The LCAPM regression is estimated without the constant to obtain the correct upside and downside betas, as discussed in Post and Van Vliett (2005), and the constant is calculated as the mean of the residuals, i.e. $\hat{\alpha}_{LCPAM} = r_{pt} - \hat{\beta}^{+} r_{mt}^{+} - \hat{\beta}^{-} r_{mt}^{-}$. The t-stats are computed using the Newey and West (1987) method to correct for heteroskedasticity and autocorrelation with 3 lags. H and L are the value and growth portfolios obtained with a 2 by 3 sort on size and book-to-market, and HML is the H portfolio subtracted from the L portfolio. Hs and Ls are the small-value and small-growth portfolios obtained with a five-by-five sort on size and book-to-market, and HMLs is the Hs portfolio subtracted from the Ls portfolio. The coefficients b_2 and b_3 are multiplied by 100 for visualisation purposes.

Panel A. January 1927 - December 2006

		H		L		HML		Hs		Ls		HMLs	
Model		Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
CAPM	*a*	0.258	2.449	-0.065	-1.128	0.323	2.690	0.364	2.888	-0.114	-0.996	0.478	3.813
	b	1.270	17.700	1.128	54.660	0.142	1.904	1.333	16.617	1.284	29.679	0.048	0.696
LCAPM	*a*	-0.224	-2.158	-0.071	-1.234	-0.154	-1.385	-0.181	-1.372	-0.173	-1.495	-0.008	-0.071
	b^{-}	1.133	18.642	1.125	39.203	0.008	0.112	1.180	17.633	1.264	23.096	-0.084	-1.292
	b^{+}	1.385	9.472	1.128	30.952	0.257	1.758	1.464	8.743	1.296	15.414	0.168	1.221
HCAPM	*a*	0.102	0.942	-0.144	-2.450	0.246	2.000	0.188	1.391	-0.279	-2.360	0.467	3.522
	b_1	1.161	20.820	1.161	43.669	0.000	-0.003	1.220	17.745	1.318	23.320	-0.098	-2.120
	b_2	0.738	3.412	0.206	1.695	0.532	2.444	0.815	2.918	0.500	2.006	0.315	1.330
	b_3	0.032	2.459	-0.011	-2.268	0.043	3.925	0.033	1.983	-0.012	-1.230	0.045	4.259

Table 2.2 (Continued)

Panel B. January 1935 - December 2006

Model		H		L		HML		Hs		Ls		HMLs	
		Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
CAPM	a	0.371	3.869	-0.138	-2.530	0.510	4.564	0.454	3.839	-0.205	-1.867	0.659	5.470
	b	1.124	25.030	1.168	86.487	-0.045	-1.028	1.192	23.242	1.330	44.230	-0.138	-3.361
LCAPM	a	0.119	1.311	-0.040	-0.753	0.159	1.587	0.213	1.831	-0.010	-0.093	0.223	2.100
	b^-	1.057	16.894	1.195	59.159	-0.138	-1.978	1.133	16.837	1.385	32.989	-0.253	-3.574
	b^+	1.199	13.138	1.139	41.556	0.060	0.751	1.267	11.350	1.274	21.131	-0.007	-0.092
HCAPM	a	0.343	2.629	-0.140	-2.363	0.483	3.600	0.481	2.940	-0.170	-1.373	0.651	4.635
	b_1	1.077	26.593	1.167	71.609	-0.090	-2.136	1.142	23.784	1.325	38.693	-0.184	-4.281
	b_2	0.412	0.683	0.017	0.129	0.395	0.736	0.144	0.196	-0.153	-0.456	0.297	0.568
	b_3	0.045	1.826	0.001	0.204	0.043	1.864	0.042	1.362	0.001	0.060	0.041	1.664

Panel C. January 1963 - December 2006

Model		H		L		HML		Hs		Ls		HMLs	
		Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
CAPM	a	0.448	4.072	-0.135	-1.754	0.583	4.415	0.593	4.030	-0.202	-1.312	0.795	5.358
	b	0.935	23.620	1.208	70.235	-0.273	-6.812	1.011	20.936	1.381	34.660	-0.369	-8.937
LCAPM	a	0.235	2.456	-0.020	-0.272	0.255	2.187	0.381	2.920	0.061	0.405	0.320	2.497
	b^-	0.881	13.205	1.240	47.903	-0.360	-5.082	0.962	11.780	1.458	25.213	-0.496	-6.425
	b^+	1.005	13.442	1.173	37.597	-0.168	-2.232	1.083	11.094	1.302	18.845	-0.219	-2.803
HCAPM	a	0.507	3.833	-0.062	-0.717	0.569	3.772	0.757	4.181	0.016	0.091	0.741	4.672
	b_1	0.919	23.645	1.220	58.244	-0.300	-7.074	0.979	21.614	1.404	32.300	-0.425	-9.792
	b_2	-0.249	-0.335	-0.462	-2.351	0.214	0.292	-0.750	-0.826	-1.319	-2.698	0.570	0.903
	b_3	0.011	0.265	-0.022	-2.201	0.033	0.837	0.017	0.353	-0.053	-2.062	0.070	2.196

Table 2-3 Descriptive statistics of value, growth and value-minus-growth portfolios in different market regimes

We model the market return with a first-order regime-switching process for the mean and volatility as in equation (2.9). We report descriptive statistics for the return on the market portfolios, as well as the value, growth and value-minus-growth portfolios as defined by the H, L and HML portfolios from a 2 by 3 sort on size and book-to-market in each of the three regimes. We also conduct a Kruskal-Wallis test of equality of median HML and market returns in each regime.

Panel A. January 1927 - December 2006	Regime 1: Transition				Regime 2: Bull Market				Regime 3: Bear Market			
	H	L	HML	Market	H	L	HML	Market	H	L	HML	Market
Monthly return (%)	0.812	0.492	0.320	0.172	2.087	1.554	0.532	1.258	0.046	-0.142	0.188	-0.591
Monthly volatility (%)	5.985	6.358	3.461	5.135	3.876	3.474	2.273	2.952	17.482	12.693	6.995	11.187
Skewness	-0.266	-0.177	0.083	-0.170	0.269	-0.042	0.654	-0.116	1.627	0.828	2.085	0.627
Kurtosis	4.246	2.700	4.971	2.666	3.710	3.140	5.010	2.740	7.395	5.022	10.004	4.806

Test: equality of median value premium across regimes: 0.001
Test: equality of median market return across regimes: <0.0001

Panel B. January 1935 - December 2006	Regime 1: Transition				Regime 2: Bull Market				Regime 3: Bear Market			
	H	L	HML	Market	H	L	HML	Market	H	L	HML	Market
Monthly return (%)	1.873	1.345	0.528	1.153	2.245	1.841	0.404	1.298	-0.532	-1.051	0.519	-1.131
Monthly volatility (%)	3.424	3.093	2.101	2.625	5.149	4.844	2.838	4.041	8.689	8.517	4.141	6.887
Skewness	0.104	-0.418	0.964	-0.336	0.232	0.046	0.320	-0.017	0.589	0.075	0.813	-0.035
Kurtosis	3.351	2.968	6.992	2.810	3.773	2.574	4.390	2.429	8.589	3.803	7.728	4.196

Test: equality of median value premium across regimes: 0.683
Test: equality of median market return across regimes: <0.0001

Panel C. January 1963 - December 2006	Regime 1: Transition				Regime 2: Bull Market				Regime 3: Bear Market			
	H	L	HML	Market	H	L	HML	Market	H	L	HML	Market
Monthly return (%)	1.608	0.945	0.663	0.761	2.356	2.560	-0.205	1.578	-0.435	-1.644	1.209	-1.538
Monthly volatility (%)	2.879	2.966	1.935	2.268	3.679	4.664	2.673	3.530	6.638	7.740	3.833	6.343
Skewness	-0.058	-0.179	0.274	-0.248	0.154	0.098	-0.671	0.033	0.186	0.043	0.066	0.091
Kurtosis	2.802	2.741	3.256	2.639	3.081	2.717	5.647	2.452	5.110	3.202	3.843	3.359

Test: equality of median value premium across regimes: <0.0001
Test: equality of median market return across regimes: <0.0001

Table 2-4 OLS regressions of CAPM value, growth and value-minus-growth portfolios in different market regimes

We estimate the regime-switching model for the market return (2.9) and then estimate the CAPM for the value (H), growth (L) and value-minus-growth (HML) portfolios conditioned on the market regime. The value, growth and value-minus-growth portfolios are created from a 2 by 3 sort on size and book to market.

Panel A. January 1927 - December 2006

	H					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.219	1.019	0.389	1.136	0.762	1.465
t-stat	1.197	20.395	3.997	27.905	1.263	14.274

	L					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	-0.129	1.185	-0.093	1.096	0.363	1.108
t-stat	-1.292	57.047	-1.337	44.115	1.431	30.804

	HML					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.348	-0.167	0.482	0.04	0.399	0.356
t-stat	1.627	-3.059	4.019	0.848	0.712	3.439

Panel B. January 1935 - December 2006

	H					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.335	1.12	0.456	1.105	0.354	1.134
t-stat	2.925	26.292	3.211	22.427	1.286	14.229

	L					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	-0.16	1.09	0.009	1.138	-0.084	1.206
t-stat	-1.879	36.471	0.101	49.727	-0.635	59.56

	HML					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.494	0.029	0.447	-0.033	0.438	-0.072
t-stat	3.581	0.542	2.708	-0.599	1.535	-0.976

Panel C. January 1963 - December 2006

	H					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.47	1.079	0.415	0.895	0.441	0.935
t-stat	2.926	22.592	2.932	19.009	1.566	15.292

	L					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	-0.27	1.18	0.104	1.221	-0.383	1.185
t-stat	-2.117	25.602	0.765	35.542	-2.316	47.757

	HML					
	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.74	-0.101	0.31	-0.326	0.824	-0.251
t-stat	3.988	-1.535	1.517	-5.914	2.557	-4.129

Table 2-5 Estimation results for the regime-switching model with alternative risk measures

We estimate the regime-switching model with alternative risk measures (2.7) for several value, growth and value-minus-growth portfolios over the period from July 1963 to December 2006. The regimes are: regime 1 = CAPM, regime 2 = LCAPM, regime 3 = HCAPM), and regime changes are governed by transition probabilities $p_{ij} = P(S_{t+1} = j \mid S_t = i)$, where S_t is a latent variable which identifies the regime at month *t*. H, L and HML are the value, growth and value-minus-growth portfolios obtained with a 2 by 3 sort on size and book-to-market. Hs, Ls and HMLs are the small-value, small-growth and small value-minus-growth portfolios obtained with a five-by-five sort on size and book-to-market. The estimation is done via a Bayesian Gibbs sampling approach and the values reported are posterior means and standard deviations of the parameters from 10000 iterations of the Gibbs sampler obtained after 30000 burn-in iterations. The coefficients b_2 and b_3 are multiplied by 100 for visualisation purposes.

		Panel A. H, L and HML portfolios						**Panel B. Hs, Ls and HMLs portfolios**					
		H		**L**		**HML**		**Hs**		**Ls**		**HMLs**	
		Mean	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**
	a	0.213	0.087	-0.035	0.074	0.253	0.115	0.338	0.118	0.028	0.140	0.308	0.121
CAPM	b	0.897	0.038	1.187	0.071	-0.161	0.072	1.369	0.049	1.470	0.106	-0.281	0.046
LCAPM	b^-	0.642	0.106	1.318	0.080	-0.691	0.155	0.677	0.143	1.277	0.115	-0.873	0.135
	b^+	0.747	0.091	1.178	0.079	-0.475	0.110	1.110	0.130	0.959	0.088	-0.489	0.124
HCAPM	b_1	1.230	0.049	1.196	0.094	-0.098	0.103	0.818	0.053	1.631	0.089	-0.132	0.106
	b_2	1.940	0.391	-0.429	0.503	3.374	0.850	-0.474	0.655	-0.820	1.067	3.491	0.147
	b_3	0.109	0.041	-0.016	0.039	0.166	0.047	0.022	0.053	-0.124	0.053	0.770	0.045
	p_{11}	0.966	0.023	0.720	0.181	0.885	0.078	0.915	0.044	0.903	0.059	0.851	0.076
	p_{12}	0.021	0.018	0.073	0.068	0.045	0.036	0.070	0.041	0.043	0.033	0.036	0.026
	p_{13}	0.014	0.014	0.208	0.175	0.069	0.071	0.016	0.015	0.054	0.052	0.113	0.073
	p_{21}	0.034	0.026	0.082	0.070	0.054	0.039	0.084	0.048	0.040	0.030	0.041	0.029
	p_{22}	0.854	0.053	0.811	0.123	0.865	0.067	0.870	0.056	0.851	0.100	0.919	0.037
	p_{23}	0.112	0.052	0.106	0.112	0.081	0.059	0.046	0.036	0.109	0.093	0.041	0.032
	p_{31}	0.015	07.016	0.212	0.175	0.090	0.077	0.017	0.016	0.042	0.050	0.141	0.078
	p_{32}	0.091	0.047	0.099	0.111	0.074	0.061	0.036	0.029	0.106	0.094	0.038	0.032
	p_{33}	0.894	0.051	0.690	0.188	0.836	0.094	0.947	0.034	0.852	0.104	0.821	0.085
	σ_1	1.639	0.242	1.376	0.425	3.581	1.018	3.716	0.817	26.510	4.956	3.151	0.682
	σ_2	8.463	1.530	6.184	1.085	11.721	1.933	19.166	3.131	4.761	1.072	16.130	2.843
	σ_3	2.211	0.580	1.567	0.492	3.861	1.291	3.074	0.642	5.382	1.042	4.929	1.441

Table 2-6 Average value-minus-growth return and market return per regime

This table reports the average HML and average excess market returns per regime. The t-statistics for average HML (excess market return) compare the HML returns (excess market return) in each regime with the average HML (excess market return) over the whole sample. We also conduct a Kruskal-Wallis test of equality of median HML and market returns in each regime. We collect returns for each regime when the probability of each regime is the highest one at time *t*. Panel A reports results using the regimes inferred by applying the regime-switching model to the HML portfolio, and Panel B reports the same test using the regimes inferred from applying the regime-switching model to the HMLs portfolio. HML is the H portfolio subtracted from the L portfolio, where H and L are the value and growth portfolios obtained with a 2 by 3 sort on size and book-to-market. HMLs is the Hs portfolio subtracted from the Ls portfolio, where Hs and Ls are the small-value and small-growth portfolios obtained with a five-by-five sort on size and book-to-market. The sample period is from 1963 to 2004.

Regime	Number of months	Average HML (HMLs) (%)	t-stat	p-value for H_0: median HML (HMLs) is equal in all regimes	Average excess market return (%)	t-stat	p-value for H_0: median market return is equal in all regimes
Panel A. Regimes inferred using HML portfolio, sample: 1963-2004							
CAPM	241	0.011	-3.820		0.562	0.298	
LCAPM	150	0.667	0.587	<0.0001	0.252	-0.578	0.569
HCAPM	131	1.021	3.180		0.577	0.305	
Whole sample	522	0.453			0.477		
Panel B. Regimes inferred using HMLs portfolio, sample: 1963-2004							
CAPM	249	-0.102	-4.140		0.668	0.658	
LCAPM	153	1.110	1.983	<0.0001	-0.136	-1.658	0.116
HCAPM	120	0.768	1.575		0.861	1.191	
Whole sample	522	0.619			0.477		

Table 2-7 Descriptive statistics of additional value, growth and value-minus-growth portfolios

We report descriptive statistics for value, growth and value-minus growth portfolios based on a decile sort on Earning-to-Price ratio over the period July 1963- December 2006, and for the value, growth and value-minus-growth portfolios obtained with a 2 by 3 sort on size and book-to-market over the January 1927-June 1963 period. We report unconditional OLS alphas (and the appropriate t-statistic adjusted using the Newey and West (1987) method) average monthly return, monthly volatility (both in %), skewness and kurtosis for each portfolio.

7

	H_EP	L_EP	HML_EP	Market
Monthly return	0.959	0.359	0.600	0.477
Monthly volatility	5.263	5.740	4.591	4.380
Skewness	-0.134	-0.200	0.402	-0.510
Kurtosis	6.052	4.466	5.552	5.135
Alpha (CAPM)	0.481	-0.212	0.693	
t-stat	3.313	-1.969	3.111	

Panel B. Portfolios formed on Size and B/M - January 1927 - June 1962

	H	L	HML	Market
Monthly return	1.306	0.936	0.371	0.854
Monthly volatility	10.110	7.237	4.267	6.462
Skewness	1.923	0.885	2.478	0.421
Kurtosis	16.589	11.544	19.444	10.570
Alpha (CAPM)	0.067	0.009	0.058	
t-stat	0.378	0.114	0.331	

Table 2-8 OLS regressions of CAPM in different market regimes for additional portfolios

We estimate the regime-switching model for the market return (2.9) and then estimate the CAPM for value, growth and value-minus-growth portfolios conditioned on the market regime. Panel A reports results for value, growth and value-minus-growth portfolios formed on a decile sort on Earnings/Price over the period from 1963 to 2006. Panel B reports results for value, growth and value-minus-growth portfolios created from a 2 by 3 sort on size and book-to-market over the 1927-1963 period.

Panel A. Portfolios formed on Earning/Price, July 1963 - December 2006

H_EP

	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.511	1.136	0.285	1.022	0.551	0.984
t-stat	2.599	17.194	1.439	17.117	1.456	13.227

L_EP

	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	-0.233	1.18	-0.252	1.27	-0.34	1.16
t-stat	-1.493	18.927	-1.325	25.016	-1.374	23.988

HML_EP

	Regime 1: Transition		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.745	-0.044	0.537	-0.249	0.891	-0.176
t-stat	2.424	-0.396	1.549	-2.534	1.621	-1.552

Panel B. Portfolios formed on Size and B/M - January 1927 - June 1962

H

	Regime 1: Great Depression		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	1.116	1.501	0.143	1.294	-0.062	1.427
t-stat	1.323	13.996	0.952	20.999	-0.143	10.329

L

	Regime 1: Great Depression		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.484	1.087	0.058	1.002	0.049	1.149
t-stat	1.23	25.52	0.782	44.379	0.234	25.002

HML

	Regime 1: Great Depression		Regime 2: Bull Market		Regime 3: Bear Market	
	a	*b*	*a*	*b*	*a*	*b*
coefficient	0.633	0.414	0.085	0.292	-0.11	0.278
t-stat	0.798	3.835	0.507	5.008	-0.277	2.521

Table 2-9 Estimation results for the regime-switching model with alternative risk measures for additional portfolios

We estimate the regime-switching model with alternative risk measures (2.7) for several value, growth and value-minus-growth portfolios over the period from July 1963 to December 2006. The regimes are: regime 1 = CAPM, regime 2 = LCAPM, regime 3 = HCAPM), and regime changes are governed by transition probabilities $p_{ij} = P(S_{t+1} = j \mid S_t = i)$, where S_t is a latent variable which identifies the regime at month *t*. Panel A reports results for value, growth and value-minus-growth portfolios based on a decile sort on Earning-to-Price ratio over the July 1963-December 2006 period, and Panel B for value, growth and value-minus-growth portfolios obtained with a 2 by 3 sort on size and book-to-market over the January 1927-June 1963 period. The estimation is done via a Bayesian Gibbs sampling approach and the values reported are posterior means and standard deviations of the parameters from 10000 iterations of the Gibbs sampler obtained after 30000 burn-in iterations. The coefficients b_2 and b_3 are multiplied by 100 for visualisation purposes.

		Panel A. Portfolios formed on Earnings/Price, July 1963 - December 2006						Panel B. Portfolios formed on Size and B/M - January 1927 - June 1962					
		H_EP		**L_EP**		**HML_EP**		**H**		**L**		**HML**	
		Mean	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**
	a	0.202	0.119	-0.050	0.097	0.171	0.184	-0.110	0.110	-0.081	0.059	-0.087	0.120
CAPM	*b*	1.000	0.078	1.502	0.128	0.052	0.110	1.324	0.038	0.891	0.016	0.202	0.041
LCAPM	b^-	0.580	0.182	1.101	0.140	-1.132	0.203	1.424	0.109	1.099	0.039	0.454	0.114
	b^+	0.816	0.123	0.894	0.168	-0.649	0.150	1.820	0.070	1.102	0.031	0.669	0.061
HCAPM	b_1	1.331	0.115	1.080	0.060	0.193	0.145	0.906	0.057	1.148	0.033	-0.078	0.068
	b_2	2.563	0.972	0.454	0.529	4.006	0.997	-0.456	0.526	1.318	0.235	0.758	0.617
	b_3	0.178	0.054	0.017	0.036	0.203	0.051	-0.047	0.033	0.129	0.020	-0.032	0.036
	p_{11}	0.908	0.065	0.681	0.188	0.899	0.081	0.907	0.038	0.928	0.029	0.906	0.050
	p_{12}	0.041	0.039	0.287	0.184	0.043	0.041	0.046	0.028	0.035	0.021	0.052	0.035
	p_{13}	0.051	0.047	0.032	0.033	0.058	0.063	0.047	0.027	0.036	0.021	0.042	0.036
	p_{21}	0.059	0.050	0.295	0.190	0.054	0.046	0.070	0.039	0.040	0.025	0.067	0.042
	p_{22}	0.760	0.080	0.666	0.189	0.792	0.093	0.912	0.043	0.893	0.046	0.913	0.046
	p_{23}	0.181	0.076	0.039	0.047	0.154	0.094	0.018	0.017	0.067	0.037	0.020	0.020
	p_{31}	0.065	0.053	0.023	0.026	0.069	0.068	0.060	0.033	0.033	0.024	0.049	0.041
	p_{32}	0.166	0.072	0.036	0.043	0.151	0.099	0.018	0.016	0.069	0.037	0.025	0.023
	p_{33}	0.769	0.082	0.941	0.050	0.780	0.114	0.922	0.037	0.898	0.040	0.927	0.048
	σ_1	4.076	0.847	4.982	1.392	9.160	2.339	3.656	0.589	0.616	0.132	3.556	0.810
	σ_2	15.578	3.349	6.272	1.310	27.560	4.252	29.791	5.383	3.896	0.656	22.304	3.808
	σ_3	3.458	1.017	2.289	0.966	12.403	4.028	2.402	0.438	1.017	0.188	3.282	0.886

Table 2-10 Average value-minus-growth return and market return per regime for additional portfolios

This table reports the average value premium and average excess market returns per regime. Panel A reports results for the value-minus growth portfolio based on a decile sort on Earning-to-Price ratio over the period July 1963- December 2006, and Panel B for the value-minus-growth portfolios obtained with a 2 by 3 sort on size and book-to-market over the period January 1927- June 1963. The t-statistics for average value premium (excess market return) compare the value-minus-growth portfolios' returns (excess market return) in each regime with the average value-minus-growth return (excess market return) over each complete sample. We conduct a Kruskal-Wallis test of equality of median value premium and market returns in each regime. We collect returns for each regime when the probability of each regime is the highest one at time *t*. HML_EP is the H_EP portfolio subtracted from the L_EP portfolio, where H_EP and L_EP are the value and growth portfolios obtained with a 2 by 3 sort on size and book-to-market. HML is the H portfolio subtracted from the L portfolio, where H and L are the value and growth portfolios obtained with a 2 by 3 sort on size and book-to-market.

Regime	Number of months	Average HML (HML_EP) (%)	t-stat	p-value for H_0: median HML (HML_EP) is equal in all regimes	Average excess market return (%)	t-stat	p-value for H_0: median market return is equal in all regimes
Panel A. Regimes inferred using HML_EP portfolio, sample period: 1963-2006							
CAPM	264	0.086	-3.058		0.840	1.547	
LCAPM	137	1.293	1.146	0.032	0.551	0.178	0.012
HCAPM	121	0.934	0.890		-0.400	-2.005	
Whole sample	522	0.600			0.477		
Panel B. Regimes inferred using HML portfolio, sample period: 1927-1963							
CAPM	196	0.169	-1.353		0.898	0.121	
LCAPM	135	1.010	1.052	0.594	1.091	0.307	0.755
HCAPM	113	-0.041	-2.210		0.495	-0.778	
Whole sample	444	0.371			0.854		

Figure 2-1 Smoothed probabilities of market regimes - Full Sample (July 1926 to December 2006)

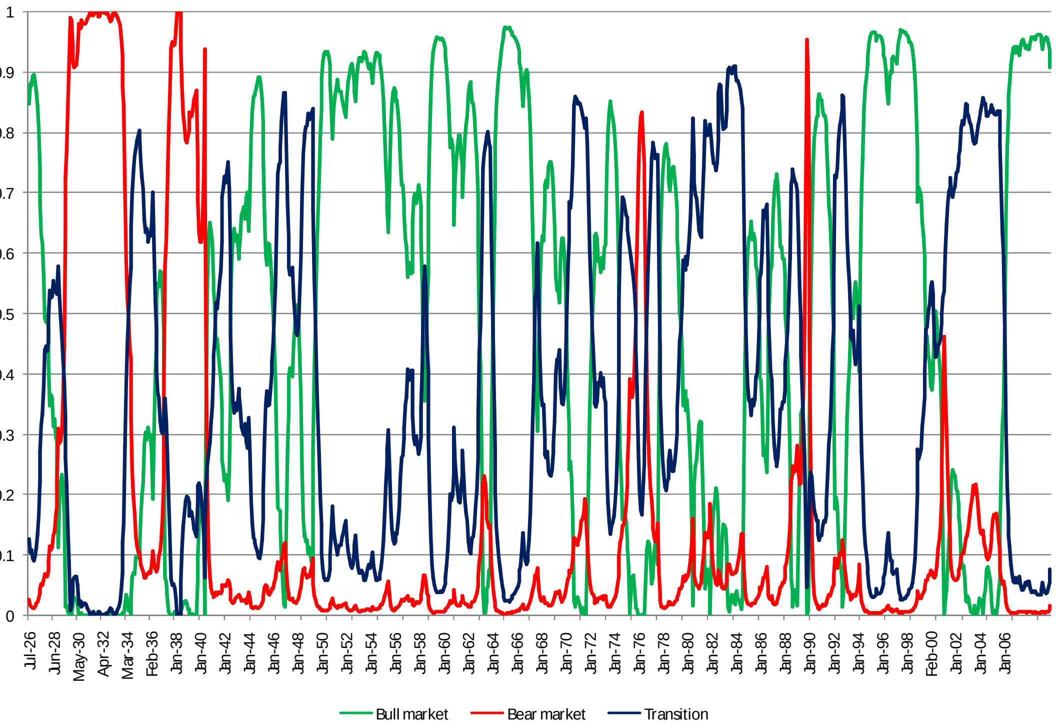


Figure 2-2 Smoothed probabilities of market regimes - Post-Depression Sample (July 1935 to December 2006)

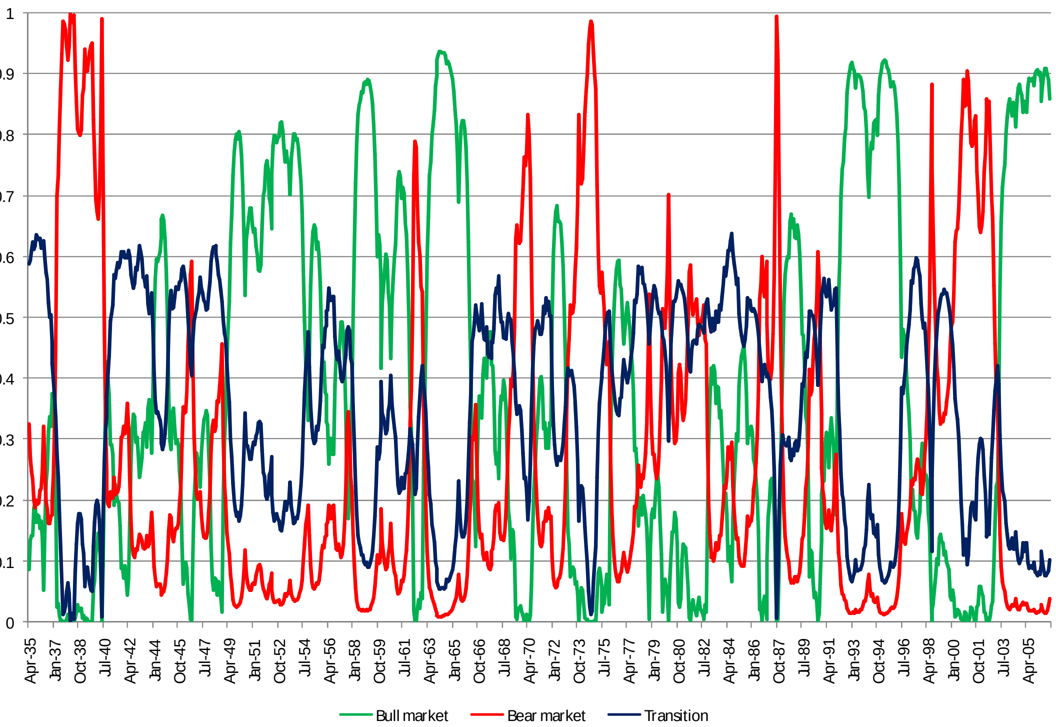


Figure 2-3 Smoothed probabilities of market regimes - Post-1963 Sample (July 1963 to December 2006)

