PERFORMANCE EVALUATION OF THE UK EQUITY UNIT TRUSTS: DOES ACTIVE MANAGEMENT ADD VALUE?

by

JIEJUN YU

A thesis submitted to

The University of Birmingham

For the degrees of

DOCTOR OF PHILOSOPHY

Department of Economics
The University of Birmingham

ACKNOWLEDGEMENTS

I would like to thank my supervisors Professor David Dickinson, Dr. Marco Barassi, my examiners Nicholas Horsewood, Dr. Fabio Spagnolo for their tuition, advice and great help. I would also want to thank my mum Yuee Hu, without whom I would not have the opportunity to study abroad.

TABLE OF CONTENTS

CHAPTER 4 SEARCHING FOR ALPHAS IN THE UK-- CONDITIONAL PERFORMANCE EVALUATION OF UNIT TRUSTS

CHAPTER 5 DOES STYLE MATTER? – STYLE PERFORMANCE AND PROFITABILITY OF ROTATION STRATEGIES

CHAPTER 1 INTRODUCTION AND THE DATA

Financial assets under management (AUM) by professional fund managers globally have increased at a rapid pace. Conventional assets managed by global fund management firms have grown to US$ 61.6 trillion in 2008 from US$ 37.7 trillion in 2002[1]. Among them, pension assets accounted for US$ 24 trillion of the total amounts, together with US$ 18.9 trillion invested in mutual funds and US$ 18.7 trillion in insurance funds. If including alternative assets (sovereign wealth funds, hedge funds, private equity funds and exchange traded funds), total assets of the global fund management industry stood at US$ 90 trillion at the end of 2008.

In the UK, based on the annual asset management survey[2] by Investment Management Association (IMA) in December 2006, the assets under management[3] by IMA members in the UK have increased from £ 2.16 trillion to £ 3.1 trillion, with 43.5% of growth rate since 2005. Including a range of funds run by non-IMA members, the survey estimates that total assets managed in the UK exceed £3.4 trillion.

Typically, managed funds in the U.S. and unit trusts in the U.K. represent some of the fastest growing types of financial intermediaries in the world economy. The problems of why the asset management industry developed so quickly and if portfolio managers have superior ability to outperform the market remain as top research interests in the field of finance research.

[1] Source: International Financial Services London

[2] The survey is based on questionnaire responses from 69 firms. As about 90% of total assets is managed by IMA members, this survey is regarded as the one of the most representative surveys of the UK asset management industry.

[3] The money the members of IMA manage is in a wide variety of investment vehicles including authorised investment funds, pension funds and stocks and shares ISAs.

There are several reasons why it is important and useful to examine the performance of managed funds. Firstly, because of different histories of funds as well as the performance track records, which could be sample-specific and driven by institutional arrangements peculiar to a specific fund, it is sensible to analyse and identify the different characteristic of fund performance during various periods.

Secondly, prominent claims about how the stock market functions can be dated back to Eugene Fama (1965), with his well-known 'Efficient Markets Hypothesis' ('EMH'). This hypothesis suggests that the stock market is efficient: a market in which asset prices always 'fully reflect' available information. Fama's claim was accepted by much of the financial community. Further, Burton Malkiel (1992) suggests 'a random walk' interpretation of the 'EMH', that is, 'stock prices follow a random behaviour and therefore no trading rules could outperform random decisions'. It suggests that financial assets fully reflect past historical information such that there is no investment strategy which can return abnormal profits based on a previous sequence of prices. In another word, asset prices are unpredictable and the market is unbeatable. Therefore, assessing the existence and persistence of fund managerial ability are important tests for the 'EMH'; the evidence of persistent ability of fund outperformance would support a rejection of 'EMH'.

Finally, it also has implications for the structure and the existence of the fund management industry. The IMA survey (2004) states that '*samples from 2003 and 2004 indicate that more assets were being managed using peer group and passive mandates.*' It suggests a trend in UK asset management industry that the assets under management are moving towards passive investments from active investments. The evaluation of UK unit trusts performance helps to understand the reasons behind such a trend by testing if active fund managers as a group can add value to the portfolios they manage or whether they merely generate wasteful transaction costs through their active management. On the other hand, at the micro level, private investors also want to know how to select a better portfolio manager, who is capable of adding value to the portfolio he/she manages. The recent financial crisis has made policymakers put financial regulatory proposals on top of their agendas, focusing on issues such as bank capital requirements, derivative instruments, and bankers' remuneration structure. Our study helps to provide empirical evidence and question portfolio managers' compensation schemes within the context of UK unit trusts.

Academic researchers have produced abundant research to explore mutual fund performance. In general, the performance evaluation literature can be divided into two strands: one tends to discover the empirical results, i.e. if and to what extent mutual funds do outperform; another focuses on the theoretical development. Although most literature looks at both aspects, the majority tends to detect if portfolio managers have the ability to outperform their counterparts, i.e. they focus on the results of empirical tests as well as the relevant interpretations to such results. To achieve this, they apply different kinds of models and are keen to seek the most accurate models with the smallest pricing errors and the least-biased results.

Another stream of literature emphasises the development and discovery of better fairly-priced asset pricing models, the application of econometric methods or other relevant theoretical research. Since Lehmann and Modest (1987), there have been studies to examine how different measures may rank funds differently. Chen and Knez (1996) provide a general theoretical framework in order to impose minimal market condition on the measurement, and to examine if the various set of measurements are compatible with so-called admissible portfolio performance measure. Ferson *et al.* (2000) apply an experiment to artificial funds to discover the extent to which the different measurement models can detect truly superior performance. Further, several authors (i.e. Pastor and Stambaugh (2001, 2002), Busse and Paul Irvine (2002), Baks, Metrick, and Wachter (2001)) have discovered the use of Bayesian estimators[4] and 'shrinkage estimator' to estimate the mutual fund alphas, with a main purpose of understanding the advantages[5] and drawbacks of the methods using the shrinkage and Bayesian estimators.

We believe both aspects of previous research as mentioned above are interesting. The thesis therefore aims to seek the empirical results of mutual fund performance on the one hand and explore the least biased pricing models on the other in order to carry out such empirical tests more accurately. The intention here however, is not to reinvent a set of new asset pricing models for performance evaluation purpose. Asset pricing theories, for example, those of Sharpe (1964), Lintner (1965), Black (1972), Merton (1973), suggest that the expected return on a financial asset is a function of its covariances (betas) with several systematic risk factors. These theories have been tested extensively in the finance literature methodologies as so-called traditional

---

[4] The Bayesian measure uses factor returns in prior periods combined with a flexible set of prior beliefs about managerial skill and the validity of certain asset pricing model to predict future fund performance. It combines prior investors' beliefs about the accuracy of the pricing model and managerial skill with the information in the data and produces posterior distribution of fund alphas.

[5] It is argued there are several significant advantages inherited in Bayesian's methods. Firstly, the conclusions based on Bayesian approach may be more informative for investors than those based on the traditional statistical methods. Secondly, Bayesian approaches take into account investors' beliefs about managerial skills in combination with managerial fees. Therefore, as suggested by Busse and Irvine (2002), such beliefs greatly affect the ability of Bayesian alphas to predict future performance. Thirdly, Bayesian approach is convenient for incorporating the information about assets with long history into the performance evaluation of funds with short return history. For example, Pastor and Stambaugh (2002) use long-horizon factor returns to provide more precise estimates of the moments of correlated short-horizon factor returns.

methodologies. According to these methodologies, a data generating process is proposed first for the returns, and then the restrictions imposed by an asset pricing model are tested as parametric constraints on the return generating process. However, this approach has a potential problem. That is, the test results could be misleading if the proposed return generating process is mis-specified. The formulation of the more recently developed Stochastic Discount Factor (SDF) model estimates the parameters and tests the pricing implications without a specified model of how the financial asset returns are generated. The main advantage of the SDF model compared to the traditional methodologies is that the SDF model is a very general methodology and requires fewer assumptions and parameters. A detailed literature review on empirical results and discussion of the methodologies employed by previous researchers can be found in chapter 2, in which we examine the merits and disadvantages of pricing methodologies employed in the past and conclude that the SDF model is the optimal candidate for empirical analysis of fund performance.

Notably, a large amount of literature has analyzed performance of US mutual funds, whereas fewer studies have investigated whether the US findings can be carried over to the other markets or to the other asset categories. That is to say, evaluation of mutual funds performance in European markets remains largely a relatively under-explored research area. More specifically, academic research with a focus on the UK market has been relatively light. As one of the most important international financial markets, the UK asset management industry certainly is in needs of more in-depth analysis.

Draper (1989) was the first to provide a review of work on the UK unit trust industry. Ashton (1996) then reviewed the power of tests of fund performance. His results suggested that the conflicts between the need to have a sample period likely to afford any statistical power and the shortness of the tenure of a fund manager mean that it could be difficult to discern superior performance. More recent work on the UK managed funds include Brown, Draper and McKenzie (1997), Quigley and Sinquefield (1998), Lunde, Blake and Timmerman (1998), Wood Mackenzie Company (1999), Giles and Worboys (2002) and Fletcher (1999) (2004) (2005) (2006).

According to author's knowledge, relatively few people have applied the SDF primitive efficient model to evaluate portfolio performance in the UK market. The main work on the UK unit trusts performance within the SDF framework is from Fletcher (2004), who examines the performance of the UK unit trusts between January 1982 and December 1996 using nine different kinds of SDF models. We therefore believe there are further room for academic exploration in this research field. This thesis aims to extend the portfolio performance evaluation literature in the U.K unit trusts market. We identify the following research topics particularly interesting and have remained under-explored: within the context of UK unit trusts and the SDF framework, the effects of Generalized Method of Moments (GMM) weighting matrices choices on performance evaluation was not well examined, however we believe it is an important issue in that the small sample distortion will give biased test results without identifying the optimal estimation estimator. Incorporating conditioning information is essential for a rigorous performance evaluation process, there is lack of research focusing on this issue using the SDF primitive efficient models, we also believe the impact of different choices of information variables on conditional evaluation is interesting; the impact of styles on performance is always intriguing as well as the question if style-rotation strategies can generate profitable returns; performance persistence analysis helps financial market participants to understand that the past track record of a portfolio manager may not be an important factor when they make fund-selection decisions. We aim to contribute to existing research by filling the gaps mentioned above. In more details, the thesis aims to shed light on the following questions:

● The effects of the use of different GMM estimators with various weighting matrices on the performance evaluation: There are several ways to construct the GMM estimators to estimate the SDF models, the question is which estimator exhibits the optimal small sample property?

● Conditional performance evaluation within the SDF primitive efficient model framework: If the UK unit trusts can outperform its benchmark? What if we take conditional measures into account?

● The impact of different choices of information variables on conditional evaluation: which information variable helps to provide the smallest pricing errors?

● Style performance and style-rotation strategies: do different types of unit trusts perform differently? Can any profit be made based on style performance patterns?

● The performance persistence tests within the SDF framework: do unit trusts perform consistently? over short term or long term? What can be the explanations?

Before a formal investigation of performance of the UK unit trusts, we firstly provide a brief introduction to the UK asset management industry, in particular, the UK unit trusts. We also explain the data that will be used in our projects and provide a brief discussion on the descriptive results of the data. Finally, we conclude and explain the structure of the thesis.

1. Overview and the trend of UK asset management industry

The UK plays an important role in international asset management industry. Along with the US and Japan, the UK is one of the largest markets in the world for fund management. It has a strong international orientation and attracts lots of overseas funds.

Within the UK, London is one of the leading international financial centres. Edinburgh and Glasgow are also important centres for fund management. In the past decades, London has remained as the centre for core asset management in Europe within three important parts of the asset management's value chain: core asset management, marketing & distribution and middle & back office, according to IMA 2005 questionnaire study[6], which examines location choices in the asset management industry. The factors contributing to the UK's success include mainly: the liquidity of the UK's capital markets, superior financial infrastructure as well as its qualified labour pool.

The asset management industry has played an important role in the UK economy. While there is no official figure for the contribution of fund management to UK GDP, IFSL (International Financial Services London, 2008) has made an estimate by applying cost margin to total asset under management in the UK. According to this

[6] See *The future of UK Asset Management: competitive position and location choice* Oxera Consulting Ltd & IMA, May 2005

measure, asset management generated around 0.65% of UK GDP or £8.4bn in 2008. It was certainly an important component of the financial sector's overall contribution of around 8.3% in that year. Asset management's wider contribution to the economy stems from its promotion of the UK's capital market and from the links the asset managers have with other financial services providers, such as banks, securities dealers and information providers.

The UK asset management industry is also highly international, with 27% of these assets (over £800 bn) are managed on behalf of overseas clients, highlighting the importance of the UK as one of the important international finance centres. The 2006 IMA survey suggests that such international opportunities are arising thanks to '*diminishing international regulatory barriers, the trend towards open architecture, the creation of new government asset pools and the gradual increase in individual savings pools.'*

The IMA survey also examines other characteristics of the industry, including degree of consolidation, its ownership, client types and asset allocation.

- **Consolidation:** As the survey suggests, the industry remains relatively unconcentrated, as the share of the ten largest firms[7] stands at 48%. In more details, 17 IMA member firms each managed in excess of £50 bn, 75 firms each managed less than £16 bn, 28 firms each managed less than £1 bn. there were also few signs of consolidation momentums in the UK industry in 2006. For example, the merger of BlackRock and Merrill Lynch Investment

[7] The 10 largest firms are: Legal& General Investment management, Barclays Global Investors; State street global advisors, M&G securities, Morley Fund management, JP Morgan asset management, Standard Life Investments, Scottish Widows Investment partnership; Insight Investment management and Blackrock Investment management.

Managers. Others include private equity involvement in the asset management companies, with the management buyouts of Gartmore from Nationwide Mutual. It may well indicate that given the essence of human capital to the asset management industry, merger and acquisition activity was more complicated than might be the case in other sectors. Organic growth (growth through the purchase of assets rather than asset management companies) may make more sense.

- **Ownership:** In terms of firm ownership, insurers remain the single largest parent group, followed by groups whose sole business is asset management.

- **Client types:** In terms of client types, institutional assets under management account for 77% of the total, with the largest segments being corporate pension funds (28%) and insurance funds (27%), followed by retail assets (21%). Private client accounts for a merely 2%.

- **Overall asset allocation:** In terms of asset allocation, the IMA survey (2006) suggests 52.4% of the assets are invested in equities, followed by 31.7% investing in fixed income funds, 8.7% in cash/money market funds, 4.8% in property and 2.4% in other asset classes. Taking into account market movement, the survey suggests that the overall asset allocation trend of shifting away from equities and into fixed income products is continuing though the samples suggest only a modest change in the position of equities as a proportion of total assets under management since December 2005. Among different equity asset classes, the equity allocation by region suggest that UK equities remain predominant (59.2%), with European ex-UK (16.2%) and US equities including North America (12.1%) as the second and third largest components respectively.

Recent trends in UK Investment Management Industry:

We illustrate several key trends in UK asset management industry, highlighting the evidence that the UK industry is currently undergoing a period of big changes, presenting both opportunities and challenges for many asset management companies.

- Separation between beta and alpha

There is a growing trend of a clearer separation between the 'beta' (market return) and the 'alpha' (value-added by active management). Evidently, over the past several years, passively managed index tracking products have enjoyed considerable success and there have also been a few innovations, such as exchange traded funds (ETFs), which have been growing rapidly in popularity. On the other hand, at the active end of the market, there is a consistent theme, the pursuit of alphas, with fund managers actively creating 'portable alphas' via 'long/short' strategy, along with increasing demand for absolute return and unconstrained strategies.

- Convergence between the hedge fund and 'main stream' asset managers:

There is an increasing demand for absolute return funds and innovative products, such as long/short funds. The techniques used to manufacture returns are increasingly eroding distinctions between traditional (long only) management techniques and hedge funds: for example, the current emphasis on new types of long/short funds (130/30, 120/20 funds[8] etc.) with strategies used to be only allowed in hedge funds, are now also emerging in the retail mutual fund market.

- Diversification of alpha:

More sources of excess returns are desired (besides conventional investments on fixed income and equity, there is increasing demand on alternative asset classes, i.e. real estate, infrastructure, commodities, private equity and foreign exchange currencies etc). However, financial turmoil since 2008 made it more and more difficult to identify uncorrelated / lowly correlated alpha sources as correlations of returns of many asset classes have been increasing dramatically when the financial markets collapse.

- Globalization:

A combination of new client and investment opportunities are provided by the gradual liberalisation of the international economy.

- Liability-Driven Investment:

The issue of pension fund deficits and the way in which schemes can be better assured of meeting future liabilities remain as the dominant theme in the UK asset management industry. (It also remains the case for many other countries. i.e. the Netherlands and Japan). A range of developments over the last few years - particularly regulatory and accounting standard changes have combined to put pressure on pension funds to address the question more precisely and comprehensively. There is growing demand on innovative asset management solutions to meet the pension liability.

- Ongoing Europeanization of the regulatory and commercial operating environment.

[8] 130/30 refers to the fund with limits of up to 130% of NAV long positions and up to -30% of NAV short positions. 120/20 refers to those with the limits of 120% long and -20% of short.

2. UK Unit Trusts

Unit trusts in the UK are one of three sets of financial institutions (the other two are investment trusts and open-ended investment companies (OEICs)), allowing market participants to buy an easily realizable stake in a diversified portfolio of marketable securities that is managed by a professional asset management company.

Unit trusts are open-ended mutual funds, legally established under trust law with trustees acting as custodians of the securities on behalf of the beneficial owners and with a separate asset management company, pursuing the investment objectives specified in the trust deeds. The investor, one party to the unit trust, is a unit holder, holding a certain number of units. A second party, the fund manager, is responsible for the running of the trust and managing the investments of the funds. The third party, the trustees are governed by the Trust Companies Act 1967. Their role is to monitor the fund manager's performance against the trust's deed. The assets of the trust are held in the name of the trustee; they are held "in trust" for the unit holders. The fund's trustee is usually a financial institution, i.e. a clearing bank or insurance company, authorised by the UK Financial Services Authority (FSA). It is the custodian of the trust and ensures that the fund is run in accordance with FSA regulations, the trust deed, and other scheme particulars.

Unit trusts are also “open-ended” because the managers can “create” or “cancel” units. As measured by the asset under management (AUM), the fund gets smaller when the investors sell their shares back to the company (redemption) and they are not bought by other investors, and the fund managers would expect some cash outflow. The fund AUM gets larger when investors buy more shares and the fund managers would expect cash inflow. In both cases (cash outflow & cash inflow), fund managers need to rebalance the portfolio by either reducing or increasing exposures across all the positions.

OEICs are quite similar to unit trusts, but constituted as companies rather than trusts. They are the established structure in many other European countries and are usually single priced. They are also open-ended so that the number of shares changes when investors increase/decrease their exposure to the OEICs.

The unit trust and OEIC providers generally calculate their prices once a day at around noon, in accordance with FSA regulations. This is where the main difference between unit trusts and OEICs arises. With a unit trust there are generally two prices, a 'bid' price and an 'offer' price. The difference between the two prices incorporates the 'initial charge'. The number of units allocated to a unit trust investor is calculated by dividing the value of his or her investment by the unit 'offer price'. The value of the units goes up and down in line with the performance of the fund's share portfolio, which will reflect the ups and downs of the underlying securities. OEICs have only one price, with the initial charge being taken as a separate commission, although recent regulatory change now permits dual pricing too, in line with unit trusts.

In recent years, many unit trust managers have converted to OEICs. The motivation for such conversion is mainly to provide a simplification to offering funds Europe-wide under EU rules. OEICs are more appealing to investors because they have a simpler structure than unit trusts and, more importantly, there is a single price for the shares which investors buy rather than units.

2.1 History[9]

The first unit trust was launched in the UK in 1931 by M&G. The rationale was to emulate the comparative robustness of the US Mutual Funds through the 1929 Wall Street crash. The first fund was called the 'First British Fixed Trust', which held the shares of 24 large companies. The fund was re-launched later as the M&G General Trust and renamed as the Blue Chip Fund. The number of unit trusts increased to around 100 trusts in the UK by around 1939, managing total funds of about £80 million.[10] By 2004, the total number of investment funds has expanded to over 2000[11].

2.2 Classification

To help investors to identify funds with similar characteristics, IMA categorized the funds using a fund classification system of over thirty sectors. The sector categories are broadly divided into funds those aim to provide an "income" and those were designed to provide "growth". Each sector consists of funds investing in similar assets, or in the same geographical region or in the same stock market sectors.

The funds can be basically divided into growth, income and specialist funds. The income funds are made up of immediate income (UK gilts, UK index linked gilt, UK corporate bond, UK other bond, Global bonds and UK equity & bond income funds) and growing income funds (UK equity income fund). Growth funds include capital protection (money market protected/ guaranteed funds) and capital growth/total return

[9] Source: Wikipedia, M&G

[10] For details of the origin of the unit trust and its relationship with American mutual funds, please see K F Sin, The Legal Nature of the Unit Trust, Clarendon Press (Oxford University Press) 1998.

[11] Source: IMA (2004)

fund. Specialist funds refer to the funds that have an investment universe that is not accommodated by the mainstream sectors.

We focus on equity type of funds, to name just a few, they are:

UK Equity Income: Funds that invest at least 80% of their assets in UK equities and aim to have a yield in excess of 110% of the FT All Share Index;

UK All Companies: (Growth funds as defined in Unit trust yearbook)

Funds that invest at least 80% of their assets in UK equities which have a primary objective of achieving capital growth;

UK Smaller Companies: Funds those invest at least 80% of their assets in UK equities of companies which form the bottom 10% by market capitalisation.

Balanced Managed: Funds would offer investment in a range of assets, with the maximum equity exposure restricted to 85% of the Fund. At least 10% must be held in non-UK equities. Assets must be at least 50% in Sterling/Euro and equities are deemed to include convertibles.

Global Growth: Funds which invest at least 80% of their assets in equities (but not more than 80% in UK assets) and which have the prime objective of achieving growth of capital.

Global Emerging Markets: Funds which invest 80% or more of their assets directly or indirectly in emerging markets as defined by the World Bank, without geographical restriction. Indirect investment, e.g., China's H shares (Companies listed in Hong Kong) should not exceed 50% of the portfolio.

2.3 The prices of the unit trusts

The unit trusts employ a dual pricing system (bid & offer price). The difference between them is known as the bid/offer spread.

The creation price and cancellation price do not always correspond with the offer and bid price. Subject to regulatory rules, these prices are allowed to differ and relate to the highs and lows of the asset value throughout the day. The trading profits based on the difference between these two sets of prices are known as the box profits[12]. The Financial Services Authority (FSA) lays down the process by which the maximum offer and minimum bid prices are calculated such as:

- *The Maximum Offer Price (Creation Price)[13]:* It is the maximum offer price that can be charged and measures the full cost of creating a unit.

- *The Minimum Bid Price (Liquidation or Cancellation Price)[14]*

  It is the minimum price at which units can be sold back to the fund managers and represents the full cost of cancelling or liquidating a unit.

- *The Maximum Bid-Offer Spread*

  It is the difference between the creation and the cancellation unit prices. The

[12] Source: Wikipedia

[13] Creation Price is calculated as follows: Value all the underlying assets in the trust at their mid-prices. Valuation takes place at a set time known as the valuation point; Add the value of any other trust property (i.e. net accrued income less fees, charges) and other expenses to calculate Net Asset Value; Add the notional dealing costs of buying the portfolio (stamp duty etc); Divide the total by the total number of the units in issue; Add the initial charge and round the sum to "four significant figures".

[14] The cancellation or liquidation price represents the full cost of cancelling or liquidating a unit and is calculated as follows: calculate the NAV of the portfolio as with the creation price; deduct the dealing costs of selling the portfolio; Divide by the number of units in issue.

spread actually charged is usually narrower because the managers decide how much of the notional dealing costs to pass on to investors.

The managers have the right to widen the spread in the direction of the creation or cancellation price in the event that they become net buyers or net sellers of units. If the fund encounters selling, then the managers may become net buyers of units (i.e. sales are greater than purchases by investors) and may move the spread in the direction of the cancellation price to protect the fund against the full costs of liquidating units. If the managers become net sellers of units on account of buying activity by investors, they may raise the offer price towards the creation price to cover the full cost of creating units. When the fund managers reduce the bid price all the way down to the cancellation price, the trust is said to be priced on the full “bid basis”. If the buying price is raised to the creation price the trust is said to be priced on the full “offer basis”. Under normal market conditions, the fund managers will accumulate units in a “box” for future resale rather than cancel units, or meet excess demand for units from the manager's “box” rather than create units. As a precaution, they may simply adjust the spread in the direction of the cancellation price or the creation price.

2.4 The expenses

Total Expense Ratios

Among the items accounting for the difference between the bid and offer prices of the unit trusts are: the initial charge (sales load), typically 5-6% depends on the asset classes, stamp duty (presently 0.5% for purchases only), dealing charges (commissions), an annual fee and the bid/offer spreads of the underlying securities.

An initial charge is an upfront fee paid by the investors when they buy into the trust: the initial charge depends on asset classes within the fund (e.g. equities, fixed income securities and cash), their market spreads and commissions paid to investment advisors; Some unit trusts forego the initial charge in favour of a sliding scale of exit charges levied on sale of the units within a given period from the date of the purchase. Notional dealing charges levied when there is net buying or selling of units by investors; the annual management charge covers the running expenses of the fund and any renewal (trail) commissions paid to investment advisors. It is about 1% - 2% as a typical fee for an equity fund retail share class; fees may be lower for an institutional share class. Annual fees for fixed income funds are in general lower than fees for equity funds.

Total Expense Ratio (TER) indicates what the true annual unit trust charges are for the individual investor. Information on historical investment management fees and expenses are not readily available. The only source for them is the annual report of each Unit Trust, many of which no longer exist. Prior to 1998, there was no industry – wide publication that collected and reported this information. From 1998, Fitzrovia[15] has published a book that includes TER, which they define as the drag on fund performance caused by all annual operating costs (including administration, custody and audit fees), not just the annual management charge.

[15] Fitzrovia International Ltd is a leading investment fund research company. Based in London, the company's research covers 40,000 investment funds and share classes worldwide. In October 2004, the company was acquired by Lipper Ltd (a wholly-owned subsidiary of Reuters Group plc)

Fitzrovia's research (2003) shows TERs of OEICs averages 1.69% and that of Unit Trusts averages 1.57%, as against the company's previous comparison four years ago, of 1.66% and 1.49% respectively. The gap has narrowed, but average annual charges overall for both types of funds have increased.

3. The data

The unit trusts

The lists of the unit trusts (currently in existence) of each category can be obtained from IMA, however, since its official classification has been changing over time, it is hard to track them back to 1975. Therefore, our dataset is based on 1975 unit trust yearbook, and select those unit trusts with consistent investment objectives throughout the whole sample period.

This study excluded all other non-equity based funds, including international, sector specialist, balanced and fixed income unit trusts. Following the Unit Trust Yearbook, we focus mainly on three categories: Income (UK Equity Income in IMA), Growth (UK All Companies in IMA*)* and General funds (who invest most of their assets in a portfolio of UK growth stocks and income stocks).

Previous UK studies of unit trust performance have used various data sources, mainly S&P Micropal and DataStream. Table 1-1 provides a review of data details of UK studies of unit trust performance. Fletcher (1997) examined the performance of 120 Unit trusts over the period from 1981 to 1989, using monthly returns based on the offer price, collected from Money Management. He also collected dividend information and ex-dividend dates from Extel UK dividend and fixed interest record;

Blake and Timmermann (1998) collected a large sample, comprised of 2300 funds over the period from Feb 1972 to June 1995. They obtained the data from Micropal and calculated net monthly returns based on bid prices and net income. Allen (1999) reported the performance of 131 unit trusts, based on information of closing price from Datastream for the period from 1989 to 1995. Quigley and Singuefile (2000) evaluated the unit trust performance based on the S&P Micropal database from 1978 to 1997. Later work can be found from Charles River Associates Limited (2002), which examined 942 funds over 1981 to 2001, with data from S&P Micropal and Money management; Fletcher *et al.* (2002, 2004) calculated returns based on offer prices and dividend (gross of the load charge and trading costs but net of management charge), collected from Finstat managed fund database, Extel UK dividend & fixed interest record and unit trust yearbook.

We retrieved the data from DataStream International Database, which was the only available data source to us at the time. Our sample is different from the other studies mentioned above in that our data are more recent, and covers a longer period of time, which is a contribution to the current research.

The sample period is from Jan 1975 to Oct 2003[16]. In order to have a long enough sample, only the unit trusts which survived throughout the whole period are selected. We therefore omit all the funds that do poorly and merge or fail. In another word, our tests might suffer survivorship bias and generate upwards performance bias.

[16] The sample did not include more recent data at the time of finishing the thesis, as I started the projects in 2001 and carried out most of the empirical tests during 2003 to 2005. The length of the sample is sufficient however, as it already covers a few full economic cycles and interesting periods such as black Monday in 1985 and IT bubble burst in 2001. These periods may provide insights for examining 2008/2009 financial market crash and envisage future economic crisis, if history can be any reference.

The reference variables[17]

It is recommended that the reference variables/ primitive assets / benchmark should at best reflect all the assets available to the investors and fund managers, while we can only use a finite sample. It is, however, not practical to measure the entire universe of investment opportunities due to problems such as the handling of large econometric systems. To include only limited reference assets, it is then essential to form the reference portfolio with the same characteristic with the assets holding by the unit trusts. That is, it should at least include the same assets that comprise the trust under evaluation, as the task is to identify the positive performer, who significantly enlarge the investment opportunity set due to managers' superior skills (or say, the dynamic trading strategies) rather than the inclusion of certain reference assets and the exclusion of others.

The choices of the references variables can be arbitrary. For example, Ferson *et al.* (2002) constructed nine portfolios, including a short-term risk free security, two long-term bond returns, and stock portfolios, which mimic large-cap, small-cap, value, growth, momentum and contrarian investment strategies. Chen and Knez (1996) constructed 12 equally weighted industrial portfolios based on the monthly returns for all individual stocks listed on the NYSE and the AMEX. Other examples use pure industrial portfolios, for example, Dittmar (2002), Ahn, Conrad and Dittmar (2003) and Fletcher and Kihanda (2005).

---

[17] Reference variables, also called the primitive assets, are used in the stochastic discount factor model, reflecting all the assets available to the investors and the fund managers. Further details can be found in chapter 3 and chapter 4.

Since we only examine the performance of UK equity unit trusts, the intension is to reflect all the possible aspects of investing opportunities available to the fund managers in the UK. Our reference assets are eight passive buy & hold stock portfolios. The seven industrial portfolios include consumer service, financials, health care, consumer goods, industrials, basic material and oil & gas[18]. We also consider a general portfolio based on all firms, that is, the Financial Times All Share-Price index (FTSE).

The information variables[19]

To be consistent with our reference variables, the unit trusts selected are only restricted to those who invest at least 80% of their assets in the UK. There are two types of unit trusts, one that distributes dividends on a regular basis, an income unit, and one that accumulates dividends inside the unit trust, an accumulation unit. Generally, when both units are available, they are like two classes of shares for the same underlying portfolio. The investors have to pay tax on dividend income earned if they did not invest in unit trusts via a share's ISA (Individual Savings Account). For income type of share class, the information of dividend payment and taxation[20] are not easily available, we therefore only select the accumulation units type of unit trusts in order to calculate the total return only based on the price.

---

[18] The sector portfolios are constructed and provided by DataStream, including DS consumer services (CNSMSUK), DS financials (FINANUK), DS health care (HLTHCUK), DS consumer goods (CNSMGUK), DS industrials (INDUSUK), DS Basic Materials (BMATRUK) and DS oil & gas (OILGSUK).

[19] Information variables are used in conditional performance evaluation and explained in Chapter 4.

[20] Tax on dividend payment for Income type of unit trusts: Unit trust equity funds receive dividends from their underlying investments net of corporation tax. Investors receive the net dividend together with a tax credit of 10%; Non-taxpayers have no further liability to income tax but cannot reclaim the tax credit.

Capital Gains Tax (CGT): unit trusts themselves are not liable for CGT on their internal realised gains. Investors are personally potentially liable for CGT on gains realised on disposal of their units.

To fulfill all the selection criteria mentioned above, the sample we collected includes totally 25 general funds, 29 income funds and 16 growth funds. Among them, 12 general funds (48%), 13 income funds (44.8%) and 5 growth funds (31.25%) have been dead or merged since October 1995. It gives us a total sample of 13 general funds, 11 growth funds and 16 income funds. Due to the stringent requirements of the dataset, our sample is inevitably quite small, compared to more than 900 equity investment funds available in the market. Though the size of our data is relatively small, the proportion of dead or merged funds among the total amount of funds is significant.

3.1 The price & its stationarity

We discuss briefly the unit trust pricing and examine the stationarity of the fund prices in this section.

Stationarity

A stationary series can be defined as one with a constant mean, constant variance and constant covariance for each given lag. There are two types of non-stationarity: one is called the random walk model with a drift, another is the trend stationary process. If a non-stationary series must be differenced *d* times before it becomes stationary, then it is said to be integrated of order *d*, i.e. it contains *d* unit roots.

It is important to test if the time series data are stationary as the use of non-stationary data can lead to spurious regressions. The stationarity or otherwise of a series can strongly influence its behaviour and properties. If the variable employed in a regression model is not stationary, it can be proved that the standard assumptions for asymptotic analysis will not be valid.

We firstly examine that stationarity of the fund prices. We initially select a fund *i* randomly from each one fund group, then we examine if the price of this selected fund has a unit root based on the Dickey-Fuller (DF) test.

The basic object of DF test is to examine the null hypothesis that

$\phi = 1$

in

$P_t = \phi P_{t-1} + \mu_t$ (1.1)

where P_t is the price of selected fund *i* at time *t*, against the one-sided alternative such as:

$\phi < 1$

It is equivalent to test

$\Delta P_t = \lambda P_{t-1} + \mu_t$ (1.2)

so that a test of $\phi = 1$ is equivalent to a test of $\lambda = 0$.

The test statistic is defined as $\frac{\hat{\lambda}}{S\hat{E}(\hat{\lambda})}$, which does not follow the usual *t*-distribution under the null hypothesis, since the null is one of non-stationarity, but rather they follow a non-standard distribution. Critical values are derived from simulation experiments in Fuller (1976).

As shown in table 1-2, the results[21] suggest that for all types of funds, the prices of randomly selected funds are not stationary as the null hypothesis of there is an unit

[21] The test is implemented using E-View program. Lag length is 0, which is decided automatically based on SIC by the program and no trend is added.

root is not rejected at 1% significance level: In all of three cases: general, growth, income funds, *t*-statistics are larger in magnitude than the test critical values. i.e. *t*-statistic ranges from -1.359899 to 1.376943 for growth funds, ranges from -1.544695 to 0.700922 for general funds and ranges from -1.894611 to 1.795564 for income funds, compared to critical value of -3.4342 at 1% significance level.

TABLE 1-2 THE RESULTS OF ADF TESTS: FUND PRICES

This table reports the *t*-statistic and *p*-value of ADF test for the prices of the funds, including 11 growth funds, 13 general funds and 16 income funds.

Augmented Dickey-Fuller test critical Values:

1% level	-3.4342
5% level	-2.8631
10% level	-2.5677

Growth Fund *P*	t-Statistic	Prob.*	General Fund *P*	t-Statistic	Prob.*	Income Fund *P*	t-Statistic	Prob.*
Growth Fund 1	-0.314	0.920	General Fund 1	-1.545	0.510	Income Fund 1	-0.804	0.816
Growth Fund 2	-1.360	0.602	General Fund 2	-0.084	0.949	Income Fund 2	-0.242	0.930
Growth Fund 3	0.847	0.995	General Fund 3	0.701	0.992	Income Fund 3	-0.741	0.834
Growth Fund 4	-0.435	0.901	General Fund 4	-0.845	0.805	Income Fund 4	1.495	0.999
Growth Fund 5	1.377	0.999	General Fund 5	-1.283	0.639	Income Fund 5	-1.403	0.581
Growth Fund 6	-0.573	0.873	General Fund 6	-0.347	0.729	Income Fund 6	-1.895	0.335
Growth Fund 7	-0.263	0.927	General Fund 7	0.291	0.978	Income Fund 7	-1.042	0.739
Growth Fund 8	0.713	0.992	General Fund 8	-0.504	0.887	Income Fund 8	-0.765	0.827
Growth Fund 9	-0.228	0.932	General Fund 9	-0.965	0.767	Income Fund 9	-0.619	0.863
Growth Fund 10	0.617	0.990	General Fund 10	-1.253	0.653	Income Fund 10	0.106	0.966
Growth Fund 11	0.020	0.959	General Fund 11	-0.183	0.938	Income Fund 11	0.836	0.995
			General Fund 12	0.372	0.982	Income Fund 12	1.097	0.998
			General Fund 13	-0.999	0.754	Income Fund 13	-0.556	0.877
						Income Fund 14	0.202	0.973
						Income Fund 15	1.796	1.000
						Income Fund 16	0.570	0.989

*MacKinnon (1996) one-sided p-values.
Exogenous: Constant;
Lag Length: 0 (Automatic based on SIC, MAXLAG=16)

3.2 The returns & the descriptive results

For the financial dataset, typically stock prices, market indices and dividend are generally not stationary, which were also confirmed by our stationarity test above. Moreover, as we will apply GMM techniques[22] to estimate the asset pricing models and the GMM distribution theory does require some statistical assumptions[23]. The

[22] We will introduce the techniques in the following chapters.
[23] Hansen (1982) and Ogaki(1993) cover them in depth.

most important assumption of these is all the variables should be stationary random variables, in another word, the sample averages of these variables must converge to population means as the sample size grows, and stationary is necessary for this result. As a result, we calculate the returns of the funds based on the prices.

The accumulation type of unit trusts do not distribute dividend. Hence there are no tax and reinvestment charges. The gross return can be calculated simply as P_t/P_{t-1}.

We also implement the ADF test[24] to examine the stationarity of the fund returns. The results are reported in table 1-3. *t*-statistics for fund returns ranges from -19.119 to -14.493 for growth funds, -18.473 to -15.418 for general funds and -18.65 to -15.642 for income funds, compared to -3.447 (the ADF critical value) at 1% level. It suggests that the returns are stationary.

TABLE 1-3. THE RESULTS OF ADF TEST FOR THE FUND RETURNS

This table reports the *t*-statistic and *p*-value of ADF test for the returns of the funds, including 11 growth funds, 13 general funds and 16 income funds.

Augmented Dickey-Fuller test critical Values:	
1% level	-3.43422
5% level	-2.86314
10% level	-2.56767

Growth Fund *R*	**t-Statistic**	**Prob.***	**General Fund *R***	**t-Statistic**	**Prob.***	**Income Fund *R***	**t-Statistic**	**Prob.***
Growth Fund 1	-18.680		General Fund 1	-16.663	0	Income Fund 1	-17.799	
Growth Fund 2	-16.868	0	General Fund 2	-18.215	0	Income Fund 2	-18.345	0
Growth Fund 3	-17.653	0	General Fund 3	-15.658	0	Income Fund 3	-15.642	0
Growth Fund 4	-14.907	0	General Fund 4	-16.666	0	Income Fund 4	-18.345	0
Growth Fund 5	-17.200	0	General Fund 5	-17.387	0	Income Fund 5	-15.012	0
Growth Fund 6	-19.119	0	General Fund 6	-18.473	0	Income Fund 6	-16.064	0
Growth Fund 7	-17.865	0	General Fund 7	-16.646	0	Income Fund 7	-17.544	0
Growth Fund 8	-18.363	0	General Fund 8	-17.454	0	Income Fund 8	-18.005	0
Growth Fund 9	-15.409	0	General Fund 9	-18.385	0	Income Fund 9	-18.65	0
Growth Fund 10	-18.435	0	General Fund 10	-16.528	0	Income Fund 10	-16.584	0
Growth Fund 11	-14.493	0	General Fund 11	-16.926	0	Income Fund 11	-18.025	0
			General Fund 12	-15.418	0	Income Fund 12	-18.214	0
			General Fund 13	-18.425	0	Income Fund 13	-17.658	0
						Income Fund 14	-16.458	0
						Income Fund 15	-16.871	0
						Income Fund 16	-18.369	0

*MacKinnon (1996) one-sided p-values.

Exogenous: Constant;

Lag Length: 0 (Automatic based on SIC, MAXLAG=16)

[24] Again, the test is implemented using E-View program. Lag length is 0, which is decided automatically based on SIC by the program and no trend is added.

The funds which died before Oct 2003 are defined as dead funds while the funds survived the whole sample period from Jan 1975 to Oct 2003 are defined as live funds. We report the descriptive results (the mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of the returns of growth, general, income funds and their corresponding dead funds separately in appendix table 1-4 to table 1-6.

For the period from Jan 1975 to Oct 2003 (a total of 346 months, with the total number of observations on return is 345 as we calculate returns from Feb 1975), live[25] growth fund has generated average monthly return of 1.06%, live general fund has average monthly return of 1.06% and live income fund has returned 1.05% monthly.

In appendix, table 1-7 to table 1-9 report the performance for the funds which have been merged or died over the period from 1975 to the time they terminated.

For the period from 01/01/1975 to 09/01/1997 (272 months), when the first dead general funds started to cease to exist, the dead general fund has average monthly return of 1.32%, compared to 1.42% of monthly return for average live general fund over the same period (see table 1-7). Average monthly return of dead income fund stands at 1.04% for the period from 1975 to Dec 1996, compared to an average monthly return of 1.36% for average live income funds (see table 1-9). From 01/01/1975 to 10/01/1995, average monthly return of dead growth fund stands at 1.43%, compared to 1.38% for average live growth funds for the same period (see table 1-8). These results suggest that in most of the cases, live funds tend to

[25] Remain alive on 01/01/08.

outperform dead funds, which is coherent with 'survivorship bias' phenomena, the only exception is growth fund, in which case the dead fund beat their surviving counterparties, if not to mention the sample size of dead growth fund is very small (5).

The table 1-10 shows the moments for the monthly gross returns of the reference portfolios. The benchmark portfolios are passive buy & hold industry portfolios. The mean and standard deviations of returns are expressed as net returns and in % per monthly. The average monthly returns of the equity sector index range from 0.9% (consumer goods sector) to 1.3% (oil & gas) per month. The standard deviations of the stock portfolios are between 5.7% (FTSE all share) to 7.89% (consumer goods sector). Moreover, the table also shows the coefficients of skewness and excess kurtosis of the portfolio returns.

4. Conclusion and the thesis structure

To conclude, in this chapter, we firstly provide a brief introduction to UK asset management industry and its latest trend. Secondly, we explain the differences between UK unit trusts and OEIC as well as the fact that unit trust has recently lost its popularity mainly because of its complicated make-up. Although there are thousands of them still, they have now been superseded by OEICs which were devised in Europe. Thirdly, we explain the dataset we use in the thesis. We apply an ADF test to examine the stationarity of the fund prices and the returns. Our results suggest that the fund prices are not stationary, but after the first-differencing, general, income and growth fund returns become stationary. The descriptive results suggest that live general and income funds on average outperformed their dead counterparties, with the only exception that dead growth funds outperformed surviving growth funds on average.

In the following chapters, we will investigate different aspects of the UK unit trust performance based on the dataset discussed. We start with a chapter providing a literature survey, descriptions of research questions and issues, methodologies employed, empirical results entailed in previous research. Chapter 3 examines the small sample properties of the GMM iterated and 2-Step estimators within the framework of the SDF primitive efficient models. Based on the results from the simulation tests, the optimal method of estimation is employed to evaluate UK unit trusts' performance, with a special focus on the role of conditioning information on performance evaluation in Chapter 4. Relative performance according to different fund styles will be investigated in Chapter 5, aiming also at identifying a winning style-rotation strategy. Chapter 6 examines the persistence of fund performance. Chapter 7 concludes.

APPENDIX CHAPTER 1

TABLE 1-1 DATA DETAILS OF RECENT STUDIES OF UK UNIT TRUST PERFORMANCE

The paper	Period and funds coverage	The data details	The return calculation
Fletcher (1997)	120 Trusts 1981-1989	Monthly offer prices (Money management) Dividend information and ex-dividend dates (Extel UK Dividend and Fixed interest Record) ;One month treasury bill (DataStream)	Monthly Returns based on offer prices
Blake & Timmermann(1998)	2300 funds 02.1972—06. 1995	Micropal Ltd.	Net Monthly returns using bid prices and net income
D. E Allen (1999)	131 funds 1989--1995	DataStream	Closing price
Quigley & Sinquefield (2000)	1978--1997	S&P Micropal (Micropal) database	The same as B&T. Expenses: 1.35% TERs
Charles River Associates Limited (2002)	942 funds (508 alive & 434 dead) 1981--2001	S&P Micropal; Dead fund (Quigley and Sinquefield 1998) Money Management (1998—2001)	
Fletcher &Forbes (2002, 2004)	253 trusts 1982---1996	Offer Prices (Finstat managed fund database: FT Business information Service and Money Management; Dividends (Finstat & Extel U.K. Dividend and Fixed Interest Record.);1-month U.K. Treasury bill returns (LSPD)*;Trading costs: C,E&K **;Investment objective, annual &load charge (Unit Trust Yearbook)	Returns based on offer prices and dividend (Gross of the load charge and trading costs but net of the management charge.)

*: London Business School Share Price Database
**Chalmers, J.M.R.,Edelen, R.M.and G.B. Kadlec, 2001, Fund returns and trading expenses: Evidence on the value of active fund management, working paper, University of Pennslvania.

TABLE 1-4. THE RETURNS OF LIVE GENERAL FUNDS

This table reports the descriptive results (mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of 13 general-fund returns based on monthly data from Jan 1975 to Oct 2003. ‘Average’ refers to the statistics of average returns of the general funds.

	1	2	3	4	5	6	7	8	9	10	11	12	13	Average
Mean	0.85%	0.92%	1.15%	1.11%	0.93%	1.08%	1.38%	1.07%	1.03%	1.03%	1.09%	0.98%	1.10%	1.06%
Median	0.86%	0.18%	1.33%	1.58%	1.30%	1.17%	1.79%	1.02%	0.40%	1.29%	1.27%	1.03%	0.22%	1.33%
Maximum	31.86%	40.26%	21.80%	25.34%	34.22%	38.71%	27.87%	28.61%	42.91%	35.83%	42.28%	45.38%	41.15%	32.68%
Minimum	-20.80%	-28.29%	-23.54%	-30.13%	-27.71%	-25.29%	-24.59%	-23.24%	-30.40%	-29.22%	-26.98%	-26.04%	-27.32%	-26.31%
Std. Dev.	0.058	0.052	0.048	0.049	0.053	0.052	0.055	0.053	0.054	0.054	0.055	0.055	0.052	0.049
Skewness	0.003	0.788	-0.234	-0.320	0.177	1.009	-0.238	0.305	0.914	0.260	0.800	1.165	0.943	0.303
Kurtosis	6.060	13.597	5.757	8.593	8.771	13.201	6.646	6.611	14.749	9.945	12.309	14.590	14.452	10.228

No. of Observation: 345

TABLE 1-5. THE RETURNS OF LIVE GROWTH FUNDS

This table reports the descriptive results (mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of 11 growth-fund returns based on monthly data from Jan 1975 to Oct 2003. ‘Average’ refers to the statistics of average returns of the growth funds.

	1	2	3	4	5	6	7	8	9	10	11	Average
Mean	0.95%	1.24%	1.13%	0.94%	1.31%	1.01%	0.98%	1.06%	1.04%	1.13%	0.89%	1.06%
Median	1.21%	1.68%	1.30%	0.76%	1.52%	0.00%	1.00%	1.48%	1.20%	1.20%	1.27%	1.47%
Maximum	25.19%	24.03%	16.72%	21.67%	25.93%	18.78%	19.29%	21.43%	41.46%	39.48%	21.49%	20.98%
Minimum	-37.07%	-27.63%	-25.87%	-20.77%	-29.10%	-31.69%	-23.16%	-24.08%	-30.07%	-32.47%	-32.38%	-28.44%
Std. Dev.	0.054	0.050	0.047	0.058	0.051	0.047	0.045	0.049	0.056	0.057	0.053	0.046
Skewness	-0.413	-0.392	-0.530	-0.216	-0.415	-0.498	-0.334	-0.359	0.862	0.287	-0.715	-0.522
Kurtosis	10.476	6.664	5.638	4.735	8.147	9.970	5.944	5.506	13.694	11.145	7.917	8.307

No. of observation: 345

TABLE 1-6. THE RETURNS OF LIVE INCOME FUNDS

This table reports the descriptive results (mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of 16 Income-fund returns based on monthly data from Jan 1975 to Oct 2003. ‘Average’ refers to the statistics of average returns of the income funds.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average
Mean	0.87%	1.29%	0.85%	1.27%	0.95%	1.08%	1.40%	1.37%	0.99%	0.73%	1.28%	1.18%	0.88%	1.83%	1.41%	1.18%	1.05%
Median	1.06%	0.31%	1.14%	1.69%	0.96%	1.42%	1.32%	1.47%	1.25%	0.92%	1.15%	1.01%	1.21%	1.84%	1.32%	1.26%	1.25%
Maximum	30.85%	20.77%	14.47%	36.19%	33.33%	30.52%	44.86%	36.41%	27.85%	11.36%	19.33%	18.82%	22.91%	44.59%	27.27%	15.54%	23.95%
Minimum	-26.99%	-27.35%	-23.79%	-24.92%	-25.97%	-21.47%	-26.03%	-23.10%	-21.79%	-22.25%	-18.24%	-29.31%	-83.19%	-30.68%	-23.78%	-24.73%	-24.07%
Std. Dev.	0.049	0.049	0.038	0.049	0.050	0.044	0.052	0.051	0.049	0.036	0.049	0.064	0.077	0.062	0.056	0.055	0.044
Skewness	-0.023	-0.239	-0.667	0.434	0.243	0.171	1.256	0.529	0.155	-1.034	-0.073	-0.585	-5.052	0.827	0.136	-0.445	-0.210
Kurtosis	9.506	7.402	7.922	10.965	9.323	9.647	17.057	9.584	6.621	7.619	4.871	5.969	57.061	13.533	5.969	5.137	7.840

No. of observation: 345 (from Jan 1975 to Oct 2003)

TABLE 1-7. THE RETURNS OF DEAD GENERAL FUNDS VS LIVE FUNDS

This table reports the descriptive results (mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of 12 dead general fund returns based on monthly data from 01/01/1975 to 09/01/1997. ‘Average Dead’ refers to the statistics of average returns of the dead general funds from 01/01/1975 to 09/01/1997. ‘AverageGeneral’ refers to for the same period, the statistics of average returns of the general funds which survived from Jan 1975 to Oct 2003.