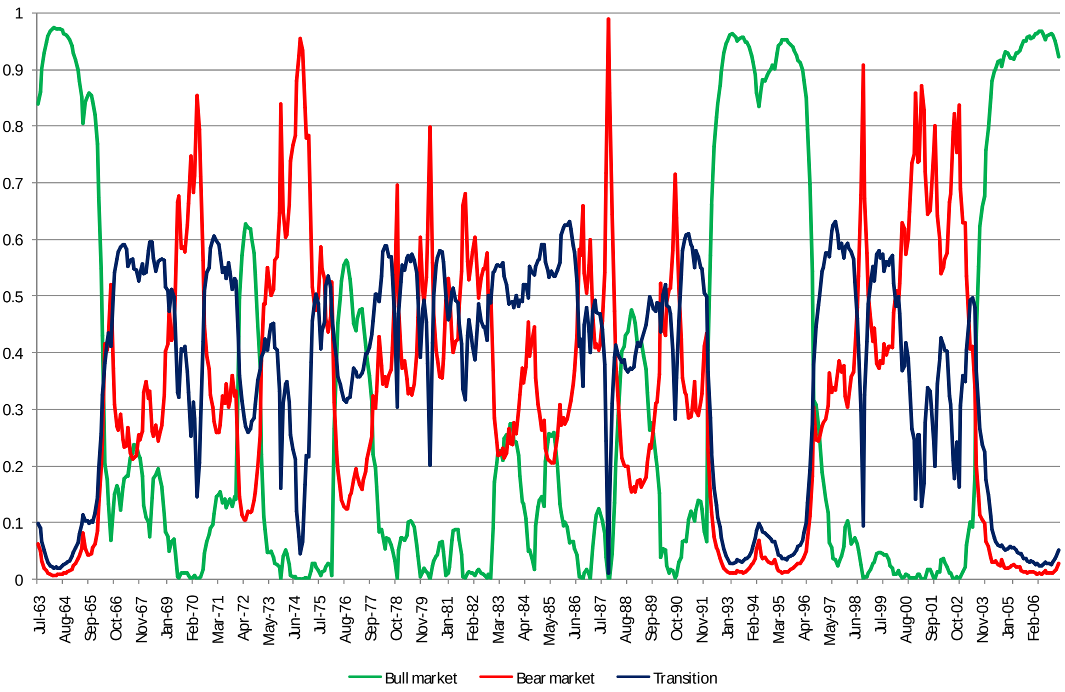


Figure 2-4 Smoothed probabilities of market regimes - Pre-1963 Sample (July 1926 to June 1963)

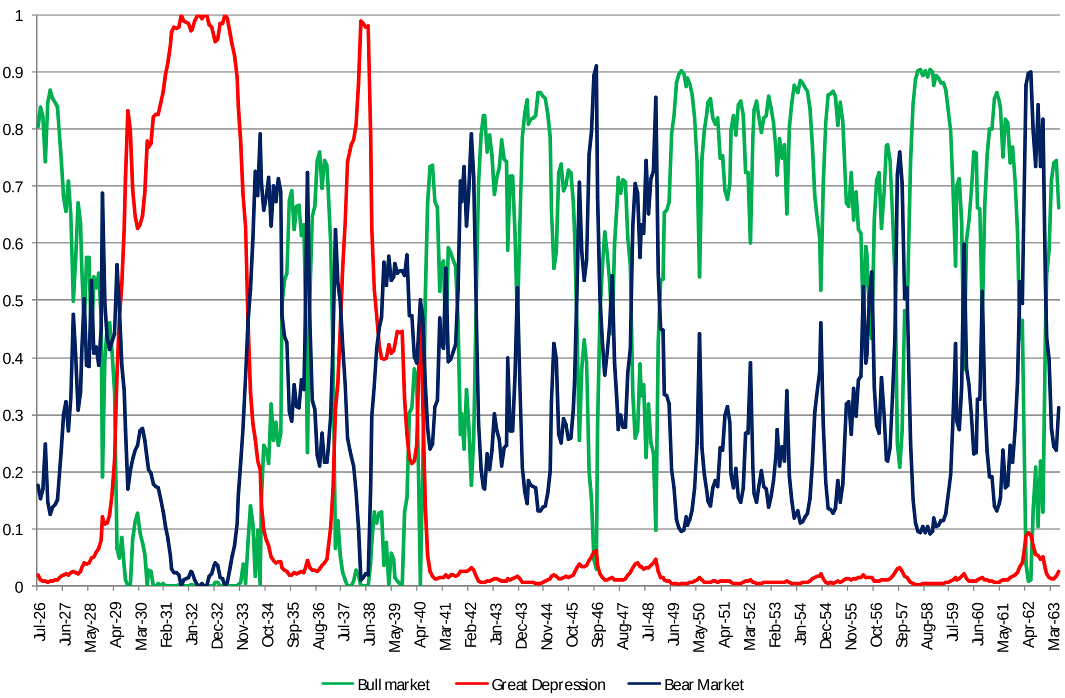


Figure 2-5 Smoothed probabilities for the regime-switching model with alternative risk measures

The graphs show smoothed probabilities of the three regimes in model (2.7) (which correspond to the CAPM, LCAPM and HCAPM models) for the H, L, HML, Hs, Ls and HMLs portfolios. The sample is from January 1963 to December 2006. The probabilities are estimated with 10000 iterations of the Gibbs-sampling algorithm.

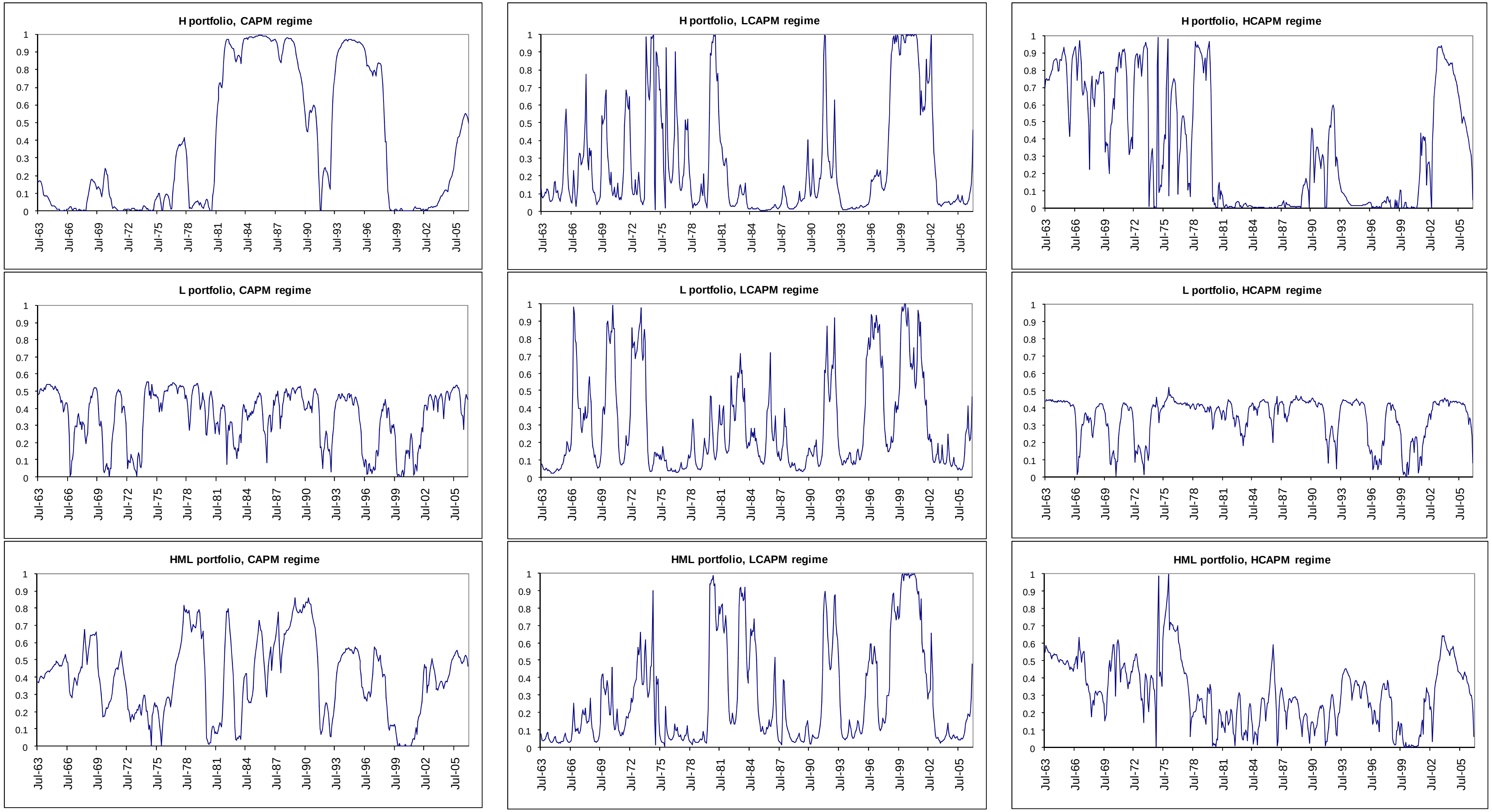
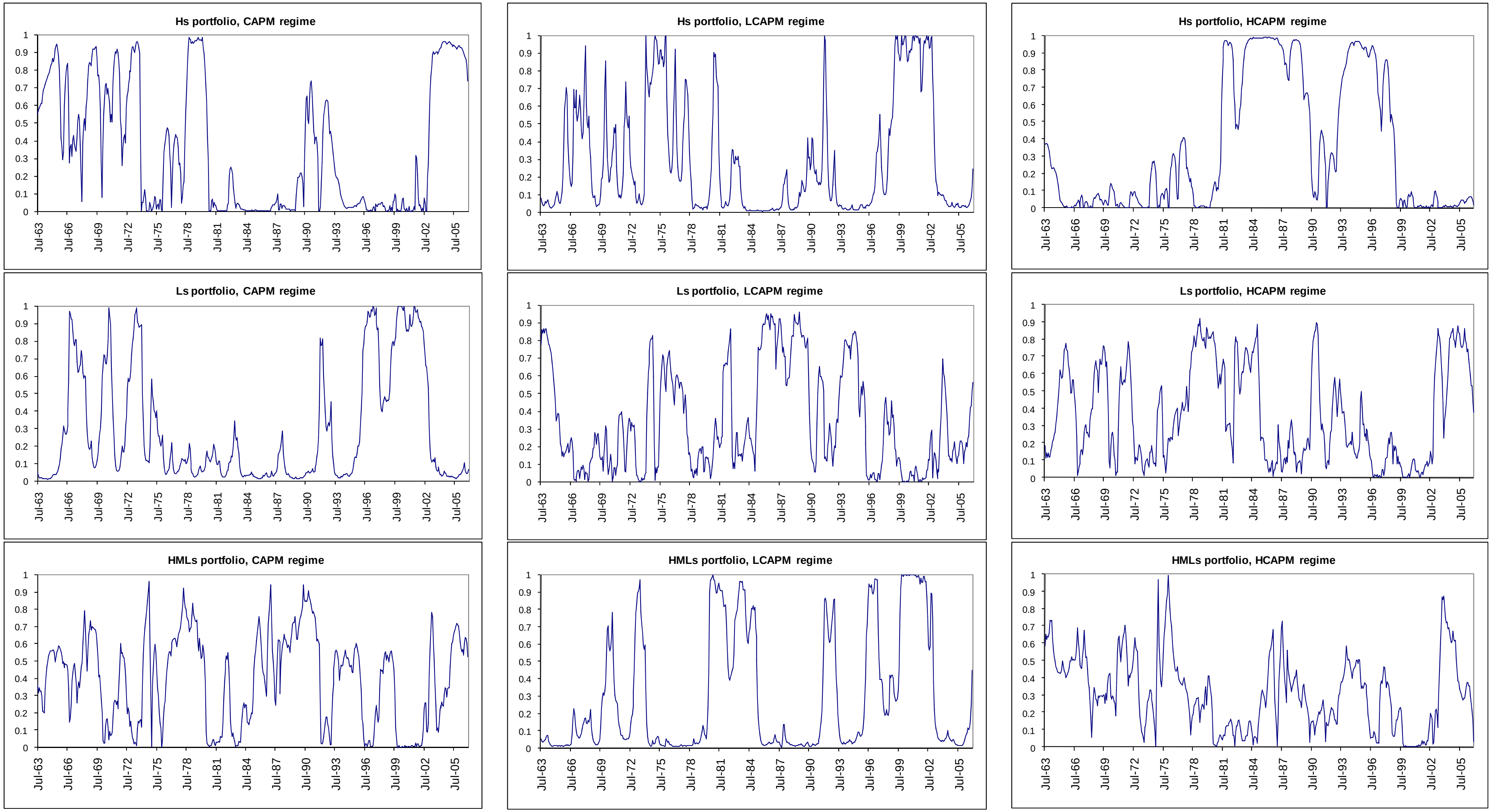


Figure 2-5 (continued)



References

Ang, Andrew, and Joseph Chen, 2007, CAPM over the long run: 1926-2001, *Journal of Empirical Finance* 14, 1-40.

Arrow, Kenneth, 1971. *Essays in the theory of risk-bearing* (Markham Publishing Co, Chicago).

Barone-Adesi, G., 1985, Arbitrage equilibrium with skewed assets, *The Journal of Financial and Quantitative Analysis* 20, 299-313.

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

Bernanke, Ben, 1983, Nonmonetary effects of the financial crisis in the propagation of the great depression, *American Economic Review* 73, 257-276.

Brennan, Michael J., and Y. Xia, 2001, Assessing asset pricing anomalies, *Review of Financial Studies* 14, 905-942.

Campbell, John Y., and Tuomo Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review* 94, 1249-1275.

Carter, C. K., and R. Kohn, 1994, On Gibbs sampling for state-space models, *Biometrika* 81, 541-553.

Chan, K.C., 1988, On the contrarian investment strategy, *The Journal of Business* 61, 147-163.

Chen, Long, R. Petkova, and L. Zhang, 2006, The expected value premium, (Ross School Of Business, University of Michigan).

Chen, Son-Nan, 1982, An examination of risk-return relationship in bull and bear markets using time-varying betas, *The Journal of Finance and Quantitative Analysis* 17, 265-286.

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, business cycle, and time-varying expected returns, *The Journal of Finance* 57, 985-1019.

Cooper, Ilan, Bruno Gerard, and Guojun Wu, 2005, (Norwegian School of Management and Ross School of Business, University of Michigan,).

Cooper, Michael J., and Stefano Gubellini, 2008, The crucial role of conditioning information in determining if value is riskier than growth, *Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1107528*.

Daniel, Kent, and Sheridan Titman, 2005, Testing factor-model explanations of market anomalies, *working paper*.

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

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

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

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns of stocks and bonds, *The Journal of Financial Economics* 33, 357-384.

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

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

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *The Journal of Finance* 63, 1653-1678.

Fama, Eugene F., and Kenneth R. French, 2006a, The value premium and the CAPM, *The Journal of Finance* 61, 2163-2185.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *The Journal of Political Economy* 81, 607-636.

Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin, 2003, Spurious regressions in financial economics?, *The Journal of Finance* 58, 1393-1414.

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

Hamilton, J. D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57, 357-384.

Hamilton, J. D., and Gang Lin, 1996, Stock market volatility and the business cycle, *Journal of Applied Econometrics* 11, 573-593.

Hamilton, J. D., and R. Susmel, 1994, Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics* 64, 307-333.

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

Harvey, C. R., and A. Siddique, 2000, Conditional skewness in asset pricing tests, *The Journal of Finance* 50.

Harvey, C. R., and A. Siddique, 2000, Time-varying conditional skewness and the market risk premium, *Research in Banking and Finance* 1, 25-58.

Hwang, S., and S. E. Satchell, 1999, Modelling emerging market risk premia using higher moments, *International Journal of Finance and Economics* 4.

Ingersoll, 1987, *Theory of financial decision making* (Rowman & Littlefield Publishers, Inc.).

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

Kim, C., and C. R. Nelson, 1999. *State-space models with regime switching* (MIT Press).

Kim, Moon K. , and J. Kenton Zwalt, 1979, An analysis of risk in bull and bear markets, *The Journal of Financial and Quantitative Analysis* 14, 1015-1025.

Kraus, A., and R. H. Litzenberger, 1976, Skewness preference and the valuation of risky assets, *The Journal of Finance* 31, 1085-1100.

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

Lewellen, J., and S. Nagel, 2006, The conditional CAPM does not explain asset-pricing anomalies, *forthcoming, Journal of Financial Economics*.

Lim, K., 1989, A new test of the three-moment capital asset pricing model, *The Journal of Financial and Quantitative Analysis* 24, 205-216.

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

Miller, Merton, and Myron Scholes, 1972, Rates of return in relation to risk: A re-examination of some recent findings, in Michael Jensen, ed.: *Studies in the theory of capital markets* (New York: Praeger).

Newey, W. K., and K.D. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Perez-Quiros, Gabriel, and Allan Timmermann, 2000, Firm size and cyclical variations in stock returns, *The Journal of Finance* 55, 1229-1262.

Petkova, R., and L. Zhang, 2005, Is value riskier than growth?, *Journal of Financial Economics* 78, 187-202.

Post, T., and P. Van Vliett, 2005, Conditional downside risk and the CAPM.

Rosenberg, B, K. Reid, and R. Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-11.

Ross, S. A., 1978, The current status of the capital asset pricing model, *The Journal of Finance* 23, 885-901.

Schwert, G. W., 1989, Why does stock market volatility change over time?, *The Journal of Finance* 44, 1115-1153.

Sears, R. S, and K. C. J Wei, 1985, Asset pricing, higher moments and the market risk premium: A note, *The Journal of Finance* 40, 1251-1253.

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

Turner, C. M., R. Startz, and C. R. Nelson, 1989, A Markov model of heteroskedasticity, risk and learning in the stock market, *Journal of Financial Economics* 25, 3-22.

Xing, Yuhang, and L. Zhang, 2005, Value versus growth: Movements in economic fundamentals, (Simon School Working Paper).

Zellner, A., 1971. *An introduction to Bayesian inference in econometrics* (John Wiley & Sons).

Zhang, L., 2005, The value premium, *The Journal of Finance* 60, 67-103.

Chapter 3

The Disappearance of Momentum

3.1. Introduction

Extensive empirical studies on asset pricing since the late 1970s have brought us a number of unexpected results, one of which is the momentum premium documented by Jegadeesh and Titman (1993, 2001). In their seminal work, Jegadeesh and Titman (1993) report that strategies which buy stocks with high returns over the previous 3 to 12 months and short stocks with low returns in the same period earn an abnormal return of approximately 1% per month over a holding period of up to 12 months. In their second paper on momentum, Jegadeesh and Titman (2001) argue that the fact that the premium exists throughout the 1990s after being documented in their earlier work with data until 1990 suggests that the anomaly is not a product of data mining. Moreover, evidence of the existence of momentum in other markets makes this possibility unlikely (see Rouwenhorst (1998)).

The size and apparent persistence of momentum profits have attracted considerable attention, and many studies have tried to explain the phenomenon. One strand of research aims to explain the momentum premium rationally. Some studies attempt to explain the premium with changes in fundamentals; Johnson (2002), for instance, proposes a model in which stock returns are more sensitive to changes in expected growth when expected growth is high, and Liu, Warner and Zhang (2005) show that winners have higher dividend, investment and sales growth rates than losers, and thus the expected returns of winners are systematically higher than those of losers. Berk, Green and Naik (1999) suggest a dynamic model for a firm's optimal investment choices which can generate the empirical pattern of momentum. Conrad and Kaul (1998) propose a risk-based explanation suggesting that the momentum premium originates from cross-sectional variation in expected returns. On the other hand, Chordia and Shivakumar (2002) argue that time-variation in the sensitivities of the momentum strategy to macroeconomic factors can account for the momentum premium.

Another view is that the momentum premium arises due to biases in the way investors behave or interpret information. Chan, Jegadeesh and Lakonishok (1996) suggest that investors underreact to information such as past earnings, whilst Daniel, Hirshleifer and Subrahmanyam (1998) argue that psychological biases in investor's perceptions create under- and overreactions which in turn are responsible for momentum and long-term reversal. Hong and Stein (1999) propose a model where information diffuses slowly, creating short-term price underreaction and allowing momentum traders to profit from following trends. These studies are supported by Jegadeesh and Titman (2001) and Cooper, Gutierrez and Hameed (2004). Zhang (2006) analyses the role that information uncertainty (i.e. ambiguity about the implication of new information for a firm's value) plays in the continuation of prices. He argues that investors underreact more strongly to public information when there is more uncertainty regarding this information.

The studies mentioned above propose various explanations about why momentum strategies are profitable, but they do not directly explain why the momentum premium has not disappeared once it became well known in the early 1990s. One would expect that such a high premium would quickly disappear in a well-developed market due to arbitrage activity, unless it is a reward for risk, or it is not exploitable due to transaction costs, restrictions on short sales, illiquidity, or other market microstructure issues. Korajczyk and Sadka (2004) argue that, although market microstructure may explain a part of the profitability of momentum strategies, it cannot fully explain the anomaly, whilst Pastor and Stambaugh (2003) estimate that half of momentum profits can be explained by a liquidity factor. The large momentum premium, if it still exists more than 10 years from the first documentation in the academic literature, is hard to explain considering the increasing number of hedge funds that try to exploit every possible arbitrage opportunity. Moreover, other well-known premia such as the size and value effects have weakened since they were documented[16].

In this study, we contribute to the debate by questioning whether momentum strategies consistently provide premia over the sample period from January 1927 to December 2006. We analyse the risk-adjusted momentum premium of several momentum strategies with a model which allows for multiple structural breaks (or change-points) in the relationship between momentum returns and the risk factors[17]. The risk factors we use are the ones in the Fama and French (1993, 1996) three-factor model.[18] The model is estimated via a Bayesian approach, which provides a

[16]There is stronger evidence that the size premium has decreased, and only mixed evidence that the value premium has weakened in recent times, see Schwert (2003) and the discussion in Jegadeesh and Titman (2001).
[17] We use the terms 'structural breaks' and 'change-points' interchangeably in this work.
[18] All the results reported in this study are obtained using the Fama-French (1996) three-factor model, but we also found that these results vary little from those obtained with the CAPM.

convenient way to determine the optimal number of change-points through the Bayes factors.

Our results indicate that, far from being a consistently profitable investment opportunity, the (risk-adjusted) premium from the momentum strategy is significantly positive only during certain periods, whilst it is close to zero or even negative during other periods. We find similar structural breaks for different momentum strategies during the period from January 1927 to December 2006. The momentum strategies appear to be anomalously profitable only during the 1940-1965 and the 1975-2000 periods. In particular, the strategies do poorly since the last structural break which occurs around 2000.

If momentum disappeared since 2000, the natural question is why it took so long for arbitrageurs to act, since momentum was first reported in the early 1990s? We seek an answer from the extraordinary boom in a few sectors such as high-tech and telecom stocks in the late 1990s. During the late 1990s, the momentum premium was driven by winners, many of which were hi-tech and telecom stocks which showed extraordinary outperformance. In fact, when these hi-tech and telecom stocks are excluded, the overall pattern in momentum profitability from the early 1990s is decreasing, and the magnitude of momentum profits in the 1990s becomes far smaller than that reported by Jegadeesh and Titman (2001). Therefore we conclude that the momentum premium has slowly begun to disappear since the early 1990s. The main reason why momentum premium still appears to exist after it was first reported by Jegadeesh and Titman (1993) is that the premium was driven by a few sectors that showed an extraordinary boom in the late 1990s when these sectors played the role of winners in the momentum calculation.

This paper is organised as follows. We introduce our methodology and briefly discuss the estimation method in Section 3.2. The explanation of the data set and empirical tests follows in Section 3.3. In Section 3.4 we discuss momentum in the context of the 1990s stock market boom. Section 3.5 concludes. A detailed description of the estimation of the methodology is given in the Appendix in Section 3.6.

3.2. Modelling Structural Changes in Premia

3.2.1. Multiple Change-Point Model

Estimates of a premium based on past data vary widely depending on sample periods and methodologies. Although estimates with longer time series data could provide a more precise perspective on the premium over a long horizon, they do not reflect time variation in the premium. Modelling time-variation in a premium may not be an easy task, since the model specification of the premium is not known. A simple method would be to assume a stochastic process for the premium. An alternative method which we use in this study is to allow regime changes in the premium over time, which is similar to the model that Pastor and Stambaugh (2001) use to estimate the equity premium. Our method consists in detecting changes in the premium with respect to other risk factors, whilst also allowing the premium to be different across regimes.

The model we use to analyse the momentum premium is based on the class of multiple change-point models proposed by Chib (1998). This framework consists of a hidden Markov regime-switching model with a restricted transition probability matrix to model the change-points. The estimation of the model is carried out in a Bayesian fashion, using Markov Chain Monte Carlo (MCMC) efficient posterior simulation methods proposed by Chib (1998), in which the parameters of the model and the change-points are estimated conditioned on the number of breaks. Chib (1998) also provides details of the calculation of the marginal likelihood, from which it is possible to calculate the Bayes factors, in order to compare models with different numbers of structural breaks.

Consider the classical linear regression model

$$y_t = \mathbf{x_t}\boldsymbol{\beta}' + \varepsilon_t\ , \quad t = 1,\ldots,T\ , \tag{3.1}$$

where in our case y_t represents the return on a momentum strategy at time *t*, $\mathbf{x_t} = \left(1\ f_{1t}\ f_{2t} \cdots f_{Kt}\right)$, where f_{jt}, *j* = 1,…,*K* is the realised value of the *j*-th asset pricing factor at time *t*, $\boldsymbol{\beta} = \left(\beta_0\ \beta_1 \cdots \beta_K\right)$ is the vector of factor sensitivities and β_0 is the intercept term, and ε_t is a zero-mean error term with variance σ^2. The parameters of model (3.1) are collected in a vector $\boldsymbol{\theta} = \left(\boldsymbol{\beta}, \sigma^2\right)$. Now suppose that there are *m* structural breaks in the model at unknown times $\Upsilon = \left\{\tau_1, \tau_2, \ldots, \tau_m\right\}$, that is, each of the intervals defined by the change-points is a different regime. In each of the *m*+1 regimes, the relationship between y_t and X_t is governed by a parameter vector $\boldsymbol{\theta}_k = \left(\boldsymbol{\beta}_k\ \sigma_k^2\right)$, that is,

$$\boldsymbol{\theta}_k = \begin{cases} \boldsymbol{\theta}_1 & \text{if} \quad t < \tau_1 \\ \boldsymbol{\theta}_2 & \text{if} \quad \tau_1 \le t < \tau_2 \\ \vdots & \vdots \qquad \vdots \\ \boldsymbol{\theta}_m & \text{if} \quad \tau_{m-1} \le t < \tau_m \\ \boldsymbol{\theta}_{m+1} & \text{if} \quad \tau_m \le t \end{cases} . \tag{3.2}$$

The regime or state is indicated by a discrete random variable $S_t \in \{1, 2, \dots, m+1\}$. Specifically, if $S_t = k$ then the model is in regime k and the density of y_t is $f\left(y_t \mid \tilde{\mathbf{y}}_{t-1}, \boldsymbol{\theta}_k\right)$, where $\tilde{\mathbf{y}}_{t-1} = \left(y_1 \cdots y_{t-1}\right)$. The variable S_t is modelled as a discrete-state Markov process with a constrained transition probability matrix, such that switches can only occur from a certain regime, say regime *k*, to the next one, $k+1$. The transition probability matrix $\mathbf{P} = \{p_{ij}\}$, where $p_{ij} = P\left(S_t = j \mid S_{t-1} = i\right)$, is therefore represented as

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ \cdots & \vdots & 0 & p_{mm} & p_{m,m+1} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} . \tag{3.3}$$

Therefore, in this model the probability of a regime change depends on the current regime. The model starts in regime 1 at $t = 1$, the first data point, and at $t = T$, the last observation, the model is in regime $m+1$. This constrained transition probability matrix enforces the ordered structure of the change-points given by (3.2).

3.2.2. Estimation

The model is estimated via a Markov Chain Monte Carlo (MCMC) Bayesian approach, using the method proposed by Chib (1998). The parameters of the model are $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}_1 \cdots \boldsymbol{\theta}_{m+1})$ and $\mathbf{P}$. In the Bayesian context, given a prior density $\pi(\tilde{\boldsymbol{\theta}}, \mathbf{P})$, the interest lies in obtaining the posterior density $\pi\left(\tilde{\boldsymbol{\theta}}, \mathbf{P} \mid \tilde{\mathbf{y}}_T\right) \propto \pi\left(\tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot f\left(\tilde{\mathbf{y}}_T \mid \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$. Chib (1998) suggests augmenting the parameter space to include the unobserved states $\tilde{\mathbf{S}}_T = (S_1, \dots, S_T)$ and then applying an MCMC sampling scheme to the posterior density $\pi(\tilde{\mathbf{S}}_T, \tilde{\boldsymbol{\theta}}, \mathbf{P} \mid \tilde{\mathbf{y}}_T)$.

3.2.2.1. Priors

One of the important benefits from using Bayesian methods is to obtain posterior probabilities for different priors, i.e. the researcher's opinion can be incorporated into the posterior probability through different priors[19]. In our study, our prior knowledge does not suggest when there were structural breaks in the momentum premium (if any) and what the momentum premia were during the regimes. Therefore we use relatively uninformative priors, allowing the data to identify the structural breaks.

We specify independent conditionally conjugate priors for the parameters of the model. The priors are given by

[19] See for instance Ang and Chen (2007), who use different priors to explain the value premium.

$$\boldsymbol{\beta}_k \sim N\left(\mathbf{b}_0, \mathbf{B}_0\right),\ \sigma_k^2 \sim IG\left(\frac{\upsilon_0}{2}, \frac{\delta_0}{2}\right),\ p_{ii} \sim Beta\left(c_0, d_0\right), \tag{3.4}$$

for $k = 1,\ldots,m+1,\ i = 1,\ldots,m$. We choose relatively uninformative priors for $\boldsymbol{\beta}_k$ with $\mathbf{b}_0 = \left(0\cdots0\right)$, and $\mathbf{B}_0 = 10\cdot\mathbf{I}_K$, where $\mathbf{I}_K$ is the identity matrix of size $K\times K$. The priors for σ_k^2 are also quite uninformative with $\upsilon_0 = \delta_0 = 0.001$ (which reflects a prior belief of a large variance). For the transition probabilities p_{ii} we follow Chib (1998) and choose a prior that favours infrequent structural breaks. Specifically, our prior reflects an *a priori* belief that all regimes have the same expected duration of $\eta = T/m$. This is achieved by choosing $c_0 = \eta/10$ and $d_0 = 0.1$. For instance, if there are T = 960 months in our sample and m = 5, the expected duration of each regime is $\eta = 192$ months, and the parameters which reflect this are $c_0 = 19.2$ $d_0 = 0.1$, which imply a mean transition probability of 0.9948[20].

3.2.2.2. Gibbs Sampling

In this subsection we briefly describe the estimation procedure, and give a detailed explanation in Section 3.6. We use an iterative Gibbs sampling scheme to obtain the augmented posterior density, $\pi(\tilde{\mathbf{S}}_T, \tilde{\boldsymbol{\theta}}, \mathbf{P} \mid \tilde{\mathbf{y}}_T)$. Given initial values $\tilde{\boldsymbol{\theta}}^{(0)}, \mathbf{P}^{(0)}$, the parameters and state vector are simulated from the following conditional distributions:

[20] Note that the expected duration of a regime, when the transition probability is p, is given by $1/(1-p)$.

1. $\tilde{\mathbf{S}}_T \mid \tilde{\mathbf{y}}_T, \tilde{\boldsymbol{\theta}}, \mathbf{P}$,
2. $\tilde{\boldsymbol{\theta}} \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_T, \mathbf{P}$, and
3. $\mathbf{P} \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_T, \tilde{\boldsymbol{\theta}}$,

where in the second step we simulate, in turn, from the conditional distributions of $\boldsymbol{\beta}_k, \sigma_k^2,\ k = 1,\ldots,m+1$. Whilst the conditional distributions in steps 2 and 3 are quite standard from Gibbs sampling methods, the simulation of $\tilde{\mathbf{S}}_T$ requires an elaborate recursive procedure. The joint distribution of $\tilde{\mathbf{S}}_T$, $p\left(\tilde{\mathbf{S}}_T \mid \tilde{\mathbf{y}}_T, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$, can be written as a recursion in reverse time order as

$$p\left(\mathbf{S}_{T-1} \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_T, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \times \ldots \times p\left(S_t \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_{t+1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \times \ldots \times p\left(S_1 \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_2, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right). \quad (3.5)$$

Given that $S_T = m+1$, all that is needed to generate each $S_t,\quad t = T-1,\ldots,1$ are the mass functions $p\left(S_t \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_{t+1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$. Chib (1996) shows that this mass function can be written as

$$p\left(S_t \mid \tilde{\mathbf{y}}_T, \tilde{\mathbf{S}}_{t+1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \propto p\left(S_t \mid \tilde{\mathbf{y}}_t, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot p\left(S_{t+1} \mid S_t, \mathbf{P}\right) \quad (3.6)$$

where $p\left(S_{t+1} \mid S_t, \mathbf{P}\right)$ is the transition probability from the Markov chain. The quantity $p\left(S_t \mid \tilde{\mathbf{y}}_t, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$ can be calculated using a set of backwards recursions which are detailed in the Section 3.6. With these values in hand, $\tilde{\mathbf{S}}_T$ can be simulated using (3.6), where the normalising constant is easy to obtain, since conditioned on S_{t+1}, each S_t can take only two values.

All our results are obtained with 10,000 iterations after 10,000 burn-in iterations, i.e. we start the MCMC sampling and allow the first 10,000 iterations to be discarded, so that convergence of the chain is guaranteed.

3.2.3. Bayes Factors

The number of structural breaks in the multiple change-point framework explained above is fixed and needs to be chosen before the model is estimated. Typically, one considers models with different numbers of structural breaks, and then calculates Bayes factors to choose between these models[21]. Consider two models $\mathcal{M}_r$ and $\mathcal{M}_s$ where *r* and *s* denote different numbers of structural breaks. The Bayes factor for comparing model $\mathcal{M}_r$ versus model $\mathcal{M}_s$ is given by the ratio of the marginal likelihoods of $\mathcal{M}_r$ to $\mathcal{M}_s$. We apply the method proposed by Chib (1995) to our multiple change-point model in order to calculate the marginal likelihood from the Gibbs sampling iterations (see Section 3.6).

Once we calculate the marginal likelihoods for any two models $\mathcal{M}_r$ and $\mathcal{M}_s$, denoted by $m\left(\tilde{\mathbf{y}}_T \mid \mathcal{M}_r\right)$ and $m\left(\tilde{\mathbf{y}}_T \mid \mathcal{M}_s\right)$, the Bayes factor comparing $\mathcal{M}_r$ versus $\mathcal{M}_s$ is given by

$$B_{rs} = \frac{m\left(\tilde{\mathbf{y}}_T \mid \mathcal{M}_r\right)}{m\left(\tilde{\mathbf{y}}_T \mid \mathcal{M}_s\right)} \tag{3.7}$$

[21] Kass and Raftery (1995) provide a review of the calculation and interpretation of Bayes factors.

where $\mathcal{M}_r$ is more strongly supported over $\mathcal{M}_s$ as B_{rs} increases. Kass and Raftery (1995) suggest the following interpretation of the values of the Bayes factor in terms of the support for $\mathcal{M}_r$ over $\mathcal{M}_s$. If $0 \le B_{rs} < 3$, there is very weak support, barely worth mentioning; if $3 \le B_{rs} < 20$, the support is moderate and significant; if $20 \le B_{rs} < 150$, there is strong and significant support and if $B_{rs} > 150$, very strong and significant evidence to support $\mathcal{M}_r$ over $\mathcal{M}_s$.

3.3. Empirical Results

3.3.1. Momentum Returns

We use various momentum portfolios in this paper for robustness. For our main results, we form portfolios using all stocks traded in the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and NASDAQ on the CRSP monthly data file for the period from July 1926 to December 2006[22]. In order to minimise the effects of the bid-ask bounce, at the end of each month *t* we sort stocks into decile portfolios based on their past five month returns from *t*-5 to *t*-1, skipping month *t* (formation period: 6 months), as in Cooper, Gutierrez and Hameed (2004). We then hold the portfolios for the next 6 months (holding period: 6 months). Following Jegadeesh and Titman (2001), we exclude all stocks with prices inferior to \$5 at the portfolio formation date and all stocks whose sizes would place them in the smallest

[22] AMEX stocks are present in the CRSP file from July 1962; NASDAQ stocks from December 1972.

NYSE decile, in order to minimise the effects of small and illiquid stocks. We call this portfolio $WML_{CRSP,6x6}$[23]. As in the momentum literature, we construct overlapping portfolios to increase the power of the tests. For example, in Jegadeesh and Titman (2001) a momentum decile portfolio in any month includes stocks ranked in that decile in any of the previous formation months. In this setting, we calculate momentum returns for the period from January 1927 to December 2006.

We explore a number of other momentum portfolios, which differ with regard to the size of the stocks allowed, the exchange in which the stocks are traded, and the length of the holding period. First, we maintain the formation and holding periods equal to 6 months, and construct momentum portfolios using only NYSE stocks, excluding those with market capitalisation smaller than the percentiles of size 10%, 20% and 30% of the whole NYSE universe at formation time (we call these portfolios $WML_{NYSE10,6x6}$, $WML_{NYSE20,6x6}$ and $WML_{NYSE30,6x6}$). The purpose of this is to analyse the impact that size might have in the profitability of the momentum strategy. We also form momentum portfolios using all CRSP stocks larger than the smallest NYSE size decile, using a formation period of 6 months and holding periods of 1 month and 3 months (we call these portfolios $WML_{CRSP,6x1}$ and $WML_{CSRP,6x3}$), in order to investigate the impact of different holding periods.

Table 3-1 reports summary statistics of the returns on the different momentum portfolios[24]. Using the whole sample from January 1927 to December 2006 (Panel A), we find an average monthly momentum return ranging from 0.73% (for the $WML_{NYSE30,6x6}$ strategy, std. error 0.15) to 1.00% (for the $WML_{CRSP,6x1}$ strategy, std. error 0.19). All momentum strategies have average returns which are

[23] Unless otherwise specified, the results reported refer to this momentum strategy.

[24] Throughout this work we use the following terminology. We refer to the raw momentum returns as 'momentum returns' or 'momentum profits', and to the risk-adjusted return using the Fama-French three-factor model (i.e. alpha) as 'momentum premium'.

close to 0.80% to 0.90% a month and similar to the magnitudes reported in the literature (see Jegadeesh and Titman (1993, 2001)). In the pre-1963 period (Panel B), the average monthly returns from all strategies are smaller but still highly significant, close to 0.60% a month for all strategies. The momentum returns from all strategies is stronger in the post-1963 period (Panel C). For example, the commonly used $WML_{CRSP,6x6}$ strategy returns a monthly 1.16% (std. error 0.20). It is noteworthy that the momentum strategies are mostly negatively skewed and have large kurtosis, particularly during the pre-1963 period. It could be argued that if investors are sensitive to these features (as for instance if investors' utility functions have the Arrow (1971) desirable properties of positive but decreasing marginal utility of wealth and non-increasing absolute risk aversion) they might not find the momentum strategies desirable despite the significant momentum profits.

As expected, average momentum profits from the momentum strategy that excludes stocks smaller than the 30^{th} percentile of NYSE stocks are lower than using all stocks in the three exchanges. In the pre-1963 period, however, excluding small stocks produces little difference in momentum.

The average momentum returns are economically and statistically significant. Popular asset pricing models, namely the CAPM and the Fama-French three-factor model, fail to explain the momentum premium, as in many previous studies (see for example Jegadeesh and Titman (1993, 2001)). In Table 3-1, we confirm that the magnitude of the estimated alphas from these models is economically and statistically significant. The Fama-French model shows slightly higher premia than the CAPM: the estimates of the Fama-French alpha are larger than those of the CAPM alpha in all six momentum portfolios. For example, with the whole sample, the estimated CAPM alpha of the $WML_{CRSP,6x6}$ strategy is at 1.03% (std. error 0.16), whilst the estimate of the Fama-French alpha is 1.15% (std. error 0.15).

3.3.2. A Preliminary Analysis: Momentum in Five-Year Subsamples

In this subsection we explore how momentum profits behave over time, by conducting an exploratory OLS analysis with non-overlapping 60-month subsamples of the data. We divide our entire 1927-2006 sample into 16 subsamples of 60 months, each of which starts in January.

We analyse the momentum premium in terms of average monthly returns and alphas from the CAPM and the Fama-French three-factor model. The results are reported in Table 3-2. From the 16 subsample periods, alphas are positive at the 5% significance level in 2 (3) subsample periods if the CAPM (Fama-French three-factor model) is used. If we are less strict and increase the significance level to 10%, then alphas are significant in 9 (5) subsample periods if the CAPM (Fama-French model) is used. Alphas do tend to be positive, i.e. they are negative in only 1 (2) subsample periods for the CAPM (Fama-French three-factor model), but a small number of significantly positive alphas could be much of a blow against the strategy. For example, for Benartzi and Thaler (1995) myopic investors, the strategy would still be too risky.

In terms of the asset pricing factors that explain momentum returns, we note that there is a great deal of time-variability. For instance, in the CAPM, beta is negative and significant (at the 10% level) in some subperiods, whilst it is positive and significant in others. This means that the betas of loser and winner portfolios vary over time; sometimes loser stocks are riskier than winner stocks, whilst during other periods the opposite happens. When the Fama-French three-factor model is used, a more consistent pattern emerges. The two additional factors SMB and HML tend to be significant, with the sensitivities to both SMB and HML generally negative. The last period, however, is remarkably different with notably high and positive sensitivity on HML, which means that there might be many value stocks in the winner portfolio during that period.

3.3.3. Momentum Premium and Structural Breaks

The results in the previous subsection show that the momentum premium changes over time; contrary to the results using long samples reported in Table 3-1, there seem to exist periods of time when momentum strategies yield insignificantly positive or even negative returns. Moreover, the sensitivities of the momentum strategy to the risk factors also change from period to period. We now investigate this issue by using the multiple change-point model in Section 3.2.1, where the risk factors are the Fama-French three factors and structural breaks are allowed to occur in the constant term, the sensitivities of the momentum portfolio to the risk factors and also in the idiosyncratic volatility[25]. We start by considering the $WML_{CRSP,6x6}$ strategy (which is similar to the momentum strategies used in most studies) and later consider the other momentum strategies.

[25] In this framework the structural breaks occur in all parameters simultaneously.

3.3.3.1. Structural Breaks in the Momentum Premium

As mentioned previously, the number of structural breaks (or change-points) in the model has to be chosen in advance. In order to select an appropriate model, we estimate models with different numbers of change-points and then compare these models using the Bayes factors. Panel A of Table 3-3 reports the log-marginal likelihoods for models with different numbers of change-points, ranging from no change-point to seven change-points[26]. Panel B reports the difference between the log-marginal likelihoods of each pair of models. Recalling from (3.7) that the Bayes factor is given by the ratio of the marginal likelihoods, it follows that the Bayes factors can be calculated by exponentiating the values in Panel B[27]. Comparing the values in the table, we find very strong evidence for the model with six change-points in terms of the Bayes factors.

In order to date the change-points, we consider the MCMC distribution of the variable $\tilde{\mathbf{S}}_{\mathbf{T}}$. From these we can calculate the smoothed probability of the regimes, which we plot in Figure 3-1. The date of the change-points is chosen as the point when the smoothed probability of one regime becomes larger than the probability of the previous regime. As the graph shows, there is little ambiguity and the regime changes are quite clear.

The parameter estimates from this model, as well as the periods comprised by each regime (as explained above) are reported in Table 3-4. The table reports posterior medians and 0.95 probability intervals for the parameters, calculated with

[26] We also considered models with more than seven change-points and in those cases the extra regimes collapsed to a single observation. We interpret that the data do not support more than seven change-points.

[27] For example, the Bayes factor of the model with 6 structural breaks versus the model with 5 structural breaks is given by exp(148.909).

the MCMC draws. The most striking result from this table is that the momentum premium (as measured by the alpha in each regime) is significant only during two of the seven regimes. Specifically, alpha is positive and significant during Regime 3 (0.73% a month), which starts in May 1940 and finishes in March 1965, and during Regime 6 (1.13% a month), which starts in April 1975 and lasts until September 1999. The last regime change occurs in October 1999. Since then, the median alpha is around 0.87% a month, but it is not statistically significant. So although - as argued by Jegadeesh and Titman (2001) - the momentum premium has persisted throughout the 1990s, the momentum premium is no longer significant since the late 1990s.

As expected, the sensitivities of the $WML_{CRSP,6x6}$ strategy to the Fama-French factors change quite dramatically from regime to regime. Whilst the sensitivity to HML is mostly negative (although it is significant in only two of the regimes), suggesting that the momentum strategy has more exposure to growth than value stocks, the sensitivity to SMB seems to alternate between negative (Regimes 1-3 and 5) and positive values (Regimes 4, 6 and 7). This suggests that the momentum strategy is tilted towards small stocks during some periods and large stocks in others, which could be related to the cyclical variations in the risk and expected return of firms of different sizes, see for example Perez-Quiros and Timmermann (2000)[28]. We note that the sensitivity of the strategy to the market return also seems to alternate from positive to negative, indicating that during some regimes, winners have higher betas than losers, with the opposite happening during other regimes.

[28] Perez-Quiros and Timmermann (2000) show that the expected return of small firms increases sharply in recessions. If this is the case, small stocks might become winners during such periods and thus enter the momentum strategy.

3.3.3.2. Robustness of Structural Breaks in the Momentum Premium

The results above were obtained with the $WML_{CRSP,6x6}$ strategy which considers all CRSP stocks. We have seen from Table 3-1 that momentum strategies constructed from different subsets of stocks and using different holding periods also earn significant premia. In this subsection we investigate whether there are structural breaks in the premia from these strategies, and whether they occur in similar dates as those we found using the $WML_{CRSP,6x6}$ portfolio.

Since there is no guarantee that the number of change-points is the same for each momentum portfolio, we proceed as before and estimate models with different numbers of change-points, and then choose the optimal number of change-points through the Bayes factors. Table 3-5 reports the log-marginal likelihoods of models with different numbers of change-points for each of the momentum portfolios. To avoid repetition, we do not report the Bayes factors for all models for each portfolio[29]. Instead, we indicate the optimal number of change-points with bold values for the corresponding log-marginal likelihood. From the table, we see that the numbers of change-points are similar, i.e. 4 to 6 breaks during the sample period, although they are not exactly the same for all momentum portfolios. For example, four change-points are optimal for the $WML_{CRSP,6x3}$, five change-points for the $WML_{NYSE10,6x6}$ and $WML_{NYSE30,6x6}$, and six change-points for the $WML_{NYSE20,6x6}$ and $WML_{CRSP,6x1}$.

We report the estimation results for the optimal models for each momentum strategy in Table 3-6. First, we notice that the regimes change in similar dates from

[29] These can be easily calculated from the log-marginal likelihood values.

one strategy to another. For examples, for all strategies there is a change-point at the beginning of 1940, and the last change-point happens around the end of 1999 or 2000. Also, for most momentum portfolios there is a regime that starts at 1975 and lasts until around 2000.

Second, as in our previous results, momentum premia are not always significantly positive. A remarkably consistent pattern seems to emerge; for all momentum strategies, the periods when the alphas are positive are between 1940 and 1970 (or around 1965 for the $WML_{CRSP,6x6}$ and $WML_{CRSP,6x1}$ strategies), and from 1975 to around 1999 or 2000[30]. Although the magnitude of the premia varies according to the momentum strategies, the momentum strategies based only on NYSE stocks show smaller premia than the other strategies. For example, the premium from the three NYSE momentum strategies (Panels A, B and C) during the period from 1975 to 2000 range from 0.77% a month for the $WML_{NYSE30,6x6}$ to 0.87% a month for the $WML_{NYSE10,6x6}$. Excluding small and illiquid stocks could significantly reduce momentum premium. Strategies that are constructed with larger stocks show smaller premia (e.g. 0.77% with $WML_{NYSE30,6x6}$) than those constructed together with smaller stocks (e.g. 1.13% with $WML_{CRSP,6x6}$)[31].

Finally, after the last change-point in the end of 2000, alphas become much smaller than those in the previous periods, from 0.06% a month for the $WML_{NYSE30,6x6}$ to 0.63% a month for the $WML_{CRSP,6x1}$, and none of these alphas are significant at the 95% level.

[30] It seems that the last change-point occurs earlier for momentum strategies based on all CRSP stocks than for momentum strategies based only on NYSE stocks: for all CRSP-based momentum strategies there is a structural break in the end of 1999, and for all NYSE-based momentum strategies the break occurs in the end of 2000.
[31] The difference in the median premium is statistically significant.

Summarising the results so far, structural breaks occur in similar dates regardless of how momentum portfolios are constructed, and the momentum premium is not always significant. During periods when the momentum strategy *is* profitable, it is profitable for all the strategies regardless of firm size as in Fama and French (2008), but our results suggest that small stocks contribute more to momentum premium than large stocks. Finally and more importantly, we confirm that momentum premium does not exist since 2000.

3.3.4. Is the Recent Period Unusual?

Our previous results indicate that there have been periods of poor momentum performance in the past. An interesting question is whether the recent momentum premium is unusual relative to other periods. In this section we analyse this issue with a bootstrap experiment.

3.3.4.1. A Bootstrap Experiment

We conduct a bootstrap experiment designed to estimate the probability of observing the mean momentum return and the momentum premium in the period after the last structural break, in view of the past distribution of momentum returns over various periods of strong and poor momentum performance. We sample from the ex-post distribution of momentum returns with replacement and then calculate the mean momentum return and also estimate the momentum premium using the Fama-French three-factor model. By doing this repeatedly, we can estimate the empirical distributions of the mean momentum return and the momentum premium over a period. We then compare the recent momentum premium (and mean momentum return) with these empirically-constructed distributions. If the recent momentum premium lies on the left tail of these distributions, this is evidence that the momentum premium since the last structural break is different from the one inferred from the historic data.

We need to choose a period of time from which to sample, which we call the sampling period, and a reference period we want to calculate probabilities for. To eliminate a possible influence of the period which contains the hi-tech bubble (see Section 3.4), we choose January 2001 to December 2006 as the reference period. We then draw 100,000 random samples of 72 months (the number of months in the reference period) from the sampling period with replacement. To account for periods of strong and poor momentum performance, we use the following sampling periods, based on the structural breaks found in Table 3-4: January 1927 to December 2000 (the whole sample up to the start of the reference period), January 1927 to April 1940 (the period corresponding to the first two regimes, which is a period of poor momentum performance), May 1940 to March 1965 (the period corresponding to Regime 3, which is a period of significantly positive momentum premium), April 1965 to March 1975 (the period corresponding to the fourth and fifth regimes, another period of poor momentum performance) and April 1975 to December 2000 (the period corresponding to the sixth regime, which is a period of very strong momentum performance).

Table 3-7 reports the results. We find that the average momentum return in the recent period is not unusual. The probabilities of observing a smaller mean momentum return based on the various sampling periods are quite large, except for the last period from April 1975 to December 2000, which is a period of very strong momentum performance. On the other hand, the risk-adjusted momentum premium over the period from January 2001 to December 2006 is very improbable relative to all sampling periods, except for the period from January 1927 to April 1940. For all other sampling periods, the probabilities of observing a risk-adjusted momentum premium inferior to the one in the recent period from 2001 is less than 1%. Even if we use the time period that includes the Great Depression, i.e. the January 1927 to April 1940 sampling period, the recent momentum premium is larger than the premia from only 14% of the bootstrap samples. However, this period includes the extremely volatile period of the Great Depression, and the standard deviation of the empirical distribution of alpha calculated from this period is almost double that of any of the other periods, and thus the momentum premium during this period is not significant at all.

We therefore conclude that the recent period when momentum has disappeared is different from previous periods in terms of momentum premium, whether we consider periods of strong or weak momentum premium, or the whole historical distribution over more than 70 years. We interpret this as evidence that the premium has been eroded away, an issue which we discuss in more detail next.

3.3.5. Discussion: Delays in the Disappearance of Momentum

Jegadeesh and Titman (2001) assert that, because momentum profits exist (and are similar to previous periods' profits) in the nine years subsequent to the Jegadeesh and

Titman (1993) sample period[32], it is unlikely that the momentum premium is a result of data snooping, and that "market participants have not altered their investment strategies in a way that would eliminate this source of return predictability" (Jegadeesh and Titman (2001), page 718).

Our results suggest that since the last structural break which occurs at the end of 1999 (or 2000, depending on the momentum strategy) the premium is not significant. Considering that momentum has been very popular in academic works as well as in practice since the seminal paper of Jegadeesh and Titman (1993), we wonder why it took such a long time for investors to erode the profit opportunity from simple momentum strategies. If the market is efficient, its participants are expected to act quickly in exploiting arbitrage opportunities if momentum is not related to priced risk.

One possibility is that not all investors can easily implement the momentum strategy. The portfolio of loser stocks contains stocks with extreme past underperformance, likely to be small cap and illiquid (see Korajczyk and Sadka (2004)) and difficult to short. Also, although Korajczyk and Sadka (2004) estimate that momentum strategies could still be profitable for funds with size up to \$4.5 to \$5.0 billion (which potentially includes many funds), fund managers might be unable to borrow shares.

Whilst these possibilities may all contribute to the delayed erosion of momentum profits, in our study we propose a different story. We notice that just after the publication of Jegadeesh and Titman (1993), an unprecedented stock market boom - driven mainly by hi-tech stocks - followed in the late 1990s. During that

[32] The sample period used in Jegadeesh and Titman (1993) ends in 1990.

period, investors pursued these stocks creating bubbles in a few sectors, and momentum profits appeared to increase since stocks in these sectors were winners and widened the gap between winners and losers. Even though momentum profits after 2000 are not significant after accounting for risk, raw momentum profits are still high in the first years after the burst of the bubble, since the previously inflated prices of hi-tech stocks were driven down and these stocks became losers. We explore this idea in the next section.

3.4. Momentum Premium During Boom and Bust

Despite the unprecedented boom in the late 1990s which increased momentum profits, the momentum strategy might not have been attractive to investors during this period, since higher returns were possible from investing in hi-tech stocks. For example, during the period from January 1995 to December 1999, the Dow Jones and the NASDAQ had average returns of 1.82% and 4.79% a month (standard errors 0.49% and 1.07% respectively), which means that investors could earn an average return of 2.98% a month during the period (standard error 0.92%) by buying hi-tech stocks and short-selling ‘old economy’ stocks. In this subsection we analyse the contribution of different sectors of the economy to the profit of the momentum strategy.

There have been some studies relating individual stock momentum to industry momentum, such as Moskowitz and Grinblatt (1999), who document a momentum effect in the industry component of stock returns, i.e. a momentum strategy appears to be significantly less profitable after they control for industry momentum, whilst industry momentum is highly profitable even after controlling for size, book-to-market and individual stock momentum effects. They claim that this industry momentum accounts for much of the individual stock momentum. However, subsequent work by Lee and Swaminathan (2000) and Grundy and Martin (2001) shows that industry effects do not explain momentum. We analyse this issue from a different perspective. Instead of calculating industry momentum, we calculate the momentum premium using stocks in all industries, and then break it down into industry-specific components. Thus we can calculate how much each sector of the economy contributes to the profits of the winner and loser portfolios.

Our definition of industries is similar to the one used by Moskowitz and Grinblatt (1999). We separate stocks into eleven sectors (Consumer Non-Durables, Consumer Durables, Manufacturing, Energy, Hi-Tech, Telecom, Shops, Health, Utilities, Financials and Others) according to their SIC codes. Every month, we create overlapping momentum portfolios as before, making a note of the industry from which each stock in the loser and winner portfolios come. Then we break down each month's loser and winner returns as the equally-weighted sum of the returns on the stocks coming from each industry. Since we are interested in the effect of the bubble on the contribution of hi-tech stocks to momentum profits, we focus on five-year subsamples such that one of the subsample isolates the bubble period. Table 3-8 reports these contributions in terms of average returns (Panel A) and as proportions of total momentum return (Panel B). The data from Panel B are plotted in Figure 3-2. The graph shows that some industries tend to contribute more to the momentum premium than others. Before the 1990s, most of the profits from the momentum strategies seem to come from traditional 'old economy' industries such as Manufacturing, Consumer Non-Durables, Shops and Financials, although Hi-Tech has a sizeable contribution as well. Also, a large part of the profits seems to come from stocks not classified in any of the groups, which we label 'Other' (meaning they do not belong to any of the other 10 industries).

Momentum profits are driven by different sectors in different periods. For example, the contribution of the Energy sector to momentum profits becomes large during the 1977-1982 period (around 24% of the profits of the momentum strategy, see Panel B of Table 3-8), which is quite different from other periods. A closer inspection reveals that this is due to the outperformance of the Energy sector that belongs to the winner portfolio as a result of the energy crisis of 1979. Another example is Financials from 1983 to 1994 when the financial sector contributes to almost 40% of the total momentum profits. Financials contribute to winners during the period from 1983 to 1988, whilst they contribute to winners as well as losers for the period from 1989 to 1994. Financials contribute to losers from the late 1980s to the early 1990s since many firms in the financial sector suffered losses from mortgage loans during the Savings and Loan crisis.

The subsample period that contains the period from 1995 and 1999 practically isolates the impact of the hi-tech bubble of the 1990s on momentum profits. The momentum strategy during this period yields the highest profit amongstst the subsample periods considered, and more than half of these profits come from high-tech and telecom stocks. Panel A of Table 3-8 reveals that, from the average raw monthly momentum return of 2.07% during the period, 1.17% comes from hi-tech and telecom stocks, which corresponds to a proportion of almost 60% of momentum profits. A similar analysis for the winner and loser portfolios (not reported) indicate that most of the profits come from the long side, i.e. from the winner portfolio: the average winner return over this period is 2.81% (of which

1.62% is from hi-tech and telecom stocks), whilst the average loser return is 0.74% (of which 0.45% comes from telecom stocks).

The following subsample, from 2001 to 2006, contains the aftermath of the hi-tech bubble. Although momentum profits during this period are lower than those during the bubble, a large proportion of the profits from the momentum strategy in this period still comes from hi-tech and telecom stocks, as shown by the last row in Panel B, where the combined contribution of the hi-tech and telecom stocks corresponds to nearly 50% of the raw momentum return. Unlike in the previous subsample, however, this profitability comes from the loser portfolio. From the average momentum return of 1.03% during the period, hi-tech and telecom stocks contribute with only 0.10% in the winner portfolio, but their negative returns in the loser portfolio of -0.38% boost momentum during this period (not reported).

One simple method to investigate how much the momentum profits are affected by the effect of the bubble in these sectors would be to separate the contribution of hi-tech and telecom stocks. Figure 3-3 shows average monthly momentum profits year-by-year, broken down by the contribution of two categories: hi-tech and telecom stocks, and all other stocks. First, it is clear that the contribution of the hi-tech and telecom group is atypical during the late 1990s. Second, a few years after the burst of the bubble (2000), momentum returns increase again (this is due to the negative returns of hi-tech and telecom stocks in the Loser portfolio). Finally, immediately after the bubble, i.e. 2001 and 2002, momentum profits are close to zero. Considering that winners and losers are calculated with past returns, the difference between winner and losers during the transition period is not as clear as that in other periods.

We analyse the impact of the hi-tech and telecom stocks on momentum profits in a few different ways. First, looking at the momentum profits without the contribution of hi-tech and telecom stocks (Figure 3-4), the overall pattern is that momentum returns have been going down with successively lower average returns since the middle 1990s. This is confirmed by looking at the cumulative profits of momentum strategies, which we plot in Figure 3-6 for various starting dates. If the profitability of momentum were constant, the graph should decline constantly with the starting date. Quite clearly, the profitability has decreased sharply in recent years. Second, we construct a momentum portfolio that does not allow hi-tech and telecom stocks at all. In this case, winners and losers are not necessarily the same as those when all firms in the stock market are considered. Figure 3-5 shows the average yearly returns of this restricted momentum strategy. The pattern seems to be very similar to the one in Figure 3-4, although it is clear that the profits tend to be larger in the late 1990s, indicating that other sectors were also affected by the bubble.