	1	2	3	4	5	6	7	8	9	10	11	12	AverageDead	AverageGeneral
Mean	1.18%	1.29%	1.31%	1.13%	1.02%	1.30%	1.25%	1.27%	1.39%	1.67%	1.50%	1.50%	1.32%	1.42%
Median	1.26%	1.37%	1.34%	1.26%	1.53%	1.45%	1.62%	1.15%	1.68%	1.64%	1.39%	1.50%	1.64%	1.66%
Maximum	28.38%	30.98%	31.75%	29.55%	24.03%	29.91%	21.35%	34.85%	30.47%	28.71%	26.27%	19.68%	26.35%	32.68%
Minimum	-28.59%	-26.83%	-26.30%	-23.41%	-93.63%	-24.50%	-23.76%	-25.68%	-28.79%	-28.46%	-28.53%	-23.67%	-26.21%	-26.31%
Std. Dev.	0.055	0.055	0.056	0.055	0.080	0.054	0.053	0.055	0.055	0.063	0.061	0.051	0.053	0.053
Skewness	-0.078	0.196	0.252	0.229	-6.164	0.227	-0.228	0.451	-0.038	0.003	-0.104	-0.409	-0.127	0.360
Kurtosis	7.869	8.338	8.167	6.753	75.345	7.273	5.520	9.721	8.714	6.193	6.180	5.947	7.240	9.933

No. of observation: 272 (from 01/01/1975 to 09/01/1997)

TABLE 1-8. THE RETURNS OF DEAD GROWTH FUNDS VS LIVE FUNDS

This table reports the descriptive results (mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of 5 dead growth fund returns based on monthly data from 01/01/1975 to 10/01/1995. 'Average Dead' refers to the statistics of average returns of the dead income funds from 01/01/1975 to 12/01/1996. 'Average Income' refers to for the same period, the statistics of average returns of the growth funds which survived from Jan 1975 to Oct 2003.

	1	2	3	4	5	AverageDead	AverageGrowth
Mean	1.42%	1.25%	1.44%	1.53%	1.51%	1.43%	1.38%
Median	1.63%	1.52%	1.72%	1.48%	1.84%	1.65%	1.61%
Maximum	37.02%	23.12%	39.02%	43.93%	35.18%	35.65%	20.98%
Minimum	-34.50%	-27.76%	-73.08%	-25.59%	-26.81%	-28.06%	-28.44%
Std. Dev.	0.07	0.06	0.07	0.06	0.06	0.05	0.05
Skewness	-0.04	-0.32	-3.95	1.16	0.34	0.25	-0.54
Kurtosis	9.31	5.64	48.59	16.50	10.53	11.41	8.13

No. of Observation: 249 (from 01/01/1975 to 10/01/1995)

TABLE 1-9. THE RETURNS OF DEAD INCOME FUNDS VS LIVE FUNDS

This table reports the descriptive results (mean, median, maximum, minimum, standard deviation, skewness and kurtosis) of 13 dead income fund returns based on monthly data from 01/01/1975 to 12/01/1996. 'Average Dead' refers to the statistics of average returns of the dead income funds from 01/01/1975 to 12/01/1996. 'Average Income' refers to for the same period, the statistics of average returns of the income funds which survived from Jan 1975 to Oct 2003.

	1	2	3	4	5	6	7	8	9	10	11	12	13	AverageDead	Average Income
Mean	1.32%	1.12%	1.15%	1.14%	1.20%	1.05%	1.18%	0.40%	0.94%	1.09%	0.72%	1.02%	1.20%	1.04%	1.36%
Median	1.18%	0.92%	0.97%	1.01%	1.38%	1.14%	1.09%	0.04%	1.03%	0.86%	0.76%	1.36%	1.42%	1.14%	1.55%
Maximum	31.17%	29.35%	33.33%	38.89%	22.41%	38.36%	30.97%	11.22%	13.00%	37.28%	23.68%	17.47%	22.06%	25.83%	23.95%
Minimum	-27.40%	-24.85%	-25.12%	-21.81%	-22.21%	-23.16%	-30.05%	-13.46%	-19.21%	-22.17%	-19.95%	-26.51%	-23.21%	-22.16%	-24.07%
Std. Dev.	0.055	0.058	0.057	0.058	0.051	0.059	0.059	0.035	0.044	0.060	0.049	0.057	0.052	0.049	0.049
Skewness	0.008	0.140	0.225	0.676	-0.096	0.615	-0.186	0.126	-0.284	0.691	0.169	-0.574	-0.084	0.005	-0.219
Kurtosis	8.239	6.442	7.617	10.088	5.659	9.509	8.088	4.498	4.316	8.348	5.788	5.590	5.376	6.962	7.183

No. of Observation: 263 (from 01/01/1975 to 12/01/1996)

TABLE 1-10. THE DESCRIPTIVE STATISTICS OF REFERENCE VARIABLES

This table shows the moments for the monthly gross returns of the benchmark portfolios. The benchmark portfolios are passive buy & hold industry portfolios. The data are obtained from DataStream. The mean and standard deviations of the returns are expressed as net returns and in % per monthly.

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
FTSE all share	1.02%	1.41%	52.40%	-27.90%	5.70%	1.515	20.508
Basic Materials	1.02%	1.31%	39.77%	-33.94%	6.91%	-0.202	7.753
Consumer goods	0.90%	0.68%	41.81%	-34.93%	7.89%	0.259	5.615
Consumer service	1.00%	1.10%	51.65%	-26.01%	6.33%	1.081	13.704
Financials	1.02%	1.23%	53.52%	-27.42%	6.54%	1.030	14.493
Health care	1.19%	1.31%	45.94%	-29.30%	5.71%	1.043	13.812
Industrials	1.11%	1.37%	49.17%	-31.00%	6.91%	0.454	9.774
Oil & Gas	1.30%	1.19%	51.39%	-30.32%	6.95%	0.958	10.458

CHAPTER 2 LITERATURE REVIEW

This chapter aims at providing a survey of the performance evaluation literature. Attention is paid in particular to the main methodologies of the thesis: the Stochastic Discount Factor (SDF) models and the conditional measure of asset pricing models. Three sections are presented to sketch the research issues in this field. The first section introduces the methods, both traditional and recently-discovered, which have been applied to performance evaluation; Section 2 provides a review of conditional performance evaluation; the issues of performance attribution and performance persistence are discussed in section 3 and section 4; Conclusions follow.

1. The Methods of Performance Evaluation

This section investigates the methodologies of performance evaluation employed by the researchers. It has two major purposes. One is to provide a broad and general view of the methodologies. It starts with traditional measures, so-called "three indices", derived from the Capital Asset Pricing Model (CAPM), followed by a discussion of Jensen's alpha within the framework of multi-factor models. We pay special attention to the implication of a more recently developed method: the Stochastic Discount Factor (SDF) model. The characteristic-based benchmark method is also introduced. In addition, we illustrate the issues associated with each methodology and review the theoretical background of conditioning performance evaluation. Finally, we discuss the methods of how to deal with the conditioning information problems.

1.1 Conventional Asset Pricing Measures

Before 1960, investors evaluated portfolio performance almost entirely based on the rate of return, although they understood that risk was a very important variable in determining investment achievements. The reason for omitting risk was the lack of knowledge of measuring and quantifying it.

It is well known that the evolution of performance measures follows the development of the asset pricing theories. The risk-adjusted performance measures only started from Sharpe (1966), Treynor (1965), and Jensen (1968), who are famous for their "three indices", namely the Sharpe ratio, Treynor's ratio and Jensen's alpha. All "three-indices" are derived from the CAPM model. When multi-factor models become more predominant, methodologies of performance measure are based on those models, i.e. those of Linter (1965), Merton (1973), Ross (1976), Fama (1993) and Carhart (1997). The essential theory underlining this strand of measure is to take a fairly-priced expected return as a baseline, where the expected return is regarded as a linear function of its covariance (or betas) with some systematic risk factors.

Besides regression-based measures, other conventional methods also include the period weighting measures of Grinblatt and Titman (1989), the inter-temporal marginal rates of substitution–based measures of Glosten and Jagannathan (1994), and the characteristic-based benchmark model of Daniel *et al.*(1997).

1.11 The CAPM and Three Indices

The CAPM predicts that the excess return on any financial asset adjusted for the risk on the specific beta of that asset, holding for all the assets, such as:

$$ER_i = r + \beta_i (ER_m - r) \qquad (2.1)$$

In the context of mutual fund performance measurement, ER_i is the expected return on mutual fund *i, r* is the risk-free return, β_i stands for the risk sensitivity of the testing stocks or portfolio *i* to the market portfolio, ER_m is the return from holding the market portfolio. This method was developed into a methodology for performance evaluation and was firstly described in Sharpe (1964) and Lintner (1965).

The index suggested by Sharpe (so called Sharpe ratio) is a reward-to-variability ratio and is defined for portfolio *i* as:

$$SH_i = \frac{EXR_i}{\sigma_i} = \frac{ER_i - r}{\sigma_i} \qquad (2.2)$$

where EXR_i is the excess return, ER_i is the return on portfolio *i, r* is the related risk-free rate; σ_i is the variance of portfolio *i*. Sharpe's index measures the slope of the transformation line and can be calculated for any portfolio using historic data.

This index captures the manager's ability to diversify away idiosyncratic risks and has been widely applied in the performance evaluation literature. Sharpe's ratio assumes that individual investors hold only the risk-free assets and a single portfolio of risky assets. In comparison, Treynor's performance index, T_i, assumes that the individual investor has a choice between the mutual fund and another portfolio of risky assets. It is also a measure of excess return per unit of risk, but the risk is measured by the beta of the portfolio, β_i, which is given by

$$T_i = \frac{ER_i - r}{\beta_i} \qquad (2.3)$$

It comes directly from the CAPM, which may be written as

$$\frac{ER_i - r}{\beta_i} = ER_m - r \qquad (2.4)$$

where ER_m is the return on market portfolio m.

Under the CAPM, the value of T_i should be the same for all portfolios of securities when the market is in equilibrium. It follows that if the mutual fund manager invests in a portfolio where the value of T_i exceeds the excess return on the market portfolio, he will be earning an abnormal return relative to that given by the CAPM.

Jensen's performance index also assumes that investors can hold either the mutual fund denoted i or a well-diversified portfolio such as the market portfolio. This index is given by the following regression:

$$ER_i - r = \alpha_i + \beta_i (ER_m - r_t) \qquad (2.5)$$

α_i in equation (2.5) is so called Jensen's alpha, which measures the deviation of a portfolio from the securities market line, and it has been most widely employed in academic empirical studies.

It is apparent that if α = 0, then we have the standard CAPM. Hence the mutual fund earns a return in excess of that given by the CAPM if it is greater than zero. In the case that the alpha is negative, the fund manager has under performed relative to the risk adjusted rate of return given by the CAPM. Hence, this alpha index based on the CAPM model actually measures the abnormal return of the portfolio.

We can compare Treynor's and Jensen's index by rearranging equation (2.5) as follows:

$$\frac{ER_i - r}{\beta_i} = \frac{\alpha_i}{\beta_i} + (ER_m - r) \qquad (2.6)$$

Here, the left-hand side of the above equation is simply the Treynor index. If β_i is positive, it is easy to see that when $T_i > ER_m - r$, α_i is greater than zero. Hence, a success with the Treynor index also implies a 'success' on Jensen's index.

However, it is argued that the results generated from different indices could be arbitrary. That is to say, a higher value of the Treynor index for fund A over fund B may be consistent with a higher value of Jensen's alpha for fund B instead of fund A. Hence the relative performance of two mutual funds depends on the index chosen.

Another measure derived from Jensen's alpha evaluates performance according to information precision. Connor and Korazxyk (1986) describe a case in which the Treynor and Black (1973) appraisal ratio, such as $\frac{\alpha_i}{S_i}$ where α_i is Jensen's alpha and S_i is the standard deviation of the error term in the regression used to obtain alpha, properly ranks managers according to their forecasting abilities. The result requires a number of assumptions before it is valid, including: no ability to forecast the market, multivariate normal returns, exponential utility as the criterion for investment for all managers, and the tradability of all assets for all managers. These restrictions appear to be stringent enough to preclude the usefulness of this ratio as a tool for performance evaluation.

Further, the choice of a benchmark portfolio is probably the most controversial issue in performance evaluation. The debate about benchmarks was initiated by Roll (1978), who noted that different benchmark portfolios would provide different risk adjustments and hence different assessments of abnormal performance. Roll's critique, which concerns the estimation of the CAPM using a sample of data, indicates that in any dataset, the relationship shown in equation (2.7) will always hold:

$$\hat{R}_i = r + \beta(\hat{R}_m - r) \qquad (2.7)$$

where $\hat{R}_i$ is the sample mean of the return on testing portfolio and $\hat{R}_m$ is the sample mean of the return on the market portfolio. Regression (2.7) shows that there are exact linear relationships in any sample of data between the mean return on portfolio i and that portfolio's beta, if the market portfolio is correctly measured. Hence, if the

CAPM was a correct description of investor behaviour, Treynor's index would always be equal to the sample excess return on the market portfolio and Jensen's index would always be zero.

It is also argued (Ferson 1996) that a proper benchmark portfolio needs to be both mean-variance efficient and mean-variance inefficient at the same time. It needs to be mean-variance efficient so that the portfolios of uninformed managers and all passive portfolios will have Jensen's alphas of zero. It needs to be mean-variance inefficient for the portfolios of managers with forecasting ability so that these portfolios can have nonzero alphas. This would seem to be an impossible task for a portfolio, but for the fact that the information sets of managers with forecasting ability differ from those without forecasting ability. Two managers with different information sets would necessarily draw different mean-standard deviation diagrams. In particular, the manager with forecasting ability would have mean-variance frontiers that are improved by dynamic portfolio strategies—strategies that weight more heavily those securities that are forecasted to have unusually high returns in a period. Managers lacking this ability cannot achieve a better mean-variance trade-off by dynamically changing their portfolio weights. Hence, their efficient frontier plots inside the efficient frontier of better-informed managers. The benchmark can be chosen as mean-variance efficient with respect to passive portfolios, but not with respect to the dynamic portfolios chosen by managers with forecasting ability.

The performance obtained with a benchmark having this property is analyzed in the models developed by Mayers and Rice (1979), Dybvig and Ross (1985), and Grinblatt and Titman (1989). In these models, the investors with superior information on individual securities expected returns (i.e. selectivity information) but not with information on the benchmark returns (i.e. timing information) achieve positive alphas if the benchmark is mean-variance efficient from the perspective of an investor without forecasting ability. For investors with only timing information, this result does not necessarily hold.

1.12 The Multi-Factor Model

The CAPM with its three-indices have kept its long popularity in the academic field. However, a practical complication remains. The benchmark portfolio is suggested lying on the *ex-ante* efficient frontier, but finding such a portfolio is not an easy task. During the late *1970s* and *1980s*, a number of market anomalies were discovered, unfavourable to the traditional CAPM. These anomalies imply that a single risk factor actually does not represent all the risks. Indeed, unless asset returns are generated by only one common factor, it is unlikely that an arbitrarily chosen diversified portfolio will be mean-variance efficient.

Under this circumstance, Fama and French (1993) initiate a three-factor model to explain the cross section of portfolio returns. Besides the factor of excess return of the market portfolio relative to risk-free return, it captures the excess return on a portfolio of small stocks relative to a portfolio of large stocks as well as the excess return on a portfolio of value stocks (high book to market ratio stocks) relative to a portfolio of growth stocks (low book/market ratio stocks). In Carhart's further studies (1997), he adds a momentum factor by including a portfolio of stocks with high returns over recent months. It is consistent with a model of market equilibrium with four risk factors. It may also be interpreted as a factor-mimicking portfolio, indicating the proportion of mean return attributable to four elementary strategies: high versus low beta stocks, large versus small market capitalization stocks, value versus growth stocks, and one-year return momentum versus contrarian stocks.

Empirical investigations of mutual fund performance consistent with multi-factor models can be found in papers by Lehmann and Modest (1987), Grinblatt and Titman (1989, 1994), Connor and Korajczyk (1991) and Elton, Gruber, Das and Hlavka (1993). The study by Lehamann and Modest was the first to evaluate mutual fund using multiple portfolio benchmarks. The formulae of the performance measures, derived from the multi-factor model, can be expressed as

$$R_{FD} - R_F = \alpha + \beta(R - R_F) + \varepsilon \qquad (2.8)$$

where $\boldsymbol{R}$ is a vector of risk factor loading and $\boldsymbol{\beta}$ is a vector of corresponding risk sensitivity. Similar to Jensen's alpha, the signs and magnitude of alpha in regression (2.8) can be used to identify the performance.

Concerning the problems mentioned above, many researchers started looking for better multiple-portfolio benchmarks based on the sensitivity of performance inferences to the choice of the benchmark portfolio. The aim was to search for a benchmark, which should exhibit more powerful explanatory ability, and provide a more precise interpretation of the various risks in discovering some specific index with specific risk characteristics. The basic idea underlying the formation of this benchmark is that various firm characteristics are correlated with their stocks' factor loadings. As a result, portfolios formed from the stocks grouped by securities' characteristics can be used as proxies for the factors. The examples include Connor and Korajczyk (1991), Gruber (1992, 1996), Chan, Chen, and Lakonishok (1999) and Lynch and Musto (2000).

Connor and Korajczyk (1991) firstly construct five portfolios using principal components analysis. Four linear combinations of these five portfolios that best mimic a set of four pre-specified macroeconomic factors were then formed to yield four new factor portfolios that correspond to the four macro-factors. A fifth factor, a residual of the regression of the value-weighted index on the previously described four macro-factor portfolios, was also used. The performance results generated with this residual were the same as those that would have been generated by including the value-weighted index as the fifth factor and they did not find evidence of negative performance.

As shown in Gruber (1996), he employs a four-index model, in which one bond index is added to Fama's three factors. Other multi-factor benchmarks include the 10-factor benchmark (see Lehmann-Modest (1988)) and the eight-portfolio benchmark *'P_8'* used in Grinblatt and Titman (1992). Grinblatt and Titman indicate that persistent biases can exist among traditional benchmarks. For example, the CAPM and APT based benchmarks favour small capitalization and high dividend-yield stocks. Thus, small-firm funds and income-oriented funds may appear persistently to outperform other funds if traditional benchmarks were used. To avoid this problem, Grinblatt and Titman create P_8, consisting of four size-based portfolios, three dividend-yield-based portfolios, and the lowest past returns portfolio.

Moreover, several studies have examined changes in fund styles measured as a factor loading in a multi-factor model. Chan, Chen, and Lakonishok (1999) find that fund styles tend to cluster around a broad market benchmark. Some fund managers have a consistent style approach. On the other hand, funds with poor past performance could be more likely to change styles. Lynch and Musto (2000) examine if mutual funds change the investment styles after a period of bad performance. They discover that the changes are larger for the funds in the bottom performance quartile than for the other funds. The changes in strategy, as well as managerial replacement among the poor performers, have led to the performance improvement.

1.13 Characteristic-Based Benchmark Method

Differing from all the regression methods mentioned above, Daniel, Grinblatt, Titman and Wermers (1997) initiate a method of using characteristic-based benchmarks in performance measurement. The characteristic-based benchmarks are constructed from the returns of over 100 passive portfolios that are matched with stocks held in the evaluated portfolio on the basis of different characteristics of these stocks (i.e. market capitalization, book-to-market ratio and prior-year return). They suggest that their direct approach can provide better *ex-ante* forecasts of the cross-sectional patterns of future returns, and exhibit more statistical power to detect abnormal performance than the factor models. More importantly, this approach allows them to decompose fund returns into several components, which could be a more accurate way to determine how the mutual funds generate returns.

Further, based on direct evidence, Kothari and Warner (2001) verify that the characteristic-based benchmark method does exhibit higher power than factor models (regression-based methods), after they compare the properties of various fund performance measures using a simulation procedure.

1.2 The Stochastic Discount Factor methods

More lately, the concept of Stochastic Discount Factor (SDF) in modern asset pricing models is used for performance evaluation. The basic idea is simple, that is, all financial asset pricing models imply that the return of any asset *i*, $R_{i,t+1}$, multiplied by a stochastic discount factor, m_{t+1}, has a constant expectation:

$$E_t(m_{t+1} R_{i,t+1})=1, \; all \; i=1,\ldots,N. \qquad (2.9)$$

The gross return $R_{i,t+1}$ is defined as $(P_{i,t+1} + D_{i,t+1})/P_{i,t}$, where $P_{i,t}$ is the price of the asset *i* at time *t* and $D_{i,t+1}$ is the amount of any dividends, interest or other payments received at time $t+1$.

Equation (2.9) implies that

$$E_t(m_{t+1} R_{i,t+1})-1=0 \qquad (2.10)$$

With this approach, one may measure the abnormality based on the expected risk-adjusted return by discounting back to the present value in a scalar random variable SDF. If the asset is fairly priced, the expectation of discounted gross return should be 1. As a natural extension of this simple rule, one may derive a SDF Jensen's alpha, which is the difference of the expectation of discounted gross return and 1 such as:

$$\alpha_i = E_t(m_{t+1} R_{i,t+1}) - 1 \qquad (2.11)$$

Mutual funds perform well when α_i is positive and badly when it is negative.

The stochastic discount factor *m* has a specific form implied by any specific model that would give the equation further empirical content. Empirical tests of asset pricing models often work directly with equation (2.9) and the relevant definition of *m*. In special cases, the SDF *m* is known as an equivalent martingale measure, a Radon-Nicodym derivative, or an inter-temporal marginal rate of substitution.

For example, in the context of a Consumption-based CAPM (C-CAPM), equation (2.9) arises as a first-order condition for a consumer-investor's optimization problem, and a benchmark variable is defined by the model. The agent maximizes a lifetime utility function of consumption, which is denoted by *U*(.). If the allocation of resources to consumption and investment assets is optimal, it is not possible to obtain higher utility by changing the allocation. Suppose an investor considers reducing consumption at time *t* to purchase more of (any) asset. The utility cost at time *t* of foregone consumption is the marginal utility of consumption expenditures C_t, denoted by $MU_t=(\partial U/\partial C_t)>0$, multiplied by the price $P_{i,t}$ of the asset, measured in the same units as the consumption expenditures. The expected utility gain of selling the security (equity *i*) and consuming the proceeds at time $t+1$ is

$$E_t\{(P_{i,t+1}+D_{i,t+1})\, MU_{t+1}\}$$

where $D_{i,t+1}$ is the dividend of security i at time $t+1$.

If the allocation maximizes expected utility, the following must hold:

$$P_{i,t}E_t\left(\frac{\partial U}{\partial C_t}\right)=E_t\left(P_{i,t+1}+D_{i,t+1}\right)\left(\frac{\partial U}{\partial C_{t+1}}\right)$$

which is equivalent to equation (2.9), with

$$m_{t+1}=\frac{\frac{\partial U}{\partial C_{t+1}}}{E_t\left(\frac{\partial U}{\partial C_t}\right)} \qquad (2.12)$$

where m_{t+1} is the intertemporal marginal rate of substitution (IMRS) of the consumer. When the equation (2.12) defines m_{t+1}, equation (2.9) is the consumer's intertemporal Euler equation. The Euler equation is a necessary condition for an individual consumer's optimization problem.

The choice of utility function is arbitrary. Since the empirical test has shown that the exponential utility function cannot support the C-CAPM framework, it could be more appropriate to adopt a habit utility function etc.

Other SDF models include linear factor models, the primitive-efficient SDF models; Long's Numeraire portfolio model and Bakshi and Chen's model. Each inherits a specific way to construct the stochastic discount factor. Recently, the models have been applied to portfolio performance evaluation.

Since Dybvig and Ingersoll (1982), the linear relation between the SDF and the market portfolio return in the CAPM has been discovered. The same relation will also hold when the single-factor model (CAPM) is expanded to multi-factor models (see Cochrane 2001). The models where the stochastic discount factor is linear in pre-specified factors are known as linear factor models, with *m* constructed as:

$$m_{t+1} = a + b' \boldsymbol{f}_{t+1}$$

where $\boldsymbol{f}_{t+1}$ is the factor vector. The CAPM is one such model in which m_{t+1} is a linear function of the market portfolio return (see Dybvig and Ingersoll, 1982).

The primitive-efficient SDF model comes from Chen and Knez (1994)[26]. The basic idea is to construct a minimum variance-efficient portfolio as a benchmark in the place of a market index. In the primitive-efficient model, the SDF is a linear function of a combination of the primitive assets available to fund managers such as:

$$\boldsymbol{m_t = R^* = 1' E_{t-1}(R_t R_t')^{-1} R_t}$$

More details of the primitive efficient SDF model will be presented in chapter 3.

In Long (1990), the SDF is found as the inverse of the gross rate of return on a "numeraire portfolio", which is a self-financing portfolio with positive value and the best conditional forecast return zero. Kang (1995), Hentschell, Kang and Long (1998)

[26] See Chen and Knez (1994), which utilized a conditional version of the method of Grinblatt and Titman (1989).

and Farnsworth, Ferson, Jackson and Todd (2000) apply a numeraire portfolio to evaluate fixed income mutual funds, international bonds and artificial mutual funds respectively. The SDF here is a non-linear function, although they still assume the weights are linear functions of the information instruments in conditional models. Unlike the primitive-efficient SDF model, the numeraire portfolio conditional model is different to an unconditional model applied to the dynamic strategy returns.

The SDF becomes an exponential of a linear function of the log returns on the primitive assets in Bakshi and Chen (1998)'s model. The formulation implies that the SDF has to be positive and it also has an implication that the non-arbitrage condition should be met in this non-linear SDF model.

Faced with various methods, the problem is to find a suitable SDF to fulfil the alpha statistic. It is important to identify the sensitivity of performance measures to the specification of the SDF. To detect the true pricing ability of various SDF models, Farnsworth, Ferson, Jackson and Todd (2000) construct artificial funds and apply a common experimental design, in which they send random signals to each fund and observe their market timing and stock picking ability according to these signals. Under the null hypothesis of no abnormal performance, they show that many of the models are biased in generating negative alpha, while there are few different results from using the various SDF models.

Among those SDF models, we will focus on the implication of the SDF primitive efficient models. There are various reasons for doing so. Firstly, the primitive-efficient SDF is a very general way, which is independent of a lot of traditional assumptions in conventional pricing theories. Secondly, practically, C-CAPM and other models have not displayed satisfactory results with various pricing anomaly, such as the risk-free puzzle, which raised doubt about the valid of this theory.

Compared with the traditional methodologies, the SDF models are more parsimonious in generating fewer estimators. The general approach to estimate the SDF is the Generalized Method of Moments (GMM, Hansen, 1982). As GMM can handle both heteroscedasticity and serial correlation in pricing errors (with appropriate weights related to various variance, as well as instrumental variables), the specific distributional assumptions of the asset returns are not required, and we do not need to work in a normal independently and identically distributed (IID) setting.

However, there are still several issues and arguments associated with the SDF model and the GMM methods:

Firstly, dubious about the explanatory power of the SDF estimators, Kan and Zhou (1999) use asymptotic theory and Monte Carlo simulations to compare the performance of the traditional and the SDF methods. They find out that, due to their volatile moment conditions, it is more difficult for the SDF linear factor model to obtain accurate estimations and exact identification tests rather than traditional methods. Meanwhile, if the mean and variance of the factor are known, GMM estimates of non-traded factors' risk price are less precise relative to those from OLS or MLE. The same could be true in non-linear SDF models, which can be always be approximated by a linear one.

Secondly, Jagannathan and Wang (2000) show if additional moments conditions, identifying the mean and factor variance, are appended to the system, the efficiency of GMM estimates of factor premiums can be identical to OLS. Further, Jagannathan and Wang (2002) argue that Kan and Zhou (1999) make an inappropriate comparison of the SDF method and the beta method for estimating the parameters related to the factor risk premium. They also show that asymptotically, the SDF method provides as precise an estimate of the risk premium as the beta method. They also demonstrate that the two methods provide equally precise estimates in finite samples as well. In a word, they show linearizing nonlinear asset pricing models and estimating risk premiums using the beta method will not lead to an increase in estimation efficiency.

Thirdly, from Ferson's (2000) experiment, they show that the pricing accuracy of the SDF models can be enhanced through ancillary moment conditions. When they impose the restrictions that the trade-factor model can price the factors, the pricing errors are reduced. The same is true, and more apparent, when they force the non-traded factor models to price the risk-free asset. The issues of GMM estimation will be discussed in chapter 3.

2. Conditional Performance Evaluation

Assuming that the market participants do not use information about the state of the economy to form their expectations, traditional asset pricing models use only unconditional expected returns, where the factor loadings are constant. This implies the linear models are right only when the systematic risk characteristics of the securities held in the portfolio remain fixed and when the portfolio weights remain fixed through time and as well. However, if the expected returns and risks vary over time, such an unconditional approach may give biased results. To remove the impact of variation risks on the biased results, it is important to incorporate information into the performance measurement. This is the, so-called, conditional method.

Since unconditional asset pricing models are unreliable, performance evaluation based on unconditional asset pricing models are also unreliable. Ignoring the changing information condition, abnormal fund performance results from unconditional measurements could be attributed to only public information. However, based on the semi-strong efficient market hypothesis, all predetermined information, such as stock dividend, interest rates, the slope of the term structure, quality spread in bond market, are assumed available to every market participant as a criterion to predicate the stock returns. Superior performance can only be achieved based on private information that is beyond public information.

As a result, conditional performance evaluation is proposed, with an aim to identify the superior investment ability based on private information from those based on public information. Unlike the unconditional approach, conditional performance evaluation compares the return of a mutual fund to the return of a dynamic strategy that attempts to match the fund's risk exposure. The difference is the conditional alpha. Conditional performance evaluation allows a fund's risk exposures and the related market premiums to vary over time with the state of the economy.

From the perspective of academics, the conditional approach plays a significant role in improving performance evaluation accuracy. From the perspective of the investment industry, it also has implications for active managers selection and can lead to better asset allocation decisions and more equitable compensation schemes, because the conditional approach helps to differentiate between passive effects, effects produced by using public information, and effects from using better than public information, after incorporating the information instruments into performance measure analyses. Conditional performance evaluation of UK unit trusts examines the semi-strong form rather than the weak form of market efficiency.

2.1 Asset Pricing Models: Conditional?

We firstly discuss why an unconditional asset pricing model does not necessarily imply a conditional asset pricing model. This is true for the conventional beta pricing models, the SDF models and also the time-varying second moments.

2.11 The Beta Pricing Models

The basic principle underlying the conditional approach is the law of iterated expectations, which states that your best forecast today of your best forecast tomorrow is the same as your best forecast today. This law also allows conditional moments to be conditioned down as unconditional moments in the way of

$$E(E_t(x))=E(x)$$

and also helps conditioning down from agents' fine information sets to a coarser set

$E[E(R \mid \Omega) \mid Z_t \subset \Omega] = E(R \mid Z_t)$

The final rule then enables us to interpret the unconditional measure as a conditional measure on a coarser information set.

For a conventional single/multiple beta pricing model, the equation when the expected return and the beta are varying with information can be expressed as follows:

$$E(R_{i,t+1} \mid Z_t) = \alpha + \boldsymbol{\beta_{it}\lambda_t} \quad (2.13)$$

where $E(R_{i,t+1} \mid Z_t)$ is the expected return of portfolio i at time $t+1$ based on information Z_t, α is the constant, $\beta_{it} = \dfrac{Cov(R_{i,t+1}, R_{b,t+1} \mid Z_t)}{Var(R_{b,t+1} \mid Z_t)}$, is the conditional beta vector of asset i with the benchmark return and

$\lambda_t = E(R_{b,t+1}) - R_f$, is the expected excess return on the benchmark portfolio over a risk-free return. When λ_t is a single factor loading (eg, excess market return), β_{it} is correspondingly a single element, equation (2.13) addresses the basic relation of the CAPM.

As argued by Jagannathan and Wang (1996), this conditional version of the CAPM need not imply the unconditional CAPM. Taking unconditional expectations of equation (2.13), we get

$$E\{E(R_{i,t+1} \mid Z_t)\} = E(R_{i,t+1}) = \alpha + E(\beta_{it})\, E(\lambda_t) + Cov(\beta_{it}\lambda_t) \quad (2.14)$$

Here $E(\lambda_t)$ is the unconditional expected excess return on the benchmark portfolio. $E(\beta_{it})$ is the unconditional expectation of the conditional beta, which needs not be the same as the unconditional beta, although the difference is likely to be small. In this case, even if the covariance term is zero, it *does* not necessarily imply that equation (2.13) and equation (2.14) are equivalent. Suppose $E(\lambda_t) = \lambda_{t,}$ a constant, but the betas vary over time. The average of a conditional beta is not the same as the unconditional beta. Even if the conditional betas are constant (for example, both $Cov(\beta_{it}\lambda_t)$ and $Var(R_{m,t+1} \mid Z_t)$ vary over time to keep their ratio constant), the unconditional beta need not to be equal to this constant since the unconditional beta is the ratio of the unconditional covariance to the unconditional variance. It seems the only chance of making no difference between equation (2.13) and (2.14) is by coincidence. That is, the betas are constant, with both the conditional covariance and variances that form them being constant.

2.12 The 'SDF'

In the context of the SDF model, let us firstly assume the conditional SDF model is equal to its unconditional version such as:

$$E\,(m_{t+1} R_{t+1} \mid Z_t) = E\,(m_{t+1} R_{t+1}) = 1 \qquad (2.15)$$

where m_{t+1} is the stochastic discount factor, R_{t+1} is the portfolio return at $t+1$.

Equation (2.15) holds only if the stochastic discount factor is constant over time. Examples include the consumption-based CAPM model with power utility ($m_{t+1} = \beta\,(C_{t+1}/C_t)^{-\gamma}$) and the log utility CAPM ($m_{t+1} = 1/R_{t+1}^{m}$).

However, for those models with time-varying parameters, this need not be the case. For example, $m=a-bR_m$ in an SDF version of CAPM, where a and b are functions of $E_t(R_m)$ and $Var_t(R_m)$. With varying $E_t(R_m)$ and $Var_t(R_m)$, the SDF is unlikely to keep constant.

2.13 The Time-Varying Second Moments

As shown above, the beta and the SDF are rarely constant in a conditional framework. In addition, the covariance between the conditional beta and the expected excess benchmark return in (2.14) is also unlikely to be zero or remain constant. It is then important to study the time-varying second moments.

One argument advocates that there is a positive covariance between betas and benchmark excess returns. Assets whose betas are high when the market risk premium is high will have higher unconditional mean returns than would be predicted by the unconditional CAPM. Jagannathan and Wang (1996) display that the high average returns on small stocks might be explained by this effect if small-stock betas tend to rise at times when the expected excess return on the stock market is high. They present some indirect evidence for this story, although they do not directly model the time variation of small-stock betas.

The spirit of this argument is similar to that of market timing models, which attempt to detect if fund managers can deliver such a positive covariance. If they can, that is, they magnify (lessen) the positive (negative) excess market return by increasing (decreasing) betas; they will be regarded as superior market timers. Therefore, one can measure the unconditional covariance term to examine if the unconditional market timing ability exists, or model a conditional covariance to measure the conditional market timing. Market timing models are discussed further in section 3.1.

On the contrary, Ferson and Schadt (1996) suggest that a negative covariance exists. They assume that fund managers prefer stable risk exposure and conclude that with the changing state of the economy, specifically financial market conditions, the betas will also become variable due to the changing style employed by the managers to keep volatility relatively stable. Time-variation in a managed portfolio may come from three different sources. First, the weights of a passive strategy, i.e. a buy-and-hold strategy, will vary as relative values change. Second, the betas of the underlying assets may change over time. Third, a manager can adjust the portfolio weights by departing from a buy-and-hold strategy. Based on this assumption, the increase in beta could imply that the factor sensitivity of underlying assets increase, or that the fund manager increases the weight in active strategy or totally departs from a passive strategy, all of which generate higher risks. In the sense, to keep stable risk exposure, fund managers will tend to increase beta when the market return is low (low market risk) and reduce beta when market return is high (high market risk). It implies that $Cov\ (\beta_{it}\lambda_t)$ in equation (2.14) would be negative.

Applying this assumption to performance evaluation, they conclude that the alpha in unconditional models, where betas are constant, will tend to have distorted results since beta and market return have such a negative relation. Therefore, the negative performance under unconditional measurement is not because of the bad investment skills but because of taking more risks when the market risk premium is low. On the contrary, the unconditional approach tends to overvalue mutual funds' performance when market premium is high. Under this circumstance, the true beta set by the manager will be lower than the constant beta set in the model. This assumption provides a good explanation for some empirical results, in which mutual funds have neutral performance by using the time-varying betas conditional model, but have a negative performance when the relevant unconditional measurement is employed.

Further, Ferson (1995) illustrated the necessity of measuring time-varying second moments in terms of a gross return-based SDF expression.

From equation (2.15), one can easily derive equation (2.16) as:

$$E(R_{i,t+1}) = \frac{1}{E(m_{t+1})} + \frac{Cov(R_{i,t+1,} - m_{t+1} | Z_t)}{E(m_{t+1})} \qquad (2.16) \qquad \text{for all } i.$$

where $R_{i,t+1}$ is the gross return of portfolio i at time $t+1$.

Given $E^{-1}(m_{t+1} | Z_t) = R_{ft}$, for any instruments Z_t, i.e. risk-free gross return R_{ft}, equation (2.16) implies the following regression equation:

$$R_{i,t+1} - R_{ft} = C_{it} + C_{it} r_{ft} + U_{i,t+1} \qquad (2.17)$$

where $C_{it} = Cov(R_{i,t+1}, -m_{t+1} | Z_t)$, $R_{ft}=1+ r_{ft,}$ and $U_{i,t+1}$ is a forecast error with conditional mean zero. If there is a benchmark variable for which the conditional covariances are constant parameters over time, then a time-series regression of excess returns on the risk free interest rate (r_{ft}) and a constant should have slopes and intercepts both equal to this constant. Fama and Schwert (1977) run a regression of asset returns on the Treasury bill rate and they find large, negative slope coefficients and positive intercepts for common stocks, using 1953-1971 data. Ferson (1989) uses more recent data and other assets and comes up with a similar conclusion. These results imply the C_{it} cannot be a constant. Therefore, a model should allow the conditional covariance of returns with a benchmark pricing variable to change as the level of the interest rate changes.

2.2 The Methods of Incorporating Conditioning Information

The arguments mentioned above provide a general motivation for empirical modelling of both timing-varying first and second moments. These empirical modelling methods include a direct approach and an instrumental variable approach.

2.21 The Direct Approach

A direct approach is to specify and estimate explicit statistical models of conditional distributions of asset returns and other variables. For example, we can express a conditional first moment as

$$E(r_{m,t+1} \mid Z_t)=f(Z_t) \qquad (2.18)$$

where the function $f(Z_t)$ determines how the average value of $r_{m,t+1}$ changes as the elements of Z_t change. Because $E(r_{m,t+1} \mid Z_t)$ is an expectation, it can be obtained from the conditional density of $r_{m,t+1}$ given Z_t by integration, summation, or a combination of the two (depending on the nature of Z_t). It follows that the conditional expectation operator has the same linearity properties as the unconditional expectation operator, and several additional properties that are consequences of the randomness of $f(Z_t)$. Most often in econometrics, a model for a conditional expectation is specified to depend on a finite set of parameters, which gives a parametric model of $E(r_{m,t+1} \mid Z_t)$.

With this procedure, in practice, we can examine all of a model's implications about conditional moments, but the number of required parameters can exceed the number of observations quickly, if we make the conditional mean, variance, and other parameters of the distribution of, say, N returns depend flexibly on M information variables. More importantly, this explicit approach typically requires an assumption that the investors use the same model of conditioning information as we do. However, we obviously do not even observe all the conditioning information used by economic agents, and we cannot include even a fraction of observed conditioning information in our models. In another words, the information set included by academic researchers in the empirical investigation is normally the subset of the information set observed by market participants. This, therefore, increases the difficulty in establishing an explicit unbiased conditional model.

2.22 The Instrumental Variable Approach

Campbell (1987) and Harvey (1989, 1991) have suggested that one can estimate the parameters by linking the first and the second moments, with the Generalized Method of Moments approach. These authors start with a model for the "market" return that makes the expected market return linear in its own variance, conditional on some vector Z_t, containing information instruments or forecasting variables:

$$E(r_{m,t+1} \mid Z_t) = \gamma_0 + \gamma_1 Var(r_{m,t+1} \mid Z_t) \qquad (2.19)$$

assuming conditional expected returns are linear in the instruments

$$E(r_{m,t+1} \mid Z_t) = Z_t \boldsymbol{b_m} \qquad\qquad and$$

$$Var(r_{m,t+1} \mid Z_t) = (r_{m,t+1} - Z_t \boldsymbol{b_m})^2$$

The errors are:

$$\mu_{m,t+1} = r_{m,t+1} - Z_t \boldsymbol{b_m}$$

$$e_{m,t+1} = r_{m,t+1} - \gamma_0 - \gamma_1 (r_{m,t+1} - Z_t \boldsymbol{b_m})^2$$

Here $\boldsymbol{b_m}$ is a vector of regression coefficients of the market return on the instruments. The error $\mu_{m,t+1}$ is the difference between the market return and a linear combination of the instruments, while the error $e_{m,t+1}$ is the difference between the market return and a linear function of variance. Equation (2.19) implies that the errors $\mu_{m,t+1}$ and $e_{m,t+1}$ are both orthogonal to the instruments Z_t.

This approach can easily be extended to a model where the expected return is a linear function of the covariance.

$$E(r_{i,t+1} \mid Z_t) = \gamma_0 + \gamma_1 Cov(r_{i,t+1}, r_{m,t+1} \mid Z_t) \qquad (2.20)$$

If there are N such assets, we can define a vector as $\boldsymbol{r}_{t+1} = [r_{1,t+1}, \ldots r_{N,t+1}]'$.

The conditional expectation of $\boldsymbol{r}_{t+1}$ is given by

$E[\boldsymbol{r}_{t+1} | Z_t] = | Z_t\boldsymbol{B}$, where $\boldsymbol{B}$ is a matrix with NL coefficients. The errors in this case are:

$$\mu_{t+1} = r_{t+1} - Z_t\boldsymbol{B}$$

$$e_{t+1} = r_{t+1} - \gamma_0 - \gamma_1 (r_{t+1} Z_t\boldsymbol{B})(r_{m,t+1} - Z_t\boldsymbol{B}),$$

Harvey (1989) further generalizes the model to allow for a time-varying price of risk. He replaces (2.20) by

$$E(r_{i,t+1} | Z_t) = \gamma_0 + \gamma_{1t} Cov(r_{i,t+1}, r_{m,t+1} | Z_t) \qquad (2.21)$$

where γ_{1t} varies through time but is common to all assets. Since (2.21) holds for the market portfolio itself,

$$\gamma_{1t} = \{E(r_{m,t+1} | Z_t) - \gamma_0\} / Var(r_{m,t+1} | Z_t) \qquad (2.22)$$

He substitutes (2.22) into (2.21), multiplies through by $Var(r_{m,t+1} | Z_t)$, and uses $E(r_{m,t+1} | Z_t) = Z_t\mathbf{b_m}$ and $E[\boldsymbol{r}_{t+1} | Z_t] = Z_t\boldsymbol{B}$ to construct a new error vector

$$\boldsymbol{V}_{t+1} = (r_{m,t+1} - Z_t\boldsymbol{b}_m)^2 (Z_t\boldsymbol{B} - \gamma_0) - (r_{m,t+1} - Z_t\boldsymbol{B})(r_{t+1} - Z_t\boldsymbol{B})(Z_t\boldsymbol{b}_m - \gamma_0)$$

Applying this model, Harvey (1989) finds some evidence that the price of risk varies when a US stock index is used as the market portfolio; and He (1991) uses a world stock index as the market portfolio, obtaining similar results.

A slightly different approach is adopted by Admati and Ross (1985). Assuming the investment agents to maximize a constant absolute risk aversion expected utility function with normal distributions; they develop a rational expectations equilibrium capital asset pricing model, and find out that the fund manager will choose the optimal portfolio weights as a linear function of the information. With CAPM as the underlying asset pricing model, beta is concerned as a function only of the information variable Z_t.

$E\,(\beta_t \mid Z_t) = f\,(Z_t)$

To approximate the function linearly using the Taylor series, beta in conditional CAPM can be expressed as an average beta plus the product of a response vector with the deviations of the information variable from the unconditional means.

$$E_t\,(\beta_t \mid Z_t) = b_{op} + \boldsymbol{B}_p \boldsymbol{z_t} \tag{2.23}$$

where $\boldsymbol{z_t}$=Z_t-$E(Z_t)$ is a vector of the deviations of Z_t from the unconditional means, and $\boldsymbol{B_p}$ is a vector with dimension equal to the dimensions of Z_t. The coefficient b_{op} may be interpreted as an "average beta". i.e. the unconditional mean of the conditional beta: $E(E_t\,(\beta_t \mid Z_t))$. The elements of $\boldsymbol{B_p}$ are the response coefficients of the conditional beta with respect to the information variables Z_t.

In the context of the SDF model, Cochrane (2000) employed an implicit pricing interpretation. Based on the basic SDF model formulae, p_t=$E_t(m_{t+1} X_{t+1})$, he multiplies the payoff and price by instrument Z_t observed at time t and names this portfolio, with payoff $x_{+1}\, Z_t$ and price $p_t\, Z_t$, the managed portfolio. Suppose instrument Z_t can be used to forecast the expected asset returns, investors might adjust their instrument according to the change of Z_t. If the investors follow a linear rule, they will put only Z_t p_t units into the asset each period and receive Z_tX_{t+1} units the next period. Based on this interpretation, all the asset pricing theory can be applied directly and there is no restriction on the linearity of $X_{t+1}Z_t$. All the nonlinear transformations of time-t instruments can be considered if the investors want to follow a non-linear rule, eg, investing $Z_t^{\ 2}+4\ Z_t^{\ 3}$ units with payoff $(Z_t^{\ 2}+4\ Z_t^{\ 3})\ x_{t+1}$. Practically, there is no need to incorporate all linear and nonlinear transformations of all variable, which will be a large amount. Cochrane argues that one needs only to scale by those instruments which can forecast returns. If adding instruments is the same thing as including potential managed portfolio, then the choice of a few instruments is the same thing as the choice of a few assets or portfolios that one makes in any test of an asset pricing model.

Cochrane then displays an example of a SDF single-factor model. In the conditional context, the parameters

$$E(m_{t+1}|\ Z_t) = E(a_t|\ Z_t) + E(b_t|\ Z_t)\, f_{t+1} \qquad (2.24)$$

in such a model may vary over time.

Using a linear model, we substitute

$E(a_t|\ Z_t) = a_0 + a_1 Z_t$ and $E(\ b_t|\ Z_t) = b_o + b_1\ Z_t$ into equation (2.24),

We can have

$E(m_{t+1}|\ Z_t) = a_0 + a_1 Z_t + b_0 f_{t+1} + b_1(Z_t f_{t+1})$

With constant coefficients, we can interpret Z_t, f_{t+1}, $Z_t f_{t+1}$ as factors. Compared with the unconditional single factor model, the conditional model is equivalent to an unconditional three-factor model, incorporating two more factors which are observed information instruments and can be used to predict asset returns. If summarizing such a three-factor model by a single-beta representation, the benchmark changed from a market portfolio in the model of CAPM to a mimicking portfolio of the three factors.

Similar multiple-factor representation can be shown as $E(m_{t+1} | Z_t) = b'(f_{t+1} \otimes Z_t)$, where $\otimes$ is the Kronecker product, indicating multiplying every element in vector f_{t+1} by every element in vector Z_t.