We test whether or not there are differences in momentum returns with and without Hi-tech and Telecom stocks. We consider three periods. The first one is from 1963 to 1994, which we name the pre-bubble sample. The second one is from 1995 to 1999 (the bubble period), and the last one is the post-bubble period from 2000-2006. We then calculate the difference between the returns on the momentum strategy which uses all stocks and the restricted momentum strategy that does not allow hi-tech and telecom stocks. Table 3-9 reports the results. During the pre-bubble period, the difference between the two strategies is negligible. The difference in average return between the two strategies is virtually zero, and not statistically significant. This indicates that before the bubble, including or not hi-tech and telecom stocks in the momentum strategy did not affect its return. During the bubble, the average monthly difference between the returns on the two strategies is around 0.6%, and statistically significant. During the post-bubble period, however, the situation seems to have come back to the pre-bubble state: during the 2000-2006 period the average difference between the returns on the unrestricted and restricted momentum strategies is around 0.11%, and not statistically significant. These results confirm that momentum is increased during the late 1990s due to the bubble in a few sectors.

3.5. Conclusion

The momentum premium from buying winner and selling loser stocks has been a defiant anomaly in the Finance literature, apparently persisting for years after being documented across a variety of markets. It has been almost 15 years since Jegadeesh and Titman's (1993) study, and although many theories have been presented, there has not been a definite explanation of the phenomenon. Not only has it proven very difficult to link momentum to a risk-based explanation, but the behavioural theories of investor under- and overreaction fail to explain why this near-arbitrage opportunity persisted for such a long period of time after being so well divulged.

Using a model with multiple structural breaks, we investigate the momentum premium over the period from 1927 to 2006. We find that the risk-adjusted momentum premia from several different momentum strategies are consistently significant and positive only during certain periods, such as the period from the 1940s to the mid-1960s and from the mid-1970s to the late 1990s. Not only has momentum disappeared during the recent period since the late 1990s, but we also find evidence that the post-1990s momentum premium is also extremely improbable compared to previous periods of weak momentum performance.

Although these results show that the momentum premium has disappeared, the question of why it persisted for so long after becoming well-known in the early 1990s still remains. We seek an answer from the bubble in hi-tech and telecommunication stocks during the late 1990s. Because of the bubble (high returns that persist for several years), hi-tech and telecom stocks entered the winner portfolio, increasing momentum profits. However, investors might not have been attracted to momentum strategies, since higher returns were possible by taking long and short positions in hi-tech firms and the so-called old economy stocks, respectively. After the bubble, momentum increasingly became the focus of exploitation by arbitrageurs in the market. Concordantly, the post-1990 period has also seen a large increase in the size and number of hedge funds - specialised agents with fewer limitations and constraints, who can undertake the kind of active trading required by momentum strategies. As far as these arbitrageurs are active in the market, the momentum profits we have observed over extended periods in the past are not likely to repeat again in the future.

Despite our interpretation that the momentum premium is not likely to be significant over long periods of time in the future, our analysis of the contribution of different industries to the profits of the momentum strategy suggests two situations in which significant momentum profits might be found. First, during periods of stock market bubbles in subsets of stocks (such as hi-tech stocks during the 1990s), because the pattern of successively increasing prices widens the gap between winners and losers. Second, when a group of stocks outperforms (underperforms) others with a series of good (bad) news over a period of time, increasing (decreasing) the share prices and making these stocks winners (losers) in the momentum strategy. Examples include energy stocks during the oil shock of 1979 and financial stocks in the 1980s. In this study we do not propose a risk-based explanation of the momentum premium, but suggest the possibility that momentum might in some cases arise due to irrational investor behaviour (in bubbles), whilst in other cases it can arise rationally (with long-lasting sequences of good or bad news for a group of stocks).

3.6. Appendix

3.6.1. Gibbs Sampling Scheme

We use an iterative Gibbs sampling scheme to obtain the augmented posterior density $\pi\left(\tilde{\mathbf{S}}_{\mathbf{T}},\tilde{\boldsymbol{\theta}},\mathbf{P}\,|\,\tilde{\mathbf{y}}_{\mathbf{T}}\right)$. Given initial values for $\boldsymbol{\theta}$ and $\mathbf{P}$, the parameters and state vector are simulated from the following conditional distributions:

1. $\tilde{\mathbf{S}}_{\mathbf{T}}\,|\,\tilde{\mathbf{y}}_{\mathbf{T}},\tilde{\boldsymbol{\theta}},\mathbf{P}$,
2. $\tilde{\boldsymbol{\theta}}\,|\,\tilde{\mathbf{S}}_{\mathbf{T}},\tilde{\mathbf{y}}_{\mathbf{T}},\mathbf{P}$ and
3. $\mathbf{P}\,|\,\tilde{\boldsymbol{\theta}},\tilde{\mathbf{S}}_{\mathbf{T}},\tilde{\mathbf{y}}_{\mathbf{T}}$.

3.6.1.1. Simulating $\tilde{\mathbf{S}}_{\mathbf{T}} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}$

The simulation of $\tilde{\mathbf{S}}_{\mathbf{T}}$ requires an elaborate recursive procedure. The joint distribution of $\tilde{\mathbf{S}}_{\mathbf{T}}$, $p\left(\tilde{\mathbf{S}}_{\mathbf{T}} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$, can be written as a recursion in reverse time order as

$$p\left(S_{T-1} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \times \ldots \times p\left(S_t \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{t+1}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \times \ldots \times p\left(S_1 \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{2}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \quad (3.8)$$

Given that $S_T = m+1$, all that is needed to generate each S_t is each mass function $p\left(S_t \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{t+1}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$. Then it is possible to generate, in turn,

- S_{T-1} from $p\left(S_{T-1} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$,
- S_{T-2} from $p\left(S_{T-2} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T-1}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$,
- $\vdots$
- S_1 from $p\left(S_1 \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{2}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$.

Chib (1996) shows that this mass function can be written as

$$p\left(S_t \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{t+1}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \propto p\left(S_t \mid \tilde{\mathbf{y}}_{\mathbf{t}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot p\left(S_{t+1} \mid S_t, \mathbf{P}\right) \quad (3.9)$$

where $p\left(S_{t+1} \mid S_t, \mathbf{P}\right)$ is the transition probability from the Markov chain. The quantity $p\left(S_t \mid \tilde{\mathbf{y}}_{\mathbf{t}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$ can be calculated using a set of backwards recursions. Given the initial value $p\left(S_1 = 1 \mid \tilde{\mathbf{y}}_1, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) = 1$, the values of $p\left(S_t \mid \tilde{\mathbf{y}}_{\mathbf{t}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right)$ from t = 2 to t = T are obtained from iterating on the two equations below:

$$\begin{aligned} p\left(S_t = k \mid \tilde{\mathbf{y}}_{\mathbf{t}-1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) &= \sum_{l=k-1}^{k} p\left(S_{t-1} = l \mid \tilde{\mathbf{y}}_{\mathbf{t}-1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot p_{lk,} \quad k = 1, \ldots, m+1 \\ p\left(S_t = k \mid \tilde{\mathbf{y}}_{\mathbf{t}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) &= \frac{p\left(S_t = k \mid \tilde{\mathbf{y}}_{\mathbf{t}-1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot f\left(y_t \mid \tilde{\mathbf{y}}_{\mathbf{t}-1}, \boldsymbol{\theta}_{\mathbf{k}}\right)}{\sum_{l=1}^{m+1} p\left(S_t = l \mid \tilde{\mathbf{y}}_{\mathbf{t}-1}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot f\left(y_t \mid \tilde{\mathbf{y}}_{\mathbf{t}-1}, \boldsymbol{\theta}_l\right)}, \quad k = 1, \ldots, m+1. \end{aligned} \tag{3.10}$$

With these values in hand, $\tilde{\mathbf{S}}_{\mathbf{T}}$ can be simulated using (3.9), where the normalising constant is easy to obtain, considering that each S_t can take only two values, conditioned on S_{t+1}.

The complete simulation of $\tilde{\mathbf{S}}_{\mathbf{T}}$ in Step 1 thus can be summarised by the following three steps.

i. Initialise $p\left(S_1 = 1 \mid \tilde{\mathbf{y}}_1, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) = 1$

ii. For $t = 2, \ldots, T$, iterate on the equations in (3.10)

iii. Set $S_T = m+1$. For $t = T-1, T-2, \ldots, 2$, compute

$$prob = \frac{p\left(S_t = k \mid \tilde{\mathbf{y}}_{\mathbf{t}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot p\left(S_{t+1} = k \mid S_t = k, \mathbf{P}\right)}{\sum_{l=k-1}^{k} p\left(S_t = l \mid \tilde{\mathbf{y}}_{\mathbf{t}}, \tilde{\boldsymbol{\theta}}, \mathbf{P}\right) \cdot p\left(S_{t+1} = k \mid S_t = l, \mathbf{P}\right)}$$

and generate S_t from a Binomial(1,*prob*).

3.6.1.2. Simulating $\tilde{\boldsymbol{\theta}} \mid \tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\mathbf{y}}_{\mathbf{T}}, \mathbf{P}$

The simulation of $\tilde{\boldsymbol{\theta}} \mid \tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\mathbf{y}}_{\mathbf{T}}, \mathbf{P}$ consists of first simulating $\boldsymbol{\beta}_{\mathbf{k}}$, $k = 1, \ldots, m+1$, conditioned on $\tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \sigma_k^2, \mathbf{P}$ and then simulating σ_k^2, $k = 1, \ldots, m+1$ conditioned on $\tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \boldsymbol{\beta}_{\mathbf{k}}, \mathbf{P}$. Recalling that the priors chosen are $\boldsymbol{\beta}_{\mathbf{k}} \sim N\left(\mathbf{b}_{\mathbf{0}}, \mathbf{B}_{\mathbf{0}}\right)$ and $\sigma_k^2 \sim IG\left(\frac{\upsilon_0}{2}, \frac{\delta_0}{2}\right)$, the posterior distributions are standard. Let $\tilde{\mathbf{y}}_{\mathbf{k}} = \{y_t : S_t = k\}$ and $\tilde{\mathbf{X}}_{\mathbf{k}} = \{X_t : S_t = k\}$ contain the values of *y* and *X* which belong to regime *k*. Then

$$\boldsymbol{\beta}_{\mathbf{k}} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \sigma_k^2, \mathbf{P} \sim N\left(\mathbf{b}_{\mathbf{1,k}}, \mathbf{B}_{\mathbf{1,k}}\right), \tag{3.11}$$

where

$$\mathbf{b}_{\mathbf{1,k}} = \left(\mathbf{B}_{\mathbf{0}}^{\mathbf{-1}} + \sigma_k^{-2}\left(\tilde{\mathbf{X}}_{\mathbf{k}}{}'\tilde{\mathbf{X}}_{\mathbf{k}}\right)\right) \cdot \left(\mathbf{B}_{\mathbf{0}}^{\mathbf{-1}}\mathbf{b}_{\mathbf{0}}{}' + \sigma_k^{-2}\tilde{\mathbf{X}}_{\mathbf{k}}{}'\tilde{\mathbf{y}}_{\mathbf{k}}\right), \tag{3.12}$$

and

$$\mathbf{B}_{\mathbf{1,k}} = \left(\mathbf{B}_{\mathbf{0}}^{\mathbf{-1}} + \sigma_k^{-2}\left(\tilde{\mathbf{X}}_{\mathbf{k}}{}'\tilde{\mathbf{X}}_{\mathbf{k}}\right)\right) \tag{3.13}$$

In turn, the posterior distribution of each σ_k^2 is

$$\sigma_k^2 \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \boldsymbol{\beta}_{\mathbf{k}}, \mathbf{P} \sim IG\left(\frac{\upsilon_1}{2}, \frac{\delta_1}{2}\right), \tag{3.14}$$

where $\upsilon_{1,k} = \upsilon_0 + \sum_{t=1}^{T} S_t = k$ and $\delta_{1,k} = \delta_0 + \left(\tilde{\mathbf{y}}_{\mathbf{k}} - \tilde{\mathbf{X}}_{\mathbf{k}}\boldsymbol{\beta}_{\mathbf{k}}{}'\right)'\left(\tilde{\mathbf{y}}_{\mathbf{k}} - \tilde{\mathbf{X}}_{\mathbf{k}}\boldsymbol{\beta}_{\mathbf{k}}{}'\right)$.

3.6.1.3. Simulating $\mathbf{P} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\boldsymbol{\theta}}$

Finally, to simulate the posterior of the transition probability matrix *P*, we only need to simulate transition probabilities $p_{kk}, \quad k = 1, \ldots, m$ (recall that $p_{m+1,m+1} = 1$). Given the prior $p_{kk} \sim Beta(c_0, d_0)$, the posterior distribution of each p_{kk} is given by

$$p_{kk} \sim Beta(c_0 + n_{kk}, d_0 + 1), \tag{3.15}$$

where n_{kk} is the number of times that the model stayed in regime *k*.

3.6.2. Calculating the Marginal Likelihood

We now show how the calculation of the marginal likelihood is done in the multiple change-point regression model following the method by Chib (1995). Chib shows that the marginal likelihood of model $\mathcal{M}_r$ may be written as

$$m(\tilde{\mathbf{y}}_{\mathbf{T}} \mid \mathcal{M}_r) = \frac{f(\tilde{\mathbf{y}}_{\mathbf{T}} \mid \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*, \mathcal{M}_r)\, \pi(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \mathcal{M}_r)}{\pi(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \mathcal{M}_r)}, \tag{3.16}$$

where $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$ is any point in the parameter space, $f(\tilde{\mathbf{y}}_{\mathbf{T}} \mid \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*, \mathcal{M}_r)$ is the likelihood function, and $\pi(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \mathcal{M}_r)$ and $\pi(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \mathcal{M}_r)$ are the prior and posterior ordinates evaluated at $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$. In the following we suppress the dependence on the model. It is convenient to calculate the log of (3.16),

$$\ln m\left(\tilde{\mathbf{y}}_{\mathbf{T}}\right) = \ln f\left(\tilde{\mathbf{y}}_{\mathbf{T}} \mid \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right) + \ln \pi\left(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right) - \ln \pi\left(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right) \qquad (3.17)$$

The method proposed by Chib (1995) consists in choosing a particular point $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$ and calculating the above quantity using the output from the Gibbs sampler. Although the choice of $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$ is in theory irrelevant, it is better to choose a high density point. Therefore we choose $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$ to be the posterior mean. Having selected $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$, the calculation of (3.16) requires the likelihood, the prior and posterior ordinates. From these, the prior can be obtained directly. We next describe the calculation of the likelihood and the posterior ordinate. Liu and Maheu (2008) provide a detailed explanation of these calculations for a very similar model.

3.6.2.1. Likelihood Function at $\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*$

The calculation of $f\left(\tilde{\mathbf{y}}_{\mathbf{T}} \mid \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right)$ is based on the decomposition

$$\ln f\left(\tilde{\mathbf{y}}_{\mathbf{T}} \mid \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right) = \sum_{t=1}^{T} \ln\left(f\left(y_t \mid \tilde{\mathbf{y}}_{\mathbf{t-1}}, \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right)\right), \qquad (3.18)$$

where

$$f\left(y_t \mid \tilde{\mathbf{y}}_{\mathbf{t-1}}, \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right) = \sum_{k=1}^{m} f\left(y_t \mid \tilde{\mathbf{y}}_{\mathbf{t-1}}, \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*, S_t = k\right) p\left(S_t = k \mid \tilde{\mathbf{y}}_{\mathbf{t-1}}, \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right). \quad (3.19)$$

The expression above is easily obtained as $f\left(y_t \mid \tilde{\mathbf{y}}_{\mathbf{t-1}}, \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*, S_t = k\right)$ is the conditional density of y_t given the regime $S_t = k$ and the term $p\left(S_t = k \mid \tilde{\mathbf{y}}_{\mathbf{t-1}}, \tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right)$ is obtained from the first equation in (3.10).

3.6.2.2. Posterior Ordinate $\pi\left(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right)$

To calculate $\ln \pi\left(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^*\right)$, we first note that

$$\pi\left(\tilde{\boldsymbol{\theta}}^*, \mathbf{P}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right) = \pi\left(\tilde{\boldsymbol{\beta}}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right) \pi\left(\tilde{\boldsymbol{\sigma}}^{\mathbf{2}*} \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\boldsymbol{\beta}}^*\right) \pi\left(\mathbf{P}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{\mathbf{2}*}\right), \quad (3.20)$$

where $\tilde{\boldsymbol{\beta}}^* = \left(\boldsymbol{\beta}_{\mathbf{1}}^* \cdots \boldsymbol{\beta}_{\mathbf{m+1}}^*\right)$ and $\tilde{\boldsymbol{\sigma}}^{\mathbf{2}*} = \left(\sigma_1^{2*} \cdots \sigma_{m+1}^{2*}\right)$. Each of these terms can be calculated through the output from the Gibbs sampler. The first term is

$$\pi\left(\tilde{\boldsymbol{\beta}}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right) = \int \pi\left(\tilde{\boldsymbol{\beta}}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\boldsymbol{\sigma}}^{\mathbf{2}}\right) p\left(\tilde{\mathbf{S}}_{\mathbf{T}}, \tilde{\boldsymbol{\sigma}}^{\mathbf{2}} \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right) d\,\mathbf{S} d\tilde{\boldsymbol{\sigma}}^{\mathbf{2}}. \quad (3.21)$$

If the number of MCMC draws is *G* and $\left\{\tilde{\mathbf{S}}^{(\mathbf{g})}, \tilde{\boldsymbol{\sigma}}^{\mathbf{2}(\mathbf{g})}\right\}_{g=1}^{G}$ are the MCMC draws of $\tilde{\mathbf{S}}$ and $\tilde{\boldsymbol{\sigma}}^{\mathbf{2}}$, then the above integral can be estimated as

$$\hat{\pi}\left(\tilde{\boldsymbol{\beta}}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}\right) = G^{-1} \sum_{g=1}^{G} \pi\left(\tilde{\boldsymbol{\beta}}^* \mid \tilde{\mathbf{y}}_{\mathbf{T}}, \mathbf{S}^{(\mathbf{g})}, \tilde{\boldsymbol{\sigma}}^{\mathbf{2}(\mathbf{g})}\right), \quad (3.22)$$

where $\pi\left(\tilde{\boldsymbol{\beta}}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \mathbf{S}^{(\mathbf{g})}, \tilde{\boldsymbol{\sigma}}^{2(\mathbf{g})}\right) = \prod_{k=1}^{m+1} \pi\left(\boldsymbol{\beta}_\mathbf{k}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \mathbf{S}^{(\mathbf{g})}, \tilde{\boldsymbol{\sigma}}^{2(\mathbf{g})}\right)$ and each density is given by (3.11).

The second term in (3.20) is $\pi\left(\tilde{\boldsymbol{\sigma}}^{2*} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*\right)$, which can be written as

$$\pi\left(\tilde{\boldsymbol{\sigma}}^{2*} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*\right) = \int \pi\left(\tilde{\boldsymbol{\sigma}}^{2*} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\mathbf{S}}\right) p\left(\tilde{\mathbf{S}} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*\right) d\,\mathbf{S} \tag{3.23}$$

To estimate the above integral we require draws from the distribution of $\tilde{\mathbf{S}} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*$ which are not readily available. To obtain those, we run an additional reduced Gibbs sampler conditional on $\tilde{\boldsymbol{\beta}}^*$. This is achieved by fixing $\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}^*$ and sampling G_1 draws from $\tilde{\boldsymbol{\sigma}}^2$, $\tilde{\mathbf{S}}$ and $\mathbf{P}$ as before. Once this is done, we estimate (3.23) as

$$\hat{\pi}\left(\tilde{\boldsymbol{\sigma}}^{2*} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*\right) = \sum_{g=1}^{G_1} \pi\left(\tilde{\boldsymbol{\sigma}}^{2*} \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\mathbf{S}}^{(g)}\right). \tag{3.24}$$

Finally, the last term of (3.20) is $\pi\left(\mathbf{P}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^2\right)$, which can be expressed as

$$\pi\left(\mathbf{P}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right) = \prod_{k=1}^{m} \pi\left(p_{kk}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right), \tag{3.25}$$

where

$$\pi\left(p_{kk}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right) = \int \pi\left(p_{kk}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\mathbf{S}}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right) p\left(\mathbf{S} \mid \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}, \tilde{\mathbf{y}}_\mathbf{T}\right) d\,\mathbf{S}. \tag{3.26}$$

To estimate this quantity, we proceed as before and sample additional values of $\tilde{\mathbf{S}}$ from $p\left(\mathbf{S} \mid \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}, \tilde{\mathbf{y}}_\mathbf{T}\right)$, that is we maintain $\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\sigma}}^2$ fixed at $\tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}$ and sample $\tilde{\mathbf{S}}$ as before. Then the estimate of (3.26) is

$$\hat{\pi}\left(p_{kk}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right) = \sum_{g=1}^{G_1} \pi\left(p_{kk}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\mathbf{S}}^{(\mathbf{g})}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right), \tag{3.27}$$

and the estimate of (3.25) is

$$\hat{\pi}\left(\mathbf{P}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right) = \prod_{k=1}^{m} \sum_{g=1}^{G_1} \pi\left(p_{kk}^* \mid \tilde{\mathbf{y}}_\mathbf{T}, \tilde{\mathbf{S}}^{(\mathbf{g})}, \tilde{\boldsymbol{\beta}}^*, \tilde{\boldsymbol{\sigma}}^{2*}\right). \tag{3.28}$$

Table 3-1 Summary statistics of portfolios sorted on past returns

This table reports summary statistics for several momentum portfolios which are formed via the following procedure. At the end of each month t, all stocks traded in the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past five month returns from t-5 to t-1 (for some momentum strategies, only NYSE stocks are considered). The momentum strategy consists in holding a long position in the Winner (highest decile) portfolio whilst holding a short position on the Loser (lowest decile) portfolio. The positions are held for a subsequent period, which is 6, 3 or 1 month depending on the momentum strategy. Standard errors are given inside brackets where appropriate.

Panel A. January 1927 - December 2006

	Portfolio						
	$WML_{CRSP,6x6}$	$WML_{NYSE10,6x6}$	$WML_{NYSE20,6x6}$	$WML_{NYSE30,6x6}$	$WML_{CRSP,6x1}$	$WML_{CRSP,6x3}$	Market
Avg. return	0.938	0.787	0.760	0.733	0.996	0.957	0.647
	(0.157)	(0.144)	(0.145)	(0.148)	(0.192)	(0.177)	(0.176)
Skewness	-1.862	-2.586	-2.533	-2.727	-1.680	-1.721	0.222
Kurtosis	20.640	22.644	22.792	25.150	19.342	20.378	10.868
Alpha (CAPM)	1.030	0.896	0.867	0.847	1.168	1.100	
	(0.156)	(0.142)	(0.143)	(0.146)	(0.188)	(0.174)	
Alpha (FF)	1.152	1.005	0.980	0.963	1.304	1.237	
	(0.151)	(0.138)	(0.139)	(0.141)	(0.182)	(0.168)	

Panel B. January 1927 - June 1962

	Portfolio						
	$WML_{CRSP,6x6}$	$WML_{NYSE10,6x6}$	$WML_{NYSE20,6x6}$	$WML_{NYSE30,6x6}$	$WML_{CRSP,6x1}$	$WML_{CRSP,6x3}$	Market
Avg. return	0.675	0.673	0.669	0.666	0.636	0.632	0.849
	(0.250)	(0.250)	(0.252)	(0.258)	(0.308)	(0.281)	(0.311)
Skewness	-3.081	-3.081	-3.070	-3.315	-2.845	-3.014	0.422
Kurtosis	23.925	23.928	24.398	26.612	21.549	23.641	10.463
Alpha (CAPM)	0.872	0.870	0.866	0.879	0.970	0.918	
	(0.242)	(0.242)	(0.244)	(0.249)	(0.285)	(0.264)	
Alpha (FF)	0.906	0.904	0.900	0.913	1.004	0.958	
	(0.231)	(0.231)	(0.232)	(0.236)	(0.276)	(0.252)	

Table 3-1 (continued)

Panel C. July 1962 – December 2006

	Portfolio						
	$WML_{CRSP,6x6}$	$WML_{NYSE10,6x6}$	$WML_{NYSE20,6x6}$	$WML_{NYSE30,6x6}$	$WML_{CRSP,6x1}$	$WML_{CRSP,6x3}$	Market
Avg. return	1.159	0.883	0.836	0.789	1.299	1.230	0.477
	(0.198)	(0.161)	(0.163)	(0.164)	(0.241)	(0.223)	(0.192)
Skewness	-0.218	-0.946	-0.842	-0.761	-0.045	0.023	-0.510
Kurtosis	14.007	6.838	6.405	6.002	13.985	13.689	5.135
Alpha (CAPM)	1.146	0.906	0.857	0.807	1.311	1.232	
	(0.199)	(0.162)	(0.164)	(0.165)	(0.243)	(0.225)	
Alpha (FF)	1.301	1.015	0.953	0.900	1.467	1.379	
	(0.199)	(0.165)	(0.167)	(0.168)	(0.242)	(0.224)	

Table 3-2 Momentum returns in 5-year non-overlapping subsamples

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past five month returns from *t*-5 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy ($WML_{CRSP\ 6x6}$) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. We report, for each subsample, the coefficients of the CAPM and the Fama-French three-factor model estimated with OLS. Two (one) * indicates significance at 5% (10%) significance level. The samples span the period January 1927 - December 2006. All subsamples have equal size of 60 months, starting in January and ending in December, except for subsample 16 which ends in December 2006 and has 43 observations.

		Market return	**Momentum return**	**CAPM**				**Fama-French Model**							
Sample	Period			$\hat{\alpha}_{CAPM}$		$\hat{\beta}_{MKT}$		$\hat{\alpha}_{FF}$		$\hat{\beta}_{MKT}$		$\hat{\beta}_{SMB}$		$\hat{\beta}_{HML}$	
1	1927-1931	-0.673	1.653	1.548	*	-0.156	**	1.001		-0.042		-0.340	**	-0.747	*
2	1932-1936	2.303	0.236	0.883		-0.281	*	0.645		0.010		0.101		-0.523	*
3	1937-1941	-0.262	-0.411	-0.544		-0.511	*	-0.443		-0.229		-0.529	**	-0.276	
4	1942-1946	1.494	0.472	0.234		0.159	**	0.267		0.188		-0.157		0.061	
5	1947-1951	1.140	0.756	0.512		0.214	*	-0.003		0.280	*	-1.175	*	0.333	*
6	1952-1956	1.330	1.130	0.990	*	0.106		1.030	*	0.090		-0.387	*	-0.444	*
7	1957-1961	0.900	0.822	0.867	*	-0.050		0.950	*	-0.075		-0.190		-0.968	*
8	1962-1966	0.226	1.131	1.143	*	-0.054		1.185	**	-0.188		0.614	*	-0.126	
9	1967-1971	0.364	0.944	0.953		-0.026		1.186		0.118		-0.394	**	-0.244	
10	1972-1976	-0.052	0.800	0.780		-0.386	*	0.514		-0.190	*	-0.694	*	0.092	
11	1977-1981	0.193	1.470	1.379	*	0.470	*	0.927		0.230	**	0.488	*	-0.302	**
12	1982-1986	0.806	1.227	1.241	*	-0.017		1.484	**	-0.099		0.216		-0.212	
13	1987-1991	0.701	1.201	1.104	*	0.138	*	0.942		0.066		-0.244	**	-0.492	*
14	1992-1996	0.835	0.776	0.481		0.354	*	0.565		0.296	**	0.300	**	-0.087	
15	1997-2001	0.509	2.208	2.151	**	0.111		2.444	**	-0.253		0.431		-0.457	
16	2002-2006	0.536	0.595	0.819	**	-0.419	*	0.141		-0.428	*	0.453		0.636	*

Table 3-3 Log-marginal likelihoods and differences in log-marginal likelihoods of models with different numbers of change-points

We estimate a multiple change-point regression model based on the methodology of Chib (1998). The left-hand variable is a momentum portfolio based on six-month formation and holding periods using all stocks in the CRSP database (which includes NYSE, AMEX and NASDAQ), excluding stocks priced at less than $5 at the beginning of the formation period and stocks in the lowest (NYSE) market cap decile. This portfolio is regressed on a constant, the excess market return and the two Fama-French factors SMB and HML. We estimate models with different numbers of change-points and calculate their marginal likelihoods, which are reported in Panel A. Panel B reports the differences in marginal likelihood across models with different numbers of breakpoints. The Bayes factor for comparing two models in given by the exponential of the values in Panel B.

Panel A. Log-Marginal Likelihood of models with different number of change-points

	Number of change-points							
	0	1	2	3	4	5	6	7
Log-Marginal Likelihood	-3845.78	-3759.69	-2971.61	-3022.92	-2837.96	-2771.16	-2622.25	-2875.26

Panel B. Differences in the marginal likelihoods across models with different number of change-points

	0	1	2	3	4	5	6	7
0	-	-87.026	-875.105	-823.790	-1008.753	-1075.551	-1224.461	-971.447
1	87.026	-	-788.080	-736.764	-921.727	-988.526	-1137.435	-884.422
2	875.105	788.080	-	51.315	-133.647	-200.446	-349.355	-96.342
3	823.790	736.764	-51.315	-	-184.963	-251.761	-400.671	-147.657
4	1008.753	921.727	133.647	184.963	-	-66.798	-215.708	37.305
5	1075.551	988.526	200.446	251.761	66.798	-	-148.909	104.104
6	**1224.461**	**1137.435**	**349.355**	**400.671**	**215.708**	**148.909**	**-**	**253.013**
7	971.447	884.422	96.342	147.657	-37.305	-104.104	-253.013	-

Table 3-4 Parameter estimates - $WML_{CRSP\ 6x6}$ portfolio

This table reports posterior medians and 0.95 probability intervals for the multiple change-point regression model where the left-hand variable is a momentum portfolio based on six-month formation and holding periods using all stocks in the CRSP database (which includes NYSE, AMEX and NASDAQ), excluding stocks priced at less than $5 at the beginning of the formation period and stocks in the lowest (NYSE) market cap decile. This portfolio is regressed on a constant, the excess market return and the two Fama-French factors SMB and HML. The model is estimated for with seven change-points, which is the optimal number of change-points based on the Bayes factors calculated from Table 3-3.The results are based on 10,000 iterations MCMC iterations.

	constant	**MKT**	**SMB**	**HML**	**σ^2**
Regime 1	0.965	-0.123	-0.303	**-0.877**	24.198
Jan-1927 : Sep-1932	(-0.177 , 2.125)	(-0.266 , 0.026)	(-0.608 , 0.002)	(-1.107 , -0.643)	(17.491 , 34.799)
Regime 2	0.435	-0.012	-0.265	-0.028	53.901
Oct-1932 : Apr-1940	(-1.035 , 1.924)	(-0.253 , 0.227)	(-0.586 , 0.051)	(-0.381 , 0.324)	(40.470 , 74.085)
Regime 3	**0.733**	**0.096**	**-0.406**	0.015	6.488
May-1940 : Mar-1965	(0.419 , 1.047)	(0.013 , 0.181)	(-0.572 , -0.237)	(-0.118 , 0.152)	(5.509 , 7.670)
Regime 4	0.870	0.068	**0.647**	-0.453	12.699
Apr-1965 : Aug-1969	(-0.238 , 1.925)	(-0.265 , 0.411)	(0.239 , 1.042)	(-0.963 , 0.046)	(8.620 , 19.991)
Regime 5	0.264	**-0.256**	**-0.824**	-0.173	14.489
Sep-1969 : Mar-1975	(-0.682 , 1.230)	(-0.467 , -0.045)	(-1.101 , -0.553)	(-0.481 , 0.151)	(10.287 , 21.057)
Regime 6	**1.129**	**0.125**	0.131	**-0.222**	7.820
Apr-1975 : Sep-1999	(0.763 , 1.479)	(0.038 , 0.211)	(-0.010 , 0.275)	(-0.379 , -0.059)	(6.470 , 9.431)
Regime 7	0.874	**-0.384**	**0.755**	-0.125	46.117
Oct-1999 : Dec-2006	(-0.805 , 2.422)	(-0.764 , -0.015)	(0.314 , 1.230)	(-0.642 , 0.436)	(34.689 , 62.998)

Table 3-5 Log-marginal likelihoods of multiple change-point models for several momentum portfolios

For each of the momentum portfolios below, we estimated a multiple change-point regression model based on the methodology of Chib (1998). Each of the momentum portfolios is regressed on a constant, the excess market return and the two Fama-French factors SMB and HML. We estimate models with different numbers of change-points and calculate their marginal likelihoods. Bold numbers indicate the best model according to the Bayes factors. For some portfolios, a '-' indicates that the model with the corresponding number of change-points collapsed to a model with fewer change-points.

	Number of change-points								
Momentum portfolio	0	1	2	3	4	5	6	7	8
$WML_{NYSE10,6x6}$	-3614.085	-3515.769	-3506.008	-3340.683	-2937.795	**-2731.933**	-3363.122	-3068.712	-3053.703
$WML_{NYSE20,6x6}$	-3634.427	-3540.673	-3466.853	-3223.696	-2910.509	-2879.645	**-2828.935**	-3120.405	-3106.869
$WML_{NYSE30,6x6}$	-3671.450	-3563.327	-3444.123	-3225.020	-2854.688	**-2787.702**	-2827.391	-3228.131	-3129.602
$WML_{CRSP,6x1}$	-4390.320	-4263.989	-3272.004	-3384.179	-3205.751	-3129.092	**-2862.161**	-	-
$WML_{CRSP,6x3}$	-4143.443	-4017.029	-3116.757	-3241.441	**-3035.281**	-3098.019	-3051.802	-	-

Table 3-6 Estimation of the multiple change-point model for various momentum portfolios

This table reports posterior medians and 0.95 probability intervals for the multiple change-point regression model where the left-hand variable is one of several momentum portfolios. Each portfolio is regressed on a constant, the excess market return and the two Fama-French factors SMB and HML. For each momentum portfolio, the model is estimated for the optimal number of change-points based on the Bayes factors calculated from Table 3-5.The results are based on 10,000 iterations MCMC iterations.

Panel A. $WML_{NYSE10,6x6}$

	constant	MKT	SMB	HML	σ^2
Regime 1 Jan-1927 : Apr-1940	0.541 (-0.525 , 1.615)	-0.103 (-0.253 , 0.051)	-0.075 (-0.303 , 0.150)	**-0.460** (-0.687 , -0.234)	49.210 (39.366 , 61.811)
Regime 2 May-1940 : May-1970	**0.876** (0.586 , 1.176)	**0.092** (0.011 , 0.172)	-0.107 (-0.246 , 0.036)	-0.098 (-0.224 , 0.029)	7.770 (6.727 , 9.060)
Regime 3 Jun-1970 : Apr-1975	0.460 (-0.513 , 1.438)	**-0.277** (-0.494 , -0.067)	**-0.918** (-1.203 , -0.640)	-0.108 (-0.445 , 0.225)	13.522 (9.506 , 20.042)
Regime 4 May -1975 : Nov-2000	**0.867** (0.496 , 1.232)	0.050 (-0.040 , 0.155)	**0.134** (0.012 , 0.249)	**-0.243** (-0.388 , -0.049)	8.772 (6.796 , 10.478)
Regime 5 Dec-2000 : Dec-2006	0.295 (-0.641 , 1.553)	**-0.317** (-0.548 , -0.092)	0.156 (-0.145 , 0.434)	0.579 (-0.264 , 0.906)	12.587 (8.916 , 23.271)

Table 3-6 (continued)

Panel B. $WML_{NYSE20,6x6}$

	constant	**MKT**	**SMB**	**HML**	**σ^2**
Regime 1 Jan-1927 : May-1940	0.537 (-0.528 , 1.609)	-0.087 (-0.240 , 0.066)	-0.047 (-0.280 , 0.178)	**-0.513** (-0.746 , -0.284)	48.682 (39.213 , 61.196)
Regime 2 Jun-1940 : Jul-1957	**0.514** (0.042 , 0.916)	**0.170** (0.059 , 0.300)	**-0.454** (-0.715 , -0.251)	0.126 (-0.058 , 0.330)	6.097 (4.964 , 7.615)
Regime 3 Aug-1957 : May-1970	**1.173** (0.309 , 1.664)	-0.088 (-0.264 , 0.100)	0.182 (-0.078 , 0.641)	**-0.450** (-0.713 , -0.142)	8.193 (6.447 , 12.308)
Regime 4 Jun-1970 : Mar-1975	0.392 (-0.569 , 1.339)	**-0.288** (-0.510 , -0.080)	**-0.836** (-1.124 , -0.549)	**-0.069** (-0.395 , 0.271)	13.953 (9.896 , 20.766)
Regime 5 Apr-1975 : Nov-2000	**0.822** (0.462 , 1.176)	0.052 (-0.037 , 0.147)	**0.166** (0.047 , 0.280)	**-0.260** (-0.399 , -0.107)	9.188 (7.568 , 10.846)
Regime 6 Dec-2000 : Dec-2006	0.159 (-0.753 , 1.178)	**-0.309** (-0.541 , -0.078)	0.165 (-0.137 , 0.467)	0.632 (-0.023 , 0.956)	12.919 (9.347 , 20.453)

Panel C. $WML_{NYSE30,6x6}$

	constant	**MKT**	**SMB**	**HML**	**σ^2**
Regime 1	0.502 (-0.609 , 1.582)	-0.108 (-0.267 , 0.046)	-0.020 (-0.257 , 0.207)	**-0.548** (-0.782 , -0.310)	50.931 (41.089 , 64.431)
Regime 2	**0.867** (0.568 , 1.171)	**0.098** (0.019 , 0.179)	-0.045 (-0.182 , 0.097)	-0.112 (-0.237 , 0.012)	7.658 (6.616 , 8.907)
Regime 3	0.375 (-0.561 , 1.325)	**-0.300** (-0.515 , -0.092)	**-0.784** (-1.065 , -0.506)	-0.011 (-0.344 , 0.309)	13.181 (9.291 , 19.603)
Regime 4	**0.774** (0.406 , 1.137)	0.047 (-0.044 , 0.136)	**0.181** (0.063 , 0.296)	**-0.280** (-0.420 , -0.144)	9.652 (8.278 , 11.367)
Regime 5	0.061 (-0.870 , 0.980)	**-0.339** (-0.580 , -0.101)	0.211 (-0.096 , 0.526)	**0.662** (0.325 , 0.988)	13.579 (9.857 , 19.453)

Table 3-6 (continued)

Panel D. $\mathbf{WML_{CRSP\ 6x1}}$

	constant	**MKT**	**SMB**	**HML**	**σ^2**
Regime 1 Jan-1927 : Jan-1940	0.399 (-0.842 , 1.682)	**-0.263** (-0.442 , -0.085)	-0.064 (-0.335 , 0.210)	**-0.530** (-0.797 , -0.263)	67.674 (54.416 , 85.218)
Regime 2 Feb-1940 : Feb-1965	**0.871** (0.493 , 1.244)	-0.015 (-0.118 , 0.086)	**-0.366** (-0.570 , -0.165)	0.140 (-0.024 , 0.304)	9.908 (8.470 , 11.744)
Regime 3 Mar-1965 : Aug-1969	1.137 (-0.049 , 2.302)	0.070 (-0.298 , 0.440)	**0.671** (0.247 , 1.101)	-0.156 (-0.710 , 0.399)	15.534 (10.379 , 24.920)
Regime 4 Sep-1969 : Feb-1975	0.383 (-0.824 , 1.567)	-0.255 (-0.522 , 0.012)	**-0.951** (-1.302 , -0.603)	-0.216 (-0.610 , 0.185)	23.388 (16.761 , 33.794)
Regime 5 Mar-1975 : Oct-1999	**1.181** (0.752 , 1.614)	0.086 (-0.023 , 0.196)	0.123 (-0.035 , 0.287)	**-0.274** (-0.458 , -0.081)	12.684 (10.741 , 15.043)
Regime 6 Nov-1999 : Dec-2006	0.630 (-1.123 , 2.437)	**-0.667** (-1.130 , -0.214)	**1.289** (0.758 , 1.775)	**0.052** (-0.580 , 0.631)	62.601 (46.897 , 86.139)

Panel E. $\mathbf{WML_{CRSP\ 6x3}}$

	constant	**MKT**	**SMB**	**HML**	**σ^2**
Regime 1 Jan-1927 : May-1940	0.564 (-0.491 , 1.649)	-0.101 (-0.251 , 0.048)	-0.070 (-0.299 , 0.155)	**-0.466** (-0.693 , -0.233)	48.713 (39.303 , 61.220)
Regime 2 Jun-1940 : Nov-1957	**0.514** (0.089 , 0.889)	**0.159** (0.057 , 0.268)	**-0.465** (-0.688 , -0.273)	0.147 (-0.022 , 0.338)	5.899 (4.852 , 7.343)
Regime 3 Dec-1957 : Oct-1999	**1.266** (0.953 , 1.575)	-0.020 (-0.100 , 0.057)	-0.101 (-0.216 , 0.016)	**-0.344** (-0.473 , -0.213)	11.626 (10.214 , 13.422)
Regime 4 Nov-1999 : Dec-2006	0.583 (-0.926 , 2.108)	**-0.411** (-0.807 , -0.020)	**0.872** (0.458 , 1.277)	-0.007 (-0.507 , 0.502)	46.587 (34.926 , 64.171)

Table 3-7 Bootstrap estimation of the probability of observing the momentum premium over the period from January 2001 to December 2006

We conduct a bootstrap experiment to estimate the probability of observing a momentum premium lower than the momentum premium in the 72-month period from January 2001 to December 2006. Based on the ex-post distribution of momentum returns over different time periods described in the "sampling period" column, we draw 100,000 random samples of size seventy-two months (with replacement). We report the probability of observing a momentum premium smaller than the one in the period from January 2001 to December 2006 using two methods. The first method consists in taking the average of each bootstrap sample and estimating the distribution of the mean momentum return over a seventy-two month period. The second method consists of calculating the distribution of alpha from the Fama-French model based on the bootstrap samples. We estimate the probability of the momentum premium over a seventy-two month period being smaller than the momentum premium from January 2001 to December 2006 as the proportion of samples with average return (or alpha) smaller than the average momentum return (or alpha) from January 2001 to December 2006.

Sampling period	**Probability of observing a momentum premium inferior to the momentum premium from January 2001 to December 2006**	
	Average Return	Alpha (Fama-French)
January 1927 to December 2000	0.208	0.005
January 1927 to April 1940	0.499	0.139
May 1940 to March 1965	0.186	0.004
April 1965 to March 1975	0.193	0.009
April 1975 to December 2000	0.066	0.000

Table 3-8 Contribution of different sectors of the economy to momentum returns

We decompose the monthly momentum return into the sum of the returns of the stocks from each of 11 industries such that each month, the return on the momentum strategy is the sum of the returns on the stocks from each industry. Panel A reports averages of the contribution of each industry for several 5-year subsamples. Panel B reports the proportion of momentum returns due to each sector. The column labelled 'Total' displays the sum of the values in each row.

Panel A. Return contribution of several sectors to momentum profits in terms of equally weighted returns

	Consumer Non-Durables	Consumer Durables	Manuf.	Energy	Hi-Tech	Telecom	Shops	Health	Utilities	Financials	Others	Total
1959-1964	0.23%	0.01%	0.19%	0.05%	0.16%	0.02%	0.07%	0.01%	0.03%	0.02%	0.27%	1.07%
1965-1970	0.13%	0.11%	0.22%	0.02%	0.12%	-0.02%	0.16%	-0.03%	0.08%	0.07%	0.22%	1.08%
1971-1976	0.09%	0.08%	0.06%	-0.02%	-0.13%	0.04%	0.30%	0.03%	-0.04%	0.10%	0.30%	0.82%
1977-1982	0.08%	0.01%	0.09%	0.38%	0.27%	0.04%	0.14%	0.08%	-0.01%	0.12%	0.44%	1.63%
1983-1988	0.04%	-0.01%	0.02%	-0.06%	0.09%	0.04%	0.13%	0.05%	0.01%	0.28%	0.11%	0.70%
1989-1994	0.06%	0.03%	0.06%	0.00%	0.06%	0.05%	0.15%	0.14%	0.00%	0.42%	0.11%	1.08%
1995-1999	0.04%	0.03%	0.12%	0.07%	0.98%	0.19%	0.13%	0.17%	-0.01%	0.15%	0.19%	2.07%
2000-2006	-0.02%	-0.02%	0.00%	0.11%	0.29%	0.19%	0.00%	0.27%	0.03%	0.03%	0.14%	1.03%

Panel B. Return contribution of several sectors to momentum profits in terms of proportion of returns

	Consumer Non-Durables	Consumer Durables	Manuf.	Energy	Hi-Tech	Telecom	Shops	Health	Utilities	Financials	Others	Total
1959-1964	21.17%	1.26%	17.69%	5.13%	14.99%	1.65%	6.74%	1.27%	2.95%	1.43%	25.74%	100.00%
1965-1970	12.35%	10.56%	20.20%	1.46%	11.18%	-1.87%	14.89%	-3.13%	7.62%	6.45%	20.29%	100.00%
1971-1976	10.89%	10.01%	7.44%	-2.47%	-16.14%	4.62%	36.91%	3.72%	-4.74%	12.62%	37.14%	100.00%
1977-1982	5.12%	0.47%	5.27%	23.45%	16.80%	2.36%	8.29%	4.69%	-0.91%	7.25%	27.20%	100.00%
1983-1988	6.33%	-0.89%	2.87%	-9.00%	13.15%	5.87%	18.61%	6.56%	0.77%	39.59%	16.13%	100.00%
1989-1994	5.79%	3.15%	5.09%	0.27%	5.34%	4.81%	14.06%	12.94%	0.15%	38.48%	9.94%	100.00%
1995-1999	1.76%	1.61%	6.03%	3.38%	47.51%	8.99%	6.40%	8.25%	-0.34%	7.26%	9.15%	100.00%
2000-2006	-2.16%	-1.92%	0.18%	10.64%	28.22%	18.95%	0.46%	25.85%	2.58%	3.18%	14.01%	100.00%

Table 3-9 The Effect of the Hi-Tech Bubble on the profitability of momentum

We calculate two momentum strategies for the period from January 1963 to December 2006. One of them is the unrestricted momentum strategy using all stocks from the CRSP universe and the other does not allow stocks from hi-tech or telecom sectors to be included. Formation and holding periods are 6 months. As before we construct overlapping portfolios and skip one month between formation and holding periods.

Period	Average monthly momentum returns			
	All stocks (A)	No Hi-tech and Telecom stocks allowed (B)	Difference (A – B)	t-stat
Jan 1963 to Dec 1994 (Pre-bubble)	1.064%	1.063%	0.001%	0.029
Jan 1995 to Dec 1999 (During bubble)	2.072%	1.483%	0.590%	2.643
Jan 2000 to Dec 2006 (Post-bubble)	0.936%	0.823%	0.113%	0.415

Figure 3-1 Posterior probabilities of regimes for the $WML_{CRSP,6x6}$ strategy

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past 5-month returns from *t*-6 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy (WML) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. The figure shows the estimated probability of structural breaks in the model, based on 10,000 iterations of the MCMC scheme.

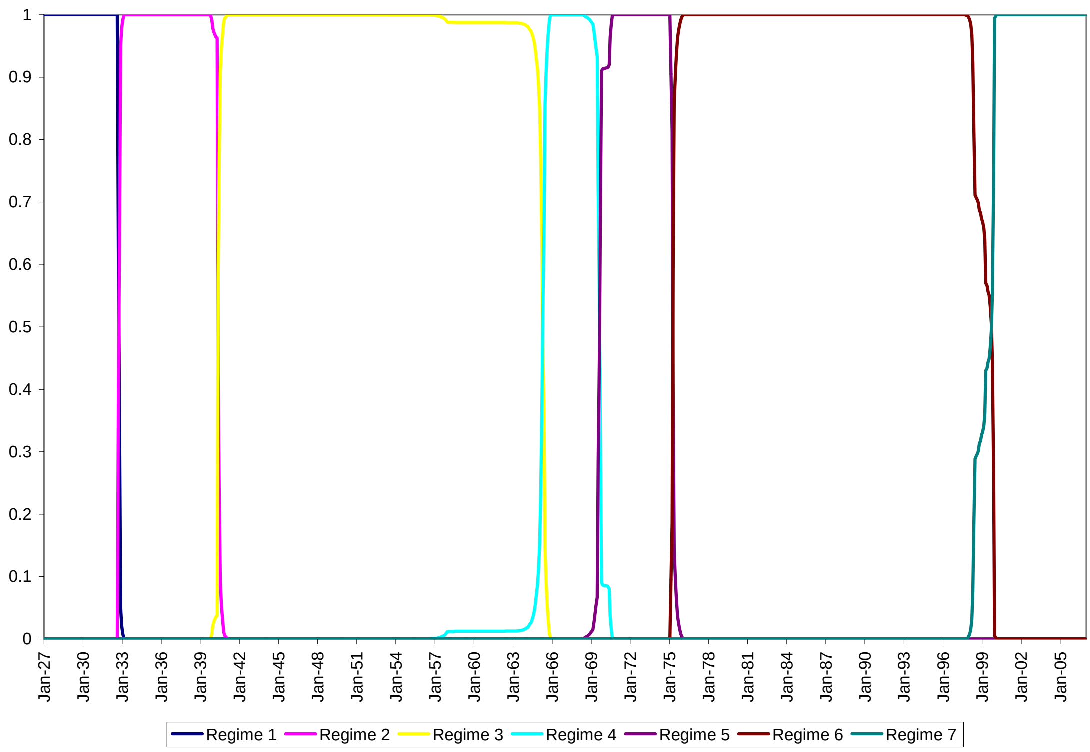


Figure 3-2 Contribution of different sectors of the economy to momentum returns

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past 5-month returns from *t*-6 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy (WML) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. We decompose momentum monthly returns into their sector-specific components, in such a way that, over a period of time, the return to any of these portfolios can be calculated as the sum of the sector-specific components. The graph displays the proportion of average annual momentum returns which correspond to each of the industry sectors over several different subsamples.

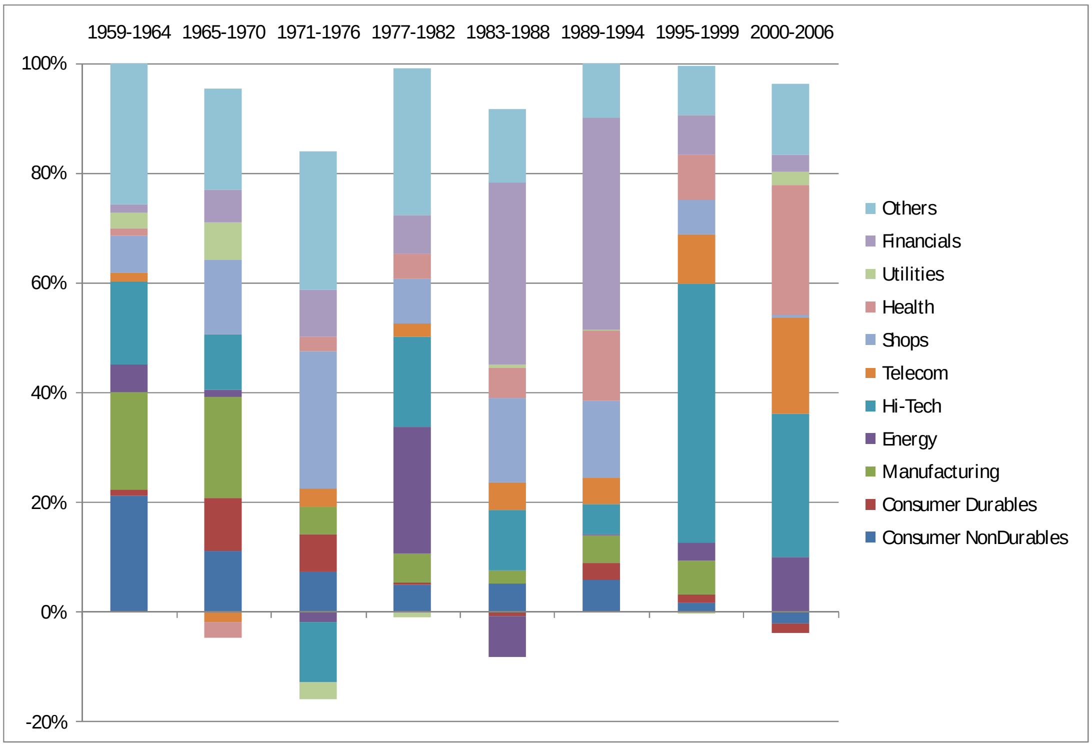


Figure 3-3 Contribution of hi-tech, telecom and all other stocks to momentum returns

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past 5-month returns from *t*-6 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy (WML) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. We decompose momentum monthly returns based on whether the stocks in the portfolios belong to one of two groups. The first group contains hi-tech and telecom stocks, and the second one contains stocks from all other industries. The decomposition is done in such a way that, over a period of time, the return to any of these portfolios can be calculated as the sum of the two components. The graph displays the average annual sector-specific returns for the momentum portfolio.

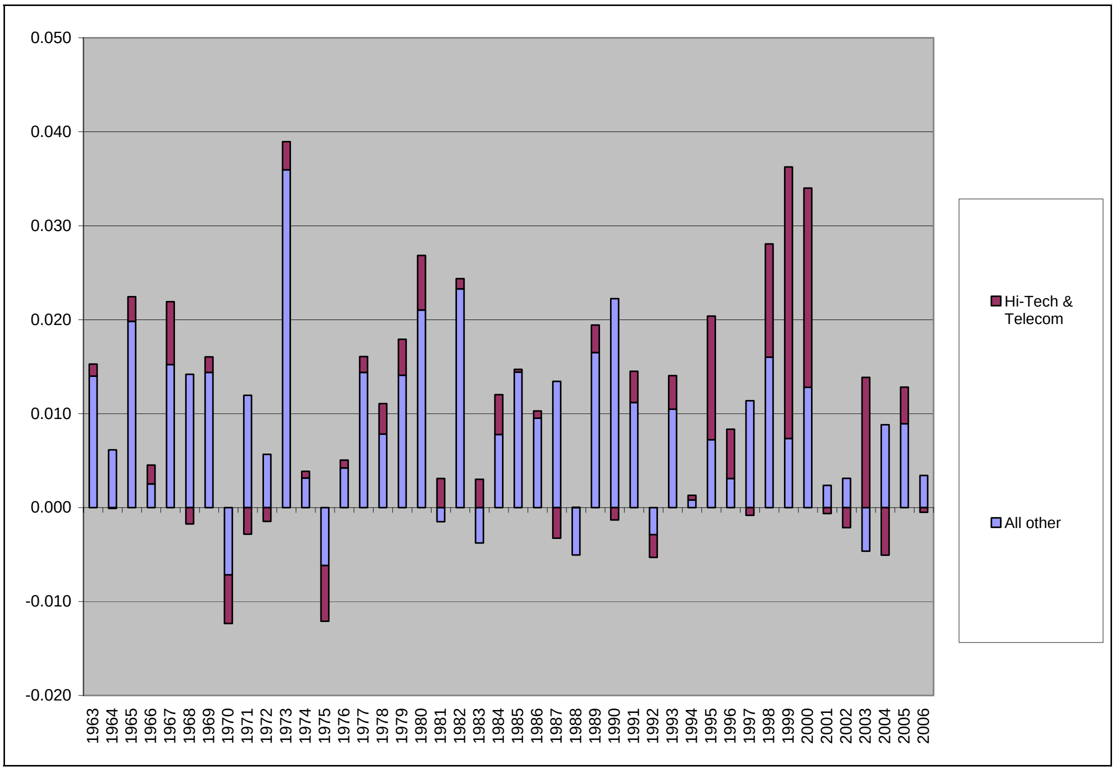


Figure 3-4 Average momentum returns excluding the contribution of hi-tech, and telecom stocks

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past 5-month returns from *t*-6 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy (WML) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. We decompose momentum monthly returns based on whether the stocks in the portfolios belong to one of two groups. The first one contains hi-tech and telecom stocks, and the second one contains stocks from all other industries. The decomposition is done in such a way that, over a period of time, the return to any of these portfolios can be calculated as the sum of the two components. We plot average momentum returns omitting the contribution of hi-tech and telecom stocks.

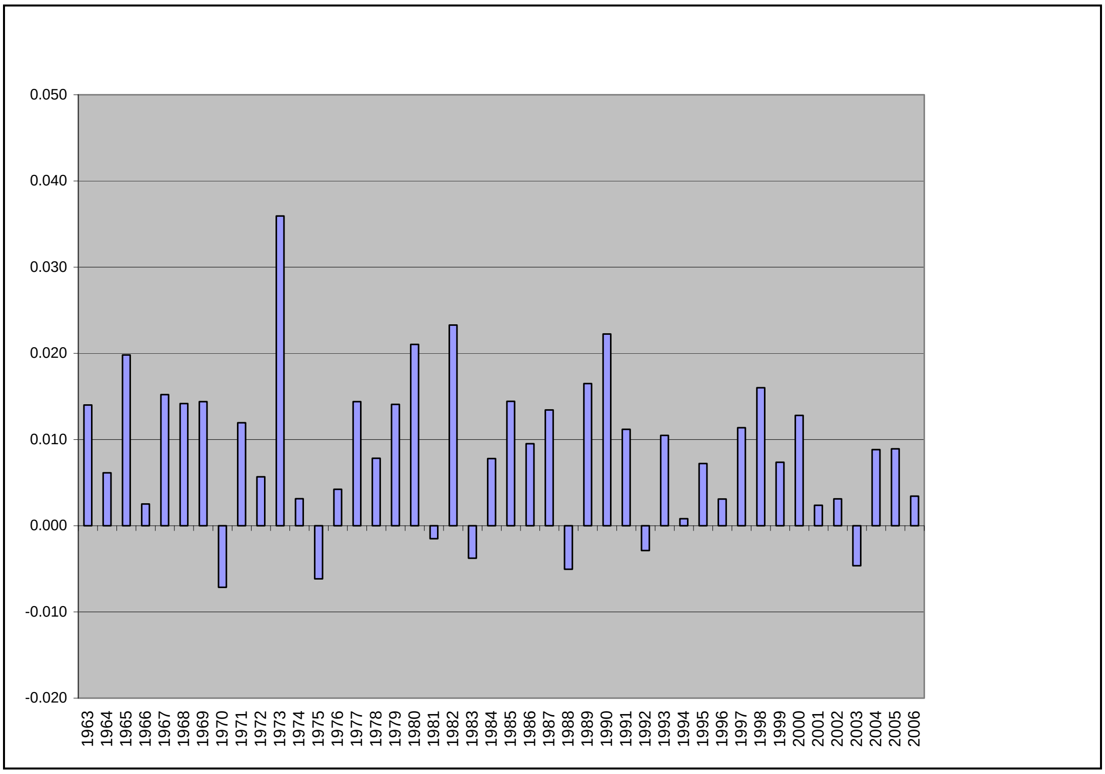


Figure 3-5 Average returns of a momentum strategy that does not allow hi-tech and telecom stocks

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file which do not belong to the Hi-tech or Telecommunications industries (according to their SIC codes) are sorted into decile portfolios based on their past 5-month returns from *t*-6 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy (WML) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. The graph shows yearly average returns of this restricted momentum strategy.