2.3 Previous Empirical Evidence

Since Jensen (1972) and Grant (1977), this problem of variation in risk and risk premium of mutual funds has been recognized. Ross and Admati (1985) use a Sharpe ratio to identify if the manager can capture better information than others do. The larger the Sharpe ratio, the better his information is. Grinblatt and Titman (1989) use predetermined information on the attributes of firms to develop performance benchmarks while they use unconditional expected returns as the baseline for measuring performance. Glosten and Jagannathan (1994) use a contingent claims approach to address nonlinearities that may arise when managers engage in dynamic strategies. Sirri and Tufano (1992) use rolling regressions for Jensen's alpha, an approach that may approximate conditional betas.

The conditional performance evaluation studies, by Chen and Knez (1996) and Ferson and Warther (1996) find that conditioning on the state of the economy is statistically significant for measuring investment performance. Conditioning also helps control biases in traditional market timing models. Jagannathan and Korajczyk (1986), Ferson and Schadt (1996) show that traditional measures of market timing can assign "negative" timing ability to a passive portfolio strategy, and earlier studies find that measures of timing ability for mutual funds are typically close to zero or negative.

Among the existing literature of conditional measure, the simple linear function forms to model time-varying betas and second moments, by Admati and Ross (1985), remains popular. This approach is attractive for two reasons. For one thing, linear betas can be motivated by theoretical models of manager behaviour; for another, the linear regression models which result from this assumption are easy to interpret. Moreover, as suggested by Ferson (1996), if interpreting the product of market return and the lagged information variables as the returns to dynamic strategies, which hold Z_t (take the lagged information vector as a dynamic weights) units of the market index and replicate time-varying risk exposure, the expression for conditional factor models can be understood as an unconditional factor model.

Chen and Knez (1996) argue that many types of dynamic portfolio strategy models cannot be applied if the weight function of the portfolio is linear. In another word, they suggest that within a linear function framework, the conditional models can be biased as all public information-conditional portfolio returns are generated by linear trading strategies rather than dynamic trading strategies. Nevertheless, it suggests the importance of a better specified model for the conditional performance measure. The question remains, such as, taking both efficiency and parsimonious specification into account, would a complicated functional form rather than a linear function be a better candidate? For instance the approach by Glosten and Jagannathan (1994), who capture the nonlinearities in managed portfolio returns with option payoffs as additional factors when calculating the alphas.

Another concern was addressed by Merton (1973), who argued that the investors might not optimally choose conditional mean-variance efficient portfolio in a dynamic model, which would make the assumptions of the CAPM invalid. In this sense, multi-factor benchmarks can be more appropriate in conditional performance measures. In addition, as argued by Longstaff (1989) and Ferson and Korajczyk (1995), the optimal investment horizon of the investors could be an endogenous variable in a dynamic model. Therefore, incorporating the return measurement interval as a dependent variable into the existing model could be a major concern in further research. One more problem is the difficulty in defining and measuring market risks with asymmetric information, especially when one considers that it must be evaluated by an uninformed observer. For this reason, there has been a great deal of controversy over if the performance measures proposed by Treynor (1965), Sharpe (1966), and Jensen (1968, 1969) can identify investors with superior information.

Given the importance of incorporating conditioning information into performance measures, we examine the conditional performance of UK Unit Trusts within the context of the SDF framework in chapter 4.

3. Performance Attribution

After assessing portfolio performance, the sequential question is how to interpret these results. The question of which factor has made a more significant contribution to the superior performance has long been important, both for academic research and for practical decision making.

From the perspective of a professional manager, it is argued that the positive alphas (superior performance) can be generated by fund managers via three sources: Superior (private) information, better information processing and behavioural biases. Professional investors who try to generate superior information are regarded as fundamental/traditional managers, and they form the majority of active investors. Investors who try to process information more effectively are called quantitative managers, who aim at developing quantitative methods to model and forecast the asset price movements. Behavioural managers therefore refer to those who try to explore situations where the securities are mispriced by the market because of behavioural factors. Sometimes, overlap exists among all three types of managers.

From the perspective of academic research, the power to detect genuine superior performance differs across the models. It is discovered that the factors that could contribute to the abnormal performance include: superior management skills, certain styles of holding portfolios, transaction costs, and survivorship bias etc. For example, Hendricks *et al.* (1993), Brown and Goetzman (1995), and Wermers (1996) attribute the superior performance to "hot hands" or common investment strategies. Grinblatt and Titman (1992) attribute long-term persistence to the manager differential information or stock-picking talents. Grinblatt, Titman and Wermers (1995) and Daniel *et al.* (1997) discover that the characteristics of the stocks held by the funds can account for much of the results. For example, the value fund (the fund with a value-investing strategy) tends to outperform because it tends to hold stocks with higher average returns than passive stock indices, whereas Carhart (1997) suggests that momentum-factor and transaction cost are better explanations. Further, Allen (2007) suggests that size (Asset under management: AUM) of the fund also contributes to fund performance. His research discovers that in capacity-constrained asset classes (Small-cap equity, High yield fixed income), all things held equally, there is a significant negative relationship between the AUM and performance. Moreover, the information ratios for these types of assets are similar or higher than their large-cap counterparts, which suggests that in less liquid asset categories, small products can have some sort of systematic advantage. However, this result does not apply to core fixed income funds and emerging market equities funds. This makes size unique among quantitative variables in its ability to explain relative performance across a broad cross-section of managers and products.

More practically, when assessing the abilities of fund managers, the conventional approach is to separate managerial skills into market timing and stock selectivity skills and try to identify if the fund managers can use any or all of these skills to time the market correctly or whether they have excellent foresights in predicting the stocks' future returns, consequently, picking the winning stocks. Besides superior skills, other factors play the equivalent important roles in interpreting the abnormal performance. The examples include style analysis. It is argued that the equities with different styles perform differently purely due to their characteristics. In this section, we provide a broad overview of interpretations for abnormal performance from the perspective of academic research. We firstly review the research on the market timing and stock selection abilities, we then provide a brief survey of style studies.

3.1 Market Timing Ability

3.11 Unconditional Market Timing Model

Market timing models aim at examining the ability of mutual fund managers to time the market, that is, to increase a fund's exposure prior to market advances and to decrease exposure prior to market declines. The idea is as follows: if the fund managers could predict bull and bear markets, they then will shift more weights into the market portfolio when the market is about to go up. This will generate a higher beta, which will then magnify the returns. On the contrary, a portfolio with a lower beta when the market is going down can help to lessen the portfolio returns. In both situations, a positive market timer will generate superior performance than otherwise.

Assuming the excess returns of a managed portfolio incorporating timing ability will be a quadratic function of excess market returns, Treynor and Mazuy (TM) suggested a regression (1966) such as:

$$r_{pt+1} = \alpha_p + \beta_p r_{m,t+1} + \gamma_{tmu}[r_{m,t+1}]^2 + u_{pt+1} \qquad (2.26)$$

where the coefficient γ_{tmu} measures market timing ability, α_p is Jensen's alpha. This framework is derived from the traditional CAPM. The CAPM assumes that the portfolio return is a linear function of the market return, while the TM measure assumes a convex relation between the portfolio return and the market return. If γ_{tmu} is positive, it implies that a mutual fund manager has market timing ability as he increases/decreases the portfolio's market exposure prior to a market increases/decreases. Following the CAPM assumptions, the TM framework assumes that the portfolio only captures the market risk while the individual risk can always be diversified away. If this assumption is invalid, one may always argue that an increased beta is associated with increased individual risk, thus even if the risk and the return have a positive correlation, the risk-adjusted return with both increased individual risk and return could remain the same level or less than otherwise.

Henriksson and Merton (HM) (1981, 1984) develop an approach with a dummy variable to estimate the timing performance. The regression used is similar to the TM regression, except that $r^*_{m,\ t+1}$ is used in place of the quadratic term on the right-hand side of the regression, as shown in equation (2.27). Like the TM regression, the model used to develop this regression is based on a narrow set of behavioural assumptions. In contrast to the linear beta adjustment of the TM framework, the portfolio beta in the HM study is assumed to switch between two betas: a high beta corresponding to a large forecasted benchmark return and a low beta corresponding to a forecasted benchmark return that is less than the risk-free return.

$$r_{p,\ t+1} = \alpha_p + \beta_p\, r_{m,\ t+1} + \gamma_{tmu}\, r^*_{m,\ t+1} + u_{p,\ t+1} \quad (2.27)$$

where $r^*_{m,\ t+1} = r_{m,\ t+1}\, D$ and D is a dummy variable that equals to 1 when $r_{m,\ t+1}$ is positive and 0 otherwise. The magnitude of γ_{tmu} in equation (2.27) measures the difference between the target betas, and is positive for a manager that successfully times the market. They suggest that the beta of the portfolio takes only 2 values. When excess market return is positive, $r^*_{m,\ t+1}$ will be 1 and zero otherwise. Hence the beta of the portfolio is $\beta_p + \gamma_{tmu}$ in bull market and β_p in bear market.

Chang and Lewellen (1984), Henriksson (1984), Grinblatt and Titman (1988), Cumby and Glen (1990), and Ferson and Schadt (1996) apply the HM technique to samples of mutual funds and do not find evidence that funds were systematically timing the market. If anything, there seems to be evidence of negative timing. The application of this technique to a multi-portfolio benchmark, in Connor and Korajczyk (1991) and Bollen and Busse (2001), reveals similar results.

In a similar spirit, Kone and Jen (1979), Kon (1983) have used switching regression models to estimate fund performance. Rather than forcing one of the betas to be zero, they assume that one of two or more unknown betas is selected and use econometric techniques to infer estimates of them and of their contribution to performance. They also conclude that there is no evidence of timing performance within funds as a group.

Although the adjustment for performance developed with either the quadratic regression or dummy variable approach provide a reasonable estimate of whether market timing exists or not, one may argue that the actual attribution of timing ability or stock selectivity ability to the portfolio return could be estimated with a bias. This is because investment behaviour is unlikely to conform to the rather narrow behavioural assumptions used in these models.

More recently, some authors, such as Daniel *et al.* (1997) and Ferson *et al.* (2000), tend to seek more sensible timing approaches. Specifically, Wermer (2000) applies Characteristic Timing Measure (CT) in timing detection. Further, by combining the market timing and fund volatility literature (see Brown, Harlow and Starks (1996), Busse (1999), Koski and Pontiff (1999)), Busse (1999) examines, in a conditional context, volatility timing ability rather than returns timing ability. That is, if managers exhibit superior market timing skills when market volatility changes over time. This idea of volatility timing measure originated from previous research on results of returns and volatility. Since empirical studies do not find a positive, simple relation between conditional market returns and conditional market volatility (Campbell (1987)), unit trust managers could have an incentive to reduce risk exposure when volatility increases. By doing so, risk-adjusted return is increased, which will have positive relation with a manager's compensation plan. This assumption of managers' behaviour is consistent with Ferson's assumption that fund managers prefer stable risk exposure. However, one may argue that the risk-adjusted return will increase with increased risk if market return and volatility have a positive correlation, which have been the consensus over years in asset pricing theories. Whether fund managers have incentives to increase risk when market risk is high or if they will prefer reduced risk exposure when the market is more risky, either due to their preference of stable risk exposure or due to their incentive of increasing risk-adjusted returns based on the assumption of no relation between expected return and exposure, is still a question.

There are several arguments associated with the market-timing literature. The first one is the matter of observation frequency. As discussed by Goetzmann, Ingersoll and Ivkovic (2000), the monthly-frequency dataset might not be able to capture the managers' timing ability because the decisions are normally made upon the market risk exposure in fewer intervals than one month. To verify this argument, Bollen and Busse (2001) test timing ability by using both a daily-frequency dataset and a monthly-frequency dataset and find out both generate positive results while the daily data seems more powerful in this case. However, argued by Scholes and Williams (1977), the daily dataset may result in biased estimates of variance, serial and contemporaneous correlation between assets and portfolios due to the infrequent trading. To accommodate infrequent trading, Dimson's correction (1979), as well as including lagged values of the risk factors as additional independent variables, may be helpful.

Secondly, Warther (1995), Ferson and Warther (1996) show that the spurious results may exist due to the relation between cash flow and the beta. Warther (1995) discovers a strong relation between a fund's cash inflows and its portfolio weight on cash. Further, Ferson and Warther (1996) find a direct negative relation between funds' betas and funds' cash flows. It is common sense that investors attempt to increase their subscriptions to mutual funds when the market returns are high, which will bring about the temporarily larger cash position, funds move more weights on cash, hence a lower beta. The natural consequence will be a biased timing ability discovery if ignoring this negative relation. In the sense that the timing coefficients tend to move downwards, even to negative levels when the market returns are high, cash flows are large and the betas are low.

Another problem associated with the spurious timing is mentioned by Jagannathan and Korajczyk (1986). They suggest that '*the spurious timing ability can be generated when portfolios hold stocks with payoffs that are more/less option-like than the market proxy. A positive timing coefficient can be found when the average stock of mutual funds is more option-like than the average stock in the market index.'* Further, to control for possible spurious results, for each fund, Bollen and Busse (2001) design a synthetic fund, which exhibits the same characteristics as the actual fund, but has no timing ability. In the sense, the new reference is as the same option-like as the actual fund. They find out that there is significant timing ability among mutual funds, as compared to the synthetic benchmark.

3.12 Conditional Market Timing Model

Based on the earlier work of Shanken (1990), Ball *et al.* (1995), Ferson and Schadt (1996) propose conditional performance evaluation. In a market-timing context, the aim of conditional timing is to distinguish timing ability as captured by a set of instrumental variables and only reflects public information, from timing based on non-public information.

Using essentially the same analysis of conditioning information as Admati *et al.* (1986), a simple conditional TM model was introduced by Ferson (1996). In a two assets model, the managers allocate funds between only the market portfolio and the risk-free asset based on information set Z_t at time t. With exponential utility and normal distribution, Adamti *et al.* (1986) show that the weight on risky asset is a linear function of the information Z_t, and the portfolio beta is just the portfolio weight on the market index, which then it is also a linear function of Z_t. Assume that the manager observes the information Z_t at time t, we can replace equation (2.26) by adding a new term C_p (Z_t $r_{m,\ t+1}$) to control for the public information effect and the regression becomes:

$$r_{p,\ t+1} = \alpha_p + \beta_p\, r_{m,\ t+1} + C_p\,(Zt\, r_{m,\ t+1}) + \gamma_{tmu}[r_{m,\ t+1}]^2 + u_{p,\ t+1} \qquad (2.28)$$

where C_p measures the response of the manager's beta to the public information Z_t. γ_{tmu} captures the manager's timing ability based on private signals. Based on this model, the timing ability based on public information and the private information are separated. In the context of conditional timing measure, the improvement of performance by adjusting beta based on public expectation of future market returns, cannot be considered as superior performance. Similar conditional timing model can be extended from the HM model.

Lately, conditional market timing ability has become a popular topic and been applied to the UK market. For example, Fletcher *et.al.* (2006) examine the conditional market timing performance of UK unit trusts between Jan 1988 to Dec 2002 and find UK unit trusts do not exhibit superior conditional market timing ability. They also provide evidence that UK unit trusts behave more like benchmark investors and have relatively high risk aversion to deviations from the benchmark. Varshney and Liu (2009) examine the stock selection and market timing ability of the UK unit trusts, under Ferson and Schadt (1996)'s conditional performance evaluation framework. Their results suggest that the conditional performance of UK unit trusts produce less negative average performance.

3.2 Stock Selection Ability

Stock selection ability refers to the professional superior skills that the fund managers exhibit in predicting the returns of individual stocks, subsequently, selecting the winning stocks. This strand of literature examines the total fund returns as well as the stock holdings by the fund managers.

In this context, the performance evaluation literature can be divided into two strands. One is to measure the returns of the funds, while another is to assess the performance of the stocks held by the funds. Two sets of the methodologies could result in disparate results. The former measure can be found in most of the studies and the main conclusions suggest that actively managed funds on average, under-perform their passively managed counterparts. Hence, there is no evidence of significant stock picking ability. Such results can be traced back to Jensen (1968), who concludes that over the period 1945 through 1964, the performance of mutual fund can actually be inferior to the performance of randomly selected portfolios with equivalent risk. Further, Gruber (1996) finds that the average unadjusted returns of mutual funds under-perform the market by 1.94% per year and the risk adjusted return is estimated to be –0.65% per year from 1985 to 1994. Carhart (1997) finds that the expense levels are negatively correlated with net returns and for actively managed funds and the expense levels are generally much higher. Worse, Carhart finds that it is likely that the more actively a fund manager trades, the lower the fund's benchmark-adjusted net return to investors. These studies do not suggest any evidence of outperformance of active mutual fund management. Instead, the studies conclude that investors could be on average better off, buying a low-expense index tracker fund. Although on the other hand, different conclusions can be found from a few studies since the 1980s. For example Chang and Lewellen (1984) and Henriksson (1984) find that during the 1970s, net returns to fund investors before load fees lie along the security market line. Ippolito (1989) discovers that from 1965 to 1984, returns before loads, but net of other expenses, are slightly above the Capital Asset Pricing Model market line, though this result may depend on the particular benchmark employed. These studies imply that fund managers with superior stock picking ability can generate higher returns to offset their expenses.

Starting from the studies by Grinblatt and Titman (1989, 1993), developed by Daniel et al (1997), the latter study about fund portfolio holding returns concludes that the managers actively trade possess significant stock-picking talents. In details, Grinblatt and Titman (1989, 1993) and Wermers (1997) conclude that the fund managers, especially those of growth-oriented funds, have the ability to choose stocks which outperform their benchmarks, before any expenses are deducted. Moreover, Daniel *et al.* (1997) and Grinblatt, Titman and Wermers (1995) discover that the characteristics of the stocks held by the funds can explain much of this performance.

It is argued that assessment based on the portfolio holding returns rather than the fund holding returns has captured several advantages. For example, the hypothetical returns generated from the portfolio holdings in Daniel's model (1997), which is gross return, do not include the fees, expenses and trading costs, hence the studies can be immune to the complicated fees calculation while the results will still be appropriate in that they use the benchmark that also ignore the transaction costs. Further, it is apparent that the assessment of portfolio holding returns can provide better and more accurate insights in detecting the funds' stock selectivity ability.

One specific measure worth mentioning is the so called 'CS' measure (Characteristic Selectivity) by Daniel *et al.*(1997). 'CS' measure is calculated as the portfolio-weighted characteristic-adjusted return of the component stocks in the portfolio. The stock portfolio is normalized so that the weights add to one. Consequently, the managers who put higher weights on the maximum-return stocks will generate higher CS values. Outperformance exists when the CS measure is positive. The advantage of this measure is significant in that it helps to decompose the portfolio return and singles out the picking talent from other management skills. As argued by Wermers (2001), there is however, a potential problem with this approach. That is, the narrowness of stocks' characteristic dimensions captured in its benchmark could lead to a biased estimation. To improve, future research may consider how to include more newly-discovered stock characteristic, such as liquidity, into the benchmark.

Another measure of stock-picking ability is constructed by Ferson *et al.* (2000). Consistent with their artificial mutual fund and noise signal designs, they use a signal function, which is a convex combination of an information term (the error term for next month from a CAPM model regression) and a noise term. The managers who can observe the error term one period ahead will be granted as having perfect foresight in stock picking. On the contrary, the managers who can only observe a noise signal capture no stock picking ability.

3.3 Styles

The growth in the number of unit trusts in the UK is a reflection of the variety of investment styles employed by the fund managers. The academic interests focus on several dominant styles as well as their relations with certain performance, the so-called "relative performance" analysis. The performance results may differ depending on different investment styles:

- Active VS Passive : if the fund manager believes outperformance can be achieved by actively trading or investing in a market index passively;
- Growth VS Value : if the unit trust select stocks with high P/E ratios or picks "cheap" stocks with low P/E and high dividend ratios;
- Technical VS Fundamental: if the fund manager advocates analyzing the charts of the stock prices and believes in price movement trends or focuses more on analyzing annual reports and meeting companies;

- Small Vs Big: if the fund manager believes that small-cap stocks can outperform large-cap stocks due to their higher potential for growth or believes that large-cap stocks can provide less volatility as well as higher return.

Furthermore, herding and trend-following/momentum-investing also belong to certain styles. The former refers to any mass movement into particular stocks for whatever reason. The latter refers to a specific type of herding by unit trusts that involve in a large group of funds chasing stocks, which have recently risen in value.

Superior performance generated from following a certain style simply should be isolated from pure stock-picking talents. This point is mentioned very clearly in the approach of Daniel, Grinblatt and Wermers (1997) and Wermers (2001), who decompose the returns into several categories, with stock selectivity generated return and style generated return as independent categories. In this strand of literature, average style measure is designed to measure the returns earned by a fund because of the fund's tendency to hold the stocks with certain characteristics. It is measured as a sum of weighted characteristic-based benchmark stock returns, which is matching each stock's characteristics mentioned above. In addition, based on Morningstar style categories, Teo and Woo (2001) conduct performance persistence test in different types of fund, relative to their peers. Their results suggest that most funds with above-average returns are clustered into certain styles and that a year-to-year variation in style returns could make it difficult to find evidence of performance persistence. however, using style-adjusted performance measures based on several models including Carhart's (1997) four-factor model, they find strong evidence of persistence.

Since the funds' styles have the potential to provide important explanations for performance evaluation, some authors focused on several related issues of styles studies. For example Falkenstein (1996), who studies the preferences of portfolio holdings in order to identify why certain funds tend to hold certain types of stocks, in a relation with the microeconomic studies of the firms as well as the studies of fund styles. The results suggest that like other firms, mutual funds aim to maximize profits by making portfolio-holding decisions. It is mainly specific comparative advantages buying certain types of assets that drive different mutual funds towards holding different types of assets. Falkenstein (1996) detects the fund ownership by applying Powell's censored least absolute deviations, a stock characteristics estimator, which relies on symmetry conditions, imposed by a censored model without the imposed assumptions of homoscedasticity or normality. His semi-parametric regression results indicate that the stock volatility, liquidity, news stories, age, size all have significant positive relations with the funds' stock demands.

Another important issue is to standardize the criterion of style classification. The example includes Brown and Goetzmann (1997), who express their doubt about the importance of a fairly classified style system and the standard equity mutual fund categories. They suggest that the standard categories are too broad to capture the exact style employed by a fund and consequently, not so useful in forecasting future performance. For example, under the same category, the mutual fund managers are given great latitude in the types of stocks to hold, the timing of buying and selling securities, the level of fund diversification, the sector allocation of the portfolio, and several other factors which determine the returns to client investments. Empirical results also suggest that the fund managers tend to misclassify themselves in order to improve *ex-post* relative performance measures. Based on these concerns, they determined style classifications empirically to avoid “window dressing” and provide a return-based style classification algorithm system to standardize the classification criterion. Consistent with Witkowski (1994) and Kim, Shukla, and Tomas (1995), Toe and Woo (2001) show that these style factors typically outperform pre-specified macroeconomic factors in out-of-sample tests on the fund returns, which implies that the style index may provide better interpretation for abnormal performance.

We examine how funds with different styles perform differently, as well as the style-rotation strategy, in chapter 5.

4. Performance Persistence

Besides performance evaluation of the mutual funds, one of the related important issues is the persistence of their performance. It also has the implications for the predictability of mutual funds performance.

The performance persistence literature on the US market is enormous. Examples include Grinblatt and Titman (1992), Hendricks, Patel, and Zeckhauser (1993), Maikiel (1995) and Sharpe (1995). Most results seem optimistic. Specifically, Brown and Goetzmann (1995) use both absolute and relative benchmarks to explore performance persistence and find evidence of performance persistence, robust to adjustment for risk. Notably, they conclude that the results are little affected even if incorporating style-adjusted risk into their measure. Further, they document a special reversal pattern in 1987, which indicates the cross-fund correlation and verifies the persistence is a group phenomenon rather than individual phenomenon. More importantly, this reversal also implies that managers' stock picking ability may not be able to account for the persistence and suggests that the researchers should concentrate on a search on common management strategies. In this sense, it is consistent with Wermers' (1994) herding behaviour identification among funds and correlated dynamic portfolio strategies, such as portfolio insurance (see Connor and Korajczyk (1991)), as well as group phenomenon conditioning upon macroeconomic variables, suggested by Ferson and Schadt (1995).

Unlike Brown and Goetzmann, who conclude that style had no significant influence on risk-adjusted returns, Teo and Woo (2001) argue that style-adjusted fund returns, rather than risk-adjusted returns, should bring about a more appropriate persistence measure. The abnormal performance generated by the mutual funds could be due to excellent management skills or due to their specific styles. Therefore, it is inappropriate to identify the superior management performance in terms of the risk-adjusted return. Further, they find evidence that there is more than 20% of probability for the funds in the best and worst deciles in terms of risk-adjusted return belonging to the best and the worst styles respectively, and the correlation between the return rank of a fund in the decile and the return rank of the style to which the funds in the decile most likely belong is strong. Theoretically and empirically, it is sensible to measure mutual fund performance in terms of certain kinds of style to eliminate any ambiguity. Using a methodology similar to Carhart (1997), they find that the difference in style-adjusted fund returns persist for up to six years.

In the UK market, quite a few academic studies have been focusing on testing the persistence of unit trusts. For more details of the literature review on this strand, please refer to chapter 6.

The typical approach to study persistence in mutual fund performance is the ranking comparison approach, i.e. an approach to estimate the performance persistence effects based on the return ranks of each fund over subsequent sample periods. The first step is to estimate for a given time interval, for each fund separately, a measure of performance, specifically risk-adjusted return (one-factor alpha or a three-factor alpha) or style-adjusted return or time-weighted return.

Secondly, the funds are separated into several deciles (portfolios of funds). In examining the performance of the deciles, the fraction of investors' capital invested in each of the funds in a portfolio could be assumed to be the same, which will generate so called equal-weighted portfolio return (Carhart 1997). Notably, Elton, Gruber and Blake (1996) construct a portfolio of funds with the optimal weights based on modern portfolio theory (MPT) techniques[27]. They also demonstrate that in using MPT, they can select a combination of actively managed portfolios that has the same risk as a passive portfolio (index-tracking funds) but has higher average return. They approve that the performance of such portfolio is improved significantly. The optimal weight (% of portfolio) to place in any fund is given by

$$c_i = \frac{\dfrac{\alpha_i}{Var(\varepsilon_i)}}{\sum \dfrac{a_i}{Var(\varepsilon_i)}}$$

[27] MPT aims at a maximum value of *E(Rp)- Var(Rp)/ T (Rp)* by choosing the optimal weight. *E(Rp)* and *Var(Rp)* are the expected portfolio return and variance of portfolio return, *T (Rp)* is the risk tolerance.

where α_i is the intercept of asset i in a multi-factor regression, and Var(ε_i) is the variance of the random error term in the same regression. This is a generalization of Treynor and Black (1973) and Elton, Gruber and Padberg (1976) criteria and is derived for a multi-factor model in Elton and Gruber (1992).

Finally, statistic metrics such as cross-product ratio based on ranking results are used for persistence tests. Other approaches include a regression-based approach. For example Mark and Titman (1992) employ a dynamic model to estimate the slope coefficient in a cross-sectional regression of abnormal returns computed from the last five years of data on abnormal returns computed from the first five years of data. Consequently, if this coefficient is positive and statistically significant, persistent abnormal performance can be accepted. Further, they argue that a cross-sectional *t*-test may not be suitable for mutual funds due to funds' correlated relations that result in a distorted *t*-distribution. Instead, they compute so-called "time-series *t*-statistics", which is a technique introduced by Fama and MacBeth (1973) to test the CAPM. Other examples include Hendricks, Paten, and Zeckhauser (1993), who test statistically if there is autocorrelation among the error terms. If the correlation coefficient is significantly positive, it implies the existence of performance persistence. Though most literature finds that there is significant persistence in both the short run and long run, the results are not well explained. We explore this issue in more depth in chapter 6.

5. Conclusion

This chapter provides a broad review of portfolio performance evaluation. Focusing on the SDF models and conditional evaluation, this chapter firstly illustrates the methodologies that are used for performance evaluation, from the conventional CAPM model, multi-factor models to conditional SDF models. It is then followed by two parts, investigating performance attribution and persistence.

CHAPTER 3 THE SDF MODELS AND SMALL SAMPLE PROPERTIES OF ALTERNATIVE GMM ESTIMATORS BASED ON THE *J*-STATISTIC

Before formally evaluating the performance of UK unit trusts, we firstly discuss the underlying asset pricing models, the Stochastic Discount Factor (SDF) models, which will be employed throughout all our tests in this thesis, as well as the estimation method, the Generalized Method of Moment (GMM).

As explained in the chapter 2 literature review, we chose the SDF approach as the underlying methodology for performance evaluation. The SDF model has many advantages compared to other asset pricing models and has attracted a lot of attention in the performance evaluation literature; GMM is regarded as a natural fit to estimate the SDF models. The popularity of the GMM estimation is based on its generality, which provides a unifying framework for the analysis of many familiar estimators, including least squares, instrumental variables and maximum likelihood as special cases.

Various weighting matrices used in the GMM estimators can have different effects on performance evaluation, as a result, it is important to examine the characteristics of the GMM estimators, and select the optimal estimator to implement the empirical tests. To do so, within the framework of the SDF primitive efficient model (with different ways to construct the GMM weighting matrices), we use simulations to probe small sample properties of the most widely used moment selection statistic of the GMM estimation, *J*-statistic. In a word, the intention of this chapter is to focus purely on the choice of an optimal estimator for the SDF models within our context of performance evaluation of the UK unit trusts. Our simulation results show that for both GMM two-step and iterated estimators, the sizes of the *J*-statistic can be seriously distorted, whist the GMM-iterated estimator exhibits superior size and power properties compared to its counterparts.

This chapter is organized as follows: In part 1, we illustrate the SDF approach, in particular, the SDF primitive efficient model. Part 2 provides an introduction to GMM estimation, where the GMM estimators with different weighting matrices and the *J*-statistic for the SDF models are explained. The investigation of small sample properties of the *J*-statistic, with the aim to identify the optimal GMM estimators for further empirical tests, is reported in part 3. The conclusions then follow.

1. The SDF Models

1.1 Introduction

The SDF model has its roots in modern asset pricing models and contingent valuation models. The basic idea of the SDF model is simple: a stochastic discount factor or pricing kernel m_t is a random variable that can be used to compute the market prices at time *t-1*, by discounting state by state the corresponding payoffs at a future date at *t* such as

$$E_{t-1}(m_t x_t) = p_{t-1} \qquad (3.1)$$

where x_t is the payoff on asset i at time t, p_{t-1} is the price of asset i at time $t-1$. Or equivalently, the gross return of any financial asset $R_{i,t}$, multiplied by the stochastic discount factor or pricing kernel m_t, should equal 1, such as

$$E_{t-1}(m_t R_{i,t}) = 1 \qquad (3.2)$$

where $R_{i,t}$ is the gross return on asset i from time $t-1$ to t.

When the financial markets satisfy the no arbitrage condition, m_t will be positive in every state of nature (Cochrane (2001)) and the stochastic discount factor will only be unique if the markets are complete.

The popularity of the SDF model lies in its wide choices of the ways of constructing the stochastic discount factor. The SDF has its specific empirical content within a specific model. For example, the SDF is the Inter-temporal Marginal Rate of Substitution of the Consumer (IMRSC) within the C-CAPM model, in which the portfolio payoffs can be modelled as bundled contingent claims to a numeraire consumption good. Alternatively, a linear factor pricing model defines the SDF as a linear valuation functional of traditional risk factor loadings.

The implication of the SDF approach to portfolio evaluation is to measure the difference between the actual and theoretical stochastic discounted gross portfolio returns such as:

$$\alpha_i = E_{t-1}(m_{t,} R_{i,t}) - 1 \qquad (3.3)$$

where $R_{i,t}$ is the return on portfolio i from time $t-1$ to t and α_i measures the relative performance of the portfolio from time $t-1$ to t.

Recent papers by Farnsworth *et al.* (2002) and Jagannathan and Wang (2002) have demonstrated the generality of this approach for performance evaluation. This approach has also been used by Chen and Knez (1996), and Farnsworth *et al.* (2002) to measure performance of the mutual funds and other actively managed portfolios.[28]

The essential underlying asset pricing models we choose for the empirical investigation is the SDF primitive efficient model, which constructs the SDF as a weighted return of the primitive efficient variables. We choose this model primarily based on its research potential while applying it to the UK professionally managed funds and the theoretical concerns relating to its counterpart SDF models. The primitive-efficient SDF model is a general approach, which does not require many of the traditional assumptions in conventional pricing theories. While it has been implemented widely to examine portfolio performance, it has been rarely applied to the UK investment management market. It is also the foundation of the linear factor model. Further, practically, C-CAPM and other SDF models have not displayed satisfactory results with various pricing anomaly such as the risk-free puzzle. I introduce briefly the methodologies of the SDF primitive efficient model in the following section.

1.2 The SDF primitive efficient model

[28] For a detailed literature review, please refer to chapter 2.

The essence of the SDF primitive efficient model is that, the SDF is represented by a portfolio of the primitive/reference assets, where the weights are estimated so that at least the primitive/reference assets themselves are correctly priced by the model. Following Chen and Knez (1996), the primitive efficient SDF can be constructed as the payoff on some constant composition portfolio with weights β in R^N, that is,

$$m_t = \beta' R_t \qquad (3.4)$$

where R_t is a vector of gross returns of primitive assets /reference variables/ benchmark and β is the weight of portfolio, β and R_t are both $N*1$ vectors. m_t is the primitive efficient SDF and it is a linear function of a minimum variance efficient portfolio (Cochrane 2000).

Based on the law of one price, which states that if two portfolios have the same payoffs (in every state of nature), then they must have the same prices. If this law does not hold, arbitrage profit may rise by selling the expensive version and buying the cheap version of the same portfolio. We can then derive the following theorems from the Law of One price[29]. There is only one discount factor which is the return of primitive-efficient portfolio that can measure all the portfolio returns by 1=E(mR) if and only if this law of one price holds.

To prove this, we construct the return space R, which is generated by portfolios of N basis returns (for example, N stocks). We organize the basis payoffs into a vector $R=[R_1\ R_2\ R_3....R_n]'$. The return space is then $\boldsymbol{R}= \{\beta' R\}$ where β is the weight. As the

[29] See Ross (1978), Rubinstein (1976) and Harrison and Kreps (1979).

theorem requires, we need a discount factor that is in the return space. Thus, the discount factor must be of the form such as $R^*= \beta^{*\prime} R$. We need to construct β^* so that R^* can price the basis assets. We want $1=E(R^*R)=E(R R' \beta^*)$. We therefore need $\beta^*=E(RR')^{-1}$ **1.** If $E(RR')$ is non-singular, then β^* exists and is unique, assuming that the markets are complete. Hence

$$m_t= R^*= \mathbf{1}' E_{t-1}(R_tR_t')^{-1}R_t \qquad (3.5)$$

The primitive efficient SDF model has been used by Chen and Knez (1996) for performance evaluation purpose; Dahlquist and Soderlind (1999) study sampling properties through simulations. He, Ng and Zhang (1998) specialize the approach to handle a larger number of primitive assets.

Within the context of performance evaluation, following Chen and Knez's approach, we impose a primitive efficient SDF model to price all the reference variables (benchmarks), which give us

$E(\mu_{1t-1})=0$

with

$$\underset{N*T}{\mu_{1t-1}} = \underset{T*N}{R_{p,t}} \underset{N*1}{R'_{p,t}\beta} - \underset{N*1}{1} \qquad (3.6)$$

where $R_{p,t}$ is a $N*T$ vector of the returns of primitive assets, and β is the weight of the portfolio, N is the number of the primitive assets, T is time.

Secondly, for a managed portfolio (the funds, unit trusts etc) with gross return $R_{s,t}$ at time t, its law of one price (LOP)-based performance value should be

$$\underset{1*T}{\mu_{2,t-1}} = \underset{T*N}{R_{s,t}}\underset{N*1}{R'_{p,t}}\underset{1*1}{\beta} - 1 \qquad (3.7)$$

The flavour of this testing approach is firstly to find an SDF that "prices" the primitive assets. The next step is to test if this SDF also prices the managed portfolio, in our case, the portfolios of UK unit trusts. The magnitude and the sign of $E(\mu_{2t-1})$ indicate the average performance of the unit trusts.

2. The GMM Estimation

The Generalized Method of Moments (GMM) was regarded as one of the most important developments in econometrics in the 1980s that revolutionized empirical work in macroeconomics. The path-breaking articles on the GMM were those of Hansen (1982) and Hansen and Singleton (1982). Since then, the GMM has been widely applied to the estimation and testing of econometric models, including the SDF pricing models. There are mainly two things supporting GMM's popularity: one, the GMM nests many common estimators and provides a useful framework for their comparison and evaluation; two, the GMM provides a "simple" alternative to other estimators, especially when it is difficult to write down the maximum likelihood estimator.

One important subject related with the GMM estimation is the choice of its weighting matrices. Various GMM estimators can be constructed with different weighting matrices. It is important to examine the properties of the GMM estimators before choosing the optimal estimator to implement the empirical tests, within the context of portfolio evaluation.

In this section, we explain the GMM estimation method, the choices of weighting matrices when constructing different GMM estimators (two-step and iterated estimators) and the over-identifying test *J*-statistic, followed by a literature review of small sample properties of the GMM estimators and an examination of the small sample properties of the GMM two-step and iterated estimators within the context of the SDF primitive efficient model based on simulation.

2.1 Introduction

The Method of Moments (MOM) is an estimation technique, which suggests the unknown parameters should be estimated by matching population (or theoretical) moments (which are functions of the unknown parameters) with the appropriate sample moments. In other words, the expected value should be chosen as close as to the true value.

For example, the variance can be expressed as:

$$\text{var}(x) = E(x^2) - (E[x])^2 \qquad (3.8)$$

Thus, the MOM estimator of the variance is

$$\widehat{\text{var}}(x) = \frac{1}{n}\sum x^2 - [\frac{1}{n}\sum x]^2 = \frac{1}{n}\sum (x-\bar{x})^2 \qquad (3.9)$$

$$\approx \frac{1}{n-1}\sum (x-\bar{x})^2 \qquad (3.10)$$

where equation 3.10 is usual unbiased estimate of the variance. The MOM estimator in equation 3.9 is biased as it divides the sum of squared deviation from the mean by n instead of by $n-1$ as in equation 3.10. However, as the sample grows larger, the difference between the two estimators nears zero, suggesting the MOM estimator is consistent.

The GMM estimator is used when the parameters' vector is over-identified by the moment conditions. In this case, there are more equations than unknowns, a vector that satisfies sample moment equals to zero cannot be found, only a vector that makes the moments as close to zero as possible can be estimated.

In the case of estimating the SDF models using GMM, formally, let β be a parameter vector of the chosen SDF model, x_t be the variable vector[30], $u_t(\beta)$ be the vector of pricing errors. Then, the moment conditions are defined as:

$$E[u_t(x_t ,\beta)]=0 \qquad (3.11)$$

As $E(u(.,.))$ cannot be observed, it is sensible to process by defining $f_t(x_t , \beta)$ as the sample mean of $u_t(x_t , \beta)$, such as

$$f_t(x_t, \beta) = T^{-1}\sum u_t(x_t , \beta) \qquad (3.12)$$

[30] In the context of the SDF models, the parameter refers to α*, a vector of portfolio weights; variables are portfolio/benchmark returns.

Since the population mean of u_t must be zero, the sample mean f_t should be small and so should a quadratic form of f_t to estimate the parameters.

When the number of moment conditions is larger than the number of parameters, one has several options. First, one could drop some of these moment conditions, which cannot be optimal in general (throwing away information rarely is). Second, by analogy to least squares, the deviations from each condition could be weighted equally in the calculations and the sum of squared deviations minimized. Third, one could weight the equations according to how precisely (measured by the variance) each of the equations is estimated[31].

Formally, we set W_t as a stochastic positive weighting matrix; β can be estimated by minimizing a weighted quadratic form of the sample moments as:

$$\beta = arg_{min\ \beta}\ Q_t(x_t\ ,\ \beta)$$

where

$$Q_t(x_t\ \beta) = f_t\ (x_t\ ,\beta)'W_tf_t(x_t\ ,\beta) \qquad (3.13)$$

The first-order conditions of minimizing $Q_t\ (x_t\ ,\beta)$ is

$$\frac{\partial f_t'}{\partial \beta} W_t f_t(x_t, \beta) = 0 \qquad (3.14)$$

also note that $Q_t\ (x_t, \beta) \geq 0$ and $Q_t\ (x_t, \beta)=0$ only if $f_t\ (x_t, \beta)=0$.

[31] The results by Hansen (1982) suggest that the optimal estimator is the third one. The choices of weighting matrices will be discussed in the next section.

Thus, $Q_t(x_t, \beta)$ can be made exactly zero in the just identified case, but is strictly positive in the over-identified case.

More generally, equation (3.13) can be expressed as:

$$a_t f_t(x_t, \beta)=0 \tag{3.15}$$

which allows you to pick arbitrary linear combinations of the moments to set to zero in parameter estimation.

In the exactly identified case, a_t=1; while in the over-identified case, according to Hansen (1982) Theorem 3.2, with the choice of $a_t = \frac{\partial f_t'}{\partial \beta} W_t$, the most efficient estimate with minimum standard error of the parameters β can be obtained.

2.2 The Choice of Weighting Matrices

W_t in equation (3.13) is a weighting matrix (or distance matrix). When the system is just identified, the GMM estimator does not depend on the choice of weighting matrices while in the case of an over-identified system, different GMM estimators are constructed with different weighting matrices.

W_t implies how much attention to pay to each moment. In the sense, it directs GMM to emphasize some moments or linear combinations of moments at the expense of others. An identity weighting matrix treats all assets symmetrically, and minimize the sum of squared pricing errors. While a non-identity matrix can be used to offset the differences in units between the moments.

As stated by Cochrane (2001), Lettau and Ludvigson (2001), the choice of the weighting matrix can have an effect on evaluating the asset pricing model, therefore, we apply estimators with different matrices (two-step and iterated GMM estimators) to examine the small sample properties of *J*-statistic (as explained in section 2.3), within the framework of the SDF primitive efficient model. The purpose is to identify a more reliable GMM estimator with better small sample properties in order to carry on the performance evaluation tests.

2.21 Optimal weighting matrix

To achieve the lowest variance of the estimator, we need the optimal weighting matrix. The most efficient estimator can be obtained by weighing each equation by the inverse of its standard deviation, denoted as Ω^{-1} since the GMM estimator with such a weighting matrix produces the smallest asymptotic variance.[32]

In the context of an asset pricing model, the interpretation of the use of such an optimal weighing matrix can be made as follows: some asset returns may have much greater variance than others. Because the sample mean will vary more from sample to sample, for these assets, the sample mean of pricing errors will be a much less accurate measurement of the population mean. Hence, it is good to pay less attention to linear combinations of the moments, about which the data set at hand has the most information.

[32] Hansen (1982) derives the optimal weighting matrix as the inverse of the asymptotic variance matrix, which may take different forms, depending on the type of heteroskedasticity and autocorrelation assumed.

Depending on how the optimal weighting matrix is estimated, there are alternative GMM estimators. The two-step GMM estimator follows two steps: it uses an identity matrix as the weighting matrix at the first stage; then the first-stage estimator is used to estimate Ω, which will be used for the second-stage GMM estimator. The alternative GMM estimator repeats this procedure until β converges or until the number of iterations attains some large value. The newly-proposed one-step (continuous-updating) estimator attempts to minimize $f_t(x_t, \beta)' \Omega^{-1} f_t(x_t, \beta)$ by allowing Ω to vary with β.

The GMM Two-step estimator

In more details, to construct a two-step GMM estimator, it normally takes two steps. It initially starts with a sub-optimal but consistent choice (an identity matrix I) of the weighting matrix W_t to provide and estimate Ω (from the residuals). The estimate of β can firstly be expressed as

$$\beta_1 = arg_{min} f_t(\beta) I f_t(\beta) \qquad (3.16)$$

Ω then will be used in the second stage (again by minimizing the criterion function) to estimate the asymptotically efficient estimator of

$$\beta_2 = arg_{min} f_t(\beta) \Omega^{-1} f_t(\beta) \qquad (3.17)$$

β_1 is said to be consistent and asymptotically normal while β_2 is consistent, asymptotically normal and asymptotically efficient [33].

The GMM Iterated estimator

It is argued that the second-stage estimate β_2 generated from the GMM two-step estimation may imply a different spectral density from the first stage. In order to keep consistency of the estimate of β and of the spectral density, it is suggested to use the GMM iterated estimator, which can be obtained by iterating between estimates of the parameter vector β and the weighting matrix W_t until convergence is attained. These estimates do not depend on the initial weighting matrix and might also provide better small-sample performance. However, although such iterations often converge, there is no fixed-pointed theorem to ensure it.

2.22 Pre-specified weighting matrix

It is argued (Cochrane (2001)) that the implementation of a pre-specified rather than the optimal weighting matrices might provide particular interests. Examples include the implication of a second-moment weighting matrix to construct the *HJ* distance measure[34], and using some weighting matrix in a particular form, such as the identity matrix or the inverse of a covariance matrix based on an *i.i.d.* assumption, might help to overcome some difficulties that occur in the numerical optimization process[35], suggested by Zhou (1994). The point is, you might prove that a pre-specified weighting matrix might provide some particular statistic features, but we cannot say it

[33] see Cochrane (2001)

[34] Initiated by Hansen and Jagannathan (1997), *HJ* distance measure is the quadratic form of the pricing errors weighted by the inverse of the second moment matrix of returns. More details can be found in chapter 4.

[35] GMM estimator cannot be solved analytically in the case of most nonlinear models, it is then necessary to apply a numerical optimization method to calculate by numerically minimizing the criterion function

is better than the optimal weighting matrix. Without any particular statistic specifics, our investigation is only restricted within the range of the implication of optimal weighting matrices.

2.3 J-Statistic for Over-identifying Conditions

The foundation of the GMM framework is a set of moment restrictions that identify the parameters to be estimated and the GMM estimation method provides a way to combine moment restrictions when there are more moments than parameters. It is natural to impose the moment condition restrictions on the GMM framework, which can be decomposed into the identifying restrictions and the over-identifying restrictions.

Since the sample analogues to the identifying restrictions are automatically satisfied by the estimated sample moments, it is therefore impossible to test the identifying restrictions, but possible to test the over-identifying restrictions. Hansen (1982) proposes a statistic (we hereafter call it *J*-statistic) to examine the over-identifying restrictions such as:

$$J=T\,[f_t(\beta)'\,\Omega^{-1}f_t(\beta)] \sim \chi^2\,(p\text{-}q) \qquad (3.18)$$

where Ω is the variance-covariance matrix of $f_t(\beta)$, p is the number of the moment conditions, q is the number of the parameters.

This statistic is the minimized pricing errors divided by their variance-covariance matrix. Under certain regularity conditions[36], sample means converge to a normal distribution, so sample means squared divided by their variances converges to the square of a standard normal, or χ^2 with the degree of freedom (p-q).

The *J*-statistic has been widely applied in the performance evaluation literature when the GMM estimation is used. Within the context of asset pricing modelling, $f_t(\beta)$ in (3.18) is the pricing error. The *J*-statistic can be interpreted as a measure of a weighted sum of squared pricing errors under the null hypothesis that the model estimated is correctly specified. Therefore, this measure can also be used to detect the abnormal portfolio performance.

3. Small Sample Properties of GMM Estimators Based on the J-statistic

Although the GMM estimators are consistent and asymptotically normally distributed under general regularity conditions, it has long been recognized that this first-order asymptotic distribution may provide a poor approximation to the finite sample distribution. In particular, the GMM estimators may be badly biased, and asymptotic tests based on these estimators may have true size substantially different from presumed nominal sizes. Given that our data sample is relatively small, we implement the size and power tests to examine the finite sample properties of the GMM estimators in order to identify a GMM estimator with superior properties to carry out our empirical research.

36 see Hansen (1982)

The finite sample properties of the GMM estimators within the context of asset pricing models have been the interest of many researchers. To name but a few, they are Tachen (1986), Kocherlakota (1990), Ferson and Foerster (1994, 1995) Hansen, Heaton and Yaron (1996) and Dahlquist and Soderlind (1999). There is no consensus on this issue, and the summarized results of their work are shown in appendix table 3-1.

The widely applied hypothesis test of the GMM estimation is Hansen's *J*-statistic, as show in equation (3.18), so it is natural to explore the finite sample properties of the GMM estimators based on the *J*-statistic. In this section, we firstly provide a literature review of small sample properties of the GMM estimators, followed by an explanation of the process of the size and power tests. Finally, we examine the small sample properties of the GMM two-step and iterated estimators within the context of the SDF primitive efficient model and the results are reported.

3.1 Literature review

The initial work can be dated back to Tachen (1986), who examined the GMM two-step estimator based on a C-CAPM model. He finds that this estimator has reasonable small sample properties, however, the GMM estimators with an arbitrary selection of instruments could perform better than such as optimal estimator. Similarly, Kocherlakota (1990) tends to analyze the iterated GMM estimator and finds evidence that the J-statistic for over-identifying restrictions tends to reject the null too frequently compared with its asymptotic size. Moreover, he concludes that the estimator with larger instrument sets tends to generate downward biases and narrow confidence intervals. Ferson and Foerster (1994, 1995) study both the GMM two-step and GMM iterated estimators, using a conditional asset pricing model (latent variable model). They believe the GMM iterated estimator has superior finite sample properties, producing approximately unbiased coefficient estimates with understated asymptotic standard errors, which, however, can be partially corrected by using simple scalar degrees-of-freedom adjustment factors for the estimated standard errors while the *J*-statistic of the GMM two-step estimators may over-reject the restriction quite often.