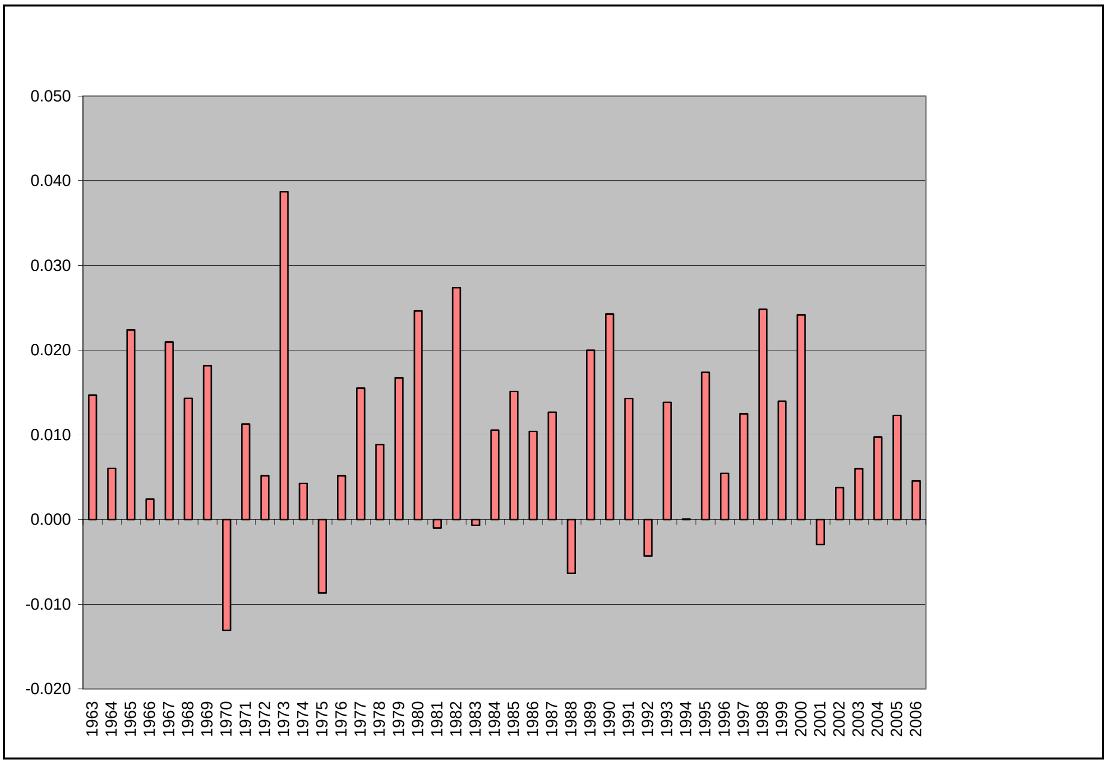
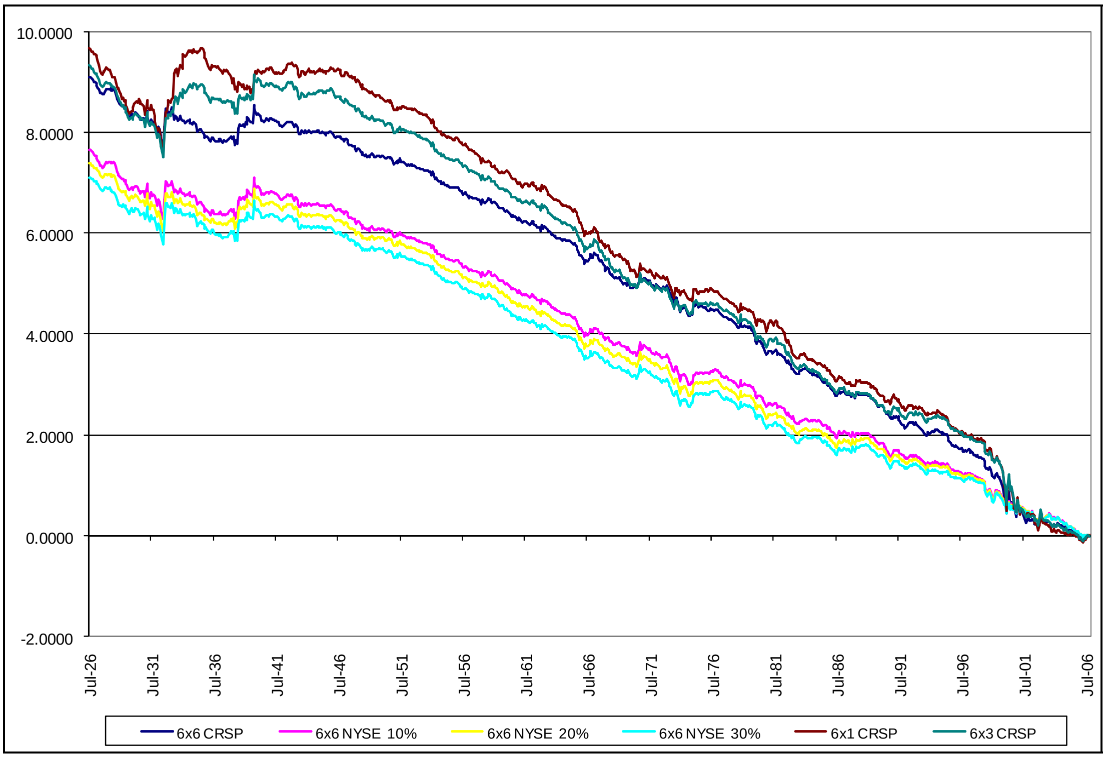


Figure 3-6 Cumulative profits of several momentum strategies through December 2006

At the end of each month *t*, all stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and NASDAQ which are listed on the CRSP monthly file are sorted into decile portfolios based on their past 5-month returns from *t*-6 to *t*-1. Positions are held for the subsequent 6 months, from *t*+1 to *t*+6. The momentum strategy (WML) consists of holding a long position in the Winner (decile 10) portfolio whilst holding a short position on the Loser (decile 1) portfolio. The figure plots the cumulative profit of several momentum strategies per $1 long through December 2006 for various starting date, i.e. at time $t = \tau$ the figure plots $\sum_{t=\tau}^{June\ 2005} r_{WML,t}$, where $r_{WML,t}$ is return of the momentum strategy at time *t* in excess of the risk-free rate of return.



References

Ang, Andrew, and Joseph Chen, 2007, CAPM over the long run: 1926-2001, *Journal of Empirical Finance* 14, 1-40.

Arrow, Kenneth, 1971. *Essays in the theory of risk-bearing* (Markham Publishing Co, Chicago).

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

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

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *The Journal of Finance* 54.

Chan, K.C., N. Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *The Journal of Finance* 51, 1681-1713.

Chib, Siddhartha, 1995, Marginal likelihood from the Gibbs output, *Journal of the American Statistical Association* 90, 1313-1321.

Chib, Siddhartha, 1996, Calculating posterior distributions and modal estimates in Markov mixture models, *Journal of Econometrics* 75, 79-97.

Chib, Siddhartha, 1998, Estimation and comparison of multiple change-point models, *Journal of Econometrics* 86, 221-241.

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, business cycle, and time-varying expected returns, *The Journal of Finance* 57, 985-1019.

Chow, G., 1960, Tests of equality between sets of coefficients in two linear regressions, *Econometrica* 28, 591-605.

Clare, A. D., R. Priestley, and S. H. Thomas, 1998, Reports of beta's death are premature: Evidence from the UK, *Journal of Banking and Finance* 22, 1207-1229.

Cochrane, John, 2001. *Asset pricing* (Princeton University Press, Princeton and Oxford).

Conrad, J., and G. Kaul, 1998, An anatomy of trading strategies, *Review of Financial Studies* 11, 489-519.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, *The Journal of Finance* 59, 1345-1366.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *The Journal of Finance* 53, 1839-1885.

Fama, Eugene F., and Kenneth R. French, 2006, Dissecting anomalies, CRSP working paper.

Grundy, Bruce D., and J. Spencer Martin, 2001, Understanding the nature of the risks and the source of the rewards to momentum investing, *The Review of Financial Studies* 14.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of Finance* 54, 2143-2184.

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

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

Johnson, Timothy C., 2002, Rational momentum effects, *The Journal of Finance* 58, 585-608.

Kass, R. E., and A. E. Raftery, 1995, Bayes factors and model uncertainty, *Journal of the American Statistical Association* 90, 773-795.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *The Journal of Finance* 59, 1039-1082.

Lee, C., and S. Swaminathan, 2000, Price momentum and trading volume, *The Journal of Finance* 55, 2017-2069.

Liu, Chun, and John M. Maheu, 2008, Are there structural breaks in realized volatility?, *Journal of Financial Econometrics* 6, 326-360.

Liu, Laura X. L., Jerold B. Warner, and L. Zhang, 2005, Momentum profits and macroeconomic risk.

Moskowitz, T. J., and M. Grinblatt, 1999, Do industries explain momentum?, *The Journal of Finance* 54, 1249-1290.

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

Perez-Quiros, Gabriel, and Allan Timmermann, 2000, Firm size and cyclical variations in stock returns, *The Journal of Finance* 55, 1229-1262.

Rouwenhorst, K. Geert, 1998, International momentum strategies, *The Journal of Finance* 53, 267-284.

Zhang, X. Frank, 2006, Information uncertainty and stock returns, *The Journal of Finance* 61.

Chapter 4

Fishing with a Licence: an Empirical Search for Asset Pricing Factors

4.1. Introduction

Which are the risk factors that should enter a linear factor asset pricing model, and what kind of fundamental, pervasive, non-diversifiable risks do they represent? This is a question that the asset pricing literature has been struggling with for a long time. The apparent failure of the CAPM of Sharpe (1964), Lintner (1965) and Black (1972) documented in the 1970s[33] and the development of the Intertemporal CAPM by Merton (1973) and the Arbitrage Pricing Theory (henceforth APT) by Ross (1976) have led to theoretically motivated linear factor models (henceforth LFM). The theory, however, is silent as to which or how many factors should be included in the model. Indeed, since the stochastic opportunity set in the ICAPM is not defined but left as a largely empirical issue, allowing researchers to choose from a wide

[33] Black, Jensen and Scholes (1972), Fama and MacBeth (1973) and Blume and Friend (1973) find that the relationship between average returns and betas is linear, but the estimated slope of the security market is too flat and the intercept is too high. However, as Hwang (2007) points out, because of methodological problems with the Fama-MacBeth procedure, specifically the migration of stocks from pre-formation to post-formation portfolios, these results may actually be misleading. Hwang finds that, after correcting for these problems, the regression slope of beta is significant even in the presence of size. Basu (1977), Banz (1981), DeBondt and Thaler (1985), and Jegadeesh and Titman (1993) find that the CAPM cannot price portfolios formed on several firm characteristics.

range of possibilities, Fama (1991) warns against using them as a "fishing license", in the sense that one might use just about any factor.

There have been many attempts to determine which factors explain asset returns. Some researchers extract the factors from assets' returns through statistical techniques such as factor analysis or principal components. Lehman and Modest (1988) and Connor and Korajczyk (1988) both undertake this approach in the context of the APT, and find evidence for between 5 to 10 factors. However, this approach has some disadvantages. The factors obtained through these techniques might not be observable, their economical interpretation is unclear, and it is difficult to distinguish between models with different numbers of factors. Moreover, these factors are likely to depend heavily on the sample, both in terms of the set of the assets chosen, as well as the sample period[34].

A different approach is to work with various risk measures or observable firm characteristics, which may or may not be economically motivated. For example, some studies have tried to explain assets returns through risk measures other than the market beta of the CAPM[35]. Others try to relate assets returns to macroeconomic variables directly; in a seminal work, Chen, Roll and Ross (1986) find evidence of five priced macroeconomic factors. However, factors can be calculated directly from cross-sectional asset returns sorted by firm characteristics; i.e. portfolios are formed by sorting stocks on firm characteristics and a factor is created by taking long and short positions on the extreme portfolios. The most famous and now widely used

[34] For instance, Brennan, Chordia and Subrahmanyam (1998) estimate the principal components factors of Connor and Korajczyk (1988) over two overlapping periods in their sample from 1963 to 1995 because of instability in the factors obtained in each subsample, compared to the whole subsample.

[35] Krauz and Litzenberger (1976) introduce coskewness as a risk measure; see also Harvey and Siddique (2000). Hwang and Satchel (1999) add cokurtosis. Bawa and Lindenberg (1977) introduce an equilibrium model with Lower Partial Moments, of which the semi-variance – or downside beta – model is a particular case. These approaches are equilibrium-based and do not pertain to the LFM literature.

firm-characteristic based model is the Fama and French (1992, 1993, 1996) three-factor model, which includes the excess market return, the size factor (small minus big, SMB) which is long small stocks and short big stocks, and the value factor (high minus low, HML) which is long high book-to-market stocks and short low book-to-market stocks[36]. Another popular factor of this kind is Jegadeesh and Titman's (2001) momentum factor.

One important issue regarding factors derived from firm characteristics is whether or not they are priced factors. This question appears repeatedly as a number of new factors are found. For instance, there is an ongoing debate in the literature regarding whether SMB and HML are related to risk[37] .

Whilst some studies attempt to find evidence that factors based on firm-characteristics are related to risk, others interpret the divergence from the CAPM as evidence of irrational behaviour by market participants, or other concerns related to market microstructure and biases in the empirical methodology. For instance, Lakonishok, Shleifer and Vishny (1994) explain the value premium with investors' overreaction of past earnings growth of firms; Jegadeesh and Titman (1993), Chan, Jegadeesh and Lakonishok (1996), Daniel, Hirshleifer and Subrahmanyam (1998), and Hong and Stein (1999) amongst others propose explanations of the momentum anomaly based on investors wrongly perceiving information due to behavioural/cognitive biases. Unfortunately, as MacKinlay (1995) points out, widely used asset pricing tests are very unlikely to detect deviations from the CAPM due to missing risk factors.

---

[36] Banz (1981) identified the size effect that small stocks earn higher returns than large stocks; Rosenberg, Leid and Lanstein (1985) identified the value effect that high book-to-market (value) stocks earn higher returns than low book-to-market (growth) stocks.

[37] See for example Khotari, Shanken and Sloan (1995), MacKinlay (1995), Lakonishok, Shleifer and Vishny (1994, Liew and Vassalou (2000), Lettau and Ludvigson (2001) and Petkova (2006).

Any LFM which uses factors from firm characteristics to explain assets' returns rationally is subject to some critiques. First, the interpretation of the firm-characteristic based factors in terms of meaningful, interpretable and economically sound risk factors may be doubtful. In other words, whichever variables one puts on the right-hand side of the asset pricing equation have to reflect the fact that investors dislike assets that do poorly in bad economic states. Even if the right-hand variables explain assets' returns statistically, if they do not help to understand how expected returns relate to risk, the model might not be very useful. Second, studies that derive factors from firm characteristics almost always use test assets obtained from grouping procedures, which potentially introduces a data-snooping bias, as pointed out by Lo and MacKinlay (1990), Ferson, Sarkissian and Simin (1999) and Berk (2000). Also, the portfolio formation procedure might make it more difficult to reject the null that a certain factor affects assets' returns, as Roll (1977) suggests. Nonetheless, many authors test LFMs using portfolios of stocks, due perhaps to a limitation of the testing procedure by Gibbons, Ross and Shanken (1989), namely that the number of assets be smaller than the number of time periods. One possible approach that circumvents this problem is the average F-test proposed by Hwang and Satchell (2007).

As a natural consequence of the empirical studies on asset prices, the number of candidate factors has increased steadily, as more CAPM-related anomalies are found. However, it is doubtful that all of these factors matter for asset pricing; it is more likely that some of them are redundant, or proxies for the same kind of fundamental risk, whilst some may just be a product of data snooping. An interesting point is argued by Lewellen, Nagel and Shanken (2007), namely that too many empirical models seem to explain the size and book-to-market effects, which is strange considering many of them have little economical common ground. Some studies have tried to determine which amongst the several extant factors best explains asset returns. In this context, the Bayesian approach is attractive in that it permits the researcher to obtain posterior probabilities of different models and factors. For instance, Hall, Hwang and Satchell (2002) use a Bayesian variable selection methodology to try to find which factors should be included in a global LFM in the context of global equity investment. Ericsson and Karlsson (2004) conduct a study of 15 potential risk factors using a Bayesian approach to estimate posterior model selection probabilities of all the 2^{15} possible LFMs. Their evidence supports a LFM with the excess market return, SMB, HML, the momentum factor and credit risk spread. However, their results might be misleading since they only use portfolios, and their set of possible factors contains mostly macroeconomic factors.

Whilst some studies seek to identify which factors should be included in a LFM, Brennan, Chordia and Subrahmanyam (1998) conduct a study of the relation between stock returns and measures of risk and non-risk characteristics. Their focus is in discovering whether non-risk characteristics have explanatory power relative to APT benchmarks using the Fama and French factors or factors obtained from the Connor and Korajczyk (1988) principal components approach. Their results suggest that, even after accounting for the risk-related factors, there is evidence of return momentum, size and book-to-market effects. They also find a significant negative relation between returns and trading volume, which they deem consistent with a liquidity premium in stock returns. Along similar lines, in a recent study Hwang and Lu (2007) consider 16 factors amongstst risk based, firm-characteristic based and macroeconomic factors for thousands of individual firms and find that the excess market return, liquidity and coskewness factors explain stock returns as well as the

Fama-French three factors with momentum. Their approach is based on whether factors explain asset returns in terms of zero alphas in a GRS-like test for individual stocks, proposed by Hwang and Satchell (2007).

In this paper we investigate the issue of selecting asset pricing factors using a Bayesian variable selection procedure from the statistics literature called Stochastic Search Variable Selection (SSSV), introduced by George and McCulloch (1993). We extend their approach to a simple multivariate case with N assets. Our procedure uses a latent dummy variable approach to identify the more promising factors as those with higher posterior probability. Our approach is different from the previous ones in several ways. First, we infer the factors that matter directly through a dummy variable vector, thus avoiding the exhaustive search over a prohibitively large number of LFMs. Second, we conduct our tests using individual stocks from the CRSP universe (although we also conduct tests on portfolios of stocks for comparison). The approach with individual stocks is important because it avoids the potential data-snooping biases which may arise when models which contain empirically motivated factors are tested on portfolios formed on the same characteristics, as explained before. Although Brennan, Chordia and Subrahmanyam (1998) and Hwang and Lu (2007) also use individual securities, our approach is quite different from theirs. We are interested in finding posterior probabilities of risk and non-risk factors, whilst Brennan, Chordia and Subrahmanyam (1998) test the explanatory powers of non-risk characteristics after adjusting for risk. Also, although we use similar factors and use individual securities, our work is distinct from Hwang and Lu (2007), who look for combinations of factors which lead to acceptance of the null of zero intercepts in time-series regressions. In contrast, our primary interest is in finding which factors or combinations of factors have high-posterior probabilities.

A by-product of our approach is that, since we treat the intercept as a factor, we can interpret high-posterior probability models which do not include the intercept as well-specified asset pricing models.

Our results provide evidence that a LFM should contain the excess market return, the size factor SMB and the liquidity factor. We find weak evidence that a factor proxying for idiosyncratic volatility should be included in the model, and also that the HML factor should be included, but these could be related to particular periods of time. Our results are robust to different prior distributions and to different subsamples during the period from 1967 to 2005.

The rest of this paper is organised as follows. We introduce the model and briefly discuss the estimation method in Section 4.2. The explanation of the data set and empirical tests follows in Section 4.3, as well as robustness tests and comparison with previous studies. Section 4.4 concludes and Section 4.5 contains detailed explanations of the model estimation.

4.2. Methodology

4.2.1. Stochastic Search Variable Selection

In this section we describe the details of the Bayesian variable selection procedure we employ. We extend the procedure known as Stochastic Search Variable Selection (henceforth SSVS) developed by George and McCulloch (1993) to a simple multivariate case. Our approach is based on the algorithm of Kuo and Mallick (1998), which significantly reduces the computational burden relative to the one suggested by George and McCulloch (1993). Parameters are estimated via a Markov Chain Monte Carlo (MCMC) method implemented with the Gibbs sampler.

The method is based on parameterising the independent variables that should be selected in a regression via a vector of indicator variables $\boldsymbol{\gamma}$, where each component is equal to 1 if the corresponding variable should be included, or 0 otherwise. We start with some prior belief about the inclusion of each variable, and then modify these with the information contained in the likelihood to form a posterior distribution for $\boldsymbol{\gamma}$. The posterior distribution of $\boldsymbol{\gamma}$ is supported on all possible combinations of variables, and therefore contains information not only about whether a particular variable should be included, but also about the probabilities of different combinations of variables. We now explain the procedure in statistical terms. Consider a system of N regressions with T observations each:

$$\mathbf{r_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \mathbf{e_i}, \text{ for } i = 1,\ldots,N \tag{3.29}$$

where $\mathbf{r_i} = \left(r_{i1}\ r_{i2} \cdots r_{iT}\right)'$ is the $\left(T \times 1\right)$ vector of observations of the i-th dependent variable, $\mathbf{X_i}$ is a $\left(T \times K_i\right)$ matrix of independent variables, $\boldsymbol{\beta}_\mathbf{i}$ is a $\left(K_i \times 1\right)$ vector of regression coefficients and $\mathbf{e_i}$ is the $\left(T \times 1\right)$ vector of error terms of the i-th regression. The system of equations (3.29) is a system of seemingly unrelated regressions (SUR) if the collected vector of error terms $\tilde{\mathbf{e}} = \left(\mathbf{e}_\mathbf{1}'\ \mathbf{e}_\mathbf{2}' \cdots \mathbf{e}_\mathbf{N}'\right)'$ is assumed to have the following distribution:

$$\tilde{\mathbf{e}} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I_T}\right) \tag{3.30}$$

where $\mathbf{\Sigma}$ is an $(N \times N)$ matrix to be estimated and $\mathbf{I_T}$ denotes the identity matrix. In this study, $\mathbf{r_i}$ represents the return on asset *i*, and $\mathbf{X_i}$ are factors in asset pricing models.

Now we define vectors $\boldsymbol{\gamma}_\mathbf{i} = \left(\gamma_{i1}\ \gamma_{i1} \cdots \gamma_{iK_i}\right)'$, where each element indicates whether a particular independent variable should be included in the i-th regression. In other words, if $\gamma_{ij} = 1$ then the j-th regressor in the i-th regression equation is included, and if $\gamma_{ij} = 0$, then it is not. The interpretation of each γ_{ij} is that it indicates whether factor j is required in the model for asset i. We can now rewrite (3.29) in matrix form in the following way:

$$\tilde{\mathbf{r}} = \tilde{\mathbf{X}}\tilde{\boldsymbol{\beta}}_\gamma + \tilde{\mathbf{e}}, \tag{3.31}$$

where

$$\tilde{\mathbf{r}} = \begin{bmatrix} \mathbf{r_1} \\ \mathbf{r_2} \\ \vdots \\ \mathbf{r_N} \end{bmatrix}, \quad \tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X_1} & 0 & \cdots & 0 \\ 0 & \mathbf{X_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{X_N} \end{bmatrix}, \quad \tilde{\boldsymbol{\beta}}_\gamma = \begin{bmatrix} \boldsymbol{\beta}_1 \odot \boldsymbol{\gamma}_1 \\ \boldsymbol{\beta}_2 \odot \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\beta}_\mathbf{N} \odot \boldsymbol{\gamma}_\mathbf{N} \end{bmatrix}, \quad \tilde{\mathbf{e}} \sim N\left(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I_T}\right),$$

and $\boldsymbol{\beta}_\mathbf{i} \odot \boldsymbol{\gamma}_\mathbf{i} = \left[\beta_{i1}\gamma_{i1}\ \beta_{i2}\gamma_{i2} \cdots \beta_{iK_i}\gamma_{iK_i}\right]'$ denotes the element-by-element multiplication of $\boldsymbol{\beta}_\mathbf{i}$ by $\boldsymbol{\gamma}_\mathbf{i}$. The above formulation is rather general, and we simplify it with some assumptions. First, we assume that the set of possible factors is the same for all assets, i.e. $\mathbf{X_1} = \mathbf{X_2} = \cdots = \mathbf{X_N} = \tilde{\mathbf{X}}$ where $\tilde{\mathbf{X}}$ contains data on K factors. We also assume that $\mathbf{\Sigma}$ is a diagonal matrix, and therefore each $\mathbf{e_i} \sim N\left(\mathbf{0}, \sigma_i^2 \mathbf{I_T}\right)$. With these assumptions the SUR model simplifies to a system of ordinary regressions. We further assume that the vectors γ_i are identical for all assets, i.e. $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2 = \cdots \boldsymbol{\gamma}_N = \boldsymbol{\gamma}$. This assumption is justified by our goal of identifying pervasive factors, and not factors which are correlated only with some subsets of stocks. With these restrictions, $\tilde{\boldsymbol{\beta}}_\gamma$ in equation (3.31) simplifies to $\tilde{\boldsymbol{\beta}}_\gamma = \tilde{\boldsymbol{\beta}} \Box \boldsymbol{\gamma}$, and $\tilde{\mathbf{X}}$ simplifies to a $(TN \times KN)$ block-diagonal matrix with $\mathbf{X}$ in the diagonal, i.e. $\tilde{\mathbf{X}} = \mathbf{X} \otimes \mathbf{I}_N$.

Our assumption of no correlation between error terms of different assets comes from our aim of using individual stocks to identify the factors. A fully-specified covariance matrix would be very difficult to estimate with thousands of individual stocks. However, studies such as Connor (1984), Dybvig (1985) and Grinblatt and Titman (1985) suggest that, if the factors are pervasive, or when the parameters of the economy are well-specified, the off-diagonal elements of $\boldsymbol{\Sigma}$ are likely to be negligible. For instance, Hwang and Satchell (2007) estimate $\boldsymbol{\Sigma}$ for a LFM containing the market return, the two Fama-French factors and momentum for individual stocks in the CRSP database, and find that about 8 percent of the correlations are significant. They also evaluate the impact of diagonality on testing LFMs and find that the off-diagonal components do not affect test statistics. We take the point of view that the simplicity gained with this assumption and the possibility of using data on thousands of individual stocks' returns compensate for the slight increase in efficiency that would be gained by accounting for the small number of correlations likely to be significant. An interesting approach, which is however outside the scope of this work, would be to allow the model to automatically select which correlations are significant, similarly to the approach proposed by Cripps, Carter and Kohn (2003).

We now turn to the specification of the priors, and relegate the derivation and details of the conditional densities used in the Gibbs sampler to Section 4.5. We follow Kuo and Mallick (1998) and choose priors for $\tilde{\boldsymbol{\beta}}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}$ independently, with $\tilde{\boldsymbol{\beta}} \sim \mathrm{N}\left(\tilde{\boldsymbol{\beta}}_0, \mathbf{D}_0\right)$, $\gamma_j \sim B\left(1, p_j\right)$ and $\sigma_j \sim IG\left(\alpha/2, 2/\eta\right)$, for $j = 1,\ldots,K$. The hyperparameters can be chosen to reflect the amount of prior knowledge about the inclusion of the factors. For the regressions coefficients $\tilde{\boldsymbol{\beta}}$ we choose a prior centred on a vector of zeros, to reflect complete lack of knowledge as to which factors should be included. We choose the prior variance-covariance matrix of $\tilde{\boldsymbol{\beta}}$ to be $\tilde{\mathbf{D}}_0 = c\left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}$ which is similar to the prior used by Smith and Kohn (1996). The choice of the constant c reflects the prior degree of uncertainty about $\tilde{\boldsymbol{\beta}}$ [38].

The prior for $\boldsymbol{\gamma}$ reflects the prior belief about the inclusion of each variable. To reflect our lack of knowledge about the inclusion of the factors, we choose $p_j = P\left(\gamma_j = 1\right) = 1/2$, $\forall j$. This prior makes any combination of factors have the same prior probability of $1/2^K$.

4.2.2. Interpreting the Posterior Results

With these priors, it is straightforward to estimate the parameters using the Gibbs sampler. Our main interest is in the posterior distribution of $\boldsymbol{\gamma}$, which is estimated based on many iterations of the Gibbs sampler. From the posterior

[38] We discuss values for c in the appendix.

distribution of $\boldsymbol{\gamma}$, several quantities of interest can be calculated, such as the marginal posterior probabilities of each factor and of particular models, i.e. combinations of factors. Since each γ_j has a prior probability of 0.5, if a factor has an estimated marginal posterior probability higher than 0.5, we can interpret that this factor should be included in the LFM. Models with high posterior probability will show up more frequently in the generated values of $\boldsymbol{\gamma}$, and thus it is possible to focus on high-posterior probability models without having to analytically compute the posterior probabilities for all the 2^K possible models analytically, which can be quite overwhelming for large values of *K*.

A comment about the intercept term in model (3.31) is in order. In the context of asset pricing models, the inclusion of the intercept is of economical importance. In this study we include the intercept as one of the regressors in the matrix $\tilde{\mathbf{X}}$ and *a priori* it has the same probability of inclusion as any of the other factors. Since all the factors we consider in this study are factor mimicking portfolios and can be considered tradable assets, the omission of the intercept in a particular high posterior probability model can be interpreted as evidence that that model is a good description of the assets returns. Thus, besides looking for models with high posterior probability, we are also interested in whether or not they include the intercept.

4.2.3. Research Design

The main objective of this paper can be described as follows. Given asset returns and a set of possible factors, our main interest is in identifying the factors that should be included in a linear factor model to explain the returns on these assets. We consider individual stocks as well as portfolios of stocks, in order to investigate whether the LFM implied by the posterior of $\boldsymbol{\gamma}$ changes significantly.

4.2.3.1. Candidate Asset Pricing Factors

The factors we test in this study are those proposed in previous studies. The candidate factors belong to one of two categories: risk-based factors and firm-characteristics (or non-risk) factors. All factors (except the excess market return) are calculated by sorting portfolios on the (risk or non-risk) characteristic and forming hedge portfolios, similarly to Fama and French's (1996) SMB and HML. We next describe the factors in each category, and refer the reader to Hwang and Lu (2007) for a detailed explanation of how they are calculated.

4.2.3.1.1. Risk-Based Factors

The five risk-based candidate asset pricing factors are the excess market return, liquidity, coskewness, cokurtosis and downside risk. The market portfolio as a factor has its justification in the CAPM. Although there is controversy as to the empirical performance of the CAPM (see for example Fama and French (1992, 1996)), an asset's covariation with a market-wide portfolio is still considered the most important determinant of asset returns, and has a very clear interpretation in terms of risk.

Some studies (Clare, Priestley and Thomas (1998), Hwang (2007)) suggest that the fact that beta does not appear to be priced in the cross-section is due to methodological problems in studies which use the Fama and MacBeth (1973) two-step methodology.

The inclusion of liquidity as a risk-based factor is motivated by Brennan and Subrahmanyam (1996), Chordia, Subrahmanyam and Anshuman (2001) and Pastor and Stambaugh (2003), amongst other studies. Brennan and Subrahmanyan report a significant and positive relation between expected returns and illiquidity which is robust to adjustment by the Fama-French factors; Chordia, Subrahamnyan and Anshuman argue that liquidity needs to be accounted for, even after controlling for size, book-to-market and momentum effects. Pastor and Stambaugh show that a large portion of momentum profits can be attributed to liquidity risk.

Coskewness as a risk factor reflects investors' attitudes towards non-normal returns; investors dislike negative skewness in their portfolios and thus are willing to pay a premium for assets that make their portfolios less negatively skewed. The seminal work of Kraus and Litzenberger (1976) laid the theoretical motivation for a CAPM with a premium for coskewness. Harvey and Siddique (2000) show that conditional skewness helps to explain the cross-section of returns even in the presence of the size and book-to-market factors. Similarly, we include a cokurtosis factor. This is motivated by assuming that investors also care about the fat tails (kurtosis) of their portfolios. Hwang and Satchell (1999) show that cokurtosis matters in emerging markets where returns are highly leptokurtic. Dittmar (2002) suggests that quadratic or cubic pricing kernels are admissible for the cross-section of industry portfolios.

Finally, the inclusion of a factor to proxy for downside risk is motivated by the downside risk literature. Starting with the theoretical developments of the lower partial moments CAPM by Bawa and Lindenberg (1977) and Harlow and Rao (1989), several authors have studied the role of downside risk. Post and Van Vliett (2005) find that when the regular CAPM beta is replaced by the downside beta, the cross-sectional risk-return relationship improves, and that this explanatory power is robust to size, value and momentum effects. Ang, Chen and Xing (2006) find that downside risk commands a premium in the cross-section of stock returns.

4.2.3.1.2. Non-Risk Factors

There is extensive research suggesting that the returns on stocks are correlated with many firm characteristics, the most famous of which are size (market equity) and the ratio of book-to-market equity (see for instance Daniel and Titman (1997) and Brennan, Chordia and Subrahmanyam (1998)). Fama and French (1993, 1996) argue that HML, which is constructed based on book-to-market ratio (B/M), represents non-diversifiable risk related to financial distress. They argue that weak firms with low earnings tend to have high book-to-market ratios, and thus pay a distress premium. It is less clear what sort of risk the small-minus-big factor SMB proxies for; some studies suggest important asymmetries in the variation of small and large firms' risks over the business cycle particularly because small firms are more affected by tighter credit market conditions than large firms, see Perez-Quiros and Timmermann (2000).

There is an ongoing debate as to whether these two anomalies are indeed related to non-diversifiable sources of risk. The same can be argued for many other factors based on firm characteristics. However, the use of these factors in LFMs has become quite common. As Brennan, Chordia and Subrahmanyam (1998) point out, the role of these characteristics might either be related to frictions within the rational pricing paradigm, or to their statistical properties as proxies for expected returns. Therefore we include several factors based on firm characteristics in this study. Since these factors are formed on stock characteristics which are not necessarily related to risk, we call them non-risk factors. However, as the discussion about the Fama-French factors illustrates, some of them may be related to risk.

The non-risk factors we consider are the Fama-French size (SMB) and value (HML) factors, momentum, asset growth, idiosyncratic volatility, volume, and long-term reversal. Momentum or short-term continuation of stock prices has been documented by many studies, starting with De Bondt and Thaler (1987) and Jegadeesh and Titman (1993), who document that past winner stocks outperform past loser stocks. The debate over the source of this anomaly is ongoing. Our results from Chapter 3 indicate that momentum has disappeared since the year 2000. However we include it as a factor since the disappearance could be linked to a risk that disappeared in the same period and thus does not necessarily suggest that momentum should not be priced.[39] Many studies in the 1980s such as De Bondt and Thaler (1987), Fama and French (1988) and Poterba and Summers (1988) report negative autocorrelations in the long-run, which Jegadeesh and Titman (2001) document as the long-term reversal effect, i.e. past winners underperform past losers on horizons longer than 12 months.

[39] The literature on momentum is enormous, see Chapter 3 for a short review and further references.

The inclusion of asset growth is motivated by Cooper, Gulen and Schill (2006), who find that asset growth rates are more important determinants of the cross-section of returns than size, B/M and several other measures. Asset growth is measured as year-on-year percentage change in total assets. Idiosyncratic volatility as a risk factor is under debate regarding whether firm-specific risk is priced in the market. The answer is not clear; for instance, whilst Goyal and Santa Clara (2003) find a positive relationship between the idiosyncratic component of stock variance and the return on the market, Bali, Cakici, Yan and Zhang (2005) argue that this is mainly due to small and illiquid stocks, and thus could be related to a liquidity premium. On the other hand, Ang, Hodrick, Xing and Zhang (2006) find an inverse relationship between idiosyncratic volatility (from the Fama-French three-factor model) and stock returns, which is robust to the size, B/M, momentum, volume and liquidity effects. The volume factor is motivated by the high-volume premium investigated by Gervais, Kaniel and Mingelgrin (2001), who document that stocks with unusually high (low) trading volume over a day or a week tend to appreciate (depreciate) over the following month.

4.3. Empirical Results

4.3.1. Data

Our main results are obtained with individual stocks from the CRSP dataset, but we also use portfolios of stocks for comparison purposes. Our data come from various sources. The data on the portfolios is obtained from Kenneth French's data library, and consists of the 25 portfolios formed on size and B/M (see Fama and French, 1996), plus 30 industry portfolios, in a total of 55 portfolios. The data covers the period from January 1967 to December 2006. This choice of portfolios was motivated by Lewellen, Nagel and Shanken (2007). They argue that it is problematic to use only portfolios formed on size and B/M to test asset pricing models, because these portfolios will have a tight factor structure by construction, in the sense that the two Fama-French factors SMB and HML (or factors which are correlated with them) explain nearly all of the time-series variation in the returns on these portfolios. In this case, the idiosyncratic component of the model will be very small, and the factors will appear to be statistically significant cross-sectionally.

The data on individual stocks' returns were obtained from the Center for Research in Stock Prices (CRSP) data file from January 1967 to December 2006. Aside from the Fama-French factors, the data on all other factors has been obtained from Hwang and Lu (2007)[40].

When dealing with individual stocks there is a potential survivorship bias, and thus we split our 480-month original sample into 4 smaller subsamples of 120 months each. In each of these subsamples, we exclude all stocks that have a price lower than U$5 in the first month of that subsample, to remove any undesirable effect from penny stocks. This approach has some advantages. First, it guarantees that a larger number of stocks will be used in each subsample period, whereas if we use the whole sample period the number of surviving stocks drops dramatically. Second, it partly allows for time-variation in the factor sensitivities of each stock across subsamples, and thus could give us some insights about whether some factors

[40] We kindly thank Chensheng Lu for providing us with these data.

may be selected only during some periods of time. To keep results comparable and consistent, we follow the same approach with the portfolios.

Table 4-1 displays summary statistics of the factors. From the average monthly returns and standard errors reported in Panel A, not all factors have average returns significantly different from zero[41]. The factors with significant average returns are the excess market return, with a monthly average return of 0.45% (std. error 0.20), liquidity with an average return of -0.51% a month (std. error 0.16), HML with an average return of 0.45% a month (std. error 0.13), momentum with an average return of 0.81% (std. error 0.19), asset growth with an average return of 0.37% (std. error 0.13) and long-term reversal with an average return of 0.36% (std. error 0.11).

In Panel B we display the Spearman correlation matrix of the factors. As expected, several of the factors have correlations which are significantly different from zero. The highest correlations in absolute value are between SMB and liquidity (-0.78), idiosyncratic volatility and downside risk (0.68), excess market return and the downside risk factor (-0.67), idiosyncratic volatility and SMB (-0.65), idiosyncratic volatility and excess market return (-0.61), HML and asset growth (0.60), and HML and downside risk (0.57). Whilst HML and asset growth are both factors constructed using the book-to-market value and thus are expected to be correlated, the fact that idiosyncratic volatility, for example, is correlated with several factors shows that this factor might contain information related to several other factors.

[41] As Hwang and Lu (2007) explain, although the studies that motivated their factors show that these factors generally have significant average returns, this does not necessarily have to be the case, due to differences in the sample period, construction method, i.e. use of pre-formation or post-formation returns, and the universe of stocks considered.

4.3.2. Exploratory Analysis of the Factors

The correlations amongst the factors suggest that some factors might be linear combinations of other factors. To investigate this issue, we conduct a preliminary analysis by regressing each factor on the remaining factors. The results are shown in Table 4-2, where each column contains the estimates of the regression of the factor in the column on all the other factors. The R^2 values show that some factors are well explained by other factors. For instance, 88% of the variation in SMB can be explained by other factors. The other Fama-French factor, HML, has an R^2 of 68% in the regression on the other factors, and all factors are significant in explaining it. Liquidity also has a high R^2 (around 80%), but the significant regressors are fewer. As expected from the correlations, idiosyncratic volatility and downside risk are also well explained by other factors with high R^2s, 79% and 74% respectively.

The results suggest that many of these factors are redundant and proxies of priced risk. Whilst these results suggest a potential for multicollinearity, the SSVS procedure is very robust in that regard, since variables with high correlation will have low posterior probability of being selected together, unless they do carry distinct information.

4.3.3. Selection of Asset Pricing Factors

This section reports our main empirical findings. First, we report the results obtained using individual stocks, and then we show the results using the portfolios. In the end of this section we analyse the robustness of our results.

4.3.3.1. Individual Stocks

The numbers of stocks in each of the four subsamples are 1141, 1146, 2159 and 2787 respectively, from Jan 1967 to December 2006. We apply our methodology to each of these subsamples, in order to obtain inference on which factors should be included in a LFM through the posterior distribution of $\boldsymbol{\gamma}$ [42]. The results are reported in Table 4-3. Panels A through D correspond to each of the four subsamples, with marginal factor posterior probabilities followed by the three highest posterior probability models in each subpanel.

Considering the four subsample periods, our results provide strong evidence that a LFM to explain individual stock returns should contain the market return, the SMB and the liquidity factor. We found weak support for the idiosyncratic volatility factor, as its posterior probability is considerably high in only one of the subsamples.

[42] The estimates of the posterior distribution of $\boldsymbol{\gamma}$ are obtained through the MCMC draws. For example, the marginal posterior probability for a factor *j* is estimated as the proportion of MCMC iterations in which γ_j is one. Likewise, the posterior probability of a certain model or combination of factors is estimated as the proportion of MCMC iterations in which the corresponding elements of $\boldsymbol{\gamma}$ are equal to one.

In most subsamples, there is little model uncertainty, and this is generally related to the inclusion of one or two factors. For instance, during the first subsample (from January 1967 to December 1976), the highest posterior probability model (with a probability of 0.78) includes the excess market return, the SMB and the liquidity factors (see Panels A.1 and A.2). The second best model (with a posterior probability of only 0.11) includes all these factors, and in addition the idiosyncratic volatility and the downside risk factors (thus these two models are selected in around 88.4% of the cases during this subsample). In the last two subsamples (Panels C.1-C.2 and D.1-D.2), the highest probability model is the CAPM. However, in the former it has a high posterior probability of 0.78, whilst in the latter it is only 0.38, indicating much model uncertainty. As to the inclusion of the intercept term, its posterior probability is low in all subsamples. This suggests that, although there is some model uncertainty, the highest posterior probability models explain the returns on the stocks adequately.

Although some studies which use portfolios of stocks have found idiosyncratic volatility to be priced in the cross-section, our results with individual stocks suggest otherwise. It is also interesting that the SMB factor is selected together with the liquidity factor, suggesting these factors contain distinct information about returns, despite their high correlation. Another striking result is that the HML factor, which appears to be ubiquitous in empirical research, has a low posterior probability in all subsamples except the last.

4.3.3.2. Portfolios of stocks

It is most common for studies to use portfolios of stocks when testing whether a particular factor should be included in a LFM. However, as Lewellen, Nagel and Shanken (2007) point out, models can be accepted too easily if portfolios are used. Although this is most obvious in cross-sectional tests of asset pricing models, it would be interesting to see how our results would be affected if portfolios of stocks are used, rather than individual stocks. In this subsection we apply our methodology to 55 portfolios of stocks, consisting of a set of 25 portfolios formed by sorting on size and book-to-market augmented by 30 industry portfolios. We follow the same scheme of dividing our original sample from 1967 to 2006 into four non-overlapping 120-month subsamples, to retain comparability with our results using individual stocks.

The results, reported in Table 4-4, suggest that a model to describe the returns on these 55 portfolios should always include the excess market return, the two Fama-French factors SMB and HML, idiosyncratic volatility and the liquidity factor. There is also some evidence for the inclusion of the downside risk factor (which is included with high probability in the first and last subsamples, see Panels A.1 and D.1). As before, in most cases there is little model uncertainty caused by the inclusion or omission of few factors. For instance, in the first subsample (Panel B.1) the two best models differ only by the inclusion of the liquidity factor, and in the third subsample (Panel C.1) the difference is the inclusion of the idiosyncratic factor. There is more uncertainty regarding the inclusion of the intercept term for the portfolios than for the individual stocks. Particularly, in the third subsample the intercept is included with certainty, suggesting that during the 1987-1996 period no combination of the factors explains these portfolios with high posterior probability. In the first and last subsamples, the probability of including the intercept is low (0.163 in the former and 0.005 in the latter), so during the 1967-1976 and the 1997-2006 periods the best models explain the returns on the portfolios.

The inclusion of the two Fama-French factors is expected, since the set of portfolios includes 25 portfolios formed on size and book-to-market. It is interesting that the idiosyncratic volatility factor has a high posterior probability for the portfolios, but not for individuals stocks. The same can be said about the downside risk factor, although it has high posterior probability in only two of the subsamples.

The fact that the two Fama-French factors SMB and HML are included with certainty for the portfolios but not for the individual stocks casts some doubts about their inclusion. To investigate this issue, in the next subsection we redo our calculations using only data on the industry portfolios.

4.3.3.3. Industry Portfolios

The difference in the results with individual stocks and portfolios indicate that HML (and potentially SMB) may have high posterior probabilities because the size and book-to-market sorted portfolios are included. In order to investigate the sensitivity of the posteriors of SMB and HML to the assets, we redo the calculations using only the 30 industry portfolios. As Daniel and Titman (2005) argue, variation in the returns on industry portfolios is less likely to be related to the B/M ratio, and more likely to covariate with other sources of risk. Because of that, the results with the industry portfolios can also be viewed as a robustness test.

The results, reported in Table 4-5, show some similarities with those from Table 4-4. The idiosyncratic volatility and the downside risk factors are in the same subsamples (the idiosyncratic volatility factor is selected in all subsamples except the third subsample from 1987 to 1996, and the downside risk factor is selected in the first and last subsamples). However, whereas the two Fama-French factors SMB and HML had a posterior probability of 1 in all subsamples for the 55 portfolios, the two factors do not show consistently high posterior probability for the 30 industry portfolios over different subsample periods: the SMB factor is selected in three subsamples whilst the HML factor has high posterior probability only in the last subsample. For instance, in the third subsample from 1987 to 1996, neither of the two factors has a high posterior probability. Indeed, only the excess market return has high posterior probability in that subsample, suggesting that the CAPM explains the returns on the industry portfolios during this period. Similarly to the results with individual stocks, the probability of the inclusion of the intercept is low.

These results suggest that the HML factor is necessary to explain the returns on the 25 size-B/M portfolios, but is not needed to explain the industry portfolios or the individual stocks. However, the size factor SMB is required for both sets of portfolios and individual stocks in most subsamples.

Overall, we conclude that using portfolios to calculate the posterior probabilities of possible factors is less desirable than using individual stocks, since the probabilities will favour factors related to the sorting variables. This suggests that the data-snooping problems which arise when asset pricing models are tested on grouped data, as argued by Lo and MacKinlay (1990), Ferson, Sarkissian and Simin (1999) and Berk (2000), also affect the calculation of posterior probabilities in the time-series setting.

4.3.4. Robustness

In this subsection we analyse the robustness of our results with regards to the priors used for the sensitivities of the assets to the factors. In the results we reported so far, the prior variance-covariance matrix of $\tilde{\boldsymbol{\beta}}$, the vector of regression coefficients, was $\tilde{\mathbf{D}}_{\mathbf{0}} = c\left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}$, where the constant c is equal to 1[43]. This reflects a prior which, although centred in zero to reflect lack of knowledge about which factors should be included, is to some extent informative, with a relatively small standard deviation.

We increased the value of *c* to 10 and 100 to check whether our results are affected in a significant way. The prior with c=100 is particularly non-informative by increasing the variance-covariance matrix of $\tilde{\boldsymbol{\beta}}$. The results of the estimation with the priors with *c*=10 and *c*=100 are reported in Table 4-6 and Table 4-7, respectively. The analysis of these results reveals that the factors that are selected are not very different from the previously reported ones though as expected, model uncertainty is higher for these less informative priors. However, we also notice that the idiosyncratic volatility factor is not selected unless we have a stronger prior in favour

[43] We could also write the prior for each $\boldsymbol{\beta}_{\mathbf{i}}$ as $\tilde{\mathbf{D}}_{\mathbf{0,i}} = c\sigma_i^2\left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}$, where σ_i is the idiosyncratic volatility of asset *i*. We opt to include a constant *c*, and since our results do not change significantly for a wide range of values of *c*, this would not influence our results.

of the factor. We also repeated our calculations with a diagonal prior for $\tilde{\mathbf{D}}_{\mathbf{0}}$, i.e. $\tilde{\mathbf{D}}_{\mathbf{0}} \propto I$. The results are similar to the ones reported above.

4.3.5. Summary and Comparison with Previous Studies

There have been several attempts to find out which factors, amongst the several available factors, should be included in a linear factor asset pricing model. The Bayesian method we use is similar to the one used by Hwang, Satchell and Hall (2002) in the context of finding style factors in global stock return models.

Our study is not easily comparable to studies such as Lehman and Modest (1988) and Connor and Korajczyk (1988), which extract the factors from assets' returns through statistical techniques such as factor analysis or principal components, because the factors obtained in that way might not be observable and have difficult economic interpretation. Also, the approach with macroeconomic variables such as Chen, Roll and Ross (1986) is difficult to compare to ours, because we do not use macroeconomic variables. In this study we focus on twelve factors which are constructed as mimicking portfolios. We chose not to include macroeconomic variables together with these factors because, as argued by Cochrane (2001), models with factor-mimicking portfolios are very likely to outperform models with real economic factors[44]. Our results suggest that amongst the two factors proposed by Fama and French (1993, 1996), only SMB matters in asset pricing. This is in contrast

[44] In tests which we do not report, we included macroeconomic factors in the set of possible factors together with the 12 factors we used. The results obtained did not change substantially, with the macroeconomic factors having very low posterior probabilities. An alternative approach is to allow only the macroeconomic variables

with many studies by Fama and French, who argue that their two factors are priced risk factors needed to explain the cross-section of return. Although it could be argued that most studies supporting the size and value factors are subject to the data-snooping bias (Lo and MacKinlay (1990), Ferson, Sarkissian and Simin (1999) and Berk (2000)), the presence of these anomalies in the cross-section of returns is very strong, even after adjusting for risk with the Fama-French factors, as shown by Brennan, Chordia and Subrahmanyam (1998), who also find a volume effect which they interpret as related to a liquidity premium. In that sense, it is surprising that we find low posterior probabilities for HML and momentum, but the finding that the liquidity factor is important agrees with their analysis (although it must be noted that their approach is cross-sectional and ours is done using the time-series of returns).

Our finding that liquidity is important is also in agreement with Hwang and Lu (2007), who find that a three-factor model with the excess market return, liquidity and the coskewness factors explains individual stocks' returns at least as well as the Fama-French factors, and that for most sample periods the excess market return and liquidity suffice. Our results, however, do not suggest that coskewness and idiosyncratic volatility play a role in describing the time series of stock returns.

4.4. Conclusion

In this work we have investigated the question of which asset pricing factors should be used in a LFM. We contribute to the literature in three ways. First, we extend a method of the statistics literature known as Stochastic Search Variable Selection introduced by George and McCulloch (1993) to a simple multivariate linear regression case with *N* assets. The resulting MCMC algorithm is simple, fast and can be applied to data on thousands of stocks. The advantage over traditional Bayesian analysis is that we do not need to calculate the posterior probabilities of all models of different sizes, which can be a daunting task since the number of possible factors is large and growing, and the number of possible models grows exponentially with the number of factors. Instead the procedure automatically focuses on the more promising models with higher posterior probabilities.

Second, we use a more comprehensive set of factors than other studies. Most studies consider either macroeconomic factors or factors derived through statistical techniques such as principal components and factor analysis, whereas we consider twelve factor-mimicking portfolios based on risk and firm characteristics, which previous studies suggest are related to several possible sources of risk or empirical irregularities.

Third, we calculate the posterior probabilities of models and factors using thousands of individual stocks. This avoids the potential for data-snooping biases which arises when the portfolios and the factors are based on the same firm characteristics, as argued by Lo and MacKinlay (1990), Ferson, Sarkissian and Simin (1999) and Berk (2000).

We found strong evidence that a LFM to explain stocks returns should include the excess market return, the SMB and the liquidity factors. We found weak evidence that the HML and idiosyncratic volatility factors matter for individual stocks. Interestingly, over the period from 1987 to 1996, we found that the simple CAPM explains stock returns with high posterior probability. Our results using portfolios of stocks differ significantly and suggest caution regarding which factors and portfolios are used. When we use a set of 30 industry portfolios that does not include portfolios formed on size and book-to-market, the posterior probability of the SMB factor remains high, but the HML factor has a low posterior probability. Moreover, these results provide stronger evidence that idiosyncratic volatility matters, and moderate evidence that liquidity and the downside risk factors are important. However, the famous HML factor does not matter.

4.5. Appendix: Detailed Explanation of Prior and Posterior Distributions

Our model can be written in the following form:

$$\tilde{\mathbf{r}} = \tilde{\mathbf{X}}\tilde{\boldsymbol{\beta}}_\gamma + \tilde{\mathbf{e}}, \tag{3.32}$$

where $\tilde{\mathbf{r}} = [\mathbf{r_1}\ \mathbf{r_2} \ldots \mathbf{r_N}]'$, $\tilde{\mathbf{X}} = \mathbf{X} \otimes \mathbf{I_N}$, $\tilde{\boldsymbol{\beta}}_\gamma = \tilde{\boldsymbol{\beta}} \odot \boldsymbol{\gamma}$, where $\odot$ denotes the element-by-element multiplication and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$ is the binary indicator variable for the inclusion of each of the K factors, and $\tilde{\mathbf{e}} = (\mathbf{e_1}'\ \mathbf{e_2}' \cdots \mathbf{e_N}')'$. The collected vector of error terms $\tilde{\mathbf{e}} \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I_T})$, with $\boldsymbol{\Sigma}$ a diagonal matrix such that each $\mathbf{e_i} \sim N(\mathbf{0}, \sigma_i^2 \mathbf{I_T})$. Thus $\boldsymbol{\Sigma}$ is completely determined by $\tilde{\boldsymbol{\sigma}} = (\sigma_1, \ldots, \sigma_N)$.

4.5.1. Prior Distributions

The parameters of the model are $\tilde{\boldsymbol{\beta}}, \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$ and $\sigma_i, i = 1, \ldots, N$. We follow Kuo and Mallick (1998) and choose priors for $\tilde{\boldsymbol{\beta}}, \boldsymbol{\gamma}$ and $\sigma_i, i = 1, \ldots, N$ independently, with $\tilde{\boldsymbol{\beta}} \sim N(\tilde{\boldsymbol{\beta}}_0, \tilde{\mathbf{D}}_0)$, $\gamma_j \sim B(1, p_j)$, $j = 1, \ldots, K$ and $\sigma_i \sim IG(\alpha/2, 2/\eta)$, for $i = 1, \ldots, N$. The parameters of the priors are chosen to reflect our lack of knowledge about $\boldsymbol{\gamma}$, i.e. about which factors should be included in model.

For $\tilde{\boldsymbol{\beta}}$ we choose a prior centred in a vector of zeros, to reflect our lack of knowledge about which factors should enter the model. The prior variance-covariance matrix of $\tilde{\boldsymbol{\beta}}$ can be chosen in several ways. One of the ways is choosing $\tilde{\mathbf{D}}_0 = c\mathbf{I}$. This choice of prior reflects an *a priori* belief on independence of the regression coefficients, and also a large variance (relative to the OLS one) so that this prior is not very informative when c is large. The other specification consists of choosing $\tilde{\mathbf{D}}_0 = c(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}')^{-1}$, which makes the prior covariance structure equal to the design covariance structure, as suggested by Zellner (1971). The parameter c can be varied to quantify the degree of prior uncertainty about the coefficients. Our main results are reported for $c = 10$, and we also use $c = 100$ to study the robustness of the results to the prior.

For the prior of p_j, the probability that factor j is included in the model, we choose an equal probability of ½ for all factors and also for the intercept. This prior reflects the lack of knowledge about the inclusion of the factors, and implies that any model, regardless of its size, has an equal prior probability of $1/2^K$.

For $\sigma_i, i = 1,\ldots,N$ we chose improper priors with $\alpha = \eta = 0$.

4.5.2. Gibbs Sampling Scheme

The main interest in this model is the posterior density $P(\boldsymbol{\gamma} \mid \tilde{\mathbf{r}})$. To do this, we use a Gibbs sampling approach. We start with initial values $\tilde{\boldsymbol{\beta}}^{\mathbf{0}}, \boldsymbol{\gamma}^{\mathbf{0}}, \boldsymbol{\Sigma}^{\mathbf{0}}$, and then sample from conditional densities to generate values $\tilde{\boldsymbol{\beta}}^{\mathbf{1}}, \boldsymbol{\gamma}^{\mathbf{1}}, \boldsymbol{\Sigma}^{\mathbf{1}}$. Similarly to Kuo and Mallick (1998), we initialise $\tilde{\boldsymbol{\beta}}^{\mathbf{0}}$ and $\boldsymbol{\Sigma}^{\mathbf{0}}$ with the individual-assets OLS estimators for the full model, and $\boldsymbol{\gamma}^{\mathbf{0}} = (1\,1 \cdots 1)$. The conditional densities we use are as follows.

4.5.2.1. Conditional Density of $\tilde{\boldsymbol{\beta}}$ given $\boldsymbol{\gamma}, \boldsymbol{\Sigma}, \tilde{\mathbf{r}}$

First, denote by $\mathbf{x_j}$ the j-th column of $\mathbf{X}$, then let $\tilde{\mathbf{X}}^* = \mathbf{X}^* \otimes \mathbf{I_N}$, where $\mathbf{X}^* = [\gamma_1 \mathbf{x_1}\ \gamma_2 \mathbf{x_2} \cdots \gamma_K \mathbf{x_K}]$. Then the conditional density of $\tilde{\boldsymbol{\beta}}$ given $\boldsymbol{\gamma}, \boldsymbol{\Sigma}, \tilde{\mathbf{r}}$ is $N\left(\tilde{\boldsymbol{\beta}}_\mathbf{1}, \mathbf{D_1}\right)$, where the posterior mean $\tilde{\boldsymbol{\beta}}_\mathbf{1} = \left(\mathbf{D_0^{-1}} + \left(\tilde{\mathbf{X}}^{*\prime} \boldsymbol{\Sigma}^{-1} \tilde{\mathbf{X}}^*\right)\right)\left(\mathbf{D_0^{-1}} \tilde{\boldsymbol{\beta}}_\mathbf{0} + \tilde{\mathbf{X}}^{*\prime} \boldsymbol{\Sigma}^{-1} \tilde{\mathbf{r}}\right)$ and the posterior variance-covariance matrix $\mathbf{D_1} = \left(\mathbf{D_0^{-1}} + \left(\tilde{\mathbf{X}}^{*\prime} \boldsymbol{\Sigma}^{-1} \tilde{\mathbf{X}}^*\right)\right)$.