Within a framework similar to C-CAPM, Hansen, Heaton and Yaron (1996) find that both the two-step and iterated GMM estimators have small sample distributions that can be greatly distorted, resulting in over-rejections of the *J*-statistic, with unreliable hypothesis testing and confidence-interval coverage. Based upon this evidence, they recommend an alternative form of the GMM estimator, the one-step estimator (the continuous updating GMM estimator). They then investigate the small sample properties of the one-step estimator and find out it typically dominates the GMM two-stage and iterated estimators, while the large sample inferences occasionally can be unreliable.

More recently, Dahlquist and Soderlind (1999) examined the GMM two-step estimator, iterated estimator and an estimator using the pre-specified weighting matrix (the inverse of the second moment matrix of managed returns). Their results provide no evidence against using a pre-specified weighting matrix and show that the size of the tests can sometimes be seriously distorted, which suggests the empirical critical values should be used. However, when they correct for the size, the power is about the same in different settings although the power properties are better in the unconditional setting compared to the conditional setting. They also highlight the case in which an economically significant excess return is needed in order to reject the null of neutral performance unless the sample is very large.

3.2 The Size of the *J*-statistic

It is common to rely on large-sample statistics to draw inferences as there are no analytical results on the finite-sample properties of estimators. However, if the number of sample observations is not large enough for the asymptotic results to provide a good approximation, it is possible that the size of the test will be incorrect. Because there is no standard sample size for which large-sample theory can be applied, it is a good practice to investigate the appropriateness of the theory. Campbell develops an ideal framework (so called *multivariate F*-test) to illustrate the problem that can arise if one relies on asymptotic distribution theory for inference.
Corresponding to each given nominal size, we can get its empirical size counterpart, which is the fraction rejected at various nominal significance levels. The fractions are taken as the portion of the 10,000 replications of an experiment in which the *J*-statistic exceeded a critical value from the Chi-square distribution. Similarly, for any given size, the empirical critical value can be found and these values can be used to compare with the nominal critical values.

The data generating process is simple. The returns R_{st} of reference variables are drawn randomly from the real data (as discussed in chapter 1). As GMM is likely to be sensitive to the moments in the data not matched by the artificial economics, we use this more direct approach, re-sampling the data in a manner similar to the bootstrap methods of Efron (1982) in order to remain many of the statistical properties of the original data.

Under the null hypothesis of a primitive efficient SDF is true, the data-generating process of the portfolio returns to be evaluated follows:

$$R_{pt}=\boldsymbol{B}\,\boldsymbol{R_t}+\varepsilon_t \qquad (3.19)$$

where B=**1**; R_t is a vector of reference variables returns in the case of testing the primitive efficient SDF model; ε_t is an *i.i.d.* normally distributed error term. The artificial data of portfolio R_{pt} is generated with Monte Carlo simulation method, using neutral performance portfolio data. The vector B can be estimated by OLS. If the models are true, R_{pt} should have no abnormal performance.

To examine the size of the *J*-statistic, the procedure is as follows. We initially collect 10,000 *J*-statistic of the trials, rank them, then find out the values which are at the position of 1%, 5% and 10% of the *J*-statistic series, and those values are empirical critical values at size of 1%, 5% and 10% respectively. On the other hand, we can also match the theoretical critical values to this *J*-statistic series, find out their closest positions in this series, the positions (in percentage terms) of the empirical critical values within the ranked *J*-statistic series are the empirical size.

The GMM iterated and two-step estimators are employed to estimate the SDF primitive efficient model. Given size of 0.10, 0.05 and 0.01, we actually find the corresponding empirical critical values for the *J*-statistic statistic, above which value alpha-fraction of the statistics simulated under the null lie.

3.3 The Power of the *J*-statistic

It is also important to consider its power when drawing inferences. Given that an alternative hypothesis is true, the power is the probability that the null hypothesis will be rejected. If the power is high, the test can be very informative but it may also reject the null hypothesis against alternatives that are close to the null in economic terms. On the other hand, low power against an alternative suggests that the test is not useful to discriminate between the alternative and the null hypothesis. We examine the power for a given size of test, which is the probability that the test statistic is greater than the critical value under the null hypothesis, given that the alternative hypothesis is true. In the context of performance evaluation, the test of power is to investigate whether it is possible (how possible it is) to reject the hypothesis of neutral performance when it is actually false.

Using the data generated from the alternative hypothesis (the excess return of R_{pt} is significantly different from zero, i.e. 5% and 10%) and within the framework of the SDF primitive efficient model, we can then test the fraction of 10,000 trials in which the *J*-statistic, using the empirical critical value, rejects the null hypothesis that the SDF model is a true model / there is no superior performance. This fraction is then the power of the given size.

3.4 The Results

We implement both power and size tests as mentioned above for the GMM two-step and iterated estimators, and we conclude the results as follows:

Size

Empirical critical values are generated for both the GMM two-step and the iterated estimators. Table 3-2 Panel A reports the empirical critical values of the *J*-statistic with the sample size of 60, 120, 240, 360 and 720, with particular attention paid to the sample size of 360 since the sample size of our monthly return data from Jan 1975 to Oct 2003 is closest to 360. The results suggest that apparently the size and critical values can be seriously distorted when the sample size is small. The summarized results of the empirical critical values are shown as below:

The Primitive Efficient SDF Iterated *over-reject*

The Primitive Efficient SDF Two-step *over-reject*

If the empirical critical values are much bigger than their theoretical counterparts, the null hypothesis can quite often be over-rejected if the tests are based on nominal critical value instead. On the contrary, if the empirical values are smaller than theoretical critical values, under-rejection will occur. For the former, in the context of performance evaluation, non-abnormal performance can be taken spuriously as abnormal performance. For the latter, superior performance can be under estimated. It is therefore quite important to use the empirical critical values to continue our empirical analysis.

It suggests that within the framework of the primitive efficient SDF, both the GMM iterated and two-step estimators generally over-reject the null hypothesis if nominal critical values are taken. In most of the cases, the GMM two-step statistic comparatively rejects the model more often compared to its iterated counterpart since the gap between the empirical and nominal critical values for the GMM two-step estimator are bigger.

These results are consistent with Kocherlakota (1990), Dahlquist and Soderlind (1999), in which over rejection of those estimators are also detected. However, it is a bit different from those of Ferson and Foerster (1994, 1995) in that from their work, over-rejection of the *J*-statistic is confirmed when using the GMM two-step estimator, but when using the GMM iterated estimator, the *J*-statistic tends to under-reject marginally relative to its nominal size.

Power

Table 3-2 Panel B reports the results of power tests for different estimators with the correct size. It shows that in all cases, the GMM iterated estimator generates higher power within unconditional primitive efficient SDF model framework. For example, the power of the GMM iterated estimators with 5% excess returns range from 0.015 to 0.282, whilst it ranges from 0.01 to 0.264 with the use of the GMM two-step estimator. It also suggests that either a significant excess return or a long sample is needed to reject the neutral performance, i.e. the power can be improved significantly by increasing either excess return or sample size.

Our results are consistent with Ferson and Foerster (1994, 1995). Apparently, the advantage of using the GMM iterated estimator is repeatedly updating the weighting matrix and searching to find new parameter estimates, which reduces the chance that the algorithm will settle on a local minimum. Hence better fitted estimations can be provided.

TABLE 3-2 SMALL SAMPLE PROPERTIES OF GMM TWO-STEP AND ITERATED ESTIMATORS

PANEL A. EMPIRICAL CRITICAL VALUES VS THEORETICAL CRITICAL VALUES

The SDF primitive efficient unconditional model is employed to test portfolio performance. GMM iterated and two-step estimators are employed to estimate the SDF models. Empirical critical values of GMM *J*-statistics are simulated for a series of small samples (sample size: 60, 120, 240, 360, 720).

Size (α) follows $P\{\chi^2(n) > \chi^2_{\alpha}(n)\} = \alpha$

GMM Iterated Estimator			
Size(α)	**0.1**	**0.05**	**0.01**
Nominal critical value of $x^2(1)$	2.70554	3.84146	6.6349
T= 60	3.20938	4.31452	7.64443
=120	3.10136	4.20771	7.56894
=240	2.82873	3.99857	7.42644
=360	2.7653	3.93682	7.01061
=720	2.64782	3.83546	6.89475
GMM 2-step Estimator			
Size (α)	**0.1**	**0.05**	**0.01**
Nominal critical value of $x^2(1)$	2.70554	3.84146	6.6349
T= 60	3.16351	4.56919	7.57317
=120	3.02267	4.13871	7.71934
=240	3.06244	4.15675	7.48175
=360	3.0751	4.10906	7.23541
=720	2.96851	3.96584	7.06584

PANEL B. POWER OF TESTS FOR DIFFERENT ESTIMATORS WITH CORRECT SIZE (USING EMPIRICAL CRITICAL VALUES)

Within the framework of the SDF primitive efficient unconditional model, fractions of 10,000 trials in which the test for non neutral performance rejects (using empirical critical values from table 3-2 Panel A) when the excess return of 5% and 10% are given respectively to generate the artificial data. The power of GMM iterated estimator, GMM two-step estimator are reported for each model

	GMM Iterated estimator			**GMM two-step estimator**		
Excess return: 5%						
Size	**0.1**	**0.05**	**0.01**	**0.1**	**0.05**	**0.01**
T= 60	0.135	0.06	0.015	0.12	0.05	0.01
360	0.176	0.08	0.03	0.15	0.06	0.02
720	0.282	0.221	0.065	0.264	0.203	0.08
Excess return : 10%						
Size	**0.1**	**0.05**	**0.01**	**0.1**	**0.05**	**0.01**
T= 60	0.214	0.12	0.127	0.205	0.118	0.08
360	0.503	0.33	0.163	0.41	0.29	0.1
720	0.705	0.611	0.404	0.667	0.58	0.41

To summary, the comparisons have been drawn between two GMM estimators by investigating the size distortion and the power to reject non-neutral performance. The results show that the GMM iterated estimator has smaller size distortion and generates higher power within our asset pricing model framework. As a result, we will apply the GMM iterated estimator to implement empirical tests for our projects.

4. Conclusion

There are two important elements throughout our empirical exercises: the underlying methodology and the way how we estimate the model. The SDF primitive efficient model is chosen as the underlying asset pricing model thanks to its superior characteristics, compared to traditional approaches. The GMM estimation, on the other hand, turns out to be a natural fit for the SDF approach, as widely applied by empirical studies. As a cornerstone of the following chapters, this chapter is dedicated to two important tasks. One is to provide a thorough explanation of the SDF primitive efficient model and the GMM estimation; another is to identify an optimal GMM estimator based on the examination of the small sample properties of the *J*-statistic.

After providing a brief introduction to the SDF models (methodologies) and the GMM estimation, we investigate the small sample properties of the GMM two-step and iterated estimators based on the *J*-statistic. Within the framework of the SDF primitive efficient models, our simulation results show that for both GMM two-step and iterated estimators, the sizes of the *J*-statistic can be seriously distorted, whilst the GMM iterated estimator exhibits superior size and power properties, compared to its counterpart.

APPENDIX CHAPTER 3

TABLE 3-1 SUMMARY OF RECENT STUDIES ON SMALL SAMPLE PROPERTIES TESTS OF GMM ESTIMATORS IN THE CONTEXT OF ASSET PRICING MODELS

Study	**Asset pricing model**	**GMM estimator**	**Major Results**
Tachen(1980)	C-CAPM	Two-stage GMM	• Reasonable small sample properties • Arbitrary selection of instruments is better than optimal selection.
Kocherlakota(1990)	C-CAPM	Iterated GMM	• J statistic tends to over-reject • Downward biases with narrow confidence intervals using larger instrument sets
Ferson and Foerster (1994,1995)	Latent variable model with single premium and two premium	Two-stage and iterated GMM	• Iterated estimator is superior with *J* statistic of two-stage estimator over-reject the restrictions • Estimated coefficients' asymptotic standard errors are underestimated.
Hansen, Heaton and Yaron (1996)	C-CAPM	Two-stage, iterated One-step GMM	• All have unsatisfactory properties with small sample size. • one-step estimator is superior

CHAPTER 4: CONDITIONAL PERFORMANCE EVALUATION OF UK UNIT TRUSTS

The asset management industry has been growing rapidly in recent years, with increasing management charges. In London alone, the annual management charges by the fund managers amount to £4.3 billion, in addition to £5.8 billion hidden costs, primarily dealing commission, taxes and market making spreads[37]. The existence of the asset management industry is partly based on the assumption that professional asset managers are capable of generating positive alphas (superior returns) consistently, compared to the pre-specified benchmarks.

Given the huge amount of fees that unit trust investors have to pay for portfolio managers' 'superior' skills, it is of great interest to investigate if the portfolio managers, in particular, those of the U.K. unit trusts have genuine abilities to outperform the market. If they do, to what extent can they beat the market?

Traditional models utilize only unconditional expected returns, where the factor loadings are constant. In another word, the linear factor models are correct only when systematic risk characteristics of the securities held in the portfolio remain fixed and the portfolio weights remain fixed through time as well. However, if the expected returns and risks vary over time, such an unconditional approach could produce biased results. For example, if the evaluation period covers a bear market, but the period going forward is a bull market, the unconditional performance evaluation can hardly have any forward looking value.

[37] According to Alan Miller in Guardian (12/10/2009) and a subsequent news story in the Times (13/10/2009)

To remove the impact of variation risks on the biased results, it is necessary to incorporate the information into the measurement based on a conditional approach. It is also interesting to examine if conditionally measured performance can be significantly different from unconditionally measured performance.

To answer these questions, we implement the empirical tests using the GMM iterated estimator, which was proved to be the optimal estimator within a small sample according to our simulation results in chapter 3. In addition, the issue of how sensitive the conditional performance measures are to the choice of conditioning variables remains under-explored. We therefore also examine this issue.

In this chapter, we firstly explain conditional performance evaluation and discuss the roles which conditioning measures play in the performance evaluation literature and the various ways of how we can incorporate conditioning information within the framework of the SDF models. Secondly, we explain the model and moment selection statistics. Thirdly, we examine the best fitted SDF model and report the results of the model and moment selection statistics based on implementation of conditioning measures on performance evaluation. Finally, the performance of 40 U.K. unit trusts is examined separately under their own style categories over a sample period from Jan 1975 to Oct 2003 and the results are interpreted. Conclusions then follow.

1. Conditional Performance Evaluation

It is well documented that in practice, there are a number of problems associated with the traditional measures of a fund manager's performance. For example, standard unconditional measures of performance do not take into account the fact that risk and expected returns could change with the state of the economy. Instead, traditional performance studies with unconditional measures assume implicitly that through the evaluation period, the risk level of the fund is stationary. As a result, many of the problems discovered in previous research in fact reflect that conventional unconditional measures are unable to deal with the dynamic behaviour of the returns. As a consequence, Ferson and Schadt (1996) suggest an approach, the conditional performance evaluation, to solve this problem. The 'conditional alpha', is defined as the difference between a fund's excess return and that of a strategy that attempts to match the fund's risk dynamics over time by mechanically trading, based on predetermined variables.

Conditional performance evaluation is consistent with the semi-strong form of market efficiency as described by Fama (1970). It is believed that if the market is seimi-strongly efficient, a fund manager utilize a trading strategy, which can be replicated based on publicly available information can not add value. In order to generate a positive conditional alpha and add value, a fund manager should produce a higher return than the strategies based on public information. Further description of rationales and methodologies of conditional measure can be found in chapter 2.

Most of conditional performance evaluation of mutual funds has been based on conventional asset pricing models. We refine the conditional measure by applying the SDF approach. In this section, we introduce the methodologies of how to incorporate conditioning information. We also explain the details of the moment conditions within the SDF framework.

1.1 The Methodologies

We adopt the instrumental variables approach to incorporate conditioning information. Following Ferson (2000), we can use the Kronecker product and include a constant 1 as the first element of the instrument vector Z_{t-1}, to construct the following orthogonality condition within the context of the SDF model[38]:

$$E\{[m_t(\beta)R_t - 1] \otimes Z_{t-1}\} = 0 \qquad (4.1)$$

where $m_t(\beta)$ is a stochastic discount factor[39] with parameter β, R_t is the portfolio return at time t and Z_{t-1} is the information variable at $t-1$.

When the information variable Z_{t-1} is restricted to be a constant, we have the unconditional model. Otherwise, it is a conditional model. The idea of the instrumental approach is straightforward. To incorporate the conditioning information, we can simply multiply the primitive assets by the lagged information instruments, and it is consistent with Cochrane's "scaled return" approach (2001). That is, the conditional model is just equivalent to a "scaled" unconditional model. It is also associated with the spirit of GMM estimation, which uses instrument variables.

Besides taking the conditioning information into account, it is also important to incorporate the dynamic strategies in the tests as the weights of asset holding in a managed portfolio can also change over time. The way to allow for dynamic strategies is to multiply the stochastic discount factor by the lagged information variables. Cochrane (2001) named it 'scaling factors', and it is different from conditional estimation, in which the 'scaling returns' have been applied. The two are distinct in that: if one had a model that predicted constant SDF over time, then it would be

[38] A detailed explanation of the SDF models can be found in chapter 3.
[39] Different restrictions can be further imposed on m for more empirical contents.

appropriate to scale the returns but not the factors. On the other hand, to examine the unconditional implications of a scaled factor pricing model, it is required to scale only the factors, but not the returns.

Within the context of the SDF primitive efficient model, as shown in equation (3.4) in chapter 3, the stochastic discount factor m_t is constructed as:

$$m_t = \beta' R_{pt}$$

To incorporate the dynamic strategy and model the scaled factor, we use

$$m_t = s_{t-1} \otimes \beta' R_{pt} \qquad (4.2)$$

where s_{t-1} is a vector of predetermined information variables. In so doing, the holding weights of portfolio change from β to $s_{t-1} \otimes \beta'$, which implies dynamic strategies of the benchmark portfolio holdings.

1.2 The Moment Conditions of Conditional SDF Models

To examine the impact of conditioning information, we investigate the following four cases:

- Unconditional Evaluation with Fixed Weights (no scaling returns, no scaling factors)
- Unconditional Model with Time-Varying Weights (with scaling factors only)
- Conditional Model with Fixed Weights (with scaling returns only)
- Conditional Model allowing for Dynamic Strategies (with both scaling returns and scaling factors).

Broadly, we have two types of models, unconditional models which do not scale the returns by the information variables and conditional models which have scaled returns. In the case of the unconditional models, we can further divide the models into two types: one when dynamic strategy is not allowed (UNFIXED) and one when such a strategy is allowed (UNVARY), Conditional models when dynamic strategies are not allowed are denoted as CONL, with large information variable sets and CONS, with small information variable sets; finally, we have CONVARY, the conditional models with time varying weights to allow for dynamic strategies.

As explained in chapter 3, we implement the tests with GMM estimation. The moment conditions for the four cases illustrated above are defined respectively as:

1.21 Unconditional Model with Fixed Weights (UNFIXED)

The first case is the unconditional model with no dynamic strategies allowed, i.e. the portfolio weights are fixed. Two systems of moment conditions can be constructed following Chen & Knez (1996) and Ferson *et al.* (2002) respectively:

Chen and Knez define the error term μ_{t-1} such as:

$$\mu_{1,t-1} = [R_{pt}\, m_t\, (\beta) - 1] \otimes Z_{t-1}$$

$$\mu_{2,t-1} = [R_{st}\, m_t\, (\beta) - 1] \otimes Z_{t-1} \qquad (4.3a)$$

Whilst Ferson defines μ_{t-1} such as:

$$\mu_{1,t-1} = [R_{pt}\, m_t\, (\beta) - 1] \otimes Z_{t-1}$$

$$\mu_{2,t-1} = [\alpha_p - R_{st}\, m_t\, (\beta) + 1] \otimes Z_{t-1} \qquad (4.3b)$$

where R_{pt} refers to the return of the benchmark at time t, R_{st} refers to the return of the portfolio s, managed by the unit trust's manager at time t. α_p and β are the parameters. Within the framework of the SDF primitive efficient model, $m_t(\beta) = \beta' R_t$ (also see equation (3.4) in chapter 3).

When the system is over-identified, it permits the use of a GMM estimator. Hence the corresponding sample moment condition is

$$f_t(\beta) = T^{1} \sum_t (\mu_{1t-1}' \mu_{2t-1}')' \qquad (4.4)$$

The number of the moment conditions for both (4.3a) and (4.3b) are equal to the number of the reference variables/primitive assets[40] plus one; the number of the parameters depends on the specification of $m_t(\beta)$, as explained in chapter 3. The main difference between (4.3a) and (4.3b) is that in Ferson's moment condition (4.3b), there is one extra parameter, α_p (therefore, one less degree of freedom) and α_p clearly indicates the magnitude of the non-neutral performance. Whilst in Chen and Knez's moment condition (4.3a), α_p is estimated *ex-post*. We calculate the *J*-statistic to examine the significance of α_p.

1.22 Unconditional Model with Time-Varying Weights (UNVARY)

The second case is also an unconditional model, but with dynamic strategies included, i.e. the portfolio can have a time-varying weight.

[40] As discussed in Chapter 1, the number of reference variables we used is eight.

We apply the moment condition (4.3b), but with

$$m_t(\beta) = S_{t-1} \otimes \beta' R_{pt} \qquad (4.5)$$

where S_{t-1} is a set of predetermined information variables[41] and Z_{t-1} remains as the constant.

We took Ferson's moment condition for this measure since it is more straightforward for estimating the alpha. To test this model, we also use iterated GMM[42]. The degrees of freedom are nine and we use a t-statistic to test the significance of α_p.

1.23 Conditional Measure with Fixed Weight (CONFIXED/CONS/CONL)

The third case is the conditional evaluation without dynamic strategies included. The moment conditions for this case are defined as shown in equation (4.3b), where

$$m_t = R_{pt}'\,\beta \qquad (4.6)$$

We use two sets of information variables for information vector Z_{t-1}. CONS denotes the model with a small set of information instruments, i.e. two information variables; CONL denotes the conditional model with the large set of information instruments, i.e. three information variables. We also implement the tests using iterated GMM[43], with 8 degrees of freedom for CONS and 16 for CONL. To examine the significance of α_p, we also report t-statistics.

[41] Information set S_{t-1}, include *FTSE* (*-1*) and *FTSE* (*-2*).

[42] Since there are 18 moment conditions/equations (2*9: 2 information variables, 8 reference variables and 1 parameter for alpha) and 9 parameters, the degree of freedom is therefore 18-9=9.

[43] *J*-statistic, with the degree of freedom of 8 (16-8) for a small information set and the degree of freedom of 16(24-8) for a large information set.

1.24 Conditional Measure with Time-Varying Weight (CONVARY)

Finally, we examine the conditional measure with dynamic strategies. The moment conditions follow equation (4.3b), but also with

$$m_t(\beta) = S_{t-1} \otimes \beta' R_{pt}$$

where S_{t-1} is the vector of information variables, as employed in UNVARY and Z_{t-1} is constructed with the small set of information variables, as in CONS. This model is also tested by the iterated GMM. In this case, the degrees of freedom of *J*-statistic are 27[44] and a *t*-statistic is used to examine the significance of α_p.

We apply the four cases mentioned above to investigate the properties of the SDF primitive efficient model within the context of different ways of incorporating information.

2. The Model and Moment Selection Statistics

Traditional asset pricing models reveal how the portfolio returns are determined and which factor affects the returns. From another point of view, the stochastic discount factors display which price is reasonable given the returns in the current period. Asset prices can be represented as inner products of payoffs and the SDFs. If asset pricing models were the true data generating process of the returns, the SDFs could price the returns perfectly. In reality, no stochastic discount factors can price financial assets perfectly as asset pricing models are approximations. The pricing errors may also occur because the empirical counterpart to the theoretical stochastic discount factor is

[44] The degree of freedom is calculated as 2*2*9-9 =27

error ridden. (e. g., see Roll's (1977) critique of the single-period capital asset pricing model). It is therefore important to measure the pricing errors produced by the SDF models so that we can evaluate these models. For this purpose, Hansen (1982) initiated the *J*-statistic for over-identifying conditions; Hansen and Jagannathan (1997) developed the *HJ* distance measure. Moreover, the moment selection criteria simply resemble the widely used likelihood-based selection criteria BIC. We discuss these selection statistics below.

2.1 *J*-Statistic for Over-identifying Conditions

As explained in chapter 3, the *J*-statistic, proposed by Hansen (1982), has been widely applied in the performance evaluation literature and can be used to detect the abnormal portfolio performance. For more details, please refer to chapter 3 section 2.3.

The *J*-statistic is applied mostly for the identification of the funds' abnormal performance when the asset pricing models are applied to funds. In addition, it can also be used to test the misspecification of the whole model in the case that the model only applies to the primitive assets.

2.2 The *HJ* distance measure

While there are many reasonable measures that can be used for model misspecification, the one introduced by Hansen and Jagannathan (1997) has gained tremendous popularity among the empirical asset pricing literature. Their proposed measure, called *HJ* distance measure, has been used both as a model diagnostic test and as a tool for model selection by many researchers. Examples include Jagannathan and Wang (1996), Jagannathan, Kubota, and Takehara (1998), Campbell and Cochrane (2000), Lettau and Ludvigson (2001), Hodrick and Zhang (2001), and Dittmar (2002), among others.

The *HJ* distance measure is a statistic to measure the distance between a true pricing kernel (the SDF) that prices all assets and the implied pricing kernel proxy of an asset pricing model. The distance between these two random variables is calculated as the square root of the expected value of the squared difference between the two variables.

In more details, as Hansen and Jagannathan (1997) note, an asset pricing model provides a pricing kernel proxy, y_{t+1}. If the model is true, $y_{t+1} \in M_{t+1}$. Suppose when the asset pricing model is false, $y \notin M$, there is a strictly positive distance between y and M. This distance is defined as:

$d = \min_{m \in L^2} \|y - m\|$ Subject to E (mR) =1, (4.7)

The problem now is to solve a Lagrangian minimization problem:

$$d^2 = \min_{m \in L^2} \sup_{\lambda \in R^n} \left\{ E(y-m)^2 + 2\lambda' [E(mR) - 1] \right\} \qquad (4.8)$$

The value of d is the minimum distance from the pricing proxy y to the set of true pricing kernels m.

The equation (4.8) can be solved to find

$$y - \tilde{m} = \tilde{\lambda}' R, \quad \text{where } \tilde{\lambda} = E(RR')^{-1} E(yR - 1) \qquad (4.9)$$

The *HJ* distance is therefore

$$d = \|y - \tilde{m}\| = \|\tilde{\lambda}' R\| = [\tilde{\lambda}' E(RR') \tilde{\lambda}]^{1/2} \qquad (4.10)$$

Substituting (4.9) into (4.10), *HJ* distance is shown as:

$$d = HJ = [E(yR-1)' E(RR')^{-1} E(yR-1)]^{1/2} \qquad (4.11)$$

where R is a given set of gross returns of the assets to be tested.

We calculate the *HJ* distance measure based on equation 4.11.

One of the key differences between the *J*-statistics and the *HJ* distance measure is the distance or weighting matrices in these two quadratic forms. The distance matrix in the *HJ* distance measure is $(ERR')^{-1}$, which is invariant to the choice of the price kernel proxy. In the case of the *J* statistic, the distance matrix is proportional to the inverse of the asymptotic covariance matrix for a central limit approximation. Equivalently, it is the inverse of the spectral density matrix for the time series process associated with the pricing error vector.

Unlike Hansen's *J*-statistic, the *HJ* distance measure does not follow a chi-squared distribution asymptotically. Instead, Jagannathan and Wang (1996) suggest that, for linear factor models, the *HJ* distance measure is distributed as a weighted chi-squared distribution asymptotically. They also suggest a simulation method to develop the empirical *p*-values of the *HJ* distance statistic.

There are mainly two reasons why the *J*-statistic has been criticized for the use of such a weighting matrix. Firstly, the *J*-statistic favours the models with highly variable pricing errors because it is inversely related to the variances of the pricing errors. Secondly, Jagannathan and Wang (1996) suggest that the *J*-statistic can not be used to compare the relative performance of different models since the statistic uses different weighting matrices for different models. Responding to these problems, the newly developed *HJ* distance measure exhibits better properties. Firstly, the *HJ* distance measure does not depend on the variances of the pricing errors; therefore, it does not reward the models with noisy pricing errors. Secondly, it uses the same weighing matrix for different asset pricing models; we therefore can compare the relative sizes of the *HJ* distances for each model to explore the relative performance of competing models.

2.3 Model and Moment Selection Criteria (MMSC-BIC)

Model and moment selection criteria (MMSC-BIC) resembles the widely used likelihood-based selection criteria BIC, with a purpose to select not only the correct model specification, but also all the correct moment conditions asymptotically.

MMSC-BIC criteria have two parts: the first is the *J*-statistic for testing over-identifying restrictions. Another term is for rewarding the use of fewer parameters for a given number of moment conditions and the use of more moment conditions for a given number of parameters. It is specified as:

$$J_n(b,c) - (|c| - |b|)\ln n \qquad (4.12)$$

where $|b|$ and $|c|$ denote the number of parameters and moments, respectively, n is the sample size[45] and it is the proper analogue of the BIC model selection criterion since it makes the same asymptotic trade-off between the ‘model fit’ and the ‘number of parameters’. (See Andrews and Lu (1999))

3. The results of model selection

Before implementing performance evaluation tests, it is essential to measure the impact of different information variables based on conditional performance evaluation models. It is also of great importance to compare the pricing errors of the SDF models which are applied to price the primitive assets. For these purposes, we initially

[45] N=344

incorporate two sets of information variables into the conditional performance evaluation of the primitive assets, with an aim of examining whether our findings are sensitive to the choice of these information instruments. We then detect the best-fitting SDF candidate model with the smallest pricing errors based on the model & moment selection statistics.

This section is organized as the following: firstly, we investigate the optimal choice of information variables within the context of the CONFIXED framework based on the model selection statistics and report the results in 3.1. Secondly, in 3.2, we measure the misspecification of the SDF primitive efficient models when they are within the context of each of the UNFIXED, UNVARY, CONS, CONL and CONVARY framework.

3.1 The Choice of Information Variables

Previous research has used a standard set of lagged variables as information variables, such as yield spreads, the level of interest rates and aggregate market dividend/price or similar ratios to measure the state of the economy. For example, Ferson and Warther (1996) and Ferson and Schadt (1996) use dividend yield and short-term interest rates. For the purposes of checking the robustness and validity of the previous results, in the context of the UK market, we include similar lagged variables, with information variable candidates as follows:

FTSE: lagged FTSE all share value-weighted market return as a proxy for equity market return

TERM: spread of a 3-month over a 1-month UK interbank rate

There is considerable empirical evidence that the term structure of interest rates contains information about future economic variables (including stock returns). The short-term risk free rates are assumed to contain information about future economic conditions and to capture the state of investment opportunities. It is also widely recognized that the term structure of interest rates, which characterizes the movements of risk-free bond yields, is determined by various state variables summarized in the pricing kernel. Therefore, the term structure of interest rates, or the yield curve helps to provide information on the asset pricing process. The yield spread between 3-month and 1-month UK interbank rate (middle rate) are then used to capture the yield curve.

DIVD: The annual dividend yield of FTSE ALL Share

It has long been recognized that the movements in the dividend yield series are related to long-term business conditions and they capture some predictable components of the returns. Researchers have been implying valuation ratios, mainly dividend yield and price-earnings ratio to examine the long-run stock market outlook. Examples include Campbell and Shiller (1998), who assume the distribution of valuation ratios is stable and therefore, should adhere to their mean reversion theory, i.e. when the prices are relatively high, the prices will eventually fall such that the ratios revert to their historical means.

Following Dahlquist and Soderlind (1999)'s approach, to examine which information variable set can provide smaller pricing errors and if the performance evaluation results are sensitive to the choice of information instruments, we estimate several conditional SDF models, with various combination of information variables to evaluate the primitive assets. We construct totally 12 sets of information variables, including 6 sets with two information variable combinations and 6 sets with 3 information variable combinations. The purpose of the study is to identify an optimal small information variable set (two variables) and an optimal large information variable set (three variables).

The *J* statistic and *HJ* distance measure are computed to measure the pricing errors of each model. The model with the smallest pricing error has the best evaluation accuracy of the primitive assets.

The results are shown in table 4-1 panel A. For each model, we report the *J*-statistic, the *HJ* distance measure and the degrees of freedom of the *J*-statistic (equals to the number of moment conditions minus the number of parameters). The results for different information variable combinations are as below:

Small set of information variables

In the case of including two information variables at a time, we have 2-lagged returns of each information variable (denoted as FTSE (-2), TERM (-2) and DIVD (-2)), and the combinations of any two information variables with one lagged return, i.e. FTSE (-1), TERM (-1); FTSE (-1), DIVD (-1) and TERM (-1), DIVD (-1). Among various proxies for the state of the economy, we find that the states of the term structure of interest rates are informative about fund performance as the model with information variables TERM (-2) delivers the smallest *HJ* distance measure and *J*-statistics. (*J*-statistic is 11.889 and *HJ* is 1.95877E-27). As a result, for all the following empirical tests, our small information variable set consists of UK interest rate spread and its lagged variables.

Large set of information variables[46]

In the case of a set of three information variables, we denote 3-lagged returns of each information variable as FTSE (-3), TERM (-3) and DIVD (-3). The combinations of information variables with different levels of lags include FTSE (-1), TERM (-1) and DIVD (-1), a combination of TERM (-2) and DIVD (-1) and a combination of TERM (-1) and DIVD (-2). Among them, the results suggest that the use of the combination of FTSE (-1), TERM (-1) and DIVD (-1) generates the smallest *J*-statistic and *HJ* distance measure. (*J*-statistic is 19.781 and *HJ* is 1.60296E-17). We therefore choose this combination as our large information variable set.

To conclude, the empirical results suggest that TERM(-2) was tested as the optimal small information variable set and we use it in model CONS; The combination of 1-lagged return of all three information variables is proved to generate the smallest pricing error among the large information variable sets and we use it to model CONL.

[46] With three information variables, there will be 24 moment conditions and the degree of freedom is 16

TABLE 4-1 THE PROPERTIES OF CONDITIONAL MEASURES

The following tables report the properties of conditional measures. Firstly, we want to select the optimal information variable sets to carry on conditional measure and the results are shown in Panel A. Panel B compares the pricing errors etc of various unconditional/conditional measures.

PANEL A. THE PRICING ERRORS OF CONFIXED SDF PRIMITIVE EFFICIENT MODELS WITH VARIOUS KINDS OF INFORMATION VARIABLE COMBINATIONS.

We report the *J* and *HJ* statistics for CONFIXED model when a broad range of information variable combinations are tested. The information variable sets we considered are up to a combination of three information variables. *x*(-*i*) stands for from *i*-lagged return of information variable *x* to 1-lagged return of variable *x*. hence the number of *x(-i)* is *i*. For information variables *x*: *FTSE* is the return of the *FTSE* all share index, *TERM* is the term structure variable; *DIVD* is the dividend yield on FTSE all share index. *df* is the degree of freedom.

SMALL SET OF INFORMATION VARIABLES			
	***FTSE*(-2)**	***TERM*(-2)**	***DIVD*(-2)**
J (df:8)	21.18131*	11.889	24.78282**
HJ	0.00000429	1.96E-27	0.00001514
	***FTSE*(-1),*TERM*(-1)**	***FTSE*(-1),*DIVD*(-1)**	***TERM*(-1),*DIVD*(-1)**
J(df:8)	22.04957**	17.28169*	15.3722*
HJ	0.000482	0.0723	0.000523
LARGE SET OF INFORMATION VARIABLES			
	***FTSE*(-3)**	***TERM*(-3)**	***DIVD*(-3)**
J (df:16)	28.3555*	28.43953*	43.80088**
HJ	0.0000425	0.00519	0.0315
	***FTSE*(-1),*TERM*(-1), *DIVD*(-1)**	***TERM*(-2),*DIVD*(-1)**	***DIVD*(-2),*TERM*(-1)**
J (df:16)	19.781	22.38991	24.9615
HJ	1.60E-17	0.00000228	0.00000386

**p<0.01

* p<0.1 based on the empirical critical values of GMM iterated estimators with a sample size of about 360. (see table 3-2 panel A)

PANEL B. THE STATISTICS OF FIVE UNCONDITIONAL/CONDITIONAL SDF PRIMITIVE EFFICIENT MODELS

It reports *HJ* and Mean square error (*MSE*) for five unconditional/conditional models within the framework of the SDF primitive efficient model. Five models are introduced, they are UNFIXED (unconditional model with fixed weight), UNVARY (unconditional model with time varying weight), CONS (conditional model with small information variable set and fixed weight), CONL (conditional model with large information variable set and fixed weight), CONVARY(conditional model with time varying weight).

	UNFIXED	UNVARY	CONS	CONL	CONVARY
E(m)	0.990412	0.98324	0.99251	0.98763	0.99713
SD(m)	0.167	0.225	0.197	0.354	0.369
HJ	2.22E-32	1.32E-25	1.96E-27	1.60E-17	3.88E-17
MSE	0.045206	0.054507	0.046996	0.055586	0.185677

3.2 Unconditional Models VS Conditional Models

The second task is to investigate the misspecification of a range of the SDF primitive models, namely UNFIXED, UNVARY, CONS, CONL, CONVARY, which incorporate the conditional information and dynamic strategies in different ways as shown in section 1.2.

Table 4-1 Panel B reports the average, standard deviation of the SDF, the *HJ* distance measure and Mean Square Error (MSE) for five models. Generally, the standard deviation of the fitted SDF is getting bigger as the complexity of the models increases. Ferson *et al.* (2002) have generated similar results and as they argued, a potential interpretation for such a result is when the number of assets increases, the minimum variance of an SDF should increase since the mean variance frontier can only expand as more assets are included.

The *HJ* distance measure ranges between 2.22E-32 and 3.88E-17 and MSE ranges between 0.045206 and 0.185677 for conditional models. Among all, UNFIXED has the smallest *HJ* and MSE. (*HJ* is 2.223E-32 and MSE is 0.045206). CONVARY generates the largest *HJ* and MSE. (*HJ* is 3.88E-17 and MSE is 0.185677). CONS has a smaller *HJ* and MSE than UNVARY; CONL generates larger pricing errors than CONS.

The improvement in performance measurement using the conditional models has been proved by quite a few publications, including Cochrane (1996), Hodrick and Zhang (2001), Lettau and Ludvigson (2001) and Fletcher (2005) among others. The main argument is: Conventional performance evaluation measures assume that economic states and investment styles remain constant. On the other hand, conditional performance evaluation compares a fund's returns with the returns of a dynamic strategy that matches the fund's time-varying risk exposures. Since conditional performance evaluation uses more information than conventional methods, it is supposed to provide more accurate performance measures.

Unlike these results, our empirical tests of the SDF primitive efficient models discover that the conditional measures do not generate smaller pricing errors, if not bigger than their unconditional counterparts. It seems difficult to explain why our results are so different from theirs as the consensus results make good intuitive sense. I would argue that firstly, it is fairly possible that different asset pricing models could generate sparse empirical results, as proved by the previous vast amount of research outcomes (can refer back to chapter 2). Secondly, our results are consistent with Ghysels (1998) and Ferson *et al.* (2002). As Ferson *et al.* (2002) argue, *'The conditional models generally produce larger unconditional HJ distances than their unconditional model counterparts. In attempting to price the dynamic strategies* *implied by the lagged instruments, the conditional models sacrifice some accuracy on the primitive returns.'*

4. The Results of Performance Evaluation

Having examined the properties of models containing different information variables and dynamic strategies, we evaluate the performance of UK unit trusts. Five models (UNFIXED, UNVARY, CONS, CONL, CONVARY) are applied to test the significant abnormal performance of the SDF primitive efficient model for each type of unit trusts. Table 4-2 reports the results of performance evaluation.

Panel A reports for each type of conditioning measure, and for each fund group particularly, the mean, standard deviation, skewness, kurtosis, minimum and maximum values of alphas, *J*-statistic (for UNFIXED CHEN only), *t*-statistics of alphas, and the *p*-value of those statistics.

Chen and Knez's moment condition is applied only to test UNFIXED and it shows that all funds group perform positively, given that the mean of alphas for growth funds is 0.546% (monthly), 0.145% for income funds and 0.1182% for general funds. Ferson's moment conditions are used for the rest of the models (UNVARY, CONS, CONL and CONVARY); it is striking to see that almost all of the models generate negative average alphas, except that the growth funds have generated positive average alpha within UNVARY framework.

Our results suggest that once the time varying nature of the unit trusts performance are corrected for, by the addition of the information variables, on average, the unit trusts generate decreased value of alphas over those based on unconditional methods of analysis. For example, the average alpha for the growth funds is 0.546% for UNFIXED; 0.0000625 for UNVARY, -0.0000388 for CONS; -0.0000432 for CONL and -0.000170 for CONVARY. Since the conditional performance evaluation approach takes the view that a managed portfolio strategy that can be replicated using information readily available to the public should not be judged as an indication of superior investment ability, only the managers who correctly use more information than is generally publicly available are considered to have potentially superior investment ability. Therefore, after incorporating conditional information, the superior abnormal performance within the unconditional framework disappeared.

To investigate the significance of non-neutral performance, we count the number of funds with significant non-neutral performance and report the results in table 4-2 panel B.

It reports for each type of funds and each type of unconditional/conditional measure, the number of funds with significant non-neutral performance/alpha, which is measured by the *p*-values of certain statistics with the null-hypothesis that there is no abnormal performance. For UNFIXED model only, CHEN refers to the measure using Chen's moment conditions while FERSON refers to the measure with Ferson's moment conditions. The null-hypothesis statistics (to examine the significance of non-neutral performance/alpha) is the *J*-statistic for CHEN and *t*-statistics for FERSON.

The results of Panel B can be explained based on the following two aspects:

Significance of non-neutral performance/alpha

For the null hypothesis that there is no abnormal performance, two tests are implemented. They are the *J*-statistic for the UNFIXED model when the moment condition is CHEN and *t*-statistic for the rest. The *p*-value is the probability of observing a value of the test statistic at least as extreme as the value actually observed, assuming that the null hypothesis is true. To test a hypothesis, we can compare the *p*-value with the significance level. If the *p*-value is smaller than the significance level, then the null hypothesis is rejected. Otherwise, it is not rejected. That is to say, concerning a 10% significant level, the *p*-value of the *J*-statistic and *t*-statistic should be less than 0.1 to reject the null hypothesis and show that there is a significant abnormal performance.

At the 10% significant level, in the case of the UNFIXED model with CHEN's moment condition, 3 growth funds (out of 11), 5 general funds (out of 13), 8 income funds (out of 16) have shown significant non-neutral performance. Among them, 3 growth funds, 5 general funds and 6 income funds have delivered positive returns. With Ferson's condition, 6 growth funds, 4 general funds, 8 income funds have significant returns. Among those, 6 growth, 4 general, 7 income funds underperformed rather than outperformed. The results with UNVARY, CONL, CONS, CONVARY measures do vary (the next section explains the various performance using different measures), but it is apparent that significant abnormal performance largely exists, i.e. with CONVARY measure, 36.36% of growth funds, 53.84% of the general funds, and 50% of the income funds show significant non-neutral performance. The results within 5% and 1% significant level are also reported in panel B.

Performance results within the framework of different models

For each unit trust, we further compare the performance results generated with different performance measures and information incorporation. The results are sparse though largely indicating the existence of significant underperformance. In the case of the UNVARY model, 4 out of 11 growth funds, 6 out of 13 general funds, and 8 out of 16 income funds generate abnormal performance. Most of them underperform significantly (i.e. 2 growth, 5 general and 6 income funds). In the case of CONL model, 2 growth funds, 5 general funds, 5 income funds have significant results. Among them, 5 general funds and 3 income funds underperform. In the case of the CONS model, 1 growth fund, 3 general funds and 6 income funds deliver significant results, with mostly underperformance (i.e. 1 growth, 2 general and 6 income funds under-performed). In the case of CONVARY model, 4 growth funds, 7 general funds and 8 income funds provide abnormal performance (3 growth funds, 7 general funds and 2 income funds underperformed).

To investigate if the performance results based on these methods are significantly different, we apply paired-sample *t*-test to compare for each fund, the means of paired-alphas, derived from two different measures, are significantly different. The paired-sample *t*-statistic computes the differences between the values of two alphas for each fund and tests whether the average differs from zero. Formally, in order to test $H_{0:}\alpha_D = 0$, we calculate a *t*-statistic such as:

$$t = \frac{\overline{\alpha}_D}{\frac{SD(\alpha_D)}{\sqrt{n}}} \qquad (4.13)$$

where α_D is the difference between corresponding alphas, and the mean of α_D is measured as: $\overline{\alpha}_D = \frac{1}{n}\sum_{i=1}^{n}(\alpha_{A,i} - \alpha_{B,i})$ with $\alpha_{A,i}$ and $\alpha_{B,i}$ as the alphas of fund i estimated based on methods A and B respectively, n is the number of the funds.

The results are reported in Table 4-2 Panel C. Ten pairs of alphas are tested for significant differences for each type of funds. We report the results following the ascending order of conditional information and dynamic strategy incorporation such as any two combinations of UNFIXED, UNVARY, CONS, CONL, CONVARY. Pair 1 to pair 7 are the comparisons of unconditional measures with conditional measures, i.e. UNFIXED/ UNVARY vs their counterparts. Pair 8 to pair 10 are the comparisons among conditional measures.

The results largely confirm the following trends:

Firstly, most of the *t*-statistics are positive, which suggest that unconditional measures produce more positive alphas than those based on conditional measures.

Secondly, from pair 1 to pair 7, the difference between paired alphas are significant whilst it is less so from pair 8 to pair 10.

With the increasing of complexity of the models (from unconditional to conditional models), the number of significant abnormal performance is decreasing whilst the number of significant negative performance is increasing. The alphas are significantly different when based on unconditional and conditional measures. While in most of the cases, they are not significantly different between the two conditional measures.

These results suggest that conditional performance evaluation can accommodate whatever information is held to be appropriate by the choice of the lagged information variables used to represent the publicly available information. By incorporating a given set of lagged information variables, the fund managers who trade mechanically in response to these information variables would not be able to outperform the benchmark. Compared to the unconditional measure, the conditional approach can raise the hurdle on managers seeking abnormal positive performance as it does not give them credit for using readily available information. Only the managers who outperform such a stringent benchmark are believed to have superior investment capabilities. Therefore, the conditional measure avoids some of the biased results generated by traditional measures.

To conclude the evaluation tests, only about half of the unit trusts within our samples (less than 50% for growth funds) performed significantly different from the benchmark. Particularly, the percentage of the unit trusts with significant superior performance is relatively few, ranging from 23% to 38.46%, with the unconditional measure and ranging from 0% to 15.38%, with conditional measures. It is also worth noting that the returns of UK unit trusts applied in our analysis are not the net returns excluding all the transaction costs and fees. In practice, portfolio managers' performance is measured by fund gross returns, but net return is certainly more meaningful for fund investors. As explained in chapter 1, the true annual charges for unit trust investors are called Total Expense Ratio (TER), which are the drag on fund performance caused not only by the annual management charge, but also by administration, custody and audit fees. The historical data of TER for each fund is difficult to find; Fitzrovia's research (2003) suggests annual TERs of UK unit trusts averages 1.57%. Nevertheless, it implies the percentage of superior performance would be even lower if taking TERs into account.

The results are not surprising. The UK equity market is a relatively mature and developed financial market. The depth of the market, the sophistication with the use of full ranges of financial instruments, the wide availability of company information, make the UK equity market one of the most efficient financial markets in the world. The fierce competitive nature of investment management industry and the fact that London attracts a pool of most talented financial professions in the world also suggest that little inefficiency in the market can be explored. As a result, very few funds can generate positive alphas, while the majority of the funds do not perform differently from the market average, some underperform occasionally. Our results may well however, support the EMH and question the existence of active asset management industry as a whole.

On the other hand, active management has its own reasons for existence. Though difficult to find, few funds did outperform over our sample period. Those are the shining stars of active management. Although the investment talents with superior investment skills only take a very small percentage, they do exist. Besides, in the eyes of active fund managers, benchmark performance numbers are simply 'ideal' figures due to the complexity of active fund management in real life. Firstly, given a portfolio size (either too small or too big), it is nearly impossible to replicate the benchmark 100% by holding all the positions it includes. Secondly, the 'ideal' benchmark does not need to manage cash inflows and outflows, which the fund manager has to face with on a daily basis. During a market sell-off, the fund manager may have to deal with large amount of redemptions by forcefully selling some positions, which would push the asset prices lower as there are only very few buyers in the market against vast majority of sellers. Thirdly, the benchmark index normally has its monthly rebalancing: the weights of certain stock in the benchmark may have changed due to the change of its market value; new comers due to IPO or a rising stock has increased its market capital sufficiently, thus made its way through to the large cap index, say, FTSE 100; some stocks may have dropped out of the index due to bad performance or be delisted etc.. Rebalancing implies transaction costs in real life. These are not, however, reflected in the benchmark performance figures.

Our results also demonstrate strong evidence of a downward performance of unit trusts after incorporating conditioning information analysis. It suggests that in the strictest and ideal way, the portfolio managers' relative performance should be measured based on conditioning information analysis rather than unconditional methods, and a fund investor should take the conditioning information into account when he/she evaluates and compares different portfolio managers. In reality, this is rarely the case. For one thing, many asset management companies measure and report their own fund performance following GIPS (Global Investment Performance Standards), a set of standardized, industry-wide ethical principles that provide investment firms with guidance on how to calculate their investment results. GIPS is a sophisticated method with strict rules on how to deal with input data, return calculation and composite construction, i.e. GIPS requires total returns must be calculated based on realized and unrealized gains and losses plus income, and time-weighted rates of return that adjust for external cash flows must be used. Nevertheless, it calculates an absolute return numbers (non-risk adjusted) unconditional measure of the portfolio against its benchmark. For another, professional practice of portfolio manager selection (according to the process by the practitioners working in multi-management, fund of funds) are more involved with the analysis of qualitative quality of the fund/portfolio managers (i.e. investment process, management philosophy, manager quality etc) and probably the simple version of risk-adjusted return numbers (i.e. Sharpe ratios). For most of the private investors, they believe that the out-performance of the portfolio manager depends on if he/she can bring the highest returns in monetary terms, possibly on a pure non-risk adjusted basis, not even to mention on a conditional measure basis. A more developed financial market requires a more sophisticated performance evaluation method. The way how conditional performance evaluation can be applied in practice needs to be explored further.

TABLE 4-2 PERFORMANCE EVALUATION OF INDIVIDUAL FUND WITHIN THE FRAMEWORK OF THE SDF PRIMITIVE EFFICIENT MODELS

The tables report the results of performance evaluation of individual fund for each type of funds separately using the SDF primitive efficient models with different unconditional/conditional measures. Panel A reports the descriptive statistic of the alphas for different types of measures. Panel B displays the paired sample *t* test results to investigate if the alphas generated from different conditional measures are significantly different. Panel C shows the degree of significant non-neutral performance.

PANEL A. DESCRIPTIVE STATISTIC SUMMARY OF ALPHAS

It reports for each type of conditioning measure, and each fund group particularly, the mean, standard deviation, skewness, kurtosis, minimum and maximum values of alphas, *J* statistic (UNFIXED only), *t*-statistics of alphas, and the *p*-value of those statistics.

UNCONDITIONAL MODELS

UNFIXED CHEN

GENERAL (13)	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum
$E(\alpha_p)$	0.001182	0.001353	0.734701	2.02187	-0.00018	0.003566
J	2.74729	2.25857	0.594393	1.65389	0.283172	6.64954
p-value	0.206111	0.187978	0.660294	2.22072	0.009918	0.594629
GROWTH (11)						
$E(\alpha_p)$	0.000546	0.001112	1.18573	3.95984	-0.00093	0.003381
J	1.70484	2.19125	1.63929	4.73513	0.001673	7.64341
p-value	0.395345	0.307495	0.420947	1.94488	0.005698	0.967378
INCOME (16)						
$E(\alpha_p)$	0.00145	0.002444	0.743148	2.80058	-0.00157	0.006535
J	3.16553	2.37893	0.521182	2.12069	0.009221	7.60171
p-value	0.195253	0.248963	1.73856	5.25317	0.005831	0.923499

UNVARY

GENERAL	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum
α_p	-0.00059	0.001257	1.47775	5.07701	-0.00226	0.002852
t-statistic	-0.89076	1.22691	0.565578	2.78633	-2.78079	1.83071
p-value	0.340872	0.326175	0.564484	1.76681	0.005423	0.959504
GROWTH						
α_p	0.0000626	0.001546	0.31657	1.76011	-0.00195	0.002694
t-statistic	-0.063	1.52089	0.316952	2.18608	-2.31854	2.91411
p-value	0.390163	0.362587	0.411247	1.46017	0.003567	0.95482
INCOME						
α_p	-0.00102	0.001833	0.458052	2.64134	-0.00403	0.002689
t-statistic	-0.8365	1.64771	0.793112	2.81367	-3.10559	2.78794
p-value	0.228449	0.221357	0.340788	1.42549	0.001899	0.579099

CONDITIONAL MODELS

CONS

GENERAL	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum
α_p	-0.00037	0.001237	1.09679	4.08151	-0.00212	0.0028
t-statistic	-0.46375	1.03773	0.565809	2.54386	-1.76816	1.81323
p-value	0.451028	0.326966	0.100422	1.2695	0.069796	0.886202
GROWTH						
α_p	-0.0000388	0.00135	0.072187	1.5843	-0.00187	0.002183
t-statistic	-0.13548	1.07281	-0.107	1.48698	-1.72214	1.40191
p-value	0.408852	0.260083	0.27933	1.37583	0.085043	0.792244
INCOME						
α_p	-0.00089	0.001704	0.576705	2.99821	-0.00374	0.003043
t-statistic	-0.75361	1.39674	0.936473	3.26524	-2.34578	2.64945
p-value	0.302103	0.27129	0.236841	1.30577	0.008062	0.684836

CONL

GENERAL	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum
α_p	-0.00092	0.001303	1.07711	3.75291	-0.00258	0.002332
t-statistic	-1.0798	1.19697	0.633075	2.48161	-2.50407	1.49408
p-value	0.276119	0.265333	0.517258	1.71213	0.012277	0.758476
GROWTH						
α_p	-0.0000432	0.001408	0.972712	2.63697	-0.00155	0.002943
t-statistic	-0.042	1.26905	0.779315	2.61867	-1.63828	2.73255
p-value	0.428489	0.320957	0.397312	1.67584	0.006285	0.964782
INCOME						
α_p	-0.00029	0.001577	0.235046	2.37055	-0.00311	0.00277
t-statistic	-0.39229	1.41418	0.38586	2.23857	-2.18813	2.57625
p-value	0.373721	0.331886	0.354994	1.64812	0.009988	0.995217

CONVARY

GENERAL	Mean	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum
α_p	-0.00054	0.001585	1.69715	5.14986	-0.00201	0.00384
t-statistic	-1.16378	1.18391	0.521937	1.84739	-2.81382	0.833181
p-value	0.277926	0.317381	0.895304	2.16895	0.004896	0.889834
GROWTH						
α_p	-0.00017	0.00174	0.0375	1.71698	-0.00294	0.002401
t-statistic	-0.19277	1.60754	0.527538	2.52198	-2.5771	3.25491
p-value	0.349992	0.347554	0.893934	2.33636	0.001134	1
INCOME						
α_p	-0.00145	0.001941	-0.11542	2.37756	-0.00521	0.002028
t-statistic	-0.74965	1.76385	0.561617	2.60165	-3.65194	2.74514
p-value	0.349945	0.423137	0.675841	1.58601	0.00026	1

PANEL B. THE NUMBER OF FUNDS WITH SIGNIFICANT NON-NEUTRAL PERFORMANCE FOR EACH FUND GROUP

It reports the non-neutral performance for each type of funds. The number in the bracket is the total number of funds tested for that type of funds. The numbers of funds with significant abnormal performance are counted. CHEN refers to Chen's moment condition; FERSON refers to Ferson's moment condition. UNFIXED, UNVARY, CONL, CONS, CONVARY are different kinds of unconditional/conditional models as mentioned above.

GENERAL FUNDS (13)						
Moment condition	**CHEN**	**FERSON**				
	UNFIXED	**UNFIXED**	**UNVARY**	**CONL**	**CONS**	**CONVARY**
p-value<0.1	5	4	6	5	3	7
p-value <0.05	4	3	2	4	0	4
p-value <0.01	0	1	1	0	0	1
Positive	5	0	1	0	1	0
Negative	0	5	5	5	2	7

GROWTH FUNDS (11)						
Moment condition	**CHEN**	**FERSON**				
	UNFIXED	**UNFIXED**	**UNVARY**	**CONL**	**CONS**	**CONVARY**
p-value <0.1	3	7	4	2	1	4
p-value <0.05	2	5	2	0	0	2
p-value <0.01	0	1	1	0	0	2
Positive	3	0	2	2	0	1
Negative	0	7	2	0	1	3

INCOME FUNDS (16)						
Moment condition	**CHEN**	**FERSON**				
	UNFIXED	**UNFIXED**	**UNVARY**	**CONL**	**CONS**	**CONVARY**
p-value <0.1	8	8	8	5	6	8
p-value <0.05	7	6	5	5	4	6
p-value <0.01	2	2	3	1	1	3
Positive	6	1	2	2	0	2
Negative	2	7	6	3	6	6

PANEL C. PAIRED SAMPLES TEST OF ALPHAS FOR UNCONDITIONAL /CONDITIONAL MEASURES

It reports for each fund group separately, the paired differences of alphas are tested to examine if the performance results are significantly different when using various conditional measures. 10 pairs of alpha differences are tested, and the purpose is to see particularly, if unconditional measure generates significantly different results from conditional measure as well as if different measures produce sparse results.

GENERAL FUNDS

		Mean	Std. Deviation	Std. Error Mean	*t*	*df*	Sig. (2-tailed)
Pair 1	UNFIXED- UNVARY	.0017742	.0023307	.0006464	**2.745**	11	**.018**
Pair 2	UNFIXED - CONS	.0015475	.0023046	.0006392	**2.421**	11	**.032**
Pair 3	UNFIXED - CONL	.0020981	.0022931	.0006359	**3.299**	11	**.006**
Pair 4	UNFIXED- CONVARY	.0017167	.0025731	.0007136	**2.406**	11	**.033**
Pair 5	UNVARY - CONS	-.000226	.0002066	.0000573	*-3.955*	11	**.002**
Pair 6	UNVARY - CONL	.000323	.0004290	.0001190	**2.722**	11	**.019**
Pair 7	UNVARY - CONVARY	-.000057	.0008367	.0002320	-.248	11	.808
Pair 8	CONS - CONVARY	.000169	.0009604	.0002663	**.635**	11	.537
Pair 9	CONS - CONL	.000483	.0023737	.001	**.734**	11	.477
Pair 10	CONL - CONVARY	-.000381	.0011051	.0003065	-1.244	11	.237

GROWTH FUNDS

		Mean	Std. Deviation	Std. Error Mean	*t*	*df*	Sig. (2-tailed)
Pair 1	UNFIXED- UNVARY	.0005506	.00041556	.0001152	**4.777**	12	**.000**
Pair 2	UNFIXED - CONS	.000584	.0021073	.0005844	**1.001**	12	.337
Pair 3	UNFIXED- CONL	.000589	.0020832	.000577	**1.020**	12	.328
Pair 4	UNFIXED - CONVARY	.000713	.0024089	.0006681	**1.068**	12	.307
Pair 5	UNVARY - CONS	.000101	.0005211	.0001445	**.701**	12	.496
Pair 6	UNVARY - CONL	.000105	.00088781	.000246	**.430**	12	.675
Pair 7	UNVARY - CONVARY	.000229	.0012261	.000340	**.676**	12	.512
Pair 8	CONS - CONL	.0000045	.000704	.0001952	**.023**	12	.982
Pair 9	CONS - CONVARY	.000128	.001185	.000328	**.391**	12	.703
Pair10	CONL - CONVARY	.0001241	.00123702	.00034309	**.362**	12	.724

INCOME FUNDS

		Mean	Std. Deviation	Std. Error Mean	*t*	*df*	Sig. (2-tailed)
Pair 1	UNFIXED - UNVARY	.0024713	.003946	.0009865	**2.505**	15	**.024**
Pair 2	UNFIXED- CONS	.002342	.003765	.0009413	**2.488**	15	**.025**
Pair 3	UNFIXED - CONL	.001741	.003749	.0009374	**1.857**	15	**.083**
Pair 4	UNFIXED - CONVARY	.002899	.003735	.0009339	**3.104**	15	**.007**
Pair 5	UNVARY - CONS	-.000557	.000727	.00018187	*-3.064*	15	**.008**
Pair 6	UNVARY - CONL	-.00073	.001039	.0002599	*-2.810*	15	**.013**
Pair 7	UNVARY - CONVARY	.000427	.00089	.0002226	**1.922**	15	**.074**
Pair 8	CONS - CONL	-.000601	.001019	.00025494	-2.357	15	**.032**
Pair 9	CONS - CONVARY	-.000129	.000439	.0001097	-1.178	15	.257
Pair 10	CONL - CONVARY	.001158	.001415	.00035379	**3.274**	15	**.005**

5. Conclusion

The SDF approach has aroused great interest of researchers within the asset pricing field in that it provides a unified framework for performance evaluation analysis, but with different asset pricing models, assuming different 'correct' SDFs. The wide choices of risk specifications give the SDF approach further rich empirical contents.

In this chapter, we apply the SDF primitive efficient model to evaluate performance of UK unit trusts for the period from Jan 1975 to Oct 2003. To incorporate the conditioning information, we applied five types of models, denoted as UNFIXED, UNVARY, CONS, CONL and CONVARY, dealing with scaling returns and scaling factors in various ways. To examine how wrong a model is and to compare the performance of different asset pricing models, we report the pricing errors based on the *HJ* distance measure, *J* statistics and mean square errors. Our results suggest mainly the following:

• TERM is proved to be the optimal information variable to construct a small information set; the combination of FTSE, DIVD and TERM forms the optimal large information set;

• Though conditional models, particularly those allowing for dynamic strategies, do not produce smaller pricing errors than unconditional models, they generally display better forecasting capabilities.

The fund performance measures derived from alternative model specifications differ depending on the number of instrument variables used to scale assets and/or factors. However, by and large, for all the cases and all types of the funds, more than half of the funds do not have significant abnormal performance. Among those demonstrating significant non-neutral performance, most of the unit trusts show inferior performance. On average, unit trusts cannot generate excess returns relative to their benchmarks that are large enough to cover their total expenses. The results also reveal that the conditional models generate negative alphas more frequently than their unconditional model counterparts. These results are further confirmed by paired sample *t*-tests. Compared to unconditional alphas, fund performance sharply deteriorates when we measure conditional alphas. Given that the stock returns are to some extent predictable based on publicly available information, conditional performance evaluation raises the benchmark for active fund managers because it gives them no credit for exploring readily available information.

CHAPTER 5: DOES STYLE MATTER? – STYLE PERFORMANCE AND PROFITABILITY OF ROTATION STRATEGIES

1. Introduction

Among equity fund managers, value (value vs growth stocks) and size (small vs large cap stocks) strategies have been widely used for discriminating relative future performance. This implementation is known as style investing, which has also attracted numerous studies. For example, Rosenberg, Reid, and Lanstein (1985), Fama and French (1992), Lakonishok, Shleifer, and Vishny (1994), and Roll (1997), examine the long-term relative performances between growth, value, small-cap, and large-cap stocks. In addition, the potential success of style-rotation strategies has also aroused immense attention from academic researchers, including e.g., Beinstein (1995), Fan (1995), Sorensen and Lazzara (1995), Leinweber, Arnott, and Luck (1997), Kao and Shumaker (1999), Levis and Liodakis (1999), Asness *et al*. (2000), Ahmed, Lockwood and Nanda (2002), Lucas, Dijk, and Kloek (2002), and Teo and Woo (2002).

This chapter evaluates if the UK unit trusts with various kinds of styles perform differently. We examine the style performance of UK unit trusts from January 1975 to October 2003, which covers two particular periods, one is around October 1987, when the historical stock market crash, 'Black Monday' occurred, and another is the most significant correction since 1985: the bursting of the IT bubble in 2000, in which equities lost half of their values. The bear market that followed lasted until early 2003, and it took seven years for the markets to attain the levels reached in 2000.

There are two main tasks: the first task is to investigate if the funds with one specific style would perform definitely better to the other types of funds or if different types of funds actually perform differently over the business cycle? The second task is, if the latter is true, does it imply potential profitable opportunities by constructing style-rotation strategies based on the patterns of their performance?

1.1 Style Consistency: Value/Size Premium

There is no universally accepted definition of value and growth stocks. However, even if they disagree on the details, practitioners and researchers do agree on broad characteristics of stocks in these two camps.

In general, value managers prefer undervalued stocks, identified by low price/earnings ratios (P/E), low market price/book value ratios (P/B), or high dividend yields. These usually include turn-around opportunities (such as companies that are experiencing problems but are expected to recover, i.e. bankruptcy restructurings stories) and unpopular stocks (such as stocks in industries considered mature and with modest growth prospects). On the other hand, growth-oriented managers invest in companies experiencing rapid growth in earnings and sales. These types of stocks are usually associated with high market price/earnings and high price/book value ratios.

It is well recognized that value stock have outperformed growth stocks in a few countries over relatively long time periods, and small capitalization stocks had also generated higher annual returns than large capitalization stocks historically. The examples include the papers of Capaul, Rowley, and Sharpe (1993), Arshanapalli, Coggin, and Doukas (1998), Fama and French (1992), La Porta (1996), Daniel and

Titman (1997), Barber and Lyon (1997) and Lewellen (1999). These are called 'value' and 'size' effects. Excess return of value stocks over growth stocks is called 'value premium' and excess return of small stocks over large stocks is called 'size premium'.

'Value' effects, along with the evidence 'value premium', have been widely documented in the literature. It suggests that several company-specific variables, like P/E or P/B ratios, have predictive power regarding the average return of a stock. Examples include Fama and French (1992) and Lakonishok *et al.* (1994), whose evidence of US markets show that the portfolios of companies with low P/B ratios have earned significantly higher returns than those with high P/B ratios. Dimson *et al.* (2003) apply the value studies to the UK market, merging accounting information with share price data. For the period of 1955-2001, they find strong evidence of value premium in the UK. They also suggest that as the small-cap equity market outside of the US is relatively illiquid, the implementation of small-cap and value strategies outside of the U.S. is rather difficult as trading costs can be high.

There is no consensus on the sources of the value and size premium. Over years, different explanations of the long-term out-performance of value and small capitalization stocks have been given.

Firstly, it is suggested that the firm variables might proxy for risk factors because firms with similar characteristics could be sensitive to the same macroeconomic factors such as growth surprises and interest rate risk. For example, Chan and Chen (1991), Fama and French (1993), Jensen, Johnson and Mercer (1997) and Lewellen (1999) suggest that the B/P ratio and size factors are proxies for distress, therefore, value and size premium can be regarded as a reward for holding stocks of firms under some distress. On the other hand, Cochrane (1999) argues that relative distress should not be viewed as firm-specific distress, which can be diversified away. Only the non-idiosyncratic component of distress is relevant since it cannot be diversified away, which implies that the size and value premium are still related to risk, i.e. business cycle risk.

A second explanation states that unexpected technological innovations are historically more related to particular equity classes, i.e. value stocks, as a result, value or size effect can be simply due to data snooping, see, e.g., Lo and MacKinlay (1990), Black (1993) and MacKinlay (1995).

Thirdly, 'Value Premium' might be explained in terms of market over-reaction leading to security mispricing. It is believed that the outperformance of value investing might be because of investor's irrational overconfidence in growth companies and the fact that investors generate pleasure and pride from owning growth stocks. For instance, La Porta (1996) provides evidence that because expectations about future growth in earnings are too optimistic, growth stocks are over-bought and over-valued, and value stocks have better fundamental values, which tend to outperform over the long run.

On the other hand, some might argue that theoretically there is no definite priority of one style over another. For one thing, value stocks are not necessarily superior to growth stocks. In a value investing case, stocks are picked for inclusion in a portfolio because they are "cheap", in another word, the P/E or P/B ratio of these stocks are low. However, the problems are: Is 'cheap' a good investment? Not necessarily so, either in absolute terms or relative to other alternative investments. A company's future growth prospects do not merely depend on P/E or P/B ratio.

Moreover, the magnitude of the value premiums also created scepticism among the financial management and academic communities: given the unprecedented extraordinary good performance of growth stocks in the UK over 1999 to 2000, (including many "new economy" stocks such as telecommunications, media and technology stocks), it is now no longer clear whether value stocks have outperformed growth stocks over long periods of time.

For another, growth stocks do not necessarily outperform in the long run. Companies with promising growth characteristics are regarded as the companies of tomorrow, and as long as the economy is continuing to grow, they should be the companies that will be rewarded in the longer term. However, growth stocks cannot be guaranteed of out-performance merely because they are currently perceived to be the industries of the future. They may have performed so well in the past that they need unrealistically high levels of growth in the future to sustain their current prices.

1.2 Style-rotation strategies

More recently, extensive research has shown that performance of value or size related investment styles may not be persistent over time, i.e. size and value premium exhibit significant short-term (directional) variation. More and more evidence states that a strategy based on hypothesis of persistent style performance may not provide long-term benefit, whilst style diversification can be the optimal solution to avoid the risk associated with pure style investing, and to capture the benefits of each separate investment styles has to offer.

Recent studies (see Kahn (1996)) are focusing on investigating if there is significant differential performance across styles under economic changing conditions. If there is, it might imply a phenomenon of so called style rotation, that is, rotating portfolio investments across stocks of different styles as economic and market conditions change as this offers an opportunity to enhance the portfolio returns.

To name just a few, Kahn (1996) reports that most funds do not systematically follow a value or growth stock orientation, but instead tend to either shift between one and the other, or adopt a blend. Indro *et al*. (1998) discover that funds that instituted both a change in their value/growth as well as small/large-capitalization stock allocation strategies were the worst-performing group of actively managed funds. Chan, Karceski and Lakonishok (1999) suggest that the regular size and value effects inverted over the period 1990 to 1998. As such, they provide a further dimension to the current debate by allowing more flexibility in the choice of investment strategies.

Even assuming that the long-run outperformance of funds with certain style does exist, in reality, professional portfolio managers may still have incentives to change their styles over time. The arguments are, in an investment management industry, both annual out-performance and intra-year variability of the out-performance is important (Roll, 1992). The returns over a multi-year period are frequently not a sufficient factor to consider a particular fixed investment style a success. Portfolio managers are often judged by the intra-year returns relative to a benchmark pre-specified by the fund prospectus. They are therefore looking for systematic patterns in the time-varying impact of value and size on returns in order to enhance their performance.

An extensive body of financial literature has been trying to explore the predictability of return spread of value stocks over growth stocks and profitability relying on a style-rotation strategy based on such predictability. The methodologies of analyzing style rotation are basically to link the performance of style portfolios to various macroeconomic factors. Different factors have been tested and the results are sparse. For example, Beinstein (1995), Fan (1995), Kao and Shumaker (1999) investigate models that forecast value spread according to measures of aggregate economic and financial conditions, focusing on variables such as the earnings yield on S&P 500, the slope of the yield curve, corporate credit spreads, corporate profits, spreads in valuation multiples, expected earnings growth spreads, and other macroeconomic measures. Levis and Liodakis (1999) attempt to identify quantitative signals that might help to predict style changes, and they show that it is important not only to exhibit skill in choosing between different styles, but also to have good timing skills, as large relative movements can happen very quickly. Cooper *et al*. (2000) suggest similar results for the US, with successful size-sorted strategies based on sufficient predictability and weaker results for value-sorted strategies.

Certain results show considerable potential that style rotation, based on forecasting ability, offers over a passive style strategy, i.e. Lucas *et al.* (2001) point out that the rotating investment styles based on firm characteristics and macroeconomic predictors can provide consistent and robust (risk-corrected) excess returns. Ahmed, Lockwood, and Nanda (2002) find that even moderate multi-style rotation gives a portfolio an excellent chance of outperforming the market index. Nalbantov (2003) shows that a perfect-forecast rotation strategy (with a long position in the higher returning asset class and a short position in the lower returning asset class) can produce 21.29% annual return during January 1993 to January 2003 in the US. Wang (2003) finds that a style momentum and a logit-based style-rotation strategy generate higher returns.

In the following sections, we investigate mainly the following two questions. The first part (section 2) is to seek the empirical evidence if style-consistent or style-rotating performance prevails in the UK market using a sample of UK unit trusts. The method is to apply the SDF primitive efficient approach, both unconditional and conditional, to measure the performance of different styles of unit trusts during different sub-sample periods.

The second part (Section 3) is to construct certain style-rotation strategies based on the results drawn from section 2 and to examine if profitable strategies can be feasible. To achieve this, we adopted several methods: the first (section 3.1) is a simple *ex-post* strategy purely based on regime switching between bull and bear markets. We then run a series of sensitivity tests of value spreads against certain economic variables, which we discuss in section 3.2. In section 3.3, we estimate logit models to predict future style performance, a trading strategy is also constructed based on the prediction results and the profits are reported. Concluding remarks are in the final section.

2. Evidence of style performance

As discussed before, a growing literature has documented that various strategies of rotating across equity styles generate significant excess returns, while some suggests value strategies outperforms growth strategies. As arguments on style performance have reached no consensus, it is interesting to investigate if a fixed style approach generates consistent performance or if indeed, different styles perform differently over time?

Unlike most of the other researchers, we examine equity unit trusts' style performance instead of equity style performance. One may argue that evaluation of fund style performance instead of equity style performance would not be able to separate style performance from fund manager skills. However, as our research is based at the aggregate level, we assume on average, fund manager for each type of funds would incorporate equivalently proficient money management skills. In addition, our research would help to provide some practical guidance on private investors' asset allocation decisions.

Due to constraints of data availability, we focus on examining the existence of 'value' effects. Consistent with our previous study, this exercise will focus only on three types of UK unit trusts: UK growth, UK income and UK general funds. Among them, UK growth funds are those which have the prime objective of achieving growth of capital. UK equity income funds principally target high dividend income and aim to have a yield that is in excess of 110% of the yield of the FT All Share Index. UK general funds refer to those having the balanced objective of pursuing growth of capital gain and high income. To test the performance of the three kinds of unit trusts, it is equivalent to draw a comparison of performance among portfolios of growth stocks, value stocks and a blend of growth/values stocks.

Catering to investment objectives, different kinds of funds aim at picking different stocks within specific characteristics. Appendix table 5-1 reports the best and worst selling unit trusts sectors based on net retails sales in 2003 and 2004 (IMA[47]).

Table 5-1 suggests the popularity of different types of funds among private investors during different market conditions varies a lot, i.e. the most popular type of funds in 2003 November was UK corporate bond, while UK all companies (growth) remain popular from December 2003 to February 2004. It implies investors' preference of different types of funds do change over time.

To explore if there is any evidence favouring a 'style consistency' or a 'style rotation' strategy, we undertake the following tests. Firstly, we provide a descriptive analysis of the risk-unadjusted returns of each type of funds using aggregate data. Secondly, to extend the analysis on a risk-adjusted basis, we apply the SDF primitive efficient model to measure, for each sub-sample period respectively, the performance of growth, value and general funds on an aggregate basis.

2.1 The descriptive results of the risk-unadjusted returns

The descriptive statistics of monthly gross returns of each type of funds at the aggregate level are reported in table 5-2 Panel A. For the sample period from Jan 1975 to October 2003, we calculate the value spread (*Rs*) as the difference between the income funds mean returns and growth funds mean returns. The summarized statistics of value spread is reported in Panel B. We also plot the twelve-month moving average of value spread in Figure 5-1. It is clear that different times favour

[47] More recent data are available from IMA website. We use 2003/2004 data as our sample data is from Jan 1975 to Oct 2003.

different types of stocks. For example, the income funds, are more profitable from late 1985 to 1988, on the other hand, growth funds are doing better during the period from 1991 to 1993.

Both Panel A and Panel B display the mean, standard deviation, skewness, kurtosis, minimum and maximum values of those variables. Panel A shows that growth funds generate the highest mean gross monthly returns at 1.0609%, associated with the highest standard deviation, at 0.0464. Panel B suggests that the number of negative value spread is higher than the number of positive spread (176 for the former compared to 169 for the latter), and the average of value spread is negative, indicating the inferior performance of income funds compared to their growth peers. The average monthly spread is -0.104%, while the maximum spread is 5.97% and the minimum spread is -3.89%.

We then run a simple Ordinary Least Square (OLS) regression, just to discover the relationship between the risk-unadjusted value spread and the current market conditions such as

$$RS_t = \alpha + \beta R_{m,t} + \varepsilon \qquad (5.1)$$

where $R_{m,t}$ is the return (at time t) of FTSE all Share Index, a proxy for market return. The results of this regression are reported in Panel C.

From Panel C, the coefficient of market return is -0.070356 and significant with t-statistics of -4.63249. It suggests that the value spread is inversely related to market returns (level), in the sense that the growth funds tend to outperform income funds

TABLE 5-2. THE DESCRIPTIVE STATISTICS OF RISK-UNADJUSTED MONTHLY RETURNS FOR EACH TYPE OF FUNDS

PANEL A. THE DESCRIPTIVE STATISTICS OF RAW RETURNS FOR EACH TYPE OF FUNDS

No. OF OBSERVATIONS: 345

	Mean	**Std.Dev.**	**Skewness**	**Kurtosis**	**Minimum**	**Maximum**
GENERAL	1.0605%	0.049294	0.303039	10.22882	-26.31%	32.681%
GROWTH	1.0609%	0.046475	-0.5224	8.30445	-28.44%	20.98%
INCOME	1.05286%	0.044	-0.21	7.87702	-24.07%	23.95%

PANEL B. SUMMARIZED STATISTICS OF INCOME / GROWTH SPREAD

	Mean	**Std.Dev.**	**Skewness**	**Kurtosis**	**Minimum**	**Maximum**	**Median**	**N0**
RS	-0.104%	0.012981	0.115934	4.79148	-0.03892	0.059749	-0.0003	345
RS≤*0*	-1.022%	0.008753	-1.32299	4.35841	-0.03892	0	-0.0074	176
RS>0	0.9022%	0.008676	2.15825	10.2622	0.0000087	0.059749	0.00621	169

PANEL C. OLS REGRESSION OF VALUE SPREAD

We report the results for OLS regression $RS_t = \alpha + \beta R_{m,t} + \varepsilon$

F (zero slopes) = 29.9362 [.000]   *R*-squared = 0.081142

Adjusted *R*-squared = 0.078431   Log likelihood = 1012.18

Variable	**coefficient**	**Standard Error**	***t*-statistic**	***p*-value**
constant	0.070310	0.015418	4.56030	[.000]
Market Return	-0.070356	0.015187	-4.63249	[.000]

Standard Errors are heteroskedastic-consistent.

during the bull market (with higher market returns) and the value funds generate superior returns during the bear market.

2. 2 The measure based on the SDF primitive efficient model

To carry out robust tests of style performance, we examine the style performance on a risk-adjusted basis. We apply the SDF primitive efficient model[48] and follow the steps such as: firstly, the whole sample period is divided into several sub-samples, separating the time when the market is generally in good condition (i.e. rising) from those when a bear market prevails. Secondly, we calculate the average returns for each group j such as $\overline{R}_j = \sum_{i=1}^{n} R_{i,j}$, with $R_{i,j}$ refers to the gross return of individual fund i of group j and the funds' performance indicator "alpha" based on $\overline{R}_j$ can be estimated using the SDF primitive efficient model. Since each sample period is potentially related to certain specific economic and market conditions, the comparison and analysis of "alphas" and other statistics among the three types of funds should provide an idea how well one style can perform associated with market conditions.

The dataset we applied in this chapter is consistent with those we used in chapter 3 and chapter 4. It is argued that weekly data exhibits certain econometric advantage by providing a larger sample though it might exhibit more noise on the other hand. We therefore examine the tests using both monthly and weekly data. In addition to the monthly data of 40 unit trusts we collected in chapter 1, we collected the weekly data for those unit trusts.

The results of the sub-sample tests based on the SDF primitive efficient models are reported in table 5-3 to table 5-5.

[48] We apply UNFIXED SDF primitive efficient model.

Table 5-3 panel A-D reports the average alphas of each type of funds over sub-sample periods. The sample period for monthly data is from February 1975 to October 2003, with totally 343 observations, which can be divided into ten 3-year sub-samples (Panel A), six 5-year sub-samples (Panel B), four 8-year sub-samples (Panel C) and three 10-year sub-samples (Panel D). The monthly data provides 60 observations for each 5-year sub-sample and 96 observations for each 8-year sub-sample. The weekly data are from Jan 1975 to Dec 2003, with totally 1515 observations, which provides 156 observations for each 3-year sub-sample, 260 observations for each 5-year sub-sample and 418 observations for each 8-year sub-sample. The alphas which are significantly different from zero are shown in bold, while the positive alphas are shown in italic.

Based on the alphas measured using the SDF primitive efficient models, we initially examine if different type of funds perform differently (over 1975 to 2003) using 3-year samples, we then rank the funds based on the magnitude of the alphas for each sub-sample period. We report the former test in table 5-4 and the latter in the table 5-5a.

In table 5-4, based on the alphas for each type of fund, derived from 3-year sub-samples[49], we report the descriptive statistics of alphas, and also paired-samples *t*-tests and one-way ANOVA procedure to examine the differences of the alphas. Paired sample *t*-tests show that the differences among three paired alphas are not significant. (t = 1.592, -0.226, -1.209 for growth-income, growth-general and income-general respectively).

[49] The number of alphas for each fund is 10 as we use 3-year sub-sample covering 1975 to 2003.

We also carry out an ANOVA test for the hypothesis that several means of the alphas from different type of the funds are equal. The - statistic is calculated such as:

$F = \frac{SSR/df_1}{SSE/df_2} = \frac{MSR}{MSE}$ where *df* refers to the corresponding degree of freedom. *MSR* is the mean square errors between the alphas; *MSE* is the mean square errors within alphas. The *F*-statistic is 0.17 with *p*-value of 0.845, indicating insignificant differences among alphas. This result is also confirmed by robust tests including the Welch and Brown-Forsythe statistic[50].

In a word, the alpha comparison tests show that over the long run, neither type of the funds could outperform consistently compared to their counterparts. i.e. style consistency does not prevail.

On the other hand, it does not mean for each sub-sample period, all funds perform equivalently well, we therefore pay closer attention to the details of performance based on the ranking test and the results are reported in table 5-5.

Based on table 5-3 and table 5-5, we discover that income funds provide significant underperformance over 1980 to 1985 while growth funds have significantly outperformed during 1995 to 2000. Most of the alphas are not significantly different from zero, which suggests there is little evidence of non-neutral performance. Among them, 72% of the unit trusts generated a negative alpha although not significant, only

[50] The *F*-statistic assumes relative homogeneity of variances of the variables. The Welch and Brown-Forsythe statistics are more robust tests of equality of means when the assumption of equal variances does not hold.

Income and general funds over 1985 to 1990, Growth and general funds over 1995 to 2000, income funds over 2000 to 2004 have shown superior performance to their benchmark counterparts.
Both monthly and weekly tests reveal a similar performance pattern. It is interesting to see that in both tests, income and general funds outperformed (though not significant) during 1985 to 1990 and growth funds outperformed during 1995 to 2000. During 2000 to 2004, although income funds also generated a negative alpha by using monthly data, it is still shown that income funds do perform superior to growth funds and general funds during such a period.

The results then display apparently that growth funds perform much better than income funds during economic growth and bull market period while income funds perform better than growth fund in the economic downturn. This performance pattern is consistent with the one we detected from our risk-unadjusted measure.

In more detail, during the period from 1975 to 1980, general funds perform the best, followed by growth funds, then income funds. From Figure 5-2, the FTSE all share Index has been flat during this period with a slight upward trend, we might regard this period as a moderate bull market. Another particular period is from 1985 to 1990, in which income funds are verified to generate the best risk-adjusted returns while growth funds are the second. This might capture the effects of the so-called black Monday in stock market history, that is, October 19, 1987, when the Dow Jones Industrial Average plunged a record 508.32 points, or 22.6%, to 1738.34 while the FTSE all-share index slipped to 951.95 from 1072.40. A record was reached on Nov

10 of a historical low 784.81. The period from 1990 to 2000 can be taken as a strong bull market, during which the growth funds were the best performers.

TABLE 5-3. THE ROBUST TESTS USING THE SDF PRIMITIVE EFFICIENT MODELS: AVERAGE ALPHA OF EACH FUND GROUP USING DIFFERENT SUB-SAMPLES

Average alphas are reported based on the SDF primitive efficient model. The results for 3-year, 5-year, 8-year, 10-year sub-samples are reported. *t*-statistic and *p*-value are presented to test if alpha is significantly different from zero. The figures in bold are those perform significantly and the figures in italic are positive.

PANEL A. ALPHAS BASED ON 3-YEAR SAMPLES

WEEKLY DATA (*t*-statistics is reported in the bracket)

	75-78	**78-81**	**81-84**	**84-87**	**87-90**
GROWTH	-0.00287	-0.00047	-0.00028	*0.000211*	-0.000084
	(-1.59308)	(-0.48901)	(-0.44372)	(0.292866)	(-0.18465)
INCOME	-0.00316	**-0.00073**	-0.00055	*0.000152*	*0.0000163*
	(-1.35312)	(-0.68631)	(-0.58712)	(0.20871)	(0.025166)
GENERAL	-0.00155	-0.00117	-0.00025	*0.000391*	-0.000013
	(-1.01391)	(-1.6236)	(-0.4134)	(0.571113)	(-0.02561)
	90-93	**93-96**	**96-99**	**99-02**	**02-03**
GROWTH	*0.0000164*	*0.000181*	*0.00067*	-0.0006	-0.00054
	(0.038576)	(0.494875)	(2.277163)	(-2.01712)	(-0.96439)
INCOME	-0.00094	*0.000248*	*0.000272*	-0.00056	-0.00025
	(-1.83093)	(0.471246)	(0.606782)	(-1.13393)	(-0.42404)
GENERAL	-0.00037	*0.0000776*	*0.000184*	**-0.0005**	-0.00013
	(-1.03197)	(0.251158)	(0.705741)	(-1.87525)	(-0.27618)

PANEL B. ALPHAS BASED ON 5-YEAR SAMPLES

WITH THE SAMPLE PERIOD INCLUDING:

1:1975:01—1979:12; 2:1980:01—1984:12; 3:1985:01—1989:12; 4:1990:01—1994:12; 5:1995:01—1999:12; 6:2000:01---2003:10 (2000:01—2003:12 for weekly data

WEEKLY DATA

	Alpha 1	***t***	***p*-value**	**Alpha 2**	***t***	***p*-value**
GROWTH	-0.0015*	-1.19665	0.231442	-0.00033	-0.601	0.547402
INCOME	-0.00155	-0.96709	0.333497	**-0.0012**	-1.754	0.079407
GENERAL	-0.00091	-0.87598	0.38104	-0.00055	-1.053	0.292085
	Alpha 3	***t***	***p*-value**	**Alpha 4**	***t***	***p*-value**
GROWTH	-0.00011	-0.25172	0.801261	-0.00014	-0.472	0.636851
INCOME	*0.00026*	0.504108	0.614185	-0.0005	-1.079	0.280504
GENERAL	*0.000113*	0.243761	0.807416	-0.00041	-1.523	0.127752
	Alpha 5	***t***	***p*-value**	**Alpha 6**	***t***	***p*-value**
GROWTH	***0.00042***	1.713011	0.086711	-0.00029	-1.035	0.300492
INCOME	-0.0002	-0.50198	0.61568	*0.000245*	0.703	0.481874
GENERAL	*0.000125*	0.636289	0.524588	-0.00016	-0.623	0.533017

*Those in bold are significant at 0.1; those italic are positive alphas

MONTHLY DATA

	Alpha 1	***t***	***p*-value**	**Alpha 2**	***t***	***p*-value**
GROWTH	**-0.00694**	-1.65179	0.098577	-0.00096	-0.555	0.578589
INCOME	**-0.0099**	-2.44642	0.014428	-0.00202	-1.255	0.209321
GENERAL	**-0.00636**	-1.9578	0.050254	-0.001	-0.907	0.363952
	Alpha 3	***t***	***p*-value**	**Alpha 4**	***t***	***p*-value**
GROWTH	-0.00021	-0.18902	0.85008	-0.00046	-0.510	0.609988
INCOME	*0.000961*	0.558569	0.576456	-0.00142	-1.483	0.137902
GENERAL	*0.000153*	0.124248	0.901119	**-0.00194**	-2.464	0.013727
	Alpha 5	***t***	***p*-value**	**Alpha 6**	***t***	***p*-value**
GROWTH	*0.0000254*	0.028473	0.977285	**-0.00277**	-2.434	0.014917
INCOME	**-0.00226**	-2.32754	0.019937	**-0.00191**	-1.860	0.062808
GENERAL	-0.00093	-1.24241	0.214084	**-0.00349**	-2.904	0.003682

PANEL C. ALPHAS BASED ON 8 –YEAR SAMPLES, INCLUDING THE FOLLOWING SAMPLE PERIODS:

1.1975.01--- 1982.12; 2.1983.01---1990.12; 3.1991.01—1998.12; 4.1999.01—2003.10

WEEKLY DATA

	Alpha 1	*t*	*p*-value	Alpha 2	*t*	*p*-value
GROWTH	-0.00135	-1.45905	0.144551	-0.000031	-0.09054	0.927856
INCOME	-0.00179	-1.49578	0.134712	-0.00016	-0.36936	0.71186
GENERAL	-0.00099	-1.294	0.195665	-0.000099	-0.0278	0.977825
	Alpha 3	***t***	***p*-value**	**Alpha 4**	***t***	***p*-value**
GROWTH	0.00031	1.359886	0.173866	-0.0005	-1.88666	0.059207
INCOME	-0.00014	-0.38656	0.699078	-0.00029	-0.80031	0.423533
GENERAL	-0.000012	-0.05841	0.953422	-0.00033	-1.47008	0.141541

MONTHLY DATA

	Alpha 1	*t*	*p*-value	Alpha 2	*t*	*p*-value
GROWTH	-0.00453	-1.48451	0.137673	-0.00058	-0.5592	0.576023
INCOME	-0.00701	-2.22295	0.026219	-0.00046	-0.31281	0.754424
GENERAL	-0.00449	-1.99242	0.046325	-0.0007	-0.71601	0.473984
	Alpha 3	***t***	***p*-value**	**Alpha 4**	***t***	***p*-value**
GROWTH	0.000255	0.363902	0.715931	-0.00215	-2.20794	0.027249
INCOME	-0.00186	-1.99534	0.046006	-0.00188	-1.9422	0.052113
GENERAL	-0.00119	-1.86142	0.062685	-0.00278	-2.72855	0.006361

PANEL D. ALPHAS BASED ON 10–YEAR SAMPLES

MONTHLY DATA

75-85	Alpha1	*t*	*p*-value
GROWTH	-0.00376	-1.54315	0.122796
INCOME	-0.00528	-2.07183	0.038281
GENERAL	-0.00314	-1.69319	0.090419
85-95			
GROWTH	-0.00015	-0.23264	0.81604
INCOME	-0.00006	-0.0612	0.951197
GENERAL	-0.00072	-0.92881	0.352986
95-2003			
GROWTH	-0.00067	-0.8934	0.371641
INCOME	-0.00076	-0.76083	0.44676
GENERAL	-0.00122	-1.7959	0.07251

TABLE 5-4 COMPARISON OF ALPHAS FOR EACH TYPE OF FUNDS BASED ON THE SDF PRIMITIVE EFFICIENT MODEL

We examine if Growth, Income, General funds performance differently using 3-year samples (over the period from 1975 to 2003) based on the SDF primitive efficient models.

THE DESCRIPTIVE STATISTICS OF ALPHAS

	N	Mean	Std. Deviation	Std. Error	Minimum	Maximum
GROWTH	10	-0.000375	0.0009595	0.0003034	-0.00287	0.00067
INCOME	10	-0.000550	0.0010117	0.0003199	-0.00316	0.00027
GENERAL	10	-0.000335	0.0006081	0.0001923	-0.00155	0.00039
TOTAL	30	-0.000420	0.0008527	0.0001557	-0.00316	0.00067

N: number of observations

PAIRED SAMPLES TEST OF ALPHAS BASED ON 3-YEAR SAMPLES

	Mean	Std. Deviation	*T*	*p-value (2-tailed)*
GROWTH - INCOME	0.000175	0.0003480	1.592	0.146
GROWTH - GENERAL	-0.00004	0.0005596	-0.226	0.827
INCOME - GENERAL	-0.00022	0.0005626	-1.209	0.257

ANOVA OF ALPHAS

	Sum of Squares	*df*	Mean Square	*F*	p-value
Between Groups	0.000	2	0.000	0.170	0.845
Within Groups	0.000	27	0.000		
Total	0.000	29			

ROBUST TESTS

	Statistic	*df1*	*df2*	*p-value*
Welch	0.161	2	16.953	0.853
Brown-Forsythe	0.170	2	23.716	0.845

a Asymptotically *F* distributed.

TABLE 5-5. SUMMARY OF STYLE ROTATION BASED ON THE SDF PRIMITIVE EFFICIENT MODEL – THE RANK OF EACH GROUP PERFORMANCE

PANEL A. 3-YEAR

YEAR		75-78	78-81	81-84	84-87	87-90
GROWTH	UNFIXED	2	1	2	2	3
	CONVARY	2	1	1	2	3
INCOME	UNFIXED	3	2	3	3	1
	CONVARY	1	3	3	3	1
GENERAL	UNFIXED	1	3	1	1	2
	CONVARY	3	2	2	1	2
YEAR		90-93	93-96	96-99	99-02	02-03
GROWTH	UNFIXED	1	2	1	3	3
	CONVARY	1	1	1	2	3
INCOME	UNFIXED	3	1	2	2	2
	CONVARY	2	2	2	2	1
GENERAL	UNFIXED	2	3	3	1	1
	CONVARY	3	3	3	1	2

PANEL B. 5-YEAR

YEAR		75-80	80-85	85-90	90-95	95-00	00-03
GROWTH	UNFIXED	2	1	3	1	1	2
	CONVARY	1	1	1	1	1	2
INCOME	UNFIXED	3	3	1	2	3	1
	CONVARY	3	3	3	2	3	1
GENERAL	UNFIXED	1	2	2	3	2	3
	CONVARY	2	2	2	3	2	3

PANEL C. 8-YEAR

YEAR		75--83	83--91	91--99	99--03
GROWTH	UNFIXED	2	2	1	2
	CONVARY	1	1	1	1
INCOME	UNFIXED	3	1	3	1
	CONVARY	3	3	3	2
GENERAL	UNFIXED	1	3	2	3
	CONVARY	2	2	2	3

PANEL D. 10-YEAR

YEAR		75--85	85--95	95--03
GROWTH	UNFIXED	2	2	1
	CONVARY	1	1	1
INCOME	UNFIXED	3	1	2
	UNFIXED	3	2	2
GENERAL	CONVARY	1	3	3
	UNFIXED	2	3	3

A more recent story of market downturn is from 2000 to April 2003, when the market disappointment made the market participants move from a massive appetite for risk to massive aversion to risk. With the catastrophic dot-com bubble bursting, economic slowdown, September 11th, accounting scandals and war on Iraq, the stock markets have been all over the charts with records being set for largest one day drops. Unsurprisingly, we found out that during this period, income funds were ranked the first while growth funds perform a little better than the general funds. During April 2003 to March 2004, there is significant market recovery due to better economic growth, on-track blue-chip companies' performance and investors' optimistic sentiment, while it is hard to use such a short sample period to verify our assumption.

The pattern of funds performance during different market conditions has been shown in appendix Figure 5-2.

Figure 5-2 illustrates how performance pattern of different styles of funds follows the FTSE all-share index and FTSE all-share dividend yield index. During the bull market, peoples' sentiment is high and the appetite for risk is strong. Investors will prefer those popular shares with high P/E ratio and bright prospective, such as high technology shares. Although the future performance of such companies have been shown in their current prices, the investors still believe the price can go even higher and they can still make big profits as long as they can liquidate their holdings before the bubble bursts. If everyone thinks in the same way, demand can be very high, with rocketing growth share prices. In the bear market, the investors are much more cautious and the low prices make those blue chips' dividend yield even higher. The investors could prefer stable and safe dividend payments rather than capital gains under such circumstances. Blue chips therefore will be much popular than high-tech shares when the market is on the downturn.

In summary, the tests (risk-unadjusted return tests and the robust tests using the SDF models) do not provide any evidence for the long-run outperformance of certain fund style; On the contrary, they all suggest that the value spread exhibits significant short-term (directional) variation according to the market/macroeconomic conditions. Since the performance pattern indicates that the growth funds are more positively affected by improving economic conditions, but more vulnerable than value stocks during economic downturns, it implies that an investor might substantially increase his/her overall returns by increasing portfolio weights in the income funds in the bear markets and investing more in the growth funds in the bull markets, depending on his/her correct prediction of future market trends.

3. Construction of style-rotation strategies

Based on our empirical tests (both raw and risk-adjusted returns) in the last section, we verified that the different types of funds performed differently over each sample period though in the long run, they did not have significantly different performances. The next question is: can we apply the trading rules based on the style performance pattern that we have discovered?

The essence of constructing the profitable trading strategies is one's forecasting ability. Within the context of style investing, two broad types of forecasting can be useful. One is to predict the timing of the bull/bear markets, given the hypothesis that the style performance pattern we have detected is true; another is to directly predict the value spread. Ideally, the rotation strategies based on the latter forecasting should also indicate an implication of the same style performance pattern.

In this section, we tend to construct style-rotation strategies based on those two broad forecasting methods, including a simple *ex-post* rotation strategy by switching portfolio holdings between the bull and the bear markets and a discussion of bull/bear market forecasting (3.1); The sensitivity tests (OLS and Logit) to detect the relationship between the value spread and various macroeconomic variables (3.2); and finally, we apply the logit approach to construct the rotation strategies depending on its prediction of value spread (3.3).

3.1 Style-rotation strategies based on regime switching between the bull and the bear markets

Our empirical results suggest growth and value funds perform differently in bull and bear markets within the framework of the SDF models. It implies that by simply changing the holding of the assets characteristics based on this pattern may enhance the portfolio returns. However, it is easily understood that there is no simple trading rule that can generate persistent profits. Suppose such a golden rule did exist and it must be widely applied by market participants and will result in the change of demand and supply of the underlying assets. Hence the profit can be soon traded away. In this section, we provide a further investigation to see if this particular style-rotation strategy is feasible?

The definition of bear and bull markets is quite diverse in the existing literature. The general definition is that a '*bull (bear) market corresponds to periods of generally* *increasing (decreasing) market prices*' (Chauvet and Potter, 2000). A more restrictive definition is provided by Lunde and Timmerman (2000), where the stock market is considered to switch "*from a bull to a bear state if stock prices have declined by a certain percentage since the previous (local) peak within that bull state*". In another word, the switch presupposes that cumulated changes exceed a certain threshold. Following Chauvet and Potter's definition, we implement a series of *ex-post* tests to construct different portfolio holding strategies during a hypothetical period of a combination of both bull markets and bear markets. Firstly, we use the *MSN MONEY* stock screener to confine the shares to have certain characteristics. Growth shares are defined as those that generate high annual income, high 5-year revenue growth and high P/E ratios, and mainly from high growth sectors, including IT, Electrical, Electronic, Machine and Equipment. Eight of them will then be randomly selected. Income stocks are randomly selected from the FTSE 100, mainly high street brand names and blue chips with high dividend yield.

Secondly, the data for each selected share are collected from Datastream, from which monthly-adjusted prices (the closing price for the requested month adjusted for all applicable splits and dividend distributions) have been obtained for 8 income shares and 8 growth shares during 1998 to 2004.

We implement two tests. One assumes our portfolio holding through September 2002 to February 2003 (bull) and June 2003 to December 2003 (bear) and the results are shown in table 5-6a panel A. Another assumes the portfolios are constructed during August 1998 to July 1999 (bull) and January 2001 to December 2001 (bear) and it is reported in table 5-6a panel B. For each test separately, we apply four style-rotation strategies and form four portfolios, then simply compare the risk-unadjusted net returns of those four portfolios. The first two strategies are either holding only income or only growth stocks through out both bear and bull markets; the third strategy is to hold income stocks during bull markets then switch to hold growth stocks instead during bear markets; the fourth strategy is to hold growth shares in bear markets and income stocks in bull markets.

The first test (Panel A) shows that growth shares perform much better than income shares through both bull and bear markets in that it provides higher positive returns in bull and less negative returns in bear markets. Therefore, the portfolio keeping growth shares during both markets is ranked first, followed by portfolio 4 while portfolio 1 performs the worst.

In Panel B, the second test displays that net average monthly return for growth shares is 2.2649% while income shares provide only 0.291% of monthly return. In the bear markets, growth shares make an average loss of 1.8% per month while income shares still provide a return of 0.6343%. Therefore, portfolio 4 generates the highest monthly return of 1.4496%, followed by portfolio 1 of 0.4626%, portfolio 3, whose holding is opposite to our style rotation, performs the worst.

TABLE 5-6A. THE PERFORMANCE OF TWO EXPERIMENTAL STYLE-ROTATION STRATEGIES

Each portfolio contains eight shares. Portfolio 1 keeps holding the same eight income shares through both bull and bear market. Portfolio 2 keeps holding eight growth shares through two periods. Portfolio 3 holds eight income shares during bull market and switch to hold eight growth shares during the bear market. Portfolio 4 holds eight growth shares in the bull market and income shares in the bear market.

PART A: Experiment 1 Bull market: June 2003 — Dec 2003

Bear market: Sept 2002 — Feb 2003

Net Average Returns (09.02-12.03)

	Growth	Income
Bull	0.058072	0.002907
Bear	-0.01748	-0.03234

Strategies:

	Bull	Bear	Net Returns	Rank
Portfolio1	INCOME	INCOME	-0.01472	4
Portfolio2	GROWTH	GROWTH	0.020298	1
Portfolio3	INCOME	GROWTH	-0.00728	3
Portfolio4	GROWTH	INCOME	0.012866	2

PART B: Experiment 2: Bull market: Aug 1998—July 1999

Bear market: Jan 2001—Dec 2001

Net Average Returns (08.98-12.01)

	Growth	Income
Bull	0.022649	0.00291
Bear	-0.0182	0.006343

Strategies:

	Bull	Bear	Net Returns	Rank
Portfolio1	INCOME	INCOME	0.004626	2
Portfolio2	GROWTH	GROWTH	0.002225	3
Portfolio3	INCOME	GROWTH	-0.00764	4
Portfolio4	GROWTH	INCOME	0.014496	1

Though the results from the second test suggest that the strategy that investing in growth funds in bull market and then switch to value funds in bear markets does outperform, it is very likely that this strategy would not work in practice. It can be extremely difficult to predict the proper timing of this trading policy due to the following reasons:

- There is always a time lag between recognition of the need for the trade and the actual implementation.
- There is always a time lag between implementation of new strategies and the impact of changes.
- Trading costs might take away large bulk of the profits.
- There is a risk of reinvestment during the gap period between bull and bear markets.

Therefore, although there is an opportunity, it could be impractical to exploit or too costly to implement. Other than the above implementation problems, another important issue remains. The essence of supposedly feasible trading strategies is the ability to predict style returns ahead of time. The concern therefore is, '*The apparent predictability gap might be due to substantial biases in many reported findings that have been obtained from a setting that benefits too much from ex post knowledge*.' (See Cooper and Gulen (2002)) Indeed, the effective implementation of the style-rotation strategy requires a realistic assessment of the manager's degree of forecasting ability. If markets are correctly forecasting the impact of future economic conditions, then we would imagine that the managers should also, in some sense, see through such factors. Hence the key point is the extent to which economic conditions change in an "unexpected" manner, and the extent to which the investors can forecast these changes before they happen.

There are quite a few studies, focusing on the timing of bull and bear markets. For example Girardin and Liu (2003) determine the actual average duration of the states utilizing the Markov-switching technique (a switch-in-the-mean plus switch-in-the-variance model). More specifically, based on the results of a *MSMH*(3)-*AR*(5) model, they classify Shanghai A-share (China's major stock index) into three states including bear, bull and speculative states.

Pagan and Sossounov (2003) apply an algorithm based on the definition of bull and bear markets to sort a given time series of equity prices (from 1835 to 1997) into periods that bull and bear market characteristics depend upon the Data Generating Process (DGP) for capital gains ($\Delta \ln P_t$). They discover that a pure random walk process provides no worse explanation of bull and bear markets than more complicated DGP models. The two essential elements for the specification of a DGP are the mean (μ) and standard error of capital gains (σ). It is likely that the probability will rise with the mean and decline with standard error. They also argue that '*regardless of the model for* $\Delta \ln P_t$*, it is clear that any theoretical model which claims to provide an explanation of historical bull and bear markets will have to be capable of reproducing the historical values of* μ *and* σ*. Since is related to the equity premium, one must therefore be able to replicate that as well as the volatility of capital gains'*.

As discussed above, there is no consensus among the researchers about the best possible method to forecast the bull/bear market. In addition, our purpose is to construct a practical rotation strategy to enhance style investing performance rather than forecasting the bull/bear markets. As a result, we are more interested at how the rotation strategy can be constructed based on various factors. We investigate the determinants of value spreads in the following sections.

3.2 The determinants of value spread

Following the performance pattern we have detected, we want to explore the potential profitable style-rotation strategies by switching the portfolio holding styles over time. In the process of style-rotation strategy construction, researchers often choose among factors, which are perceived to have explanatory power for the value spread in general. It is strongly believed that the variation pattern of value spread is related to the changes in the economic climate and equity market conditions.

Extensive research has focused on how to build a model to predict the relevant style spreads over time and then construct a style-rotation strategy that adjusts portfolio weights according to the prediction of relative style performance. The examples include Shumaker (1999), who documents the forecasting ability of variables, such as the yield spread (term structure), real bond yield and earnings yield gap. Levis and Liodakis (1999) report the significant explanatory power of the inflation rate, and Sorensen and Lazzara (1995) identify the industrial production factor.

3.21 The equity return attribution model

Our ultimate goal is to identify the determinants of value spread. Since the selection of variables can be *ad hoc*, we tend to choose the independent variables following a basic equity attribution model. Our search of those information signals starts from a simple equity return attribution model, as introduced by Grinold and Kroner (2002). Generally, equity returns can be decomposed into three components: dividend yield, earning growth and change in P/E such as:

$$R_{equity} = (\frac{D}{P} - \Delta S) + (i + g) + \Delta(\frac{P}{E}) \qquad (5.3)$$

where $\frac{D}{P} - \Delta S$ is the income component (with ΔS as the change in the number of outstanding shares; (i+g) is the nominal earnings growth component, with an inflation term i and a real growth in earnings term and $\Delta(\frac{P}{E})$ is the percentage change in the *P/E* ratio.

Based on Equation (5.3), multivariate OLS regression is developed for each type of unit trust *j* (income & growth fund), with the risk-unadjusted unit trusts returns at time *t*, $R_{j,t}$ as the dependent variable, such as:

$$R_{j,t} = \alpha + \beta_1 DY_{t-1} + \beta_2 INFLA_{t-1} + \beta_3 TERM_{t-1} + \beta_4 WAGE_{t-1} + \beta_5 INDPROD_{t-1} + \varepsilon_t \qquad (5.4)$$

where the independent variables with potential predictive ability at time *t-1* are:

- *DY*: the lagged dividend yield on the market index FTSE 100.
- *INFLA*: UK Inflation Rate Index
- *TERM:* a term structure variable defined as the monthly yield difference between the three-month and one-month Treasury bill
- *WAGE:* Labour Income Growth Rate. We apply Average UK Earnings Index (whole economy) as the proxy for labour income and Wage is the growth rate of this Index.
- *INDPROD:* Industrial production growth, We measure the growth rate of the UK Industrial Production Index

The monthly data of all those variables are collected from Thomson DATASTREAM. A justification for the variables included in these regression equations follows.

DY is corresponding to the first component of equation (5.3), and it has long been widely applied to forecast equity returns and is known to vary in response to changes in business conditions, examples like Fama and French (1988) and Campbell and Shiller (2002). The information of change in the number of outstanding shares is not available, we regard it as negligible.

INFLA, TERM, WAGE are the variables applied to mimic the characteristics of the second component. Recent evidence has shown that the nominal stock returns are significantly related to inflation in the US (see Canova (2000)). And it is well known that the term structure of interest rates, which characterizes the movements of risk-free bond yields, is determined by various state variables summarized in the pricing kernel. Therefore, by observing the term structure of interest rates, we learn information on the shares that are priced using the pricing kernel. It is also widely recognized the importance of including human wealth returns as part of the market return (e.g. Shiller (1993), Campbell (1996), and Jagannathan & Wang (1996)).

The last term in equation (5.3) is the re-pricing component of the equity return, which is the most uncertain. The surprise productivity gains favour a higher long-term P/E. Other factors influencing a long-term P/E include inflation, financial innovation, war and environmental costs. The results are reported in table 5-6. Both the growth funds and the income funds produce similar results. The coefficients of *DY, TERM*, and *WAGE* are significant. Among them, for both growth and income funds, the coefficients of *DY* are positive, and the coefficients of *TERM* and *WAGE* are negative. The coefficients of *INFLA* and *INDPROD* are not significant. Inflation has mixed effects, related with the second and the third term of equation (5.3).

It also shows that the *R*-squared is relatively small, suggesting that our factors only partially explain the equity returns due to the complexity and unpredictability of the market.

TABLE 5-6 THE DETERMINANTS OF FUNDS RETURNS

OLS regression is run to investigate the determinants of funds returns, particularly, for growth funds and income funds.

	Multivariate OLS **Growth funds**	**Income Funds**
	R-squared :0.061661 *F*:4.42904 (0.001) Log likelihood: 569.954	*R*-squared :0.051745 *F*: 3.65611 (0.001) Log likelihood: 584.912
C	1.45901	1.43767
	(4.47930)	(4.95736)
DY(-1)	0.008829	0.006512
	(3.49885)	(2.78214)
INFLA(-1)	-0.011565	-0.010625
	(-0.537763)	(-0.553201)
TERM(-1)	-0.017964	-0.022142
	(-1.75196)	(-2.44980)
WAGE(-1)	-6.19805	-5.88325
	(-1.91712)	(-2.03166)
INDPROD(-1)	0.146862	0.144580
	(0.192044)	(0.834457)

3.22 The value spread attribution

After the tests of equity return attribution, we apply similar sensitivity tests to investigate the determinants of value spread, i.e. the differenced returns of value funds and income funds.

Besides OLS, logit models have also been widely applied by researchers[51] to predict the relevant style spreads over time. It is argued that the sign of style spreads is related to a number of economic and market characteristics, thus forecasting the sign of the style spread may be sufficient for a successful style-rotation strategy. Instead of investigating the magnitude of the spread, a logit regression helps to predict the sign of the value spread by classifying each month as 1 or 0 based on the sign of the style spread, i.e. if in a particular month value stocks perform better than growth stocks, we can classify this month as 1; otherwise we set it to 0.

Formally, let $Growth_t$ be the average of time-t returns on growth funds and let $Value_t$ be the average of time-t returns on income funds, we have

$$y^*_t = Value_{t+1} - Growth_{t+1} = a_1 x_t + u_i \qquad (5.5)$$

What we are interested is not the exact magnitude of y^*_t, but the probability of the style spread to be positive, which can be defined as a following dummy variable:

$y_i = 1$ if $y^*_t > 0$ or

0 otherwise

The probability for the value spread to be positive is:

$$\begin{aligned} \hat{p}_t &= prob(y_i = 1) = prob(Value_{t+1} > Growth_{t+1} | x_t) = prob(u_i > -a_1 x_t) \\ &= 1 - F(-a_1 x_t) = F(a_1 x_t) \end{aligned} \qquad (5.6)$$

It is necessary to specify the probability distribution of u_i. Since it might reasonably be assumed that there are many independent factors which might combine additionally to this random error term, the central limit theorem can be used to justify

[51] Examples include Levis and Liodakis (1999) and Wang (2005). Levis and Liodakis (1999) were those firstly used the Logit model to construct style-rotation strategies in the UK market and they discover that during the thirty-year period 1968 through 1997, value and small-cap stocks in the U.K. outperformed their growth and large-cap counterparts. Wang (2005) also applied a logit approach, in which Fama-French three factors are considered.

the assumption that u_i is normally distributed. This leads to a probit model. However, due to its computational advantages and since the differences between the models are slight, we assume u_i follows the logistic distribution. As Greene (1996) argues, '*Other distributions have been suggested, but in econometric applications the probit and logit models have been used almost exclusively. It is difficult to justify the choice of one distribution or another on theoretical grounds… in most applications it seems not to make much difference.*'

If the cumulative distribution of u_i is logistic, we therefore have the logit model as:

$$\hat{p}_t = \frac{\exp(a_1 x_t)}{1+\exp(a_1 x_t)} \text{ or}$$

$$\ln \frac{\hat{p}_t}{1-\hat{p}_t} = a_1 x_t \qquad (5.7)$$

where $x_t = (1\ F_t)$ with F_t as a vector of explanatory variables, a_1 is $1 * N$ parameter vectors. (N equals one plus the number of the variables).

The model is estimated to maximize the following likelihood function as:

$$L = \prod_{yi=1} P_t \prod_{yi=0} (1-P_t) \qquad (5.8)$$

We examine the 'goodness-of-fit' of our models by the following measures: firstly, we report the 'The Fraction of Correct Predictions', which is a statistic calculated by identifying for each sampled decision maker the alternative with the highest probability, based on the estimated model, and determining whether or not this was the alternative that the decision maker actually chooses. The percentage of sampled decision makers for which the highest probability alternative and the chosen alternative are the same is called the percent correctly predicted.

We then report the value of 'the log-likelihood function' at the estimated parameters as *LLu* (without any restrictions at the parameters). Since the perfect prediction makes the likelihood function one when the probability of observing the choices correctly is one, therefore, the closer the log-likelihood ratio *LLu* is to zero, the better is the goodness-of-fit of the model.

The third measure is R^2. When the explained variable y takes on only two values, there is a problem with the use of conventional R^2-type measures.[52] Several R^2-type measures have been suggested for models with qualitative dependent variables. The most popular one is called McFadden's R^2 such as:

$$R^2 = 1 - \frac{LL_U}{LL_R} \qquad (5.9)$$

where *LLu* and *LLr* are the unrestricted and restricted log-likelihood ratios respectively. Finally, we also implement the hypothesis tests to measure the significance of the coefficients by using *LR* statistic such as

$$LR = -2(LLr - LLu) \qquad (5.10)$$

where *LLr* and *LLu* are the maximums of the likelihood function where, respectively, the tested parameters are constrained to equal zero and where there are no restrictions on all the parameters. *LR* is distributed as χ^2_n, where n is the number of parameters estimated in the unrestricted model.

[52] See Maddala, Limited-Dependent, pp.37-41.

The choices of forecasting factors mainly follow the test of equity attribution, with independent variables such as *DY, INFLA,* R_f*, TERM, WAGE* and *INDPROD*. In addition, since the style performance pattern identified by risk-unadjusted return and then reinforced by the SDF risk-adjusted approach is: growth funds tend to outperform in the bull market while income funds are doing better in the bear market. We also include market returns R_m as one of the independent variables.

Initially, we examine the effects of a single variable on value spread by running univariate OLS and logit models. We include both current value (at time *t*) and one-lagged value (at time *t-1*) for each independent variable. The intention is to see if the level or one-lagged values of independent variables have higher power to explain the value spread and to see if the one-lagged value can be employed for forecasting purpose. The results are reported in table 5-7 Panel A.

For univariate OLS, it reports the coefficient value, followed by *t*-statistics and *p*-value in the second and third rows, and *R*-squared. For the logit model, the fraction of correct predictions and goodness-of-fit statistics, including likelihood ratio and *R*-squared are also reported. The results for lagged variables are shown in italic and the significant coefficients are shown in bold. For variables with current values, R_m, R_f are significant based on OLS, among lagged independent variables, only $DY_{(t-1)}$, $R_{f(t-1)}$ are significant. Besides those significant variables, *INFLA, WAGE* and $WAGE_{(t-1)}$ are also significant using a logit regression. Particularly, the market return proxy R_m is negatively correlated with the value spread (-0.070356 using the OLS, and -8.50655 with the Logit model), implying that the higher market return (the bull market) can lead to lower returns of the income funds, but higher returns of the growth funds, which is consistent with the performance pattern we discovered in previous sections. What is more, the regression with R_m as an independent variable has the highest explanatory power based on 'Goodness-of-Fit' tests. (R-squared is 0.081142 with the OLS test, and for the logit model, the likelihood ratio is the smallest, at -230.045 and the R-squared is the highest, at 0.042123). However, R_m is inappropriate to be used as a forecasting variable, since its lagged level $R_{m(t-1)}$ did not show those characteristics.

More systematically, based on Panel A, we can derive its explanatory results by relating the variables with the implied market economy conditions, given the hypothesis that the performance pattern (i.e. the worse the market condition, the higher the value spread---value funds outperform in the bear market) is true. Shown in Panel B, by only examining the identified significant independent variables, the OLS results reveal a positive correlation of $R_{m(t)}$, $DY_{(t-1)}$, $DY_{(t)}$, $R_{f(t-1)}$, $R_{f(t)}$, $WAGE_{(t)}$ and the prevailing market conditions, given the hypothesis is true. In addition, the logit results demonstrate $INFLA_{(t-1)}$ is negatively correlated with the market condition and $WAGE_{(t-1)}$ is negatively correlated with the market condition. It is apparently that our hypothesis is not rejected since one would expect lower dividend yield, lower risk-free interest rate, higher inflation rate, lower average wage at time t will lead to worse equity market performance at time t. Also it is apparent that the enhanced market return, dividend yield, risk-free interest rate, average wage at time t will result in an improved market performance. We therefore confirm that our OLS and Logit approaches produce consistent results as the style performance pattern we have detected.

Based on the univariate results, we tend to pool the significant lagged variables together and run multivariate regressions to explore the explanatory power by increasing the number of dependent variables. Table 5-7 Panel C and Panel D report the results of multivariate OLS and logit estimation respectively.

We have tested models with all kinds of combinations of those variables, but only report those with better model statistics compared to other candidates. The only significant variable in both model 1 and model 2 of panel C is $DY_{(t-1)}$ and model 1 has higher *R*-squared, showing that model 2 has made no improvement by adding more variables, which also are insignificant.

Similarly, Panel D displays that both *INFLA* is significant in model 2. *TERM* is significant in model 3. In addition, model 3 exhibits the best model statistics among three models, with the highest *R*-squared (0.048 compared to 0.044 for model 2 and 0.04 for model 1), fraction of correct predictions (61.22% compared to 58.3% for model 2 and 58.89% for model 1) and *LR* (zero slope) ratios (17.05 compared to 15.72 for model 2 and 14.24 for model 1). Further, we can compare the significance of those models by examining the *LR* ratios following equation (5.10), with the null hypothesis as the coefficients of the additional variables are zero, when we compare model 2 with model 1, *LR* is 1.47; *LR* is 1.338 when we compare model 3 with model 2 and it is 2.808 when we compare model 3 with model 1, suggesting the insignificance of those additional variables.

Just as commented by Bauer (2004), '*the interaction with style portfolio returns can be direct, indirect or non causal*.' The small value of *R*-squared (for Multivariate OLS model 1, R^2= 0.038464, model 2 = 0.023313; for logit model 1, R^2 = 0.040468, model 2 = 0.043949, model 3 = 0.048351) suggests that the exact nature of the information contained in these variables spans different areas and is difficult to unravel.

TABLE 5-7. SENSITIVITY TESTS OF RETURN SPREAD (BASED ON RISK-UNADJUSTED RETURNS)
PANEL A. UNIVARIATE OLS RESULTS (1975 – 2003)

	Univariate OLS		Univariate Logit		
Variable	coefficient	R2	Coefficient	Fraction of Correct Predictions	Goodness of Fit
$R_{m(t-1)}$	-0.00782	0.00127	1.28324	0.536443	-237.43
	(-0.514296*)		(0.699871)		0.001523
	[0.607]		[0.484]		
$R_{m(t)}$	**-0.070356**	0.081142	**-8.50655**	**0.58309**	**-230.045**
	(-4.63249)		(-3.69862)		**0.042123**
	[0.000]		[0.000]		
$DY_{(t-1)}$	**-0.0020266**	0.037043	**-0.291002**	**0.565598**	**-232.327**
	(-3.06163)		(-3.20837)		**0.030505**
	[0.002]		0.001		
$DY_{(t)}$	**-0.00150198**	0.0201	**-0.230299**	0.559767	-234.336
	(-2.23097)		-2.55464		0.01941
	[0.026]		0.011		
$INFLA_{(t-1)}$	0.00122667	0.00120764	**1.97051**	0.533528	-235.585
	(0.221873)		1.98547		0.011744
	[0.825]		0.047		
$INFLA_{(t)}$	-0.00769683	0.004758	-0.291645	0.533528	-237.629
	(-1.23427)		-3.12368		0.0002927
	[0.218]		0.755		
$R_{f\,(t-1)}$	**-0.000404621**	0.011066	**-0.066546**	0.530612	-235.558
	(-1.96747)		-2.04662		0.012232
	[0.05]		0.041		
$R_{f\,(t)}$	**-0.000410411**	0.011475	**-0.068159**	0.54519	-235.437
	(-2.0031)		-2.10314		0.012927
	[0.046]		0.035		
$TERM_{(t-1)}$	-0.00376663	0.00537551	-0.102702	0.516035	-237.649
	(-1.06091)		-0.239617		0.0001656
	[0.289]		0.811		
$TERM_{(t)}$	0.000335847	0.000042	0.282255	0.521866	-237.463
	(0.128573)		0.654715		0.001287
	[0.898]		0.513		
$WAGE_{(t-1)}$	-0.066159	0.001829	**-26.3624**	0.510204	-235.629
	(-0.705122)		-1.98909		0.011533
	[0.481]		0.047		
$WAGE_{(t)}$	-1.62877	0.010965	**-21.8797**	0.533528	-236.269
	-1.9827		-1.65787		0.0080816
	0.048		0.097		
$INDPROD_{(t-1)}$	0.00857728	0.00007396	3.95055	0.510204	-237.565
	0.156778		0.474835		0.0006677
	0.876		0.635		
$INDPROD_{(t)}$	-0.071816	0.0051835	-13.6084	0.533528	-236.381
	-1.54627		-1.58324		0.0079160
	0.123		0.113		

* *t*-statistics are calculated based on heteroskedastic-consistent standard errors.

PANEL B. EXPLANATORY RESULTS OF PANEL A

Variables	The OLS/Logit Results (The relation between variables and value spread)					Implied Market Economy Condition At time *t*
	OLS	Sig.	logit	Sig.	Value Spread	
$R_{m(t-1)}$	lower	N	higher	N	higher	worse
$R_{m(t)}$	lower	Y	lower	Y	higher	worse
$DY_{(t-1)}$	lower	Y	lower	Y	higher	worse
$DY_{(t)}$	lower	Y	lower	Y	higher	worse
$INFLA_{(t-1)}$	higher	N	higher	Y	higher	worse
$INFLA_{(t)}$	lower	N	lower	N	higher	worse
$R_{f\,(t-1)}$	lower	Y	lower	Y	higher	worse
$R_{f\,(t)}$	lower	Y	lower	Y	higher	worse
$TERM_{(t-1)}$	lower	N	lower	N	higher	worse
$TERM_{(t)}$	higher	N	higher	N	higher	worse
$WAGE_{(t-1)}$	lower	N	lower	Y	higher	worse
$WAGE_{(t)}$	lower	Y	lower	Y	higher	worse
$INDPROD_{(t-1)}$	higher	N	higher	N	higher	worse
$INDPROD_{(t)}$	lower	n	lower	n	higher	worse

PANEL C. MULTIFACTOR OLS ESTIMATIONS

Model 1			
Variable	Parameter Estimate	*T*	*p*-value
Constant	0.00797387	2.69385	0.007
$DY_{(t-1)}$	-0.00240344	-3.11607	0.005
$R_{f\,(t-1)}$	0.00019988	0.808247	0.42
Model statistics			
F(zero slopes)=6.76040 [0.001]		Schwarz B.I.C.=-995.694	
R-squared=0.038464		log likelihood=1004.44	

Model 2			
Variable	Parameter Estimate	*T*	*p*-value
constant	-0.022053	-0.21186	0.832
$R_{m(t-1)}$	-0.0064771	-0.43807	0.662
$DY_{(t-1)}$	-0.0022353	-2.58195	0.01
$INFLA_{(t-1)}$	0.00089373	0.164529	0.869
$R_{f\,(t-1)}$	0.00009727	0.324749	0.746
$TERM_{(t-1)}$	-0.0035394	-0.90193	0.368
$WAGE_{(t-1)}$	0.035175	0.333842	0.739
$INDPROD_{(t-1)}$	0.00029520	0.005692	0.995
Model statistics			
F(zero slopes)=2.15938 [0.037]		Schwarz B.I.C.=-981.996	
R-squared=0.023313		log likelihood=1005.32	

PANEL D. MULTIFACTOR LOGIT ESTIMATION (FOR THE SAMPLE PERIOD FROM FEB 1975 TO OCT 2003)

Model 1			
Variable	**Coefficient**	**Std. Error**	***z*-Statistic**
C	0.027428	0.170704	0.160678
DY	1.912204	0.503667	3.796565
INFLA	-0.019194	0.019745	-0.972104
Model statistics			
Fraction of correct predictions=0.588921	Unrestricted Log likelihood(LLu1)=-230.555		
R-squared=0.040468	Schwarz B.I.C.=239.312		
LR(zero slopes)=14.2458 [.001]			
Model 2			
Variable	**Coefficient**	**Std. Error**	***z*-Statistic**
C	0.074	0.199	0.372
DY	1.929	0.505	3.818
INFLA	-0.020	0.020	-0.991
WAGE	-0.149	0.327	-0.456
Model statistics			
Fraction of correct predictions=0.58309	Unrestricted log likelihood(*LLu2*) =-229.820		
R-squared=0.043949	*LR2*(compared to model 1)=1.47*		
LR(zero slopes)=15.7160 [.003]	schwarz B.I.C.=244.414		
Model 3			
Variable	**Coefficient**	**Std. Error**	***z*-Statistic**
C	0.07	0.20	0.33
DY	1.93	0.52	3.71
INFLA	-0.02	0.02	-0.96
WAGE	-0.14	0.33	-0.42
TERM	0.03	0.41	0.06
INDPROD	0.01	0.02	0.64
Model statistics			
Fraction of correct predictions=0.612245	Unrestricted log likelihood(*LLu3*) =-229.151 *LR3*(Compared to Model1)=2.808		
R-squared=0.048351	*LR3*(compared to model 2)=1.338		
LR(zero slopes)=17.0536[0.017]	schwarz B.I.C.=244.414		

* *LR2*(compared to model 1)= -2(*LLu1*-*LLu2*)
LR3(Compared to Model1)=-2(*LLu1*-*LLu3*)
LR3(compared to model 2)=-2(*LLu2*-*LLu3*)

3.3 Profitability of style-rotation strategies based on the logit model

Based on the logit model we applied in the last section, we can obtain the estimated probability[53] $\hat{p}_t$, which gives the likelihood of a positive value spread over the next month.

We initially use data from January 1975 to December 1989 to estimate the parameters for logit model 1 to logit model 6 respectively. We then calculate the estimated probability based on equation 5.7. The forecasting period is from January 1990 to October 2003 (with total number of forecasting of 166). We use recursive window to estimate the parameters and calculate the estimated probabilities. For example, the estimated probability for Jan 1990 is calculated based on the results using data from Jan 1975 to Dec 1989; the estimated probability for Feb 1990 is calculated based on the results using data from Jan 1975 to Jan 1990; the estimated probability for March 1990 is calculated based on the results using data from Jan 1975 to Feb 1990.

To construct rotation strategies, we compare $\hat{p}_t$ with a trading threshold P_b.

The rotation strategies basically consist of two strategies based on the estimated logit probabilities.

---The Predicted Winner (the PW portfolio):

The PW strategy refers to a portfolio with the combination of the predicted winners from both value and growth funds. It involves a strategy of holding 100% value funds if $\hat{p}_t \geq P_b$ and switching to growth funds if $\hat{p}_t$ is less than P_b.

---

53 $\hat{p}_t$ is calculated based equation 5.7.

--- The Predicted Loser (the PL portfolio):

The second portfolio refers to a long position in predicted losers, which is the opposite of the PW portfolio. That is, the combination of the predicted loser from both value[54] *(* $\hat{p}_t < P_b$ *) and growth (* $\hat{p}_t > P_b$*) funds.*

A series of logit models with different explanatory variables: denoted as model 1-6, are applied to estimate the probabilities.

Table 5-8 Panel A to Panel C reports the prediction estimation results and profitability of different trading strategies. To examine if a style-rotation strategy can outperform a static strategy, we report the performance of buy & hold strategies (i.e. holding 100% of growth funds and 100% of income funds for the whole sample period) and perfect foresight strategy (a portfolio consists of only the best-performing fund for the month and measure the *ex-post* returns) in panel A.

The results suggest that a buy & hold strategy for income fund has generated a mean return of 5.01% (annualized return based on monthly average return), and a buy and hold strategy for the growth fund has generated a mean return of 5.23% annually. Perfect foresight strategy has undoubtedly outperformed, with a mean return of 10.57%. The number of fund type changes with this strategy is 66 times, which result in a net mean return of 10.01% and a net Sharpe ratio of 0.3662 given transaction cost

---

54 when $\hat{p}_t < 0.5$, that is, there is less than 50 percent of chance for value funds to do better than the growth funds or, more than 50 percent of chance for value funds to do worse than the growth funds.

of 0.05%. It shows that even taking transaction costs into account, the perfect foresight strategy can generate significantly higher net returns than buy & hold strategies. i.e. on a net basis, perfect foresight strategy outperformed buy and hold income strategy by 5% and beats the buy & hold growth strategy by 4.78% annually.

Panel B presents the results of the PW portfolio. For a PW portfolio, we initially set the probability threshold P_b as 0.5. A probability value $\hat{p}_t$ above 0.5 indicates that a month that favours value shares is to occur with a possibility of more than 50%, while a probability value below 0.5 indicates a preference for growth shares. Therefore, we assume that if in a particular month, the spread is positive and the probability is above 0.5, or if the spread is negative and the probability is below 0.5, the prediction has been successful; otherwise, the prediction has failed. *TP* measures the total percentage of the correct predictions as defined above; N_1 is the number of the months when the spread is positive and the predicted probability is also above P_b. N_2 is the number of the months when the spread is negative and the predicted probability is below P_b. P_1 is the probability for the actual spread to be positive and the predicted probability to be above P_b; P_2 is the probability for the spread to be negative and also the predicted probability to be below P_b. We also report the results when the probability threshold is 0.6 and 0.7.

In the case of a probability threshold of 0.5, mean returns for all the models range from 5.45% to 6.12%. With an assumption of 0.05% transaction cost [55] per

[55] 5 bps per transaction is realistic only in the context of derivative trading. Transaction costs with funds is higher and also practically getting in or out of the mutual fund would take more than 1 business working day and the fund price is only available at the end of the day at best, which would most possibly miss the signal already.

transaction, the net return is calculated as the gross return less than the total transaction costs, given the number of the transactions can be calculated based on the number of the strategy shifts. Net annualized mean returns range from 5.31% to 5.88%. The best performing model is model 2 and model 3 in terms of mean return (= 6.12%) and model 6 in terms of net mean return (= 5.88%). Given a low transaction cost (0.05% per transaction), all the models outperformed buy and hold income and growth strategies. i.e. net Sharpe ratios for model 1 to model 6 range from 0.0213 to 0.0621, and significantly higher than buy and hold income strategy (= 0.0011) and buy and hold growth strategy (= 0.0158). Further, given the number of shifts within a strategy, we can calculate the breakeven per transact transaction costs against a buy and hold strategy. The breakeven costs range from 0.151% to 0.327% for income funds, and range from 0.074% to 0.232% for growth funds. Total percentages of correct prediction (*TP*) are ranging from 54.2% (model 5) to 59% (model 1 and model 3). In both cases of a probability threshold of 0.6 and 0.7, net Sharpe ratios are higher than buy and hold strategies, but the outperformance is smaller, suggesting 50% probability threshold provides the best results.

For a PL portfolio (as a reversed case of a PW portfolio), we only report the results when the probability threshold is 0.5. In this case, the portfolio consists of the growth funds when the probability of value spread is larger than 0.5 and the value funds when the probability of value spread is less than 0.5. Unsurprisingly, the PL portfolios underperformed the buy & hold strategy significantly, with negative Sharpe ratios for all the models. Mean return range from 4.13% to 4.8% and the Sharpe ratios range from -0.06 to -0.01. The results suggest that the PL portfolio can be used as a short candidate against a buy & hold portfolio.

TABLE 5-8. PREDICTION ESTIMATION AND POSSIBLE PROFITABLE TRADING STRATEGIES

The following tables report the prediction estimation results. The variables used in the models are shown below: Model 1, 2, 3 are those explained and tested in multifactor logit models. Model 4, 5, 6 are the models with variables corresponding to those in model 1, 2 and 3, but with 1 lagged level.

Model	Variables
Model 1	*C, DY, INFLA*
Model 2	*C, DY, INFLA,WAGE*
Model 3	*C, DY, INFLA,WAGE, TERM, INDPROD*
Model 4	$C, DY_{(t-1)}, INFLA_{(t-1)}$
Model 5	$C, DY_{(t-1)}, INFLA_{(t-1)}, WAGE_{(t-1)}$
Model 6	$C, DY_{(t-1)}, INFLA_{(t-1)}, WAGE_{(t-1)}, TERM_{(t-1)}, INDPROD_{(t-1)}$

PANEL A. PERFORMANCE OF BUY & HOLD STRATEGIES AND PERFECT FORESIGHT STRATEGY

This table reports the mean return, Sharpe ratio, net mean return, net Sharpe ratio for buy & hold strategies (income & growth funds respectively), perfect foresight models. We also report the number of changes for perfect foresight strategy, with transaction cost at 0.05% per transaction.

	Buy & hold Income	Buy& hold growth	Perfect foresight
Mean Return	5.01%	5.23%	10.57%
Sharpe Ratio	0.0011	0.0158	0.41
No. of Changes	0	0	66
Net Mean Return	5.01%	5.23%	10.01%
Sharpe Ratio (Net)	0.0011	0.0158	0.3662
Transaction cost	-	-	0.05%

PANEL B. THE FORECASTING RESULTS OF PW PORTFOLIOS

PW is the strategy that buys the predicted winner. This table reports the returns of PW portfolio trading strategies constructed from the estimated logit probabilities for the period from Jan 1990 to Oct 2003. *TP* measures the total percentage; N_1 is the numbers of months when the spread is positive and the probability is also above the threshold (situation 1). N_2 is the numbers of months when the spread is negative and the probability is below the threshold (situation 2). P_1 and P_2 are the corresponding percentage for the first and second situation. (total *N*: 161)

Probability threshold: 0.5

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mean Return	5.98%	6.12%	6.12%	5.77%	5.45%	6.08%
Sharpe Ratio	0.07	0.08	0.08	0.05	0.03	0.07
No. of Changes	80	78	70	32	40	56
Net Mean Return	5.69%	5.84%	5.86%	5.66%	5.31%	5.88%
Sharpe Ratio (Net)	**0.0496**	**0.0606**	**0.0621**	**0.0457**	**0.0213**	**0.0607**
Transaction cost	0.05%	0.05%	0.05%	0.05%	0.05%	0.05%
Breakeven: Income	0.166%	0.197%	0.218%	0.327%	0.151%	0.263%
Breakeven: Growth	0.128%	0.158%	0.174%	0.232%	0.074%	0.208%
TP	59.0%	58.4%	59.0%	57.8%	54.2%	57.2%
P_1	22.9%	22.9%	23.5%	11.4%	11.4%	15.1%
N_1	38	38	39	19	19	25
P_2	36.1%	35.5%	35.5%	46.4%	42.8%	42.2%
N_2	60	59	59	77	71	70

Probability threshold: 0.6

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mean Return	5.49%	5.71%	5.53%	5.40%	5.45%	5.39%
Sharpe Ratio	0.03	0.05	0.04	0.03	0.03	0.03
No. of Changes	31	37	35	3	5	5
Net Mean Return	5.38%	5.57%	5.41%	5.39%	5.43%	5.37%
Sharpe Ratio (Net)	**0.0259**	**0.0395**	**0.0279**	**0.0264**	**0.0290**	**0.0250**
Transaction cost	0.05%	0.05%	0.05%	0.05%	0.05%	0.05%
TP	54.2%	56.6%	55.4%	53.0%	53.6%	53.0%
P_1	6.6%	9.0%	8.4%	0.6%	1.2%	1.2%
N_1	11	15	14	1	2	2
P_2	47.6%	47.6%	47.0%	52.4%	52.4%	51.8%
N_2	79	79	78	87	87	86

Probability threshold: 0.7

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mean Return	5.52%	5.42%	5.42%	5.23%	5.23%	5.23%
Sharpe Ratio	0.04	0.03	0.03	0.02	0.02	0.02
No. of Changes	9	11	11	1	1	1
Net Mean Return	5.49%	5.38%	5.38%	5.23%	5.23%	5.23%
Sharpe Ratio (Net)	**0.0330**	**0.0258**	**0.0258**	**0.0156**	**0.0156**	**0.0156**
Transaction cost	0.05%	0.05%	0.05%	0.05%	0.05%	0.05%
TP	54.2%	53.6%	53.6%	52.4%	52.4%	52.4%
P_1	2.4%	2.4%	2.4%	0.0%	0.0%	0.0%
N_1	4	4	4	0	0	0
P_2	51.8%	51.2%	51.2%	52.4%	52.4%	52.4%
N_2	86	85	85	87	87	87

PANEL C. THE FORECASTING RESULTS OF PL PORTFOLIOS

This table reports the returns of PL trading strategies constructed from the estimated logit probabilities. The PL portfolio refers to the strategy that buys the predicted loser. We only report the results with probability threshold of 0.5.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mean Return	4.27%	4.13%	4.13%	4.48%	4.80%	4.17%
Sharpe Ratio	-0.05	-0.06	-0.06	-0.04	-0.01	-0.06
No. of Changes	80	78	70	32	40	56
Net Mean Return	3.99%	3.84%	3.88%	4.36%	4.66%	3.97%
Sharpe Ratio (Net)	-0.0725	-0.0830	-0.0805	-0.0472	-0.0256	-0.0769
Transaction cost	0.05%	0.05%	0.05%	0.05%	0.05%	0.05%
TP	59.0%	58.4%	59.0%	57.8%	54.2%	57.2%
P_1	22.9%	22.9%	23.5%	11.4%	11.4%	15.1%
N_1	38	38	39	19	19	25
P_2	36.1%	35.5%	35.5%	46.4%	42.8%	42.2%
N_2	60	59	59	77	71	70

Appendix Figure 5-3 plots the trading performance of a PW portfolio based on six models and forecasted probabilities. Appendix Figure 5-4 plots the trading performance of PL portfolio based on six models. The trading threshold is 0.5 for all the cases.

To summary, our results mainly suggest the following: One, the rotation strategy can beat buy & hold strategies, but outperformance of the rotation strategy is subject to the level of transaction costs. Higher transaction costs can partly outweigh the benefit of a better forecasting strategy with a superior gross return, and the profitability diminishes when the transaction costs grow. This is consistent with Bauer *et al.* (2004), who reveal that in Japan, sufficient predictability only exist under low transaction cost levels. Under high transaction costs scenarios, it is more difficult to obtain incremental benefits. Practically, trading with equity derivatives (index futures) instead of with funds can help to achieve positive net returns as the transaction cost is lower and it is less time consuming to complete the derivative trades.

Two, perfect foresight strategy (100% accuracy) significantly outperformed against buy & hold strategies in both gross return and net return terms. Meanwhile, all of our profitable rotation strategies have only less than 60% of forecasting accuracy level, thus it indicates that there is huge room for further research on improving the model's forecasting accuracy in order to enhance the net returns. As suggested by Levis and Liodakis (1999), it requires more than an 80% of correct prediction of value spread to beat a passive strategy based on their simulation results.

4. Conclusion

In recent years, average return differences among styles, such as the difference between growth and value stocks, have become the focus of many studies, associated with various debates. Among them, two dominant views include one, value funds can provide the long-term benefits consistently; two, each style offers its own benefit and style diversification is the optimal solution to avoid the risk associated with pure style investing. We extend the debates by examining style performance of UK unit trusts. Our data are different from the others as we use active equity portfolio data (UK unit trusts) instead of equity indices. Hence, our results also reflect the portfolio management skills and institutional characteristics of unit trusts, in addition to fund style characteristics.

The arguments of style consistency and style rotation are illustrated initially. Employing both risk-unadjusted and risk-adjusted returns (as measured by the SDF primitive efficient models), the unit trusts performance with three types of styles are examined respectively. We compare and rank these funds over different sub-sample periods based on evaluation of their performance. It is striking to see that income funds provide no superior performance than growth funds and indeed, unit trusts performance of different styles vary over time. More specifically, growth funds perform better in bull markets while income funds do better in bear markets.

To explore the potential explanations and implications for the test results, we employ different methods to construct style strategies, including simple strategy by switching holdings during bull and bear markets; the sensitivity tests of the value spread and certain macroeconomics variables, using both OLS regression and a logit approach, in which the dependent variable is a binary variable that takes the value of one if the spread is positive and zero if the spread is negative. The trading strategies based on the logit models are constructed and the profits are measured, compared to the passive and perfect-forecasted returns.

The sensitivity tests suggest that some of those factors are significantly correlated with return spreads. The results based on logit models show that winning strategies and profits over a passive portfolio is possible based on around 60% prediction accuracy rate of future value spread, given a low level of transaction costs.

In a word, among UK unit trusts, the performance of growth and income funds vary over time, while over a long investment horizon, there is a lack of differential performance among them. On the other hand, controlled style-rotation strategies based on the underlying fundamental characteristics of the relevant macroeconomic factors can potentially enhance values, given relatively low levels of transaction costs.

APPENDIX 5

TABLE 5-1. BEST AND WORST SELLING SECTORS – NET RETAIL SALES

Year	Month	Best	Worst
2004	October	UK Equity Income	UK All Companies
	September	UK Equity Income	UK All Companies
	August	UK Equity Income	UK All Companies
	July	UK Equity Income	Europe Excluding UK
	June	UK Equity Income	Europe Excluding UK
	May	UK Equity Income	Far East Excluding Japan
	April	UK All Companies	Europe Excluding UK
	March	UK Equity Income	North America
	February	UK All Companies	UK Equity & Bond Income
	January	UK All Companies	UK Smaller Companies
2003	December	UK All Companies	UK Smaller Companies
	November	UK Corporate Bond	Global Growth
	October	UK Corporate Bond	Europe Excluding UK

Source: The Investment Management Association (IMA), formerly the Association of Unit Trusts and Investment Funds (AUTIF).

FIGURE 5-1. 12 MONTHS MOVING AVERAGE OF VALUE SPREAD OF RISK-UNADJUSTED RETURNS

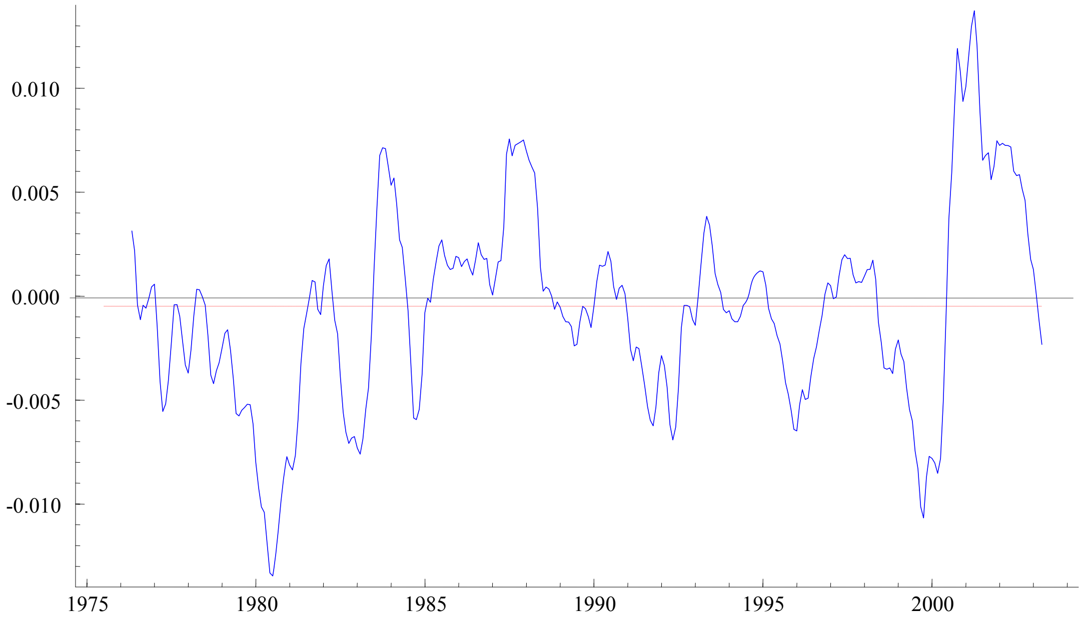
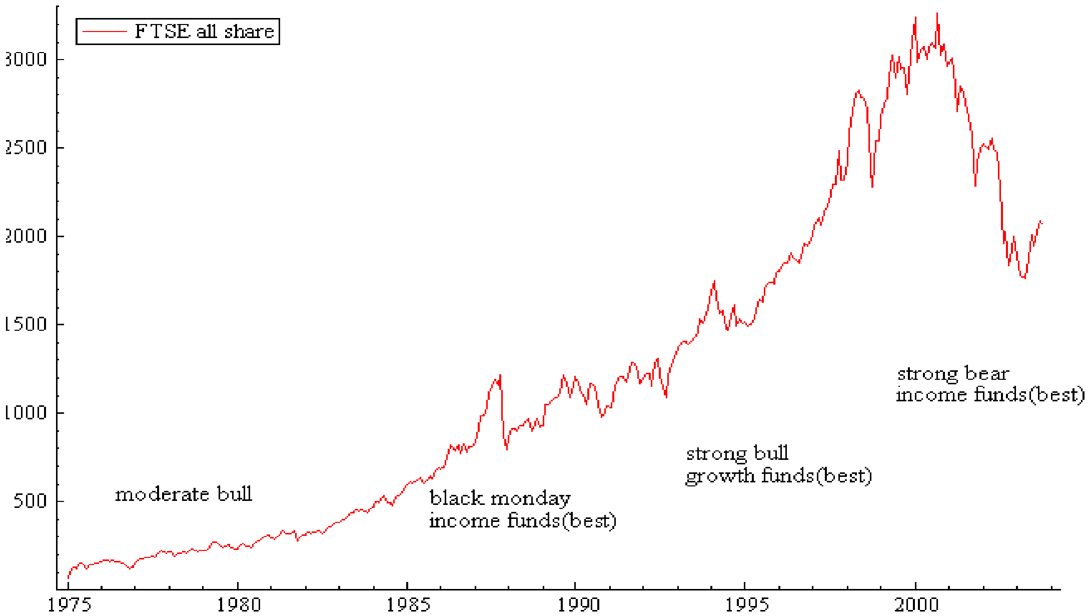
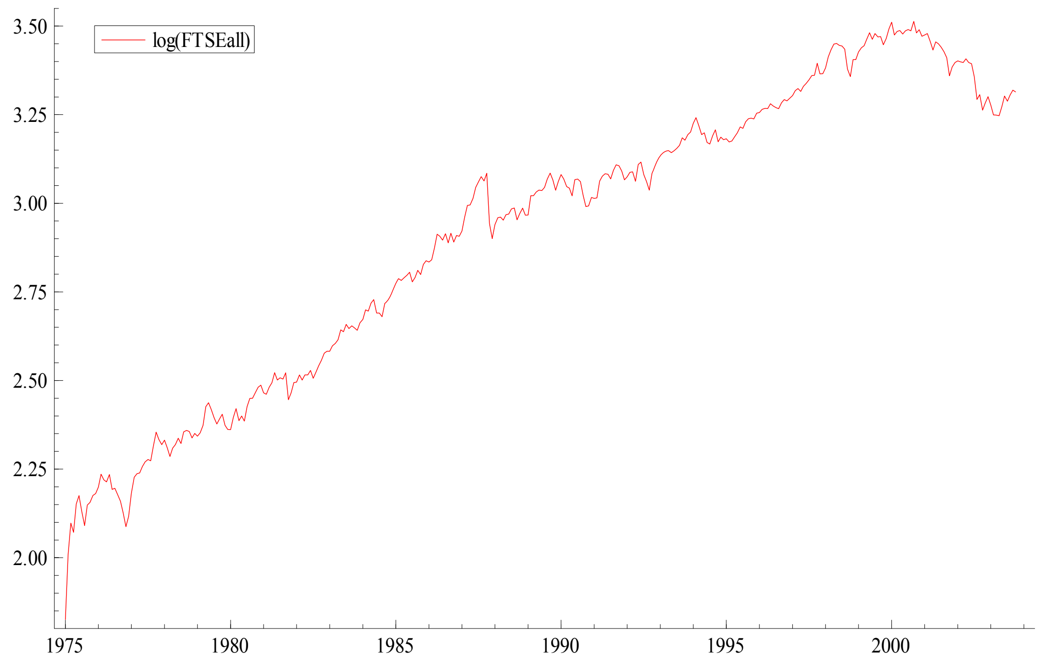


FIGURE 5-2. STYLE ROTATION WITH FTSE ALL SHARE INDEX AND FTSE ALL SHARE DIVIDEND YIELD



Log UK Stock Prices 1975/1 –2003/10


log(FTSEall)
3.50
3.25
3.00
2.75
2.50
2.25
2.00
1975
1980
1985
1990
1995
2000

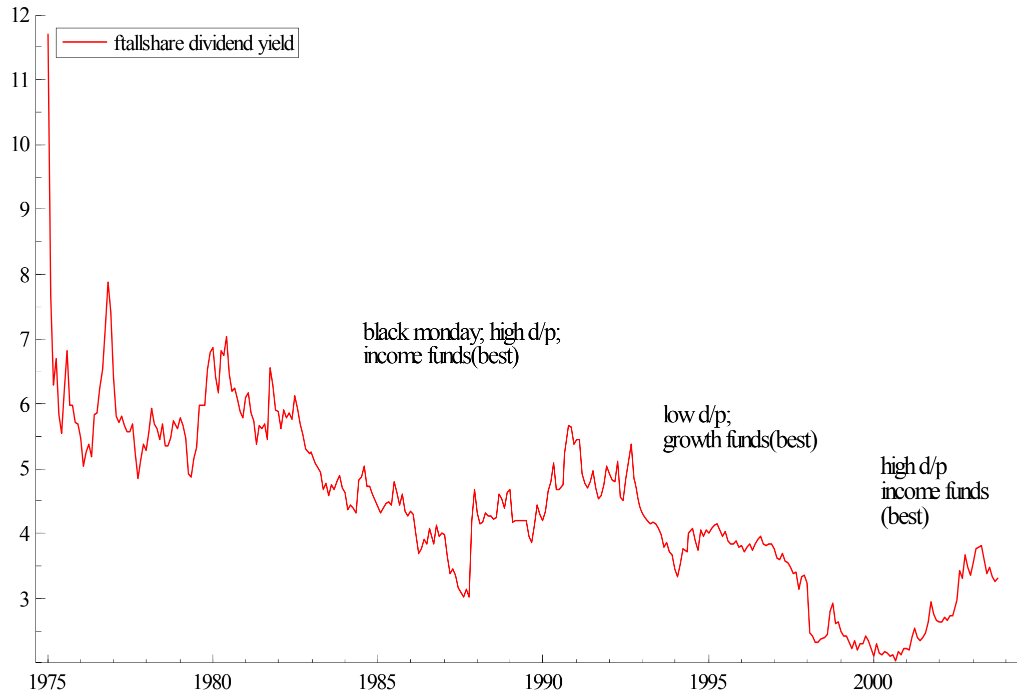
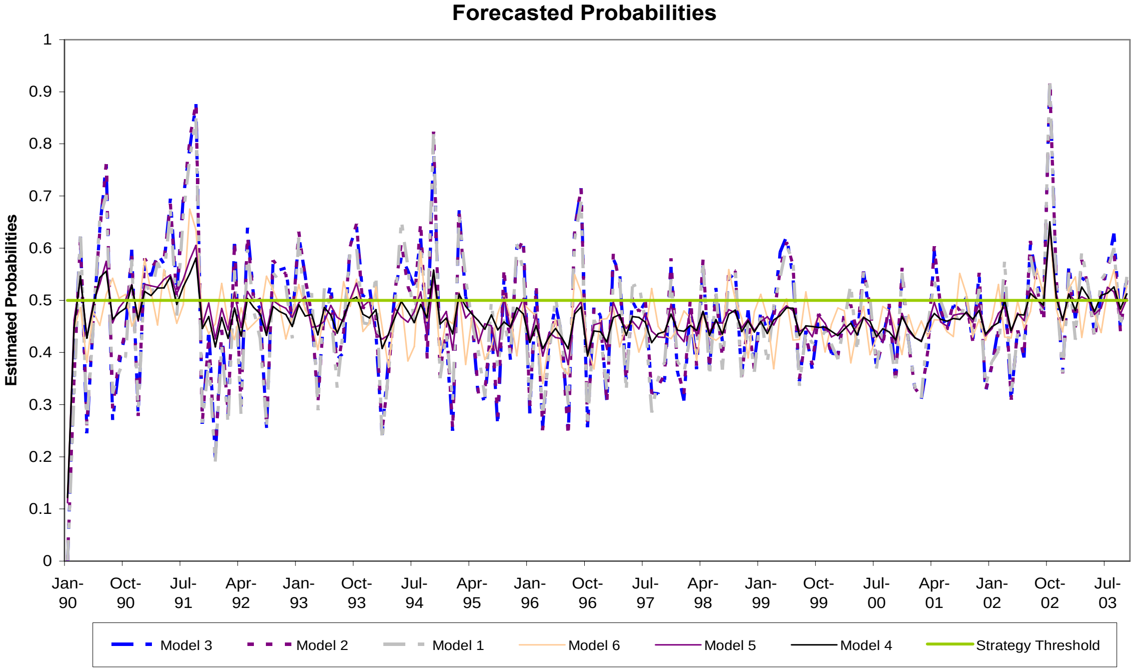
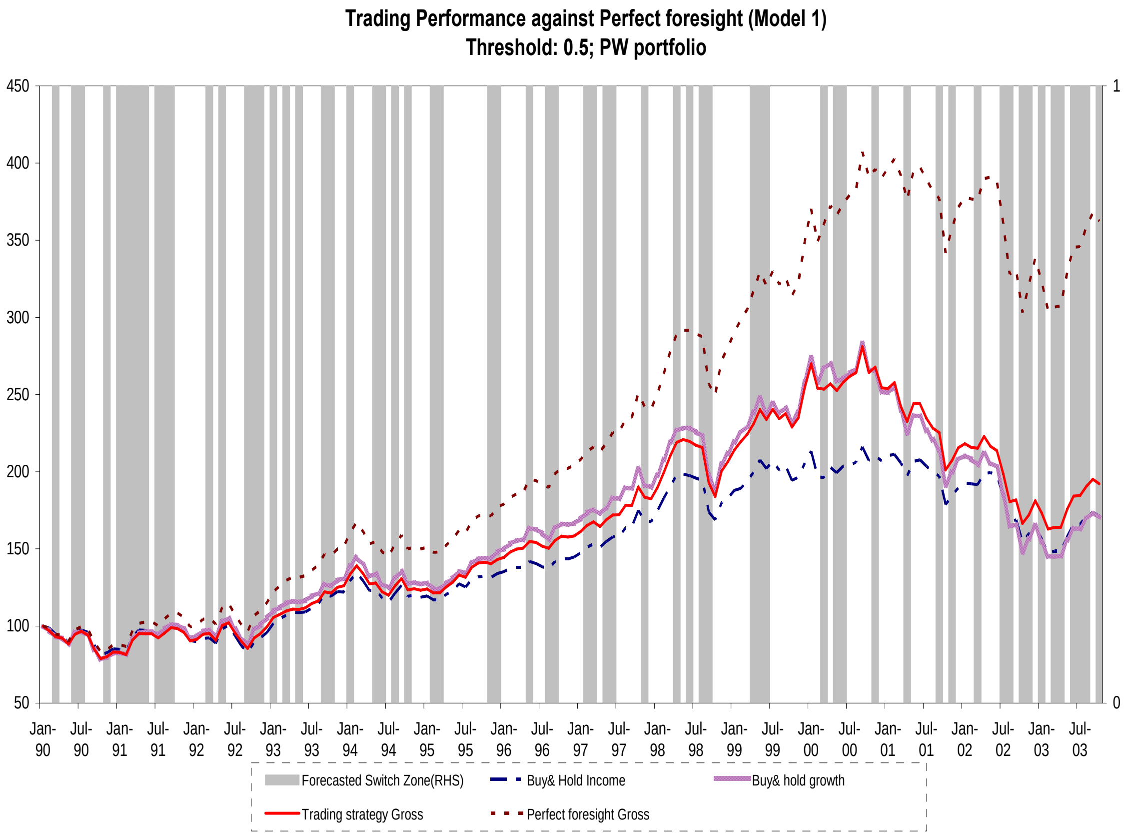
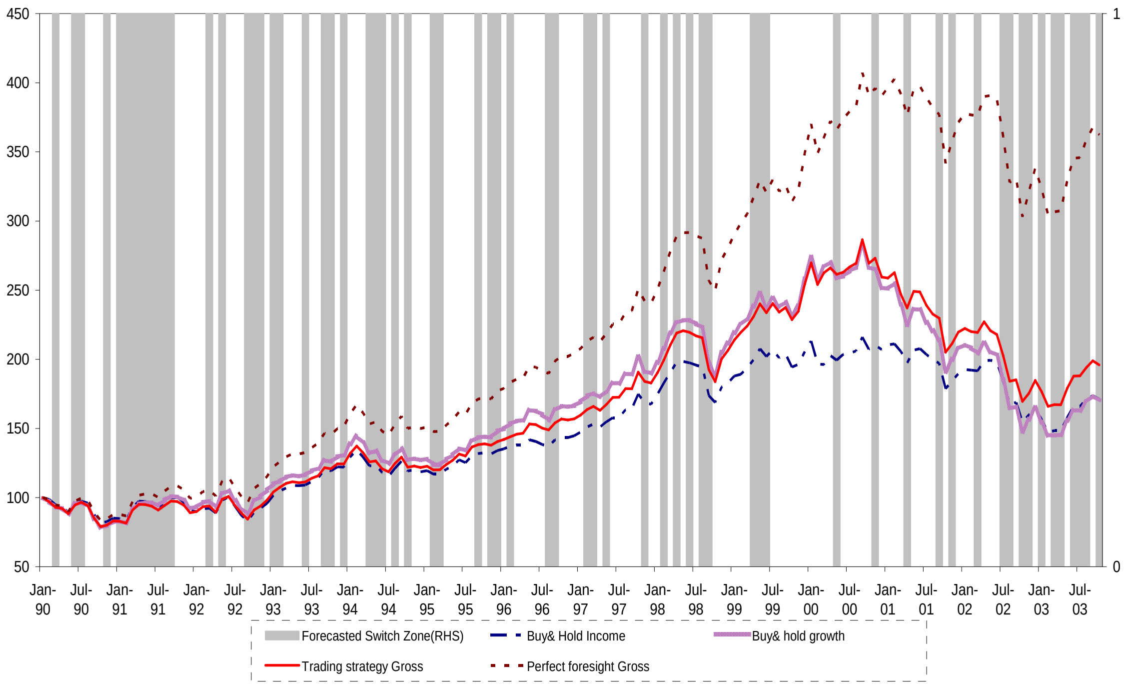

ftallshare dividend yield
12
11
10
9
8
7
6
5
4
3
black monday; high d/p;
income funds(best)
low d/p;
growth funds(best)
high d/p
income funds
(best)
1975
1980
1985
1990
1995
2000

FIGURE 5-3. TRADING PERFORMANCE BASED ON THE LOGIT MODEL TRADING RULES (PW PORTFOLIO)

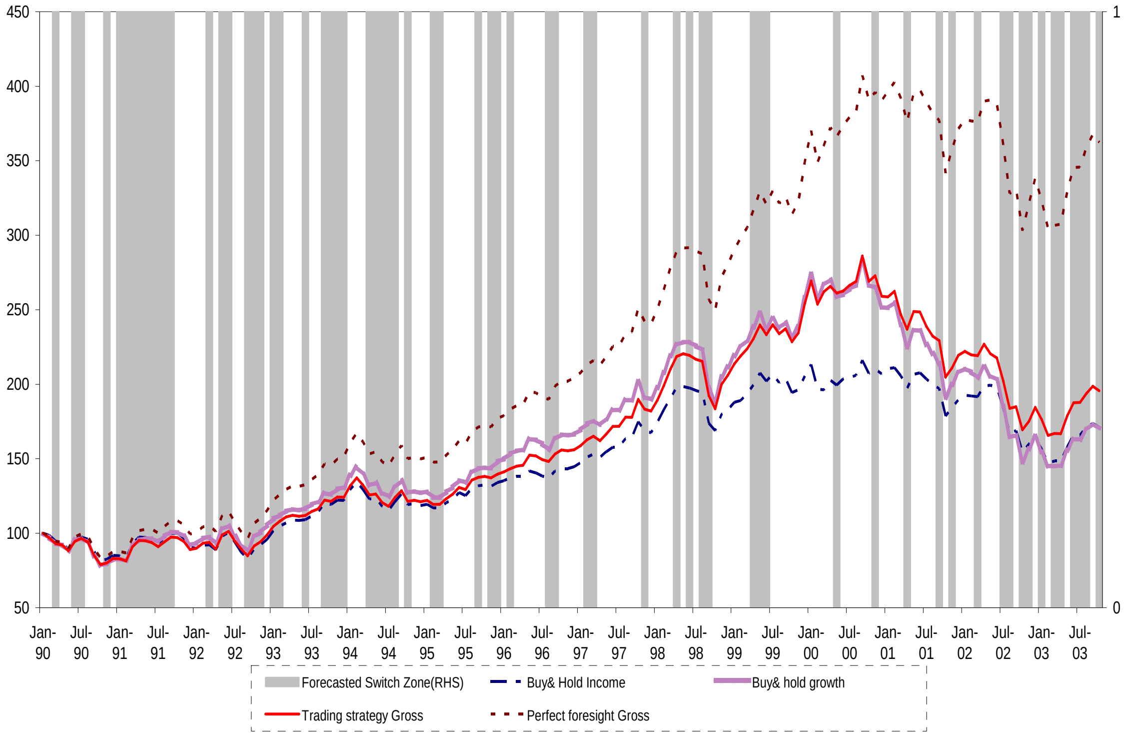




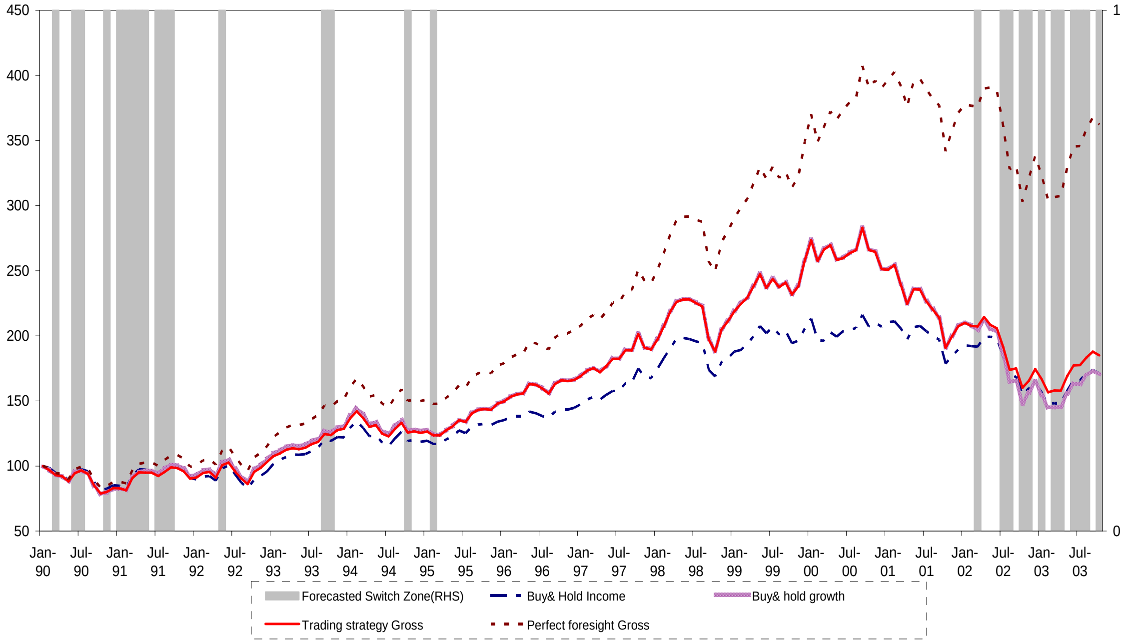
Trading Performance against Perfect foresight (Model 2)
Threshold: 0.5; PW portfolio



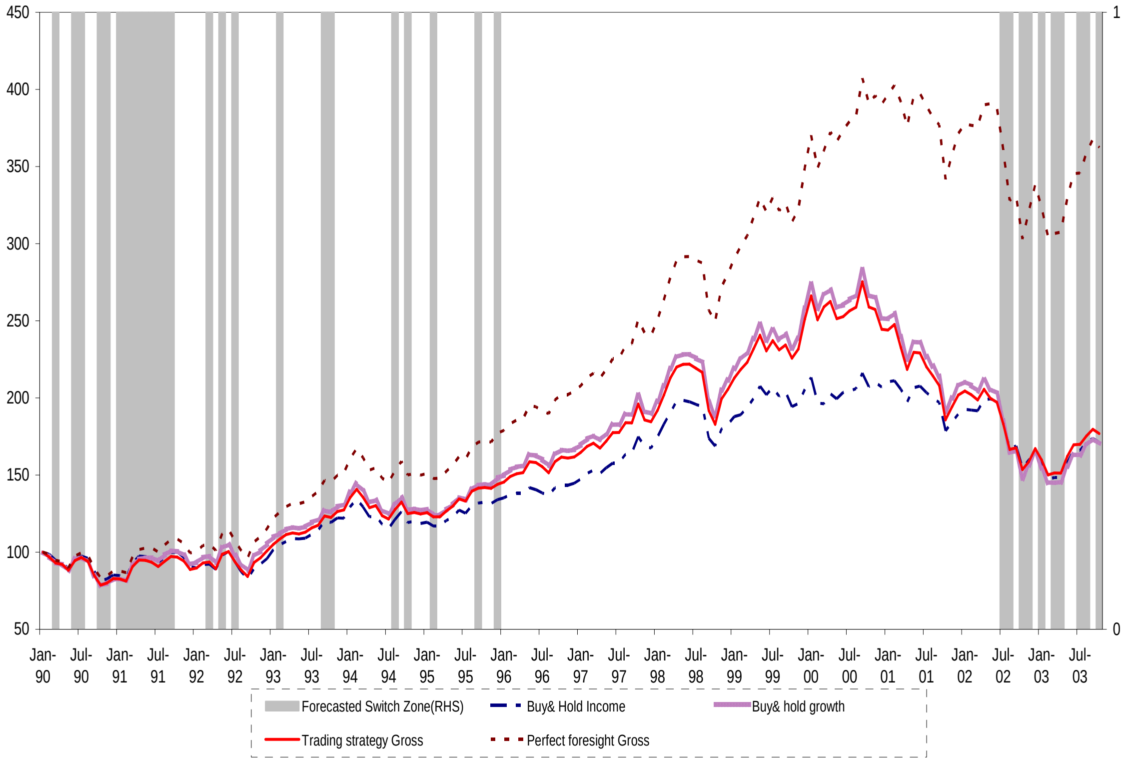
Trading Performance against Perfect foresight (Model 3)
Threshold: 0.5; PW portfolio



Trading Performance against Perfect foresight (Model 4)
Threshold: 0.5; PW portfolio



Trading Performance against Perfect foresight (Model 5)
Threshold: 0.5; PW portfolio



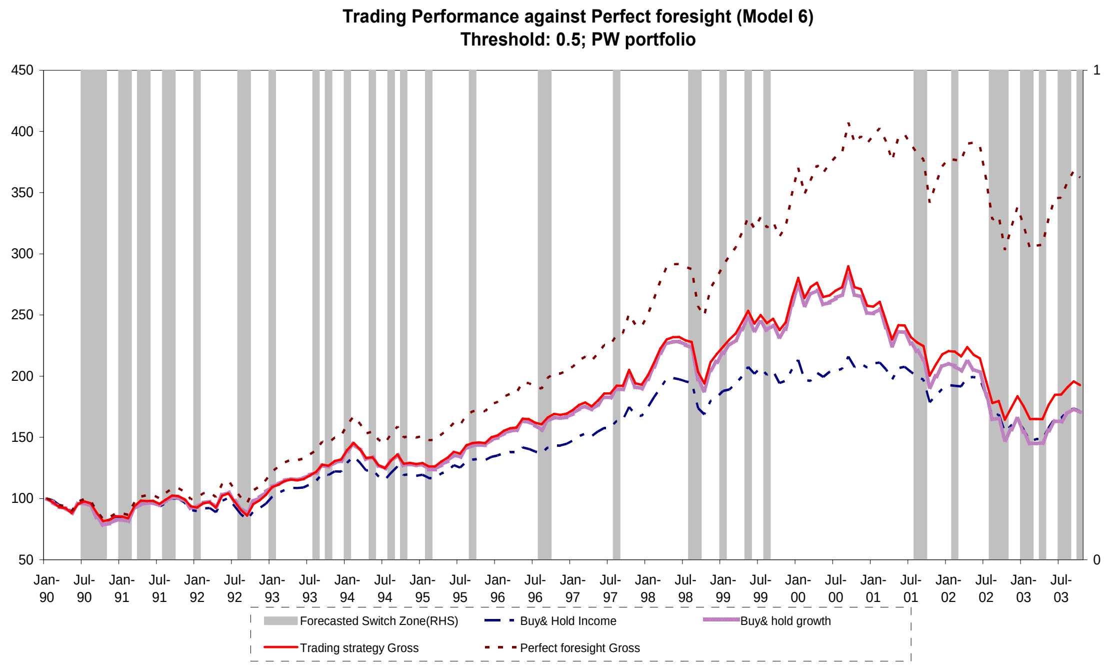
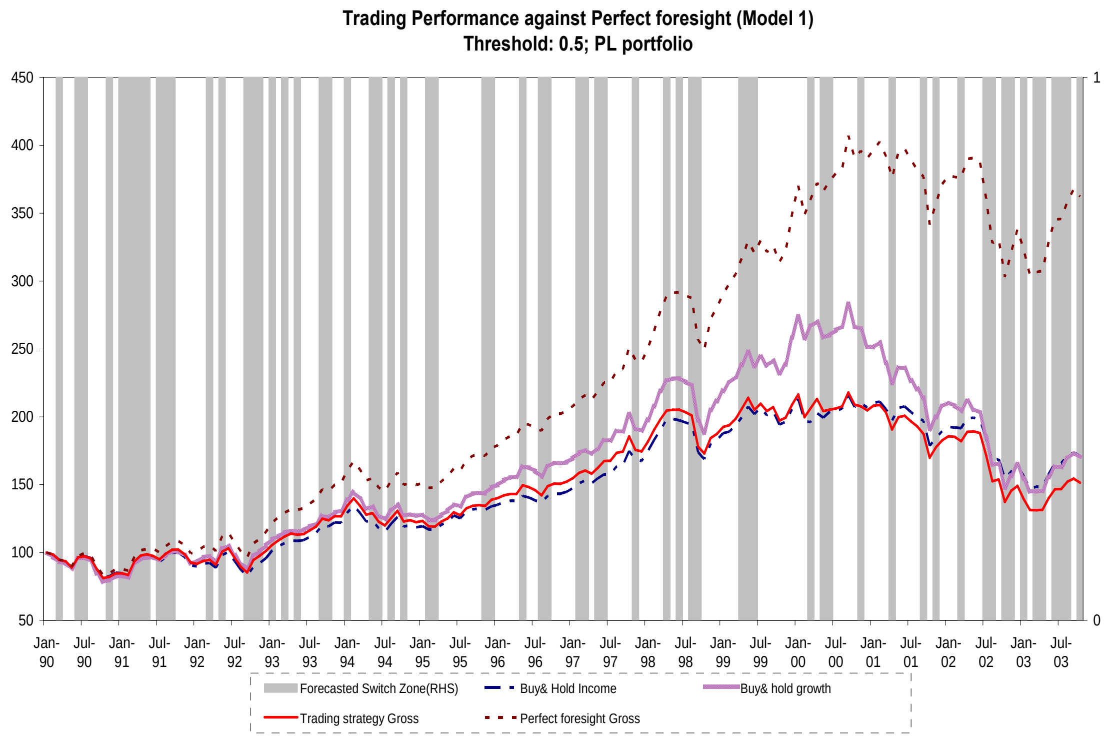
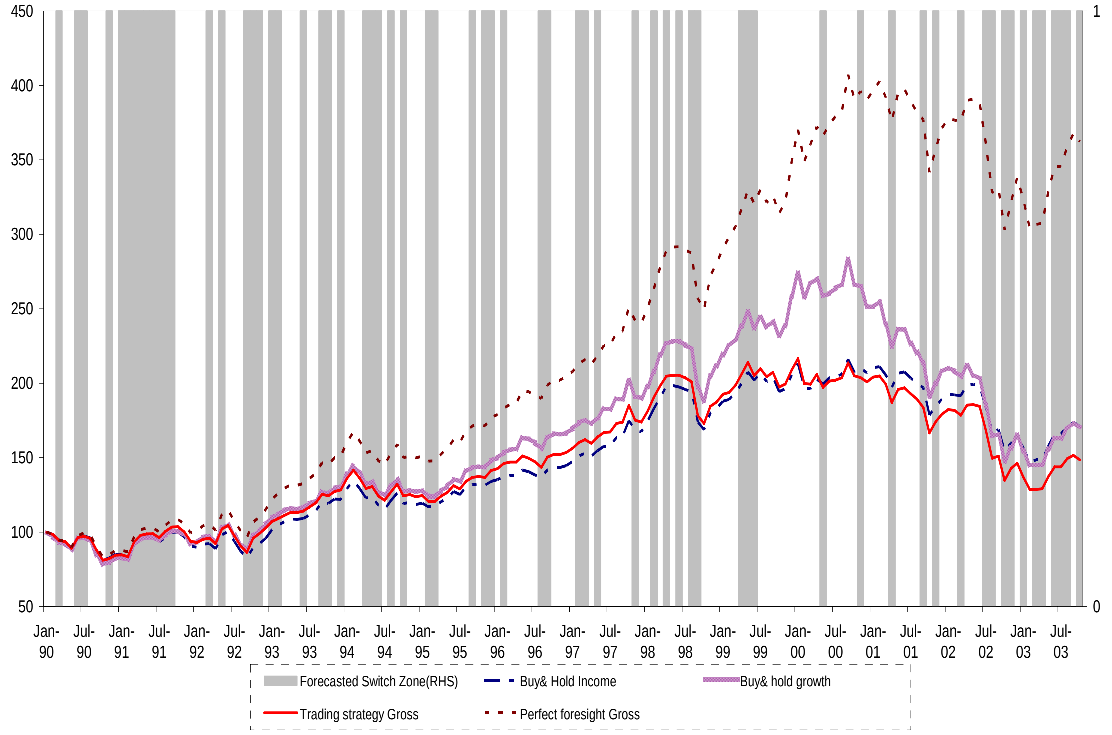


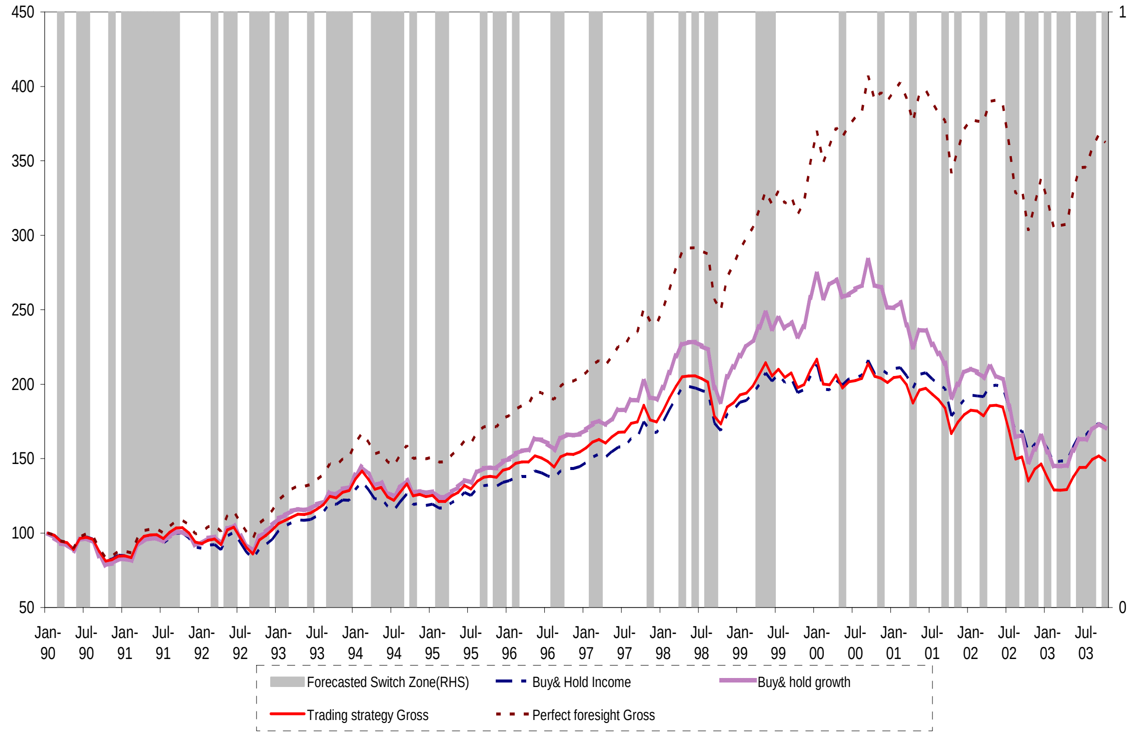
FIGURE 5-4. TRADING PERFORMANCE BASED ON THE LOGIT MODEL TRADING RULES (PL PORTFOLIO)



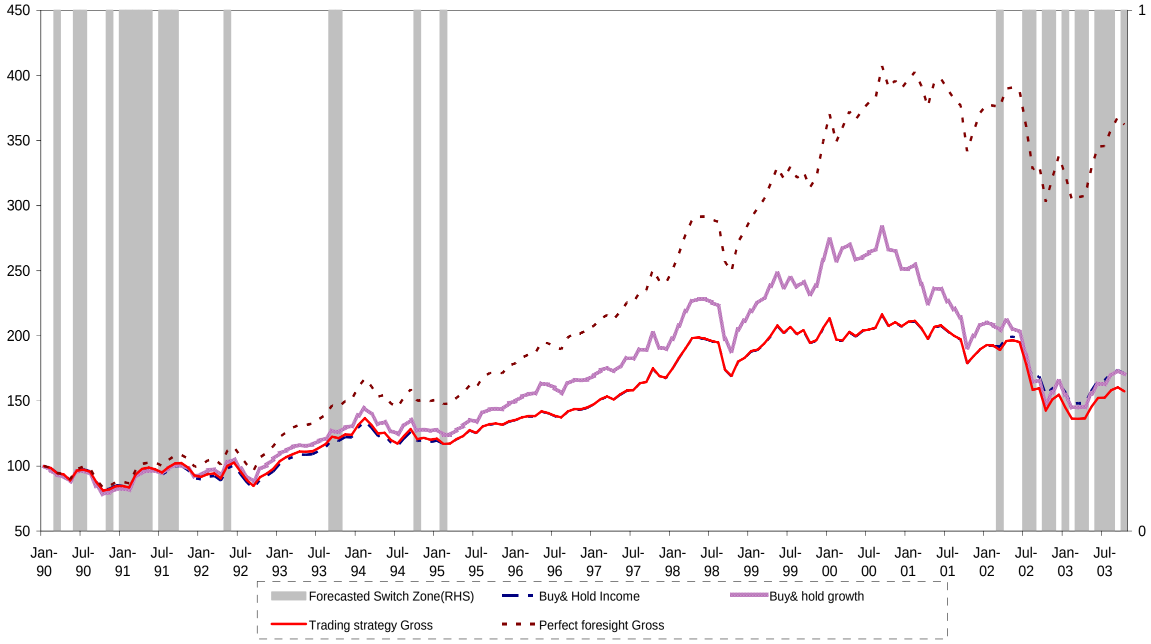
Trading Performance against Perfect foresight (Model 2)
Threshold: 0.5; PL portfolio



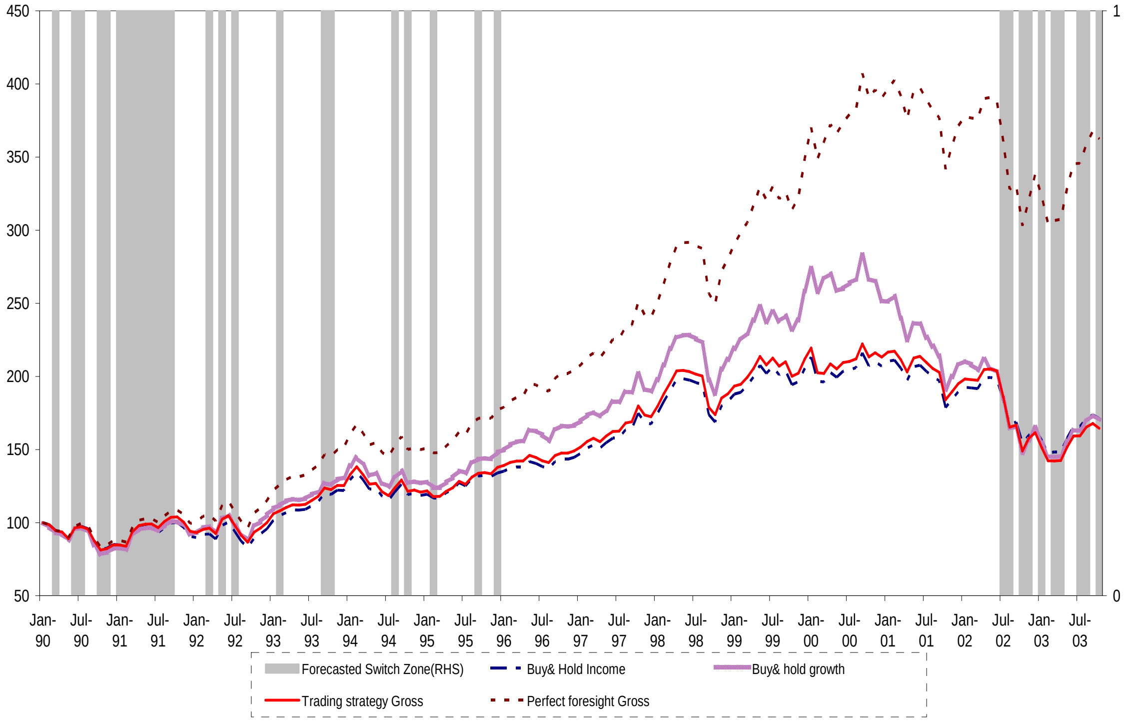
Trading Performance against Perfect foresight (Model 3)
Threshold: 0.5; PL portfolio



Trading Performance against Perfect foresight (Model 4)
Threshold: 0.5; PL portfolio



Trading Performance against Perfect foresight (Model 5)
Threshold: 0.5; PL portfolio



CHAPTER 6: PERFORMANCE PERSISTENCE OF UK UNIT TRUSTS

There is no doubt that investors chase past performance

—Martin J. Gruber. 1996 AFA Presidential Address

1. Introduction

We have been so far focusing on the economically interesting questions of whether, by skills or other means, fund managers have been able to earn significant returns in excess of the risks being carried. It is equally important to see if such performance can persist. If it does, past performance of funds can be useful information for market participants' investment decisions.

Performance persistence refers to such a question of whether past performance of a fund has any correlation with future performance. As argued by Rhodes (2000), '*As markets become more efficient it will become more difficult for any fund manager to beat the peers consistently, to any significant degree. The gains to be made from conducting ever more thorough research will diminish. Therefore theoretically at least, it seems unlikely that a given fund manager could maintain a meaningful outperformance of her peers for a long period of time*'. As the above arguments is largely purported by 'Market Efficiency Hypothesis' believers, past performance or 'track record' has been practically taken as a major consideration for fund selection. According to a report of the task force on past performance (FSA, 2001), its consumer survey suggest that consumers choose a fund based on its past performance in addition to brand reputation and editorials in the financial press. '*All respondents* *agreed that past performance and brand was essential to the choice of investment products. The less sophisticated investors placed great faith in the brand…… Past performance was seen as fundamental, and if not present would be sought elsewhere by many. Consumers were using performance figures as a reasonable indicator of future performance and benchmarking their (minimum) expectations against perceived savings account rates*.'

Can portfolio managers of actively-managed unit trusts beat the market persistently? Evidence of persistent investment performance would suggest that some fund managers are able to outperform their peers consistently, which would imply that the funds' managers must either have access to information that is not widely available or can make better use of publicly available information than most of other managers. On the contrary, evidence of a lack of performance persistency would help to provide empirical arguments for supporting the passive investment approach.

1.1 Literature Review

Many authors have directly examined persistence of U.S. mutual funds performance, the examples include Grinblatt and Titman (1992), Hendricks, Patel, and Zeckhauser (1993), Maikiel (1995) and Sharpe (1995). Vast empirical evidence has shown the existence of performance persistence in the U.S. mutual fund market. For instance, Carlson (1970) finds evidence that the funds with above-median returns over the preceding year typically repeat their superior performance. Elton and Gruber (1989) cite a 1971 Securities and Exchange Commission (SEC) study that indicates similar persistence in risk-adjusted mutual fund rankings. Lehmann and Modest (1987) report some evidence of persistence shown on mutual fund alphas, and Grinblatt and Titman

(1988, 1992) display that the effect of persistence is statistically significant. Goetzmann and Ibbotson (1994) conclude that the performance persistence phenomenon is present in both raw and risk-adjusted returns to equity funds at observation intervals from one month to three years.

More recently, performance persistence of UK Unit Trusts have also drawn many researchers' attention. Fletcher (1997) investigates a sample of 101 UK unit trusts and examines five portfolios based on both a ranking of five-year risk-adjusted performance windows and a two-year performance window. Survivorship bias was partly allowed for by the continuation of funds through name changes or changes in management groups, though mergers are treated as terminations. At the end, evidence of persistent performance was not found.

Brown, Draper and McKenzie (1997) analyze UK pension fund performance using the data from the World Market Company. They examine risk-adjusted returns using the market model over the period 1986-1992 with a sample of 409 funds. Applying one-year windows and contingency table analysis of performance persistence across the quartiles of the sample, they discover limited evidence of persistence.

Quigley and Sinquefield (1998) use a similar approach by constructing portfolios, ranked by deciles, on the basis of relative performance in a given year. They then compare the performance of each of these portfolios in the next year. They use a large sample taken from the Micropal database of all equity UK unit trusts that were in existence between 1978 and 1997 with a total of 752 funds. Only those trusts that are classified as having objectives of Growth and Income, Growth, Equity income or smaller companies are included. They construct tests of performance persistence both before and after adjusting for risk. The difference between the averages of the portfolio's performance at extremes of the deciles is positive over subsequent years. However, adjustment for transactions costs eliminates any gains. A variety of market and factor-based risk adjustments are also applied which wipe out any positive gains but lead to the conclusion that only poor performance persists.

Lunde, Blake and Timmerman (1998) use risk-adjusted returns to create portfolios of returns over 3-year periods using a large data-set of 2,300 UK unit trusts obtained from Micropal data. They construct performance measures based on bid prices and net income without any adjustment for expenses. Their analysis is based on inter-quartile fund performance over three year periods and they report the existence of performance persistence.

Fletcher (1999) examines 85 UK unit trusts with an US investment orientation between 1985 and 1996 and reports no evidence of performance persistence. Similarly, the Wood Mackenzie Company (1999) apply a technique of estimating inter-quartile transition probabilities across five year windows for a sample of UK income and growth funds and find no evidence of performance prediction, but report evidence of the top quartiles' performance persisting in the next year.

Allen and Tan (1999) find evidence of performance persistence for the period of 1989-1995 in a sample of 131 UK funds. Their study employs a UK sample data set of weekly returns of all equity mutual funds existing each year and available on the Datastream database. They examine the relative performance of the funds and test if a good past-performance can be indicative of the fund's future performance. Unlike previous studies, which compare fund performance with a benchmark, their study compares the relative performance of the sample funds themselves. They investigate the persistence in performance in both the short and long run using four empirical tests, including Ordinary Least Squares regression analysis of CAPM risk-adjusted excess returns, contingency table analysis of winners and losers, Chi-squared tests on these tables, and Spearman Rank Correlation coefficient analysis of successive period performance rankings.

To investigate the persistence of UK managed funds and if consumers can use this information to make their investment decisions, Rhodes (2000) applies an approach allowed for a more consistent examination of performance over the longer time, and he concludes that there is no persistence in the performance of managed funds in the UK after 1987 and states it would be misleading to suggest that retail investors could use past information to predict the future performance nowadays.

With a similar purpose but from the perspective of fund investors, Giles *et al* (2002), from economic consultancy Charles River Associates, conduct a research project to examine if performance persists in UK equity-based unit trusts. They obtained data from the S&P Micropal database on all equity unit trusts that existed at any time between 1981 and 2001. In addition, they also collected dead fund data from a number of other sources, with a total sample of 942 unit trusts. With the purpose of investigating if the past information of unit trusts can be useful to retail fund investors, their tests are based mainly on raw returns rather than risk-adjusted returns. Their results find evidence of performance persistence broadly — both good and bad.

It implies past performance information can be used by retail investors to make profitable investment decisions. However, they also discover that persistence depends on the time horizon and the sector, i.e. performance persistence is only significant for equity income and smaller companies funds in the short term; whilst it is significant for UK all companies and UK equity income funds over all time horizons.

1.2 Performance Persistence VS Performance Reversal

'Persistence' phenomenon and Momentum Strategies

If 'Persistence' exists during the short term, it is called 'short-term trend' or 'momentum phenomenon'. Many authors have reported strong evidence of momentum in stock prices and suggest that both upward and downward stock price movements that persist over a short time are followed typically by future movements in the same directions. It is broadly suggested by many researchers that momentum strategies based on 'momentum phenomenon' could produce significant profits. Examples including Jegadeesh and Titman (1993), who examine different types of momentum strategies and show that one of the strategies can make profits of about 1% per month for the following year. The strategy involves basically buying stocks with high returns over the previous months (say, three to twelve months) and selling stocks with poor returns over the same periods. Other examples including Rouwenhorst (1998), who reports that the momentum profits can also be obtained in the European markets. Further, Chui, Titman and Wei (2000) provide evidence that with the notable exception of Japan and Korea, momentum strategy also works in Asian markets.

If ‘persistence’ prevails over the long run, it would suggest that fund performance is consistent and the individual investors can predict performance of the funds purely based on their past performance, in the sense that past performance of a fund is the essential information for investors’ investment decision making.

‘Reversal’ and Contrarian Strategies

On the contrary, a negative persistence or ‘return reversal’ describes a situation where the fund reverses rather than repeats its past performance, i.e. the past winners will become the current losers and vice versa. It is suggested that equity returns follow such a pattern only in the long run. For example, DeBondt and Thaler (1985, 1987 & 1990) provide the evidence of long-term reversal of returns in financial markets. They compare the performance of two groups of companies: extreme `winners', i.e. companies with a few years of good performance and extreme `losers', i.e. companies with several years of poor financial results. In their studies, they find that extreme winners (losers) tend to generate on average relatively poor (relatively good) subsequent returns. If this is true, ‘contrarian strategies’ by investing in the past losers and/or selling the past winners could be profitable.

In this chapter, we explore the degree of both ‘persistence’ and ‘reversal’ of UK unit trusts over the sample period from January 1975 to December 2003 in this chapter. Section 2 illustrates the methodologies we applied for persistence tests; Section 3 provides a description of the major results; Section 4 explores the interpretation and implication of the results. Conclusions then follow.

2. Methodologies

Performance persistence can be examined using various methodologies. They can be divided broadly into two strands. One type of studies has investigated performance persistence with the use of regression analysis. Using this approach, future performance is regressed against a measure of performance in the past, the examples include Grinblatt and Titman (1993) and Bers and Madura (2000). The regression results with a significant and negative slope coefficient indicate performance reversal while a positive and significant slope coefficient indicate performance persistence. Another broadly applied approach is the ranking approach. For example, Hendricks *et al.* (1993) and Carhart (1997) sort the funds based on returns over previous periods and evaluate the performance of the resulting portfolios. Elton *et al.* (1996) rank funds by past performance to examine whether the rankings are consistent over time, while Goetzmann and Ibbotson (1994), Brown and Goetzmann (1995), Malkiel (1995) and Khan and Rudd (1995) evaluate persistence through the use of contingency tables based on return rankings.

Of the various methodologies used to evaluate performance persistence, the contingency table approach is more appropriate for the situation when the sample of the funds is limited. Therefore, we adopt the contingency table approach.

2.1 The Risk-adjusted Returns

The initial step of our tests is to define the returns of unit trusts, for a given time interval and for each fund separately. The existing literature on performance persistence has employed both raw returns (e.g., Giles *et al.* (2002)) and risk-adjusted returns, derived from various kinds of asset pricing models. On average, we would expect a fund invests in high-risk stocks to have a higher raw return than a fund that invests in low-risk stocks. It therefore also implies that persistence in raw returns is to be expected unless there is evidence that either the risk carried by each fund varies unpredictably over time or the risks carried by all funds are the same. A more relevant method is to consider the risk-adjusted returns. However, these risk-adjusted methods are mostly confined to traditional pricing models, i.e. CAPM, multi-factor models.

As to the author's knowledge, there is limited work on performance persistence with the SDF model as an underlying model. Therefore, we apply the SDF primitive efficient model to measure the persistence. In addition, more recent research[56] has suggested that it is better to examine fund persistence based on style-adjusted returns because some abnormal performance could be due to the funds' specific styles rather than the fund managers' stock picking skills. It suggests that it is more sensible to examine persistence of unit trusts within their own style categories.

2.2 The Contingency Table Approach

As applied by many researchers, (see US: Hendricks, Patel and Zeckhauser (1993), Malkiel (1995), Brown and Goetzmann (1995) and Kahn and Rudd (1995); UK: Blake and Timmermann (1998)), the contingency table shows the probability of a fund in one quartile being in the same quartile in the following period. Assuming pure random performance, you would expect these probabilities to be 25%. That is, there is an equal chance of a top quartile fund ending up in any of the four quartiles in the

[56] For example, Teo and Woo (2001) examine persistence in fund performance relative to their peers in the Morningstar style categories. They argue that most funds with good returns are clustered into certain well-performing styles and that a large year-to-year variation in style returns may preclude finding persistence. Indeed, they find a strong evidence of persistence in style-adjusted performance measures based on several models including a four-factor model of Carhart (1997).

subsequent investment period as the past assessment period effectively has no effect on the future period.

For each category of the funds respectively, we apply contingency tables based on performance evaluation results (alphas) to measure the degree of persistence. Since the number of funds in our dataset is small, it is more sensible to divide all the funds into only two categories, Winner and Loser. Formally, we define the contingency table approach as a way used to identify the frequency with which funds are defined as winners and losers over successive time periods. The funds are ranked as two groups, namely Winner (*W*) and Loser (*L*). A winner (loser) decile is normally defined as having achieved a rate of return over the calendar year that exceeds (is lower than) the median fund return. *WW, LL, WL, LW* are defined based on the flowing table such as:

Winner/ Loser Contingency Table

period	*T+1*	
T	Winner	Loser
Winner	*WW*	*WL*
Loser	*LW*	*LL*

WW refers to the number of being both the winner this period and also the winner the following period; *LL* is the number of being a loser this period and the next period; *WL* is the number of being a winner this period followed by being a loser the next period and *LW* is the number of being a loser this period, then a winner next period.

Three types of statistical measures are used to test performance persistence of the funds.

Firstly, we adopt a non-parametric method called cross-product ratio (*CPR*), which was initiated by Brown (1995) and favoured by many researchers such as Teo and Woo (2001). The basic idea is based on performance evaluation, form the *CPR*, which reports the odds ratio of the number of repeat performers to the number of those that do not repeat, and test the statistical significance of *CPR*.

In more details, the *CPR* can be calculated as a ratio of 'Persistence' (*WW & LL*) versus 'reversal' (*WL & LW*) following the formulae such as:

$$CPR = \frac{WW \times LL}{WL \times LW} \qquad (6.1)$$

The standard error of the natural logarithm of the *CPR* is given by[57]

$$\sigma_{\ln(CPR)} = \left(\frac{1}{WW} + \frac{1}{WL} + \frac{1}{LW} + \frac{1}{LL}\right)^{1/2} \qquad (6.2)$$

We want to test the null hypothesis that there is no significant persistence, which should correspond to a *CPR* of one. This is because under the null hypothesis, the probability of winning or losing in each period equals one-half and is independent of the return horizon, in the sense *WW, LL, WL, LW* each has 25% of the funds.

To test:

[57] For more details, see Christensen, Ronald (1990) 'Log-linear models' (Springer-Verlag, New York).

H_0: $CPR = 1$ or $\ln CPR = 0$

A Z-statistics can be implemented such as:

$$Z_1 = \frac{\ln(CPR)}{\frac{\sigma_{\ln(CPR)}}{\sqrt{n}}} \sim N(0,1) \qquad (6.3)$$

A Z-statistic of 1.96 corresponds with a 5% significance level, i.e. when the Z-statistic is greater than 1.96, the null hypothesis of no persistence is rejected at 5% significance.

To measure the significance of 'persistence' or 'reversal', we can see if *CPR* is higher or lower than one. If *CPR* is significantly above one (equivalent to a positive t-statistic), it suggests 'Persistence', i.e. winners follow by winners; losers follow by losers.

On the contrary, a *CPR* below one (equivalent to a negative t-statistics while $WW*LL$ is less than $WL*LW$), reveals 'reversal'; i.e. winners followed by losers or losers followed by winners, in the sense that 'reversal' refers to a 'return reversal' situation.

The *CPR* ratio tests the persistence of both repeat winners (*WW*) and repeat losers (*LL*). To examine the performance of *WW* and *LL* separately, our second test focuses only on one quadrant of the contingency table *WW* or *LL* separately. We adopt the repeat winner approach of Malkiel (1995). This test shows the proportion of repeat winners (*WW*) to winner-losers (*WL*). Malkiel (1995) suggests that if p is the probability that a winner in one period continues to be a winner in the subsequent period, a value of p less than or equal to ½ indicates no persistence.

Thus, a binomial test of p >1/2 can be used to test the significance of the proportion of

WW to (WW+WL) as follows:

$$Z_2 = \frac{WW - (WW + WL) * p}{\sqrt{(WW + WL) p (1 - p)}} \qquad (6.4)$$

The test statistic is approximately normally distributed with mean of zero and standard deviation of one, when n is reasonably large. Thus, a percentage of WW to (WW+WL) above 50% and a Z-statistic above zero is indicative of performance persistence of the winners, while a percentage value below 50% and a Z-statistic below zero indicates a reversal in performance of the winners.

Similarly, we can apply this test to examine if the percentage of LL to (LL+LW) is above 50% as follows:

$$Z_3 = \frac{LL - (LL + LW) * p}{\sqrt{(LL + LW) p (1 - p)}} \qquad (6.5)$$

The results indicate a persistence or reversal performance of the losers. There are certain arguments regarding if short-term or long-term persistence tests are more sensible. Arguing that the real persistence should come from superior stock selection and market timing skills rather than luck, Toe and Too (2001) prefer testing longer-term persistence. Indeed, one-year fund performance rather than three-year performance could possibly contain more possibility of luck. In testing long-term persistence, it is possible to separate luck from potential managerial abilities. While Nicolas and Busse (2002) suggest that it is more appropriate to detect the short-term persistence because the superior performance is short-lived due to the competitive nature of the mutual fund industry or due to managerial turnover. Further, they provide sufficient evidence for this hypothesis by evaluating mutual funds several times a year. We believe both the examinations of short-term and long-term persistence are important, as the former has an implication for short-term momentum trading strategy, and the latter helps to investigate if the funds experience truly superior performance or are simply due to luck. In addition, we also implement a medium-term test to provide a complete picture of the funds' performance persistence.

3. The Results

We apply a sample with weekly data, including 11 growth funds, 12 general funds and 15 income funds. We employ continuous 12-month, 2-year and 5-year samples for short-term, medium-term and long-term persistent tests respectively. There are totally thirty 12-month, fifteen 2-year and six 5-year sub-samples over our sample periods from 1975 to 2003.

Firstly, for each sub-sample and each fund category separately, we employ weekly data to estimate the alphas using the SDF primitive efficient model. Based on the ranking of alphas, we measure the numbers of *WW, LL, WL* and *LW*. We then calculated *CPR, Z1, Z2* and *Z3*. The detailed results for each subsequent sample period and three kinds of funds are reported in Table 6-1. Panel A reports the short-term persistence (12 months), panel B reports the medium-term persistence (2 years), and panel C presents the results of the long-term persistence (5 years).

TABLE 6-1. PERSISTENCE TESTS BASED ON THE SDF PRIMITIVE EFFICIENT MODELS

This table reports the persistence tests of unit trusts using *CPR* ratios with 1-year, 3-year and 5-year sample period. Panel A reports the results of short-term persistence tests (1-year), Panel B reports those of medium term persistence (3-year) and Panel C reports those of long-term persistence (5-year). *LL*: Number of loser-loser within two continued sample periods; *WW*: Number of winner-winner within two continued sample periods; *LW*: Number of loser-winner within two continued sample periods; *WL*: Number of winner-loser within two continued sample periods; *CPR*: cross product ratio; σ: standard error of ln(*CPR*); *Z1* is the statistics for *CPR* ratio; *Z2* and *Z3* are z-statistics to test the significance of repeated winners and repeated losers.

Panel A. A test of Short-Term Persistence (1-year sample period)

Growth funds

	LL	*WW*	*LW*	*WL*	*CPR*	*σ*	*z1*	*z2*	*z3*
1&2	2	4	1	4	2	1.414	1.626	0	0.667
2&3	3	3	3	2	1.5	1.225	1.098	0.447	0
3&4	4	3	2	2	3	1.258	2.896	0.447	0.667
4&5	3	3	2	3	1.5	1.225	1.098	0	0.4
5*6	3	3	2	3	1.5	1.225	1.098	0	0.4
6&7	2	3	2	4	0.75	1.258	-0.758	-0.38	0
7&8	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
8&9	1	3	2	5	0.3	1.426	-2.8	-0.71	-0.67
9&10	3	3	3	2	1.5	1.225	1.098	0.447	0
10&11	2	3	2	4	0.75	1.258	-0.758	-0.38	0
11&12	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
12&13	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
13&14	3	2	4	2	0.75	1.258	-0.758	0	-0.29
14&15	3	3	2	3	1.5	1.225	1.098	0	0.4
15&16	3	3	2	3	1.5	1.225	1.098	0	0.4
16&17	2	3	1	5	1.2	1.426	0.424	-0.71	0.667
17&18	3	3	2	3	1.5	1.225	1.098	0	0.4
18&19	3	3	2	3	1.5	1.225	1.098	0	0.4
19&20	2	4	2	3	1.3333	1.258	0.758	0.378	0
20&21	3	3	2	3	1.5	1.225	1.098	0	0.4
21&22	4	2	3	2	1.3333	1.258	0.758	0	0.286
22&23	5	2	2	2	2.5	1.304	2.331	0	0.857
23&24	2	3	2	4	0.75	1.258	-0.758	-0.38	0
25&26	3	3	2	3	1.5	1.225	1.098	0	0.4
27&28	3	3	3	2	1.5	1.225	1.098	0.447	0
29&30	3	3	3	2	1.5	1.225	1.098	0.447	0
Aggregate	74	74	60	78	1.1701	0.238	2.191	-0.32	0.209

General funds

	LL	*WW*	*LW*	*WL*	*CPR*	*σ*	*z1*	*z2*	*z3*
1&2	4	2	3	3	0.8889	1.19	-0.343	-0.45	0.286
2&3	3	4	3	2	2	1.19	2.017	0.816	0
3&4	4	3	3	2	2	1.19	2.017	0.447	0.286
4&5	4	3	3	2	2	1.19	2.017	0.447	0.286
5&6	3	3	2	4	1.125	1.19	0.343	-0.38	0.4
6&7	4	4	2	2	4	1.225	3.921	0.816	0.667
7&8	3	2	3	4	0.5	1.19	-2.017	-0.82	0
8&9	4	2	4	2	1	1.225	0	0	0
9&10	3	2	5	2	0.6	1.238	-1.429	0	-0.5
10&11	4	3	3	2	2	1.19	2.017	0.447	0.286
11&12	5	2	3	2	1.6667	1.238	1.429	0	0.5
12&13	3	4	3	2	2	1.19	2.017	0.816	0
13&14	5	2	2	3	1.6667	1.238	1.429	-0.45	0.857
14&15	3	2	5	2	0.6	1.238	-1.429	0	-0.5
15&16	4	2	2	4	1	1.225	0	-0.82	0.667
16&17	4	2	4	2	1	1.225	0	0	0
17&18	3	3	3	3	1	1.155	0	0	0
18&19	4	3	2	3	2	1.19	2.017	0	0.667
19&20	3	3	3	3	1	1.155	0	0	0
20&21	4	3	2	3	2	1.19	2.017	0	0.667
21&22	3	2	4	3	0.5	1.19	-2.017	-0.45	-0.286
22&23	3	2	2	5	0.6	1.238	-1.429	-1.13	0.4
23&24	2	3	2	5	0.6	1.238	-1.429	-0.71	0
25&26	3	3	2	4	1.125	1.19	0.343	-0.38	0.4
27&28	3	4	3	2	2	1.19	2.017	0.816	0
29&30	5	3	2	2	3.75	1.238	3.698	0.447	0.857
Aggregate	93	71	75	73	1.206	0.228	2.849	-0.17	0.214

Income funds

	LL	*WW*	*LW*	*WL*	*CPR*	*σ*	*z1*	*z2*	*z3*
1&2	3	4	3	5	0.8	1.0567	-0.818	-0.333	0
2&3	4	3	3	5	0.8	1.0567	-0.818	-0.707	0.2857
3&4	4	3	2	6	1	1.118	0	-1	0.6667
4&5	4	3	3	5	0.8	1.0567	-0.818	-0.707	0.2857
5&6	5	4	2	4	2.5	1.0954	3.2396	0	0.8571
6&7	4	3	2	6	1	1.118	0	-1	0.6667
7&8	3	4	3	5	0.8	1.0567	-0.818	-0.333	0
8&9	4	3	5	3	0.8	1.0567	-0.818	0	-0.222
9&10	5	4	3	3	2.2222	1.0567	2.9266	0.378	0.5
10&11	3	4	2	6	1	1.118	0	-0.632	0.4
11&12	3	4	3	5	0.8	1.0567	-0.818	-0.333	0
12&13	4	5	3	3	2.2222	1.0567	2.9266	0.7071	0.2857
13&14	4	4	2	5	1.6	1.0954	1.6617	-0.333	0.6667
14&15	4	3	2	6	1	1.118	0	-1	0.6667
15&16	5	3	2	5	1.5	1.1106	1.414	-0.707	0.8571
16&17	4	3	3	5	0.8	1.0567	-0.818	-0.707	0.2857
17&18	5	3	2	5	1.5	1.1106	1.414	-0.707	0.8571
18&19	4	5	2	4	2.5	1.0954	3.2396	0.3333	0.6667
19&20	5	4	3	3	2.2222	1.0567	2.9266	0.378	0.5
20&21	4	3	2	6	1	1.118	0	-1	0.6667
21&22	4	3	3	5	0.8	1.0567	-0.818	-0.707	0.2857
22&23	3	5	2	5	1.5	1.1106	1.414	0	0.4
23&24	4	5	3	3	2.2222	1.0567	2.9266	0.7071	0.2857
25&26	3	5	3	4	1.25	1.0567	0.8178	0.3333	0
27&28	3	4	2	6	1	1.118	0	-0.632	0.4
29&30	5	3	2	5	1.5	1.1106	1.414	-0.707	0.8571
Aggregate	103	97	67	123	1.2124	0.2075	3.5935	-1.753	0.4235

Panel B. A test of Medium Term Persistence (2-year sample period)

	LL	*WW*	*LW*	*WL*	*CPR*	*σ*	*z1*	*z2*	*z3*
Growth Funds									
1&2	4	2	3	2	1.3333	1.258	0.758	0	0.286
2&3	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
3&4	4	3	2	2	3	1.258	2.896	0.447	0.667
4&5	2	3	2	4	0.75	1.258	-0.758	-0.38	0
5*6	3	4	2	2	3	1.258	2.896	0.816	0.4
6&7	4	3	2	2	3	1.258	2.896	0.447	0.667
7&8	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
8&9	2	3	2	4	0.75	1.258	-0.758	-0.38	0
9&10	1	3	2	5	0.3	1.426	-2.8	-0.71	-0.67
10&11	2	3	2	4	0.75	1.258	-0.758	-0.38	0
11&12	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
12&13	2	3	3	3	0.6667	1.225	-1.098	0	-0.4
13&14	3	2	3	3	0.6667	1.225	-1.098	-0.45	0
14&15	4	3	2	2	3	1.258	2.896	0.447	0.667
Aggregate	40	38	34	42	1.0644	0.323	0.64	-0.45	0.162
General Funds									
1&2	2	2	4	4	0.25	1.225	-3.92	-0.82	-0.667
2&3	5	4	1	2	10	1.396	5.712	0.816	1.333
3&4	4	3	3	2	2	1.190	2.017	0.447	0.286
4&5	3	3	3	3	1	1.155	0	0	0
5*6	2	2	4	4	0.25	1.225	-3.92	-0.82	-0.667
6&7	5	5	1	1	25	1.549	7.198	1.633	1.333
7&8	1	1	5	5	0.04	1.549	-7.2	-1.63	-1.333
8&9	4	4	2	2	4	1.225	3.921	0.816	0.667
9&10	4	3	3	2	2	1.190	2.017	0.447	0.286
10&11	3	3	3	3	1	1.155	0	0	0
11&12	2	2	4	4	0.25	1.225	-3.92	-0.82	-0.667
12&13	2	2	4	4	0.25	1.225	-3.92	-0.82	-0.667
13&14	3	3	3	3	1	1.155	0	0	0
14&15	4	3	2	3	2	1.190	2.017	0	0.667
Aggregate	44	40	42	42	0.9977	0.309	-0.03	-0.22	0.047
Income Funds									
1&2	4	3	4	4	0.75	1.041	-1.07	-0.378	0
2&3	4	3	4	4	0.75	1.041	-1.07	-0.378	0
3&4	4	3	4	4	0.75	1.041	-1.07	-0.378	0
4&5	3	5	2	5	1.5	1.111	1.414	0	0.4
5&6	3	4	4	4	0.75	1.041	-1.07	0	-0.286
6&7	5	3	4	3	1.25	1.057	0.8178	0	0.2222
7&8	3	2	5	5	0.24	1.111	-4.977	-1.134	-0.5
8&9	4	5	4	2	2.5	1.095	3.2396	1.1339	0
9&10	5	4	3	3	2.2222	1.057	2.9266	0.378	0.5
10&11	4	3	4	4	0.75	1.041	-1.07	-0.378	0
11&12	6	4	3	2	4	1.118	4.8023	0.8165	0.6667
12&13	4	4	3	4	1.3333	1.041	1.0705	0	0.2857
13&14	4	3	4	4	0.75	1.041	-1.07	-0.378	0
14&15	5	4	2	4	2.5	1.095	3.2396	0	0.8571
Aggregate	58	50	50	52	1.1154	0.277	1.5294	-0.198	0.1481

Panel C: A test of long-term persistence (5-year sample period)

	LL	*WW*	*LW*	*WL*	*CPR*	*σ*	*z1*	*z2*	*z3*
Growth Funds									
1&2	2	3	3	3	0.667	1.225	-1.098	0	-0.4
2&3	3	3	3	2	1.500	1.225	1.098	0.447	0
3&4	2	3	3	3	0.667	1.225	-1.098	0	-0.4
4&5	3	2	3	3	0.667	1.225	-1.098	-0.45	0
5*6	2	2	3	4	0.333	1.258	-2.896	-0.82	-0.4
Aggregate	12	13	15	15	0.693	0.542	-2.242	-0.38	-0.22
General Funds									
1&2	3	4	3	2	2.000	1.19	2.017	0.816	0
2&3	2	3	3	4	0.500	1.19	-2.02	-0.38	-0.4
3&4	2	3	2	5	0.600	1.238	-1.43	-0.71	0
4&5	3	2	4	3	0.500	1.19	-2.02	-0.45	-0.286
5*6	3	3	3	3	1.000	1.155	0	0	0
Aggregate	13	15	15	17	0.765	0.519	-1.79	-0.35	-0.143
Income Funds									
1&2	4	4	3	4	1.333	1.0408	1.0705	0	0.2857
2&3	3	2	3	7	0.286	1.1443	-4.24	-1.667	0
3&4	4	3	4	4	0.750	1.0408	-1.07	-0.378	0
4&5	3	4	4	4	0.750	1.0408	-1.07	0	-0.286
5&6	3	3	4	5	0.450	1.0567	-2.927	-0.707	-0.286
Aggregate	17	16	18	24	0.630	0.4675	-3.833	-1.265	-0.057

Table 6-2 reports the summarized aggregate results, for each type of funds separately. It shows that for the short-term persistence test (1-year), based on the risk-adjusted return measure by using the SDF primitive efficient model, there is strong evidence of short-term persistence for all three types of funds. (i.e. for growth funds, *CPR* = 1.1701, *Z1* = 2.191; for general funds, *CPR* = 1.206, *Z1* = 2.849; for income funds, *CPR* = 1.2124, *Z1* = 3.5935).

Moreover, the repeat winner percentage out of the combination of *WW* and *WL* is mostly less than 50%, whereas the repeat loser percentages are all higher than 50%, which implies that bad performers have a higher chance of remaining at the bottom whilst for the winners, the chances for them to remain top are slightly less than the probability for them to be a loser in the next period. This result also implies that the short-term performance persistence is mainly due to the persistence of the bottom performers.

For the medium-term persistence test (2-year), there is little evidence of performance persistence though *CPR* ratios are above one in most of the cases (the *CPRs* statistics are: Growth fund: 1.0644 ($Z1$= 0.64); General fund: 0.9977 ($Z1$ = -0.03); Income fund: 1.1154 ($Z1$ = 1.5294)). The repeat winner percentage and the repeat loser percentage are all around about 50% and the $Z2$ and $Z3$ statistics are not significant most of the time.

For the long-term persistence test (5-year), the *CPR* ratios are negative for all the cases and both risk-adjusted measures. Except for general funds, $Z1$ statistics show the null hypothesis is rejected at the 5% significance level. ($Z1$= -2.242 for growth fund and -3.833 for income fund), indicating very strong evidence of return reversal. $Z2$ and $Z3$ statistics are mostly negative, and repeating percentage are less than 50% for both *WW* and *LL* tests. It indicates that both the winners and the losers tend to generate reversed performance over the long run.

To summarise, our results suggest significant evidence of short-term persistence and long-term reversal. The interpretation and implication of the results are discussed in the following section.

TABLE 6-2 SUMMARIZED RESULTS OF PERFORMANCE PERSISTENCE

The table report the aggregate results of performance persistence test of table 6-1. The significant values of those tests are shown in bold.

Period of Evaluation	1-year	2-year	5-year
Growth funds			
LL	74	40	12
WW	74	38	13
LW	60	34	15
WL	78	42	15
CPR	1.170	1.064	0.693
Z1	**2.191**	0.64	**-2.242**
Repeat Winner %	48.68	47.5	46.43
Z2	-0.32	-0.45	-0.38
Repeat Loser %	55.22	54.05	44.44
Z3	0.209	0.162	-0.22
General funds			
LL	93	44	13
WW	71	40	15
LW	75	42	15
WL	73	42	17
CPR	1.206	0.9977	0.7647
Z1	**2.849**	-0.03	-1.79
Repeat Winner %	49.305	48.78	46.88
Z2	-0.17	-0.22	-0.35
Repeat Loser %	55.357	51.162	46.42
Z3	0.214	0.047	-0.143
Income funds			
LL	103	58	17
WW	97	50	16
LW	67	50	18
WL	123	52	24
CPR	1.212	1.115	0.630
Z1	**3.5935**	1.5294	**-3.833**
Repeat Winner %	44.09	49.02	40
Z2	-1.753	-0.198	-1.265
Repeat Loser %	60.58	53.7	48.57
Z3	0.4235	0.1481	-0.057

4. The Results Interpretation and Implication

The interpretation of 'short-term momentum' and 'long-term reversal' can be broadly divided into three groups: one group believes that these anomalies are not real evidence against 'Market Efficiency Hypothesis'. The other two groups believe those anomalies indicate the equity market is inefficient. One assures that the persistent top-performed funds have truly superior ability of selecting winning shares whilst the bottom funds have consistent inferior ability. They are referred as fund managers with 'hot hand' and 'icy hand' respectively. Another group suggests that profits from anomalies arise because the investors interpret information with certain biases.

Other explanations also include the belief that the empirical results, both with or without 'performance persistence' evidence, can be spurious due to reasons such as 'survivorship bias', the existence of expense ratios and change of a portfolio manager.

The details of each type of interpretations and implications follow.

4.1 The Efficient Market

The first group firmly believes that the market is still efficient even though evidence has been found against it. As represented by Lo and MacKinlay (1990), Jegadeesh and Titman (1993) and Conrad and Kaul (1998), who support the idea that the profitability of 'momentum' and 'contrarian' strategies might simply be generated as the compensation for risk. In their view, these superior returns do not survive risk adjustment.

Other examples also include Fama and French (1988) and Ball, Kothari, and Shanken

(1995), who investigate the profitability of contrarian strategy and claim that the risk-price disparity and the asymmetric reverting pattern reflect the pricing of stocks by investors who react rationally to changing volatility. Conrad and Kaul (1998) also suggest the profitability from momentum strategies resulted entirely from cross-sectional variation in expected returns. Carhart (1997) argue that the superior performance of top funds can be explained by 'momentum factor'. Alternatively, Fama (1998) believes that anomalies happen only by chance by stating that, '*the expected value of abnormal returns is zero, but chance generates apparent anomalies that split randomly between over-reaction and under-reaction.*' In a sense that if a sufficient number of empirical tests are performed on a complex system, some will naturally contain surprising results due to chance. However, it might also be difficult to decide the matter definitively, given the tools currently available for econometric studies. While such arguments and the way of searching for a reliable 'risk' explanatory factor can be endless, we address the interpretations from other angles.

4.2 'Hot Hand' And 'Icy Hand'

In 1990s, two papers claim to have isolated a "hot hand" phenomenon based on the study focused on growth funds persistence. Hendricks, Patel and Zeckhauser (1993) and Goetzmann and Ibbotson (1994) suggest that previous mutual fund performance can be used to predict future mutual fund returns and this performance persistence phenomenon appears also robust to different types of risk adjustment measures. They attribute persistence to the managerial skills of the fund managers. The portfolio managers with superior managerial ability are regarded as having "hot hand", with which the funds can generate excess returns year after year. While the losing portfolio managers are believed to have "icy hand" in that their lost persistence is significant.

These results imply that historical performance can be used to predict future funds performance and suggest that investors could earn significant risk-adjusted returns by purchasing funds with good performance track record.

Since the 'Persistence' we identified only exist over the short term, and significance of persistence declines as we extend the length of the testing period. (Medium-term persistence is not significant). It implies that fund managers have not got superior stock picking abilities because they may coincidently hold last-year high return stocks or just follow the momentum strategy. It is similar to the results discovered by Wermer (1995), who designs a statistic to detect the extent to which momentum phenomeon exists and how this phenomenon correlates with fund performance. He finds that the relation between the tendency of buying last year winners and performance is strong. He also provides apparent evidence for herding tendency among mutual funds.

4.3 Irrational Behaviour

The field of behavioural finance merges concepts from financial economics and cognitive psychology in an attempt to construct a more detailed model of human behaviour in financial markets. 'Behaviour Finance' is the study of how investors systematically make errors in judgment, and has been applied widely to help explain 'financial empirical anomalies'. Examples include Chopra, Lakonishok, and Ritter (1992), Jones (1993), Balvers, Wu, and Gilliland (2000), and Nam, Pyun, and Avard (2001), who support DeBondt and Thaler's (1985, 1987) over-reaction hypothesis, affirming systematic mispricing and attendant risk adjusted excess returns that can be exploited by contrarian strategists.

Barberis, Shleifer and Vishny (1998), and Hong and Sei (1999) also employ behavioural models to interpret momentum phenomenon. Their results suggest that abnormal returns could arise over the holding period, because of a delayed over-reaction to information that pushes the prices of winners/losers above/below their long-term values. Therefore, their models predict, during the post-holding period, the returns of losers would exceed the returns of winners when the stock prices of the winners and the losers revert to their fundamental values.

Furthermore, Jegadeesh and Titman (2001) evaluate various explanations for the profitability of momentum strategies documented previously by Jegadeesh and Titman (1993). Their evidence support that momentum profits are mainly due to delayed over-reactions that are eventually reversed.

The 'Behavioural Finance' approach might shed some new light on the interpretation of our results. In this section, we explore the interpretation of 'short-term momentum' and 'long-term reversal' from this aspect.

4.31 The types of behavioural biases

The behavioural biases are broadly divided into two categories: non wealth-maximizing behaviour and heuristic decision processes.

Non Wealth-Maximizing Behaviour

It is suggested that sometimes, investors could behave in a way to maximize not their wealth, but some other more important things. Examples include "window dressing" at the end of a quarter or a year. Selling stocks just before the end of a quarter/a year which have been big losers and buying stocks that have been big winners will not raise the portfolio's return and the associated transaction costs of trading may actually lower the return. To reduce exposures to highly visible "losers" in the portfolio just before the quarterly/annual meeting, portfolio managers can have an easier time, avoiding difficult questions from superiors and investors.

Prospect Theory (see Kahneman and Tversky (1979)) groups another class of problems. It describes several states of mind that can be expected to influence an individual's decision making processes. The key concepts addressed by the theory include loss aversion, regret aversion, mental accounting and lack of self control.

Loss aversion refers to the situation that the penalty associated with a given loss mentally is larger than the reward from a gain of the same size mentally. If investors are loss adverse, it is possible that they may be reluctant to realize the losses.

Regret aversion indicates the situation that for a lot of investors, the mental penalty associated with losses exceeds the mental pleasure of gains. As a result, people would want to avoid feeling the pain of regret associated with a poor investment decision. It is normally more than the pain of financial loss and includes the pain of feeling responsible for the investment decisions they made which has led to the financial loss. It can firstly encourage investors to hold poorly performing shares. Without in fact selling these shares, it helps to avoid the recognition of the associated losses. As a result, it prevents the investors to make rational decisions with existing losing positions. It is also not helpful for making new investment decisions as the investors may be less willing to invest in investments or markets which have performed poorly in the recent past.

Mental accounting is the fact that investors tend to treat each element of their investment portfolio separately and put things into separate 'accounts' mentally. Mental accounting can lead to irrational decision making and make people vary in their attitudes to risk between their mental accounts. For instance, investors may be risk adverse in their downside protection accounts and risk seeking in their more speculative accounts.

Lack of self-control is also quite common. As noted by Thaler and Shefrin (1981), investors look for tools to improve self-control as most of them are subject to temptation of consumption. Investors however, can control their urge to over-consume by separating their financial resources into capital and 'available for expenditure' pools mentally.

Heuristic Decision Processes

Heuristics (mental shortcuts or rules of thumb) is what the human brain uses to made decisions in uncertain and complex environments. Heuristics are usually useful problem solving tools. However, heuristics can cause people to make systematic mental mistakes if used in the wrong situation. Typical examples resulting from the use of heuristics include:

Representativeness refers to the situation that people tend to make decisions based on stereotypes. People have the tendency to estimate the probability of an uncertain event in the future by the degree to which it is similar to the events observed recently. This can also be related to the situation when investors tend to avoid stocks which have performed poorly in the recent past and seek to buy recently outperformed shares. According to DeBondt and Thaler (1985), such irrational behaviour could help to explain why investors over-react to new information, i.e. they become overly pessimistic about past losers and too optimistic about past winners. Such bias could cause prices to deviate from their fundamental level.

Overconfidence People are grossly overconfident regarding their knowledge and ability. It can cause investors to under-react to new information and leads investors to overestimate their predictive skills. Studies have shown that one side effect of investor overconfidence is excessive trading. Besides individual investors, overconfidence might also occur to fund managers and professional analysts, who are shown to be slow to revise their previous assessment of a company's likely future performance.

Saliency It is well recognized that for events which occur infrequently, if people have recently observed such an event, they tend to overestimate the frequency of such an event occurring again in the future. The examples include the case of airplane crashes and terrorist attacks. Saliency can also cause investors to overreact to new information.

Anchoring arises when a value scale is fixed or anchored by recent observations. It has been documented by psychologists that people's estimates can be heavily influenced by previous values of the item when people make quantitative estimates, which might lead investors to expect a company's earnings to be in line with the historical trend or for a share to continue to trade in a defined range.

Availability bias refers to the situation when people place more weight and importance on easily available information when they make a decision.

Gamblers' fallacy emerges when people anticipate that a trend will end and reverse without strong evidence. This bias could lead investors to predict the end of a run of good /poor market returns inappropriately. It can also be considered to be a belief in regression to the mean.

4.32 The Explanations

As mentioned above, the investors, including professional investors, can suffer from non-wealth maximizing behaviour or may occasionally under- or over-react to information due to heuristic biases; 'short-term momentum' and 'long-term reversal' can be partly explained by such irrational behaviour.

Other researchers who have investigated the similar issues include Barberis, Shleifer and Vishny (BSV 1998) and Daniel, Hirshleifer and Subramanyam (DHS, 1997). Their interpretations are basically summarized in the following table. As commented by Fama (1998) and Hong and Stein (1997), the predictions made by both DHS and BSV models are close, and they also share the same empirical successes and failures such as the prediction of long-term return reversal does not capture the range of long-term results observed in the literature.

Model	Psychological Bias	Biased Behaviour	Effects on Shares Price Pattern
BSV (Barberis, Shleifer & Vishny 1998)	Representative	Too much weight on recent patterns	Overreact--long-term **Reversal**
	Conservatism	Slow updating of models with new evidence	Under react--Short-term **Momentum**
DHS(Daniel, Hirshleifer & Subramanyam 1997)	Overconfidence	exaggerate the precision of private signals about values	Overreact to private information; under react to public information—**Momentum** and long-term **Reversal**
	Biased self-attribution	Down-weight public signals about values	

The over-reaction hypothesis suggests that security prices could overreact to consistent patterns of information pointing in the same direction. For example, securities that have had a long record of good news tend to become overpriced and have low average returns afterwards. Securities with a few years of outperformance receive extremely high valuations, and these valuations, on average, return to the mean. It is argued that over-reaction can result in return 'long-term reversal' and probably the source of alphas for most 'contrarian strategies'.

The representativeness and saliency heuristics can cause investors to overreact to new information. According to BSV (1998), investors tend to perceive that there are two earnings regimes; one assumes earnings are trending while the other believes earnings are reverting. The former suffers from representative bias by giving too high weights to recent patterns in the data. They believe the trending regime holds, incorrectly extrapolate the trend and therefore the stock price over-reacts. Once the over-reaction is exposed by future earnings, the long-term reversal occurs. Whilst DHS (1997) adopt a different behavioural foundation by assuming there are two types of investors, informed, who determine the share prices and subject to two biases, overconfidence and biased self-attribution and uninformed, who are not subject to judgment biases. The informed tend to overreact to the private information of a share's value and under-react to public information of a share's value. This behaviour tends to produce short-term momentum of stock returns but long-term reversals as public information eventually overwhelms the behavioural biases.

'Regret theory' can also be used to explain contrarian investment strategies based on the belief of 'long-term reversal', since investing in poor performing stocks is not prudent and consequently is not undertaken.

BSV (1998) explains that a stock's price tend to under-react to a change in earnings if the investors perceive the earnings are mean reverting because investors could believe the change is likely to be temporary. When such expectation is not confirmed by later earnings, stock prices show a delayed response to earlier earnings. Therefore, consistent under-reaction would generate momentum evidence, as described above, since the short-horizon trend in returns may reflect slow incorporation of information into stock prices.

4.4 The Spurious Results

The persistence might also be regarded as a spurious phenomenon. Brown *et al.* (1992) argue that survivorship bias provides a potential explanation for short-term persistence. That is, 'survivorship bias' can bring about too optimistic results because poorly performing funds disappear more frequently from the mutual fund universe.

Moreover, it is argued that the existence of expense ratios that vary over the universe of funds can also produce some persistence in returns. Suppose a number of funds hold something approximating the market portfolio, then the fund with the lowest expense ratio is likely persistently to outperform high expense funds with the same investment strategy. In the case the persistent effect is entirely driven by different charges, an efficient investment strategy would be simply to select those funds with the lowest charges. It might be argued that expenses/charges are significant correlated with the risk. Giles *et al*. (2002) examines the relationship between risk and charges for UK equity Unit Trusts and their results show that the degree of charges had a low level of explanatory power in the case of UK All companies (growth), and UK Equity Income (income). On the other hand, Chevalier and Ellison (1999) discover that systematic differences in the jobs held by different types of managers, which result in having different expense ratios, can account for the major part of short-term persistence. All of these might suggest that we should take expense ratios into account, possibly as an explanatory variable for risk. We therefore do not rule out the possibility that differential expenses provide a reason of our results.

On the other hand, it is argued that the following elements could contribute to a lack of ‘persistence’. Firstly, it could be possible that the same portfolio manager applies a different investment process as he/she believed previous investment process did not work. Secondly, it can be because of the internal change of the portfolio managers or the departure of a key investment decision maker. Thirdly, it is understood that different trading strategies and investment skills may suit portfolios with different sizes (total asset under management) better, and the sizes of the portfolios normally evolve over the years.

5. Conclusion

The aim of 'Performance persistence' tests is to examine if a fund manager can generate superior returns relative to comparable funds/benchmark, for consecutive time periods. The existence of persistence implies that fund selection decisions can be made based on the funds' performance track record. It therefore helps private investor to make fund selection decision. It also has implication for possible profitable trading strategies applied to the funds, such as well documented 'momentum' and 'contrarian' strategies.

Based on a most-widely applied ranking approach, we examine the short-term, medium-term and long-term persistence of UK unit trusts using the 'alphas' derived from the SDF primitive efficient model. Our results suggest modest evidence of short-term 'persistence' and long-term 'reversal'. We interpret our results from several different angles. Firstly, we believe these results can hardly be explained by 'market efficiency' theory, though it might suggest a hidden risk factor(s) can fully explain our results while the searching for such a risk factor(s) can be infinite. Secondly, our evidence is not sufficient to support the existence of 'hot hand' or 'icy hand'. Thirdly, perceiving that investors' irrationality has significant effects on markets and price movements, and with attempts to better understand and explain how emotions and cognitive errors influence investors and the decision-making process, 'Behaviour Finance' approach provides new insights on our investigation. Finally, it is also possible that evidence of persistence could be spurious due to survivorship bias, evidence of expense ratios and portfolio manager turnover.

CHAPTER 7: CONCLUSION

Can superior returns be generated by active managers? This has been a popular question at both academic and institutional levels. The thesis looks at this matter within the context of the UK market. It focuses on the implication of various SDF models to evaluate the performance of UK unit trusts. The unit trusts managers, known as the portfolio managers, are expected to have superior ability to generate above market average returns for a given risk level, and to diversify the portfolio to eliminate unsystematic risks. The purpose of the study is to see if better returns can be generated by portfolio managers who are able to collect and interpret information that helps to forecast the securities' returns.

We are interested in finding out if the portfolio managers as a group can add values to the portfolios they manage, or whether it is the case that they only generate wasteful transaction costs through active management. It is especially relevant given recent financial crisis since June 2007, when two of Bear Stearns' largest hedge funds collapsed which triggered the worst global economic crisis since the Great Depression of 1930s. Post crisis, policy makers around the globe introduced a series of regulatory proposals, including consumer protection, bank capital requirements, executive pay and expanded regulation of the shadow banking system. The regulatory proposals on the shadow banking system (investment banks, hedge funds, mutual funds) have focused on the extensive usage of derivatives and the remuneration scheme for bankers working in the financial industry. Our empirical examination of mutual fund performance would help to raise the same question on if the portfolio managers' above-average remuneration is justified.

It is also an examination of the semi-strong form of market efficiency hypothesis. It is believed that if the market is efficient, a fund manager can not add value if he/she only trade based on publicly available information. To be able to add value and generate a positive conditional alpha, a fund manager will have to produce a higher return than the strategies based on public information.

Our studies can also be useful for individual fund investors as the study can shed light on how to select a portfolio manager with the ability to add value to the portfolio he/she manages, i.e. evaluating fund performance with taking conditioning information into account; fund styles may explain the dispersed fund performance; previous performance track records are not sufficient to judge the capability of the fund managers as the funds may not perform consistently.

The essential methodology underlying the whole thesis is the SDF asset pricing approach, estimated by GMM. The SDF is best known for its following characteristics: the value of a financial security equals the expected value of the product of the payoff on the asset and the SDF. The SDF model estimates the asset-pricing model with its SDF representations based on GMM estimation. A pricing model identifies an SDF model with specification of model parameters and observable variables, for example, the consumption-based CAPM uses the inter-temporal marginal rate of substitution (IMRS) of the consumer, a linear factor pricing model identifies a linear function of the factors as an SDF.

Extensive research has evaluated the US mutual fund performance, which have led to mixed results. However, the research (especially those based on the SDF primitive models) on the UK market remains relatively sparse, with the main focuses on performance persistence. The performance of the UK unit trusts is therefore a particularly interesting topic because of the potential of this research area and the important role of this market. This thesis provides an in-depth analysis of performance evaluation in the context of the UK unit trusts. It extends the existing portfolio performance evaluation literature on the UK unit trusts market, with focuses on a few under-explored areas: a pure econometric matter, the effects of weighting matrices choices on performance evaluation; conditional performance evaluation within the SDF primitive efficient models; the impact of different choices of information variables on conditional evaluation; the impact of styles on performance; style-rotation strategies and performance persistence analysis. Though our dataset stops in 2003, our analysis is still relevant for current policymakers, financial market participants and academic researchers.

The first chapter illustrates the purpose of the study and provides an overview of UK asset management industry and introduces the dataset. The second chapter gives a detailed literature review of performance evaluation, with a focus on the SDF models and conditional evaluation methods. It firstly explains the methodologies, from a traditional CAPM model to the multi-factor models, widely implemented in previous research. It then provides a detailed survey of the literature on performance attribution and performance persistence.

The main part of the thesis consists of four projects. We believe that before the formal examination of mutual fund performance, it is essential to search for the optimal estimator with the best small sample properties, hence the first project deals with this issue. As discussed in chapter 3, it tests the small sample properties of the GMM iterated and 2-step estimators within the framework of the SDF primitive efficient models. Based on the small sample properties results from chapter 3, the second project, presented in chapter 4, employs the optimal estimation method for each SDF model to evaluate the performance of UK unit trusts, the effects of conditioning information are also investigated. The following chapter examines performance attribution based on style analysis. It examines relative style performance and investigates if style consistency or style rotation prevails. With evidence of rejecting the existence of style consistency, different kinds of style-rotation strategies are constructed and tested. Furthermore, we examine how persistent the performance can last in chapter 6.

In detail, the first project focuses on a pure econometric issue. The GMM technique is chosen as the estimation method for our projects. Since our dataset is relatively small, it is essential to test the econometric properties of the GMM estimators in the case of small samples. Therefore, before starting portfolio evaluation projects formally, we firstly investigate the small sample properties of GMM iterated and 2-step estimators within the framework of the SDF models, with an aim to identify the best GMM estimator for the asset pricing models. It is striking to see for small sample sizes, the empirical critical values are necessary for a reasonable rejection of the null hypothesis. We simulate the empirical critical values of GMM *J*-statistic for a series of small sample sizes for each evaluation model. The comparisons have been drawn between two GMM estimators by investigating the power to reject non-neutral performance. Our simulation results suggest that for both GMM 2-step and iterated estimators, the sizes of *J*-statistics can be seriously distorted, whilst the GMM iterated estimator generates superior size and higher power properties compared to its counterpart. Based on the results, the optimal estimators for the asset pricing models and empirical critical values for small sample sizes are employed for all the empirical tests in the following several chapters.

For the period from Jan 1975 to Oct 2003, the second project applies the SDF primitive efficient models to evaluate the performance of UK unit trusts. Dealing with scaling returns and scaling factors in various ways, we applied five types of models to incorporate the conditioning information. We also report the pricing errors based on the the *HJ* distance measure, *J*-statistic and mean square errors in order to examine how accurate the pricing models are and to compare the performance of different asset pricing models. The fund performance measures derived from alternative model specifications differ depending on the number of instrumental variables used to scale assets/factors. However, by and large, for all the cases and all types of funds, more than half of the funds do not have significant abnormal performance. Among those demonstrating significant non-neutral performance, most of the unit trusts show inferior performance. On average, unit trusts can not generate excess returns relative to their benchmarks that are large enough to cover their total expenses. We also confirm that conditional models make the average performance of UK equity funds look worse than the performance derived based on the traditional unconditional methods. The results reveal that conditional models generate negative alphas more frequently than their unconditional model counterparts. These results are further confirmed by paired sample *t*-tests. Compared to unconditional alphas, fund performance deteriorates when we measure their conditional alphas. Given that based on public information, the share prices/returns are to some extent predictable, conditional performance evaluation brings up the level of reference/benchmark for active fund managers because the conditional measures do not give any credit for exploring publicly available information. Our results are useful for policy issues relating to UK unit trusts, i.e. recent discussion on bankers' (including portfolio managers') fair remuneration. Before taking all the transaction costs into account, on average, UK equity portfolio managers do not have superior ability to generate better returns than the benchmark, especially on a conditional evaluation basis. Such empirical evidence helps to challenge the justification for bankers' remuneration being much higher than the British average.

To investigate the attributes of mutual fund performance, the third project examines if the styles of UK unit trusts can explain the divergence in performance. In recent years, many studies and debates have focused on investigating the average return differences between styles, such as the difference between growth and value stocks. The arguments mainly focus on if each style offers its own benefits or if value funds can provide the long-term benefits consistently? Is style diversification the optimal solution to avoid the risk associated with pure style investing? We extend the debates by examining style performance of the UK unit trusts. Employing both risk-unadjusted and risk-adjusted returns based on the SDF primitive efficient models, the unit trusts performance with three types of styles are examined respectively. We compare and rank the performance of these funds over different sub-sample periods. The results suggest that among the UK unit trusts, the performance of the growth and income funds vary over time. While over a long investment horizon, there is a lack of differential performance among them. More specifically, growth funds perform better in the bull markets while income funds do better in the bear markets. Furthermore, we employ a few methods to construct the style strategies, including a simple market-switching strategy and the trading strategies based on the sensitivity test results, using both OLS regression and the logit models. The profits are measured, compared to the passive and perfect-forecasted returns. The sensitivity tests suggest that some of these factors are significantly correlated with the return spreads. The trading strategy results based on logit models suggest that the winning strategies and profits over a passive portfolio is achievable based on around 60% prediction accuracy rate of future value spread, given a low level of transaction costs. It suggests controlled style-rotation strategies, based on the underlying fundamental characteristics of the relevant macroeconomic factors, can potentially enhance values, given relatively low levels of transaction costs.

Besides performance evaluation and style analysis, it is also interesting to examine how long the performance, good or bad, will last? The final project examines performance persistence of the UK unit trusts. The test of 'performance persistence' refers to an examination of the ability of a fund to attain returns above the median, relative to comparable funds, for consecutive time periods. It not only has implications for private investors' investment decision making as the existence of persistence imply fund selection decisions can be made based on the funds' past performance. It also has implication for possible profitable trading strategies applied to the funds, such as the well documented 'momentum' and 'contrarian' strategies. Based on a most-widely applied ranking approach, we examine the short-term, medium-term and long-term persistence of UK unit trusts using the 'alphas' derived from the SDF primitive efficient model. Our results suggest modest evidence of short-term 'persistence' and long-term 'reversal'. The interpretation of such results follows:

firstly, we believe these results can be hardly explained by 'market efficiency' theory, though it might suggest that a hidden risk factor(s) can fully explain the results while the search for such a risk factor(s) can be infinite. Secondly, our evidence is not sufficient to support the existence of 'hot hand' or 'icy hand'. Thirdly, we believe the 'Behavioural Finance' approach, which examines the relationship of the irrational behaviour of the investors and the security price movements, can provide new insights on our investigation. Finally, it is also possible that evidence of persistence could be spurious because of survivorship bias, different levels of transaction costs and portfolio manager turnover.

Due to the limitation of data resources and timeframe, this thesis has only used the samples of UK Equity unit trusts, within the framework of the SDF primitive efficient models. Hence there is abundant scope for further research exploration:

Firstly, different asset classes can exhibit different degrees of market efficiency, contributed by various elements, for example, the breadth of the asset class and market, analyst coverage, accessibility of derivatives, transparency, liquidity levels and trading costs. For example, Asian real estate has low accessibility and liquidity of derivative compared to EMU government bonds, which would suggest a higher probability for market inefficiencies to be discovered in Asia real estate than in EMU government bonds. Active management of the US equity could be quickly discovered by the extensive analysts' coverage, whilst limited analyst coverage on emerging market corporate debt makes this asset class more attractive for alpha seekers. Therefore, it is interesting to examine the market efficiency of different asset classes. Within equity funds, it is also interesting to examine the equity fund performance outside of the developed markets, i.e. emerging market equities. Compared to UK equities, higher alpha / significant outperformance could be expected in less efficient markets (i.e. Asia real estate, commodities, emerging market local debt) more frequently, thanks to the characteristics of these asset classes. The difficulty of this type of analysis could be a shortage of a long sample dataset as some of these asset classes only emerged in the past few years.

Recently, more research has focused on markets outside of the UK, the US, and other asset classes other than equities. For example, Tripathy (2005) evaluates market timing abilities of fund managers in India over 1994 to 1995 and 2002 to 2003. The results suggest that the portfolio managers have not got significant timing ability; instead, they are timing the market in the wrong direction. Ferson *et al.* (2005) utilize the stochastic discount factors from continuous-time term structure models to examine the performance of US government bond funds for the period from 1986 to 2000. They find that the empirical factors that include time average of the underlying state variables, contribute explanatory power in factor model regressions and produce smaller pricing errors.

Moreno and Rodriguez (2006) examine the performance of Spanish mutual funds between 1999 and 2003 with the SDF model, adding a third co-moment of asset returns to the SDF. They explore the effects of incorporating a co-skewness factor, both in an unconditional and conditional framework. They suggest that the omission of a co-skewness factor may lead to erroneous evaluations of a fund's performance.

Afza and Rauf (2009) evaluate Pakistani mutual funds from 1999 to 2006, using a Sharpe ratio with pooled time series and cross sectional data and examine the performance attributes, including fund size, expenses, age, liquidity, turnover etc. Their results suggest that liquidity had significant impact on fund performance.

Secondly, our analysis only uses the SDF primitive efficient model. It is interesting to expand the analysis using other asset pricing models and compare the characteristics of each model. More challenging tasks would be to explore new asset pricing models with more advanced techniques involved hence offering more insightful implication to performance evaluation.

Finally, the thesis examines the performance attribution only via 'Style Analysis'. Market timing ability, which consists of a large part of attribution literature, was not discussed. It therefore leaves room for further research, i.e. conditional market timing tests, with an application of the SDF models.

LIST OF REFERENCES

Afza, T. and A. Rauf, 2009, Performance evaluation of Pakistani mutual funds, *Pakistan Economic and Social Review, Volume 47, No. 2, pp. 199-214*

Allen, D.E. and M.L. Tan, 1999, A test of the persistence in the performance of UK managed funds , *Journal of Business Finance and Accounting, Vol 26(5+6),559-594.*

Anderson, R., 1997, A Large versus small capitalization relative performance model. *Market Timing Models. Burr Ridge, IL: Irwin.*

Annaert J., J.V. Broeck and R.V. Vennet, 2001, Determinants of mutual fund performance: A bayesian stochastic frontier approach, *Ghent university, working paper, 2001/103.*

Arshanapalli, B., T.D. Coggin and J. Doukas, 1998, Multifactor asset pricing analysis of international value investment strategies. *Journal of Portfolio Management, 10-23.*

Ashton, D.J., 1996, The power of tests of fund manager performance, *Journal of Business Finance and Accounting, Vol 23(1), 1-11.*

Bacquero, G