4.5.2.2. Conditional Density of γ_j given $\boldsymbol{\gamma}_{\text{-j}}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\Sigma}, \tilde{\mathbf{r}}$

Let $\boldsymbol{\gamma}_{-j} = \left(\gamma_1 \cdots \gamma_{j-1}\, \gamma_{j+1} \cdots \gamma_K\right)$ respresent the vector of binary indicator variables with the j-th entry removed. Then the conditional probability distribution of γ_j given $\boldsymbol{\gamma}_{\text{-j}}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\Sigma}, \tilde{\mathbf{r}}$ is $B\left(1, p_j^*\right)$, with $p_j^* = c_j \big/ \left(c_j + d_j\right)$, where

$$c_j = p_j \exp\left[-\frac{1}{2}\left(\tilde{\mathbf{r}} - \tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right)\right)' \boldsymbol{\Sigma}^{-1} \left(\tilde{\mathbf{r}} - \tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right)\right)\right], \tag{3.33}$$

and

$$d_j = \left(1 - p_j\right) \exp\left[-\frac{1}{2}\left(\tilde{\mathbf{r}} - \tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^{**}\right)\right)' \boldsymbol{\Sigma}^{-1} \left(\tilde{\mathbf{r}} - \tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^{**}\right)\right)\right]. \tag{3.34}$$

In the expressions above, $\boldsymbol{\gamma}_{\mathbf{j}}^*$ represents the vector $\boldsymbol{\gamma}$ with the j-th entry replaced by 1, and likewise $\boldsymbol{\gamma}_{\mathbf{j}}^{**}$ represents the vector $\boldsymbol{\gamma}$ with the j-th entry replaced by 0. Expressions (3.33) and (3.34) can be simplified in the following way. First, we note that

$$\tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) = \begin{bmatrix} \mathbf{X} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \square \boldsymbol{\gamma}_{\mathbf{j}}^* \\ \boldsymbol{\beta}_2 \square \boldsymbol{\gamma}_{\mathbf{j}}^* \\ \vdots \\ \boldsymbol{\beta}_{\mathbf{N}} \square \boldsymbol{\gamma}_{\mathbf{j}}^* \end{bmatrix} = \begin{bmatrix} \mathbf{X}\left(\boldsymbol{\beta}_1 \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \\ \mathbf{X}\left(\boldsymbol{\beta}_2 \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \\ \vdots \\ \mathbf{X}\left(\boldsymbol{\beta}_{\mathbf{N}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \end{bmatrix}. \tag{3.35}$$

Let QF_1 denote the quadratic form in equation (3.33). Then we can rewrite it as

$$QF_1 = \left(\tilde{\mathbf{r}} - \tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right)\right)' \boldsymbol{\Sigma}^{-1} \left(\tilde{\mathbf{r}} - \tilde{\mathbf{X}}\left(\tilde{\boldsymbol{\beta}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right)\right) =$$

$$\begin{bmatrix} \mathbf{r_1} - \mathbf{X}\left(\tilde{\boldsymbol{\beta}}_1 \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \\ \mathbf{r_2} - \mathbf{X}\left(\tilde{\boldsymbol{\beta}}_2 \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \\ \vdots \\ r_N - \mathbf{X}\left(\tilde{\boldsymbol{\beta}}_{\mathbf{N}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \end{bmatrix}' \times \begin{bmatrix} \sigma_1^{-2}\mathbf{I_T} & 0 & \cdots & 0 \\ \vdots & \sigma_2^{-2}\mathbf{I_T} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^{-2}\mathbf{I_T} \end{bmatrix} \times \begin{bmatrix} \mathbf{r_1} - \mathbf{X}\left(\tilde{\boldsymbol{\beta}}_1 \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \\ \mathbf{r_2} - \mathbf{X}\left(\tilde{\boldsymbol{\beta}}_2 \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \\ \vdots \\ \mathbf{r_N} - \mathbf{X}\left(\tilde{\boldsymbol{\beta}}_{\mathbf{N}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right) \end{bmatrix}$$

$$QF_1 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left(\mathbf{r_i} - \mathbf{X}\boldsymbol{\beta}_{\mathbf{i}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right)' \left(\mathbf{r_i} - \mathbf{X}\boldsymbol{\beta}_{\mathbf{i}} \square \boldsymbol{\gamma}_{\mathbf{j}}^*\right), \tag{3.36}$$

so QF_1 can be calculated as the sum of the residuals for each asset, weighted by the reciprocal of their variances, when variable j is included in the regression. Likewise, we have

$$QF_2 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left(\mathbf{r_i} - \mathbf{X}\boldsymbol{\beta}_{\mathbf{i}} \square \boldsymbol{\gamma}_{\mathbf{j}}^{**}\right)' \left(\mathbf{r_i} - \mathbf{X}\boldsymbol{\beta}_{\mathbf{i}} \square \boldsymbol{\gamma}_{\mathbf{j}}^{**}\right), \tag{3.37}$$

the sum of the residuals for each asset, weighted by the reciprocal of their variances, when variable j is not included in the regression. Using (3.36) and (3.37), we can rewrite p_j^* as

$$p_j^* = \frac{p_j e^{-0.5QF_1}}{p_j e^{-0.5QF_1} + \left(1 - p_j\right)\left(e^{-0.5QF_2}\right)}. \tag{3.38}$$

The expression above, although analytically correct, can present some problems during implementation, since QF_1 and QF_2 can be quite large numbers. To avoid this computational nuisance, we take the inverse of p_j^* twice and get

$$p_j^* = \left(\frac{p_j e^{-0.5QF_1} + \left(1-p_j\right)\left(e^{-0.5QF_2}\right)}{p_j e^{-0.5QF_1}} \right)^{-1} = \left(1 + \frac{1-p_j}{p_j} e^{-0.5\left(QF_2 - QF_1\right)} \right)^{-1}. \quad (3.39)$$

With equation (3.39), the elements of $\boldsymbol{\gamma}$ can be simulated from a binomial distribution $B\left(1, p_j^*\right)$ one at a time, preferably in random order.

4.5.2.3. Conditional Density of $\boldsymbol{\Sigma}$ given $\boldsymbol{\gamma}, \tilde{\boldsymbol{\beta}}, \tilde{\mathbf{r}}$

The matrix $\boldsymbol{\Sigma}$ is completely determined by the volatilities of the error terms of each asset. Given the priors $\sigma_i \sim IG\left(\alpha/2, 2/\eta\right)$, for $i = 1, \ldots, N$, the conditional densities of σ_i, given $\tilde{\boldsymbol{\beta}}, \boldsymbol{\gamma}, \tilde{\mathbf{r}}$ are $IG\left(\frac{\alpha+\eta}{2}, \frac{2}{\eta + \left(\mathbf{r_i} - \mathbf{X\beta_i} \square \boldsymbol{\gamma}\right)' \left(\mathbf{r_i} - \mathbf{X\beta_i} \square \boldsymbol{\gamma}\right)} \right)$.

Table 4-1 Summary statistics of candidate factors

Our data covers the period from January 1967 to December 2005. The excess market return, (the value-weighted return on all NYSE, AMEX and NASDAQ firms from CRSP minus the one-month Treasury bill rate), the Fama and French factors SMB and HML, the momentum factor (Mom.) and the Long-term reversal factor (Rever.) are taken from Professor Kenneth French's Data Library. The data on the Liquidity, Coskewness, Cokurtosis, Downside-risk, Idiosyncratic Volatility and Volume factors are the same from and have been kindly provided by Hwang and Lu (2007). We report average returns and standard errors of the candidate factors on Panel A. Values in bold are significantly different from zero. Panel B reports the Spearman's rank correlation matrix. Bold values are significantly different from zero at 5% significance level.

Panel A. Descriptive Statistics

	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
Average	**0.448**	0.247	**0.448**	**0.812**	**0.366**	0.294	-0.042	**0.356**	**-0.514**	0.175	0.182	0.006
Std error	**0.202**	0.149	**0.134**	**0.185**	**0.132**	0.281	0.091	**0.113**	**0.157**	0.116	0.130	0.224

Panel B. Spearman's rank correlations

	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
Market	1.00											
SMB	**0.29**	1.00										
HML	**-0.40**	**-0.22**	1.00									
Momentum	-0.05	0.00	**-0.09**	1.00								
Asset Growth	**-0.41**	**-0.17**	**0.60**	0.05	1.00							
Idiosyn.Volat.	**-0.61**	**-0.65**	**0.47**	-0.02	**0.43**	1.00						
Volume	-0.06	0.02	0.05	-0.03	0.00	0.06	1.00					
Long-term	**-0.12**	**0.15**	**0.39**	-0.03	**0.35**	0.06	0.02	1.00				
Liquidity	**0.09**	**-0.78**	-0.12	**-0.09**	**-0.14**	**0.32**	-0.01	**-0.21**	1.00			
Coskewness	**0.28**	**0.09**	-0.06	-0.01	**-0.14**	**-0.23**	-0.06	**-0.12**	0.02	1.00		
Cokurtosis	**-0.31**	0.06	**0.33**	**0.17**	**0.28**	**0.18**	0.04	**0.13**	**-0.27**	**-0.21**	1.00	
Down. Risk	**-0.67**	**-0.28**	**0.57**	**-0.12**	**0.50**	**0.68**	0.05	**0.25**	-0.06	**-0.24**	**0.35**	1

Table 4-2 Exploratory Analysis of the Factors Using OLS

Our data covers the period from January 1967 to December 2006. The excess market return, (the value-weighted return on all NYSE, AMEX and NASDAQ firms from CRSP minus the one-month Treasury bill rate), the Fama and French factors SMB and HML, the momentum factor and the Long-term reversal factor are taken from Professor Kenneth French's Data Library. The data on the Liquidity, Coskewness, Cokurtosis, Downside-risk, Idiosyncratic Volatility and Volume factors are the same from and have been kindly provided by Hwang and Lu (2007). Each column of the table reports the results of the regression of a factor into a constant and all the remaining factors. The table shows the regression coefficients, with bold values representing significance at the 5% significance level.

	Market	SMB	HML	Momentum.	Asset Growth	Idiosyn. Volatility	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
constant	**0.006**	0.000	0.002	**0.008**	0.000	**0.004**	0.000	0.001	-0.001	0.002	-0.001	-0.002
Market		**0.097**	**0.074**	-0.033	-0.009	**-0.113**	-0.021	0.012	**0.150**	0.048	-0.001	**-0.220**
SMB	**0.702**		**-0.459**	**-0.456**	**-0.157**	**-0.976**	0.154	**0.405**	**-1.044**	-0.079	0.094	0.196
HML	**0.260**	**-0.221**		**-0.549**	**0.381**	**0.254**	**0.118**	**0.396**	**-0.293**	**0.268**	**0.214**	**0.197**
Momentum	-0.024	**-0.045**	**-0.113**		**0.088**	0.040	0.000	0.026	**-0.063**	-0.044	**0.149**	**-0.129**
Asset Growth	-0.023	**-0.053**	**0.271**	**0.303**		0.098	-0.029	0.087	**-0.098**	-0.054	0.013	**0.159**
Idiosyn.Volatility	**-0.141**	**-0.168**	**0.091**	0.070	0.049		0.027	**-0.072**	-0.032	-0.052	-0.006	**0.429**
Volume	-0.054	0.054	**0.086**	-0.001	-0.030	0.054		-0.064	0.050	-0.044	0.021	0.020
Long-term rev.	0.029	**0.133**	**0.270**	0.088	0.084	**-0.138**	-0.060		**0.081**	**-0.126**	-0.069	**0.200**
Liquidity	**0.689**	**-0.658**	**-0.385**	**-0.400**	**-0.181**	-0.119	0.091	**0.156**		-0.066	-0.138	-0.029
Coskewness	0.095	-0.021	**0.152**	-0.120	-0.043	-0.083	-0.035	**-0.105**	-0.029		**-0.231**	**-0.179**
Cokurtosis	-0.001	0.025	**0.117**	**0.394**	0.010	-0.010	0.016	-0.055	-0.057	**-0.223**		**0.197**
Downside Risk	**-0.340**	0.042	**0.087**	**-0.277**	**0.099**	**0.529**	0.012	**0.130**	-0.010	**-0.140**	**0.159**	
R^2	0.509	0.876	0.683	0.184	0.538	0.798	0.034	0.342	0.823	0.245	0.377	0.742

Table 4-3 Marginal posterior probabilities of factors and model posterior probability for individual stocks

Our data covers the period from January 1967 to December 2005. We divide it into four subsamples of 120 months each and apply our extension of the Stochastic Search Variable Selection Markov Chain Monte Carlo algorithm of Kuo and Mallick (1998) to all NYSE, AMEX and NASDAQ firms from CRSP. We exclude firms whose sizes are smaller than the 5th percentile of the whole NYSE universe at the beginning of each subsample and stocks with price inferior to U$5. We calculate marginal posterior probabilities of each factor as the mean of the latent variable corresponding to the factor. We also report the three models with highest posterior probabilities. Panels A and B report the marginal posterior probabilities of factors and the highest posterior probability models, respectively, with each subsample identified by a number following the panel letter.

Panel A.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	1.000	0.010	0.001	0.002	0.222	0.001	0.005	1.000	0.001	0.001	0.115

Panel A.2 Ten highest posterior probability models, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	0	0	0	1	0	0	0	0.778
0	1	1	0	0	0	1	0	0	1	0	0	1	0.106
0	1	1	0	0	0	1	0	0	1	0	0	0	0.106

Panel B.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986, number of stocks = 1146

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	1.000	0.007	0.001	0.001	0.696	0.000	0.001	1.000	0.001	0.001	0.031

Panel B.2 Ten highest posterior probability models, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	1	0	0	1	0	0	0	0.664
0	1	1	0	0	0	0	0	0	1	0	0	0	0.304
0	1	1	0	0	0	1	0	0	1	0	0	1	0.025

Table 4-3 (continued)

Panel C.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1987-1996, number of stocks = 2159

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	0.216	0.004	0.001	0.002	0.017	0.001	0.003	0.216	0.000	0.000	0.002

Panel C.2 Ten highest posterior probability models, sample period 1987-1996, number of stocks = 2159

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	0	0	0	0	0	0	0	0	0	0	0	0.783
0	1	1	0	0	0	0	0	0	1	0	0	0	0.196
0	1	1	0	0	0	1	0	0	1	0	0	0	0.014

Panel D.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1997-2006, number of stocks = 2787

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.000	1.000	0.576	0.625	0.002	0.019	0.376	0.001	0.002	0.556	0.004	0.016	0.067

Panel D.2 Ten highest posterior probability models, sample period 1997-2006, number of stocks = 2787

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	0	0	0	0	0	0	0	0	0	0	0	0.375
0	1	1	1	0	0	1	0	0	1	0	0	0	0.309
0	1	1	1	0	0	0	0	0	1	0	0	0	0.180

Table 4-4 Marginal posterior probabilities of candidate factors and model posterior probabilities for 55 portfolios

Our data covers the period from January 1967 to December 2005. The data on all portfolios have been obtained from Kenneth French's website. The portfolios used are the 25 portfolios formed on market equity and book-to-market plus 30 industry portfolios. We divide our sample into 4 subsamples of 120 months each and apply our extension of the Stochastic Search Variable Selection Markov Chain Monte Carlo algorithm of Kuo and Mallick (1998) to the portfolios. The prior probability of inclusion of each factor is 0.5. Our results are based on 10000 iterations. We calculate marginal posterior probabilities of each factor as the mean of the latent variable corresponding to the factor. Panels A and B report the marginal posterior probabilities of factors and the highest posterior probability models, respectively, with each subsample identified by a number following the panel letter.

Panel A.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1967-1976

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.163	1.000	1.000	1.000	0.002	0.031	1.000	0.001	0.028	0.651	0.006	0.005	1.000

Panel A.2 Ten highest posterior probability models, sample period 1967-1976

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	1	0	0	1	0	0	1	0	0	1	0.496
0	1	1	1	0	0	1	0	0	0	0	0	1	0.340
1	1	1	1	0	0	1	0	0	1	0	0	1	0.118

Panel B.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986, number of stocks = 1146

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	1.000	0.007	0.001	0.001	0.696	0.000	0.001	1.000	0.001	0.001	0.031

Panel B.2 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
1	1	1	1	0	0	1	0	0	1	0	0	0	0.817
1	1	1	1	0	0	1	0	0	1	0	0	1	0.161
1	1	1	1	0	0	1	0	0	1	1	0	1	0.008

Table 4-4 (continued)

Panel C.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1987-1996

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.409	1.000	1.000	1.000	0.002	0.026	0.481	0.002	0.058	1.000	0.013	0.002	0.055

Panel C.2 Ten highest posterior probability models, sample period 1987-1996, number of stocks = 2159

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	1	0	0	0	0	0	1	0	0	0	0.406
1	1	1	1	0	0	1	0	0	1	0	0	0	0.229
0	1	1	1	0	0	1	0	0	1	0	0	0	0.176

Panel D.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1997-2006

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.005	1.000	1.000	1.000	0.015	0.619	1.000	0.002	0.015	1.000	0.012	0.164	1.000

Panel D.2 Ten highest posterior probability models, sample period 1997-2006, number of stocks = 2787

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	1	0	1	1	0	0	1	0	0	1	0.456
0	1	1	1	0	0	1	0	0	1	0	0	1	0.379
0	1	1	1	0	1	1	0	0	1	0	1	1	0.139

Table 4-5 Marginal posterior probabilities of candidate factors and model posterior probabilities for 30 industry portfolios

Our data covers the period from January 1967 to December 2005. The data on all portfolios have been obtained from Kenneth French's website. The portfolios used are 30 industry portfolios. We divide our sample into 4 subsamples of 120 months each and apply our extension of the Stochastic Search Variable Selection Markov Chain Monte Carlo algorithm of Kuo and Mallick (1998) to the portfolios. The prior probability of inclusion of each factor is 0.5. Our results are based on 10000 iterations. We calculate marginal posterior probabilities of each factor as the mean of the latent variable corresponding to the factor. Panels A and B report the marginal posterior probabilities of factors and the highest posterior probability models, respectively, with each subsample identified by a number following the panel letter.

Panel A.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1967-1976

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.092	1.000	1.000	0.399	0.003	0.019	0.913	0.000	0.022	0.197	0.001	0.006	0.756

Panel A.2 Ten highest posterior probability models, sample period 1967-1976

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	1	0	0	0	0	0	1	0.322
0	1	1	1	0	0	1	0	0	0	0	0	1	0.157
0	1	1	0	0	0	1	0	0	0	0	0	0	0.109

Panel B.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986, number of stocks = 1146

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.441	1.000	1.000	0.315	0.023	0.003	1.000	0.003	0.009	0.711	0.015	0.003	0.119

Panel B.2 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	1	0	0	1	0	0	0	0.295
0	1	1	0	0	0	1	0	0	0	0	0	0	0.179
1	1	1	0	0	0	1	0	0	1	0	0	0	0.136

Table 4-5 (continued)

Panel C.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1987-1996

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.034	1.000	0.365	0.255	0.005	0.045	0.266	0.002	0.128	0.153	0.006	0.007	0.020

Panel C.2 Ten highest posterior probability models, sample period 1987-1996, number of stocks = 2159

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	0	0	0	0	0	0	0	0	0	0	0	0.519
0	1	1	0	0	0	1	0	0	0	0	0	0	0.073
0	1	1	0	0	0	0	0	0	0	0	0	0	0.063

Panel D.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1997-2006

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.056	1.000	0.859	1.000	0.026	0.057	0.918	0.001	0.005	0.849	0.032	0.434	1.000

Panel D.2 Ten highest posterior probability models, sample period 1997-2006, number of stocks = 2787

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	1	0	0	1	0	0	1	0	1	1	0.351
0	1	1	1	0	0	1	0	0	1	0	0	1	0.318
0	1	0	1	0	0	1	0	0	0	0	0	1	0.132

Table 4-6 Marginal posterior probabilities of factors and model posterior probability for individual stocks ($c = 10$)

Our data covers the period from January 1967 to December 2005. We divide it into four subsamples of 120 months each and apply our extension of the Stochastic Search Variable Selection Markov Chain Monte Carlo algorithm of Kuo and Mallick (1998) to all NYSE, AMEX and NASDAQ firms from CRSP. We exclude firms whose sizes are smaller than the 5%-tile of the whole NYSE universe at the beginning of each subsample. We also exclude stocks with price inferior to U$5. The prior probability of inclusion of each factor is 0.5. Our results are based on 10000 MCMC iterations. We calculate marginal posterior probabilities of each factor as the mean of the latent variable corresponding to the factor. We also report the three models with highest posterior probabilities. Panels A and B report the marginal posterior probabilities of factors and the highest posterior probability models, respectively, with each subsample identified by a number following the panel letter.

Panel A.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	0.899	0.013	0.001	0.002	0.293	0.001	0.002	0.653	0.003	0.001	0.029

Panel A.2 Ten highest posterior probability models, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	0	0	0	1	0	0	0	0.414
0	1	1	0	0	0	1	0	0	1	0	0	0	0.209
0	1	1	0	0	0	0	0	0	0	0	0	0	0.192

Panel B.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986, number of stocks = 1146

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	1.000	0.008	0.001	0.000	0.221	0.001	0.001	1.000	0.003	0.000	0.091

Panel B.2 Ten highest posterior probability models, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	0	0	0	1	0	0	0	0.726
0	1	1	0	0	0	1	0	0	1	0	0	0	0.183
0	1	1	0	0	0	0	0	0	1	0	0	1	0.053

Table 4-6 (continued)

Panel C.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1987-1996

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	0.260	0.004	0.001	0.002	0.031	0.000	0.001	0.262	0.001	0.001	0.004

Panel C.2 Ten highest posterior probability models, sample period 1987-1996, number of stocks = 2159

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	0	0	0	0	0	0	0	0	0	0	0	0.736
0	1	1	0	0	0	0	0	0	1	0	0	0	0.227
0	1	1	0	0	0	1	0	0	1	0	0	0	0.025

Panel D.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1997-2006

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	0.626	0.526	0.006	0.026	0.699	0.001	0.000	0.608	0.007	0.013	0.071

Panel D.2 Ten highest posterior probability models, sample period 1997-2006, number of stocks = 2787

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	1	0	0	1	0	0	0	0.283
0	1	0	1	0	0	1	0	0	0	0	0	0	0.223
0	1	0	0	0	0	0	0	0	0	0	0	0	0.126

Table 4-7 Marginal posterior probabilities of factors and model posterior probability for individual stocks (c = 100)

Our data covers the period from January 1967 to December 2005. We divide it into four subsamples of 120 months each and apply our extension of the Stochastic Search Variable Selection Markov Chain Monte Carlo algorithm of Kuo and Mallick (1998) to all NYSE, AMEX and NASDAQ firms from CRSP. We exclude firms whose sizes are smaller than the 5%-tile of the whole NYSE universe at the beginning of each subsample. We also exclude stocks with price inferior to U$5. The prior probability of inclusion of each factor is 0.5. Our results are based on 10000 MCMC iterations. We calculate marginal posterior probabilities of each factor as the mean of the latent variable corresponding to the factor. We also report the three models with highest posterior probabilities. Panels A and B report the marginal posterior probabilities of factors and the highest posterior probability models, respectively, with each subsample identified by a number following the panel letter.

Panel A.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.000	1.000	0.735	0.011	0.001	0.001	0.229	0.000	0.008	0.641	0.000	0.001	0.175

Panel A.2 Ten highest posterior probability models, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	0	0	0	1	0	0	0	0.378
0	1	0	0	0	0	0	0	0	0	0	0	0	0.244
0	1	1	0	0	0	1	0	0	1	0	0	1	0.129

Panel B.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1977-1986, number of stocks = 1146

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	0.834	0.003	0.002	0.001	0.282	0.000	0.002	0.831	0.003	0.000	0.159

Panel A.2 Ten highest posterior probability models, sample period 1967-1976, number of stocks = 1141

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	1	0	0	0	0	0	0	1	0	0	0	0.435
0	1	1	0	0	0	1	0	0	1	0	0	0	0.237
0	1	0	0	0	0	0	0	0	0	0	0	0	0.166

Table 4-7 (continued)

Panel C.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1987-1996

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.000	1.000	0.291	0.007	0.001	0.004	0.026	0.001	0.002	0.289	0.002	0.001	0.004

Panel C.2 Ten highest posterior probability models, sample period 1987-1996, number of stocks = 2159

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	0	0	0	0	0	0	0	0	0	0	0	0.708
0	1	1	0	0	0	0	0	0	1	0	0	0	0.262
0	1	1	0	0	0	1	0	0	1	0	0	0	0.017

Panel D.1 Marginal posterior probabilities of the inclusion of each factor, sample period 1997-2006

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk
0.001	1.000	0.458	0.200	0.002	0.025	0.330	0.000	0.001	0.443	0.009	0.018	0.044

Panel D.2 Ten highest posterior probability models, sample period 1997-2006, number of stocks = 2787

Intercept	Market	SMB	HML	Momentum	Asset Growth	Idiosyn. Volat.	Volume	Long-term reversal	Liquidity	Coskewness	Cokurtosis	Downside Risk	Posterior probability
0	1	0	0	0	0	0	0	0	0	0	0	0	0.481
0	1	1	0	0	0	1	0	0	1	0	0	0	0.135
0	1	1	0	0	0	0	0	0	1	0	0	0	0.118

References

Ang, Andrew, Joseph Chen, and Yuhang Xing, 2006, Downside risk, *Review of Financial Studies* 19, 1191-1239.

Ang, Andrew, Bob Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, High idiosyncratic volatility and low returns: International and further U.S. Evidence, (Columbia Business School).

Bali, Turan G., Nusret Cakici, Xuemin (Sterling) Yan, and Zhe Zhang, 2005, Does idiosyncratic risk really matter?, *The Journal of Finance* 60, 905-929.

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

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

Berk, Jonathan B., 2000, Sorting out sorts, *The Journal of Finance* 55, 407-427.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *The Journal of Business* 45, 444-455.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected returns, *Journal of Financial Economics* 49, 345-373.

Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441-464.

Chan, K.C., N. Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *The Journal of Finance* 51, 1681-1713.

Chen, Na Fui, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *The Journal of Business* 59, 383-403.

Chordia, Tarun, Avanidhar Subrahmanyam, and R. V. Anshuman, 2001, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3-32.

Cochrane, John, 2001. *Asset pricing* (Princeton University Press, Princeton and Oxford).

Connor, Gregory, 1984, A unified beta pricing theory, *Journal of Economic Theory* 34, 13-31.

Connor, Gregory, and Robert A. Korajczyk, 1988, Risk and return in equilibrium apt: Application of a new tests metrology, *Journal of Financial Economics* 21, 255:289.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2006, What best explains the cross-section of stock returns? Exploring the asset growth effect, AFA 2007 Chicago Meetings Paper.

Cripps, Edward, C. K. Carter, and R. Kohn, 2003, Variable selection and covariance selection in multivariate regression models, Technical report 2003-14 (Statistical and Applied Mathematical Sciences Institute).

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *The Journal of Finance* 53, 1839-1885.

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

Daniel, Kent, and Sheridan Titman, 2005, Testing factor-model explanations of market anomalies, *working paper*.

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

Dittmar, R. F., 2002, Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns, *The Journal of Finance* 52, 369-404.

Dybvig, P., 1985, An explicit bound on individual assets' deviations from apt, *Journal of Financial Economics* 12, 483-496.

Ericsson, Johan, and Sune Karlsson, 2004, Choosing factors in a multifactor asset pricing: A Bayesian approach, (Stockholm School of Economics).

Fama, Eugene F., 1991, Efficient capital markets, *The Journal of Finance* 52, 1-33.

Fama, Eugene F., and Kenneth R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246-273.

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

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns of stocks and bonds, *The Journal of Financial Economics* 33, 357-384.

Fama, Eugene F., and Kenneth R. French, 1996, The CAPM is wanted, dead or alive, *The Journal of Finance* 51, 1947-1958.

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

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *The Journal of Political Economy* 81, 607-636.

Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin, 1999, The alpha factor asset pricing model: A parable, *Journal of Financial Markets* 2, 49-68.

George, Edward I., and Robert E. McCulloch, 1993, Variable selection via Gibbs sampling, *Journal of the American Statistical Association* 88, 881-889.

Gervais, Simon, Ron Kaniel, and Dan H. Mingelgrin, 2001, The high-volume return premium, *The Journal of Finance* 56, 877–919.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov, 2005, There is a risk-return trade-off after all, *Journal of Financial Economics* 76, 509-548.

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

Goyal, Amit, and Pedro Santa-Clara, 2003, Idiosyncratic risk matters!, *The Journal of Finance* 58, 975-1007.

Grinblatt, M., and Sheridan Titman, 1985, Factor pricing in a finite economy, *Journal of Financial Economics* 12, 497-507.

Hall, Anthony D., S. Hwang , and S. E. Satchell, 2002, Using Bayesian variable selection methods to choose style factors in global stock return models, *Journal of Banking and Finance* 26, 2301-2325.

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

Harvey, C. R., and A. Siddique, 2000, Conditional skewness in asset pricing tests, *The Journal of Finance* 50.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of Finance* 54, 2143-2184.

Hwang , S., and C. Lu, 2007, Too many factors! Do we need them all?, (Sir John Cass Business School, London UK).

Hwang , S., and A. Rubesam, 2007, Disappearance of momentum, (Sir John Cass Business School, London UK).

Hwang , S., and S. E. Satchell, 1999, Modelling emerging market risk premia using higher moments, *International Journal of Finance and Economics* 4.

Hwang , S., and S. E. Satchell, 2007, Testing linear factor models on individual stocks using the average F test *working paper*.

Hwang, S., S. E. Satchell, and Anthony D. Hall, 2002, Using bayesian variable selection methods to choose style factors in global stock return models, *Journal of Banking and Finance* 26, 2301-2325.

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

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

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

Kraus, A., and R. H. Litzenberger, 1976, Skewness preference and the valuation of risky assets, *The Journal of Finance* 31, 1085-1100.

Kuo, Lynn, and Bani Mallick, 1998, Variable selection for regression models, *Sankhya* 60, 65-81.

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

Lehman, Bruce N., and David M. Modest, 1988, The empirical foundation of the arbitrage pricing theory, *Journal of Financial Economics* 21, 213:254.

Lettau, M., and S. Ludvigson, 2001, Resurrecting the (c)CAPM: A crosssectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238-1287.

Lewellen, J., S. Nagel, and Jay Shanken, 2007, A skeptical appraisal of asset pricing tests.

Liew, Jimmy, and Maria Vassalou, 2000, Can book-to-market, size and momentum be risk factors that predict economic growth?, *Journal of Financial Economics* 57, 221-245.

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

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

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

Perez-Quiros, Gabriel, and Allan Timmermann, 2000, Firm size and cyclical variations in stock returns, *The Journal of Finance* 55, 1229-1262.

Petkova, R., 2006, Do the Fama-French factors proxy for innovations in predictive variables, *The Journal of Finance* 61, 581-612.

Post, T., and P. Van Vliett, 2005, Conditional downside risk and the CAPM.

Poterba, J. M., and L. H. Summers, 1988, Mean reversion in stock returns: Evidence and implications, *Journal of Financial Economics* 22, 27-59.

Roll, Richard W., 1977, A critique of asset pricing theory's tests, part 1: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129-176.

Rosenberg, B, K. Reid, and R. Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9-11.

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

Smith, M., and R. Kohn, 1996, Nonparametric regression via Bayesian variable selection, *Journal of Econometrics* 75, 317-344.

Zellner, A., 1971. *An introduction to Bayesian inference in econometrics* (John Wiley & Sons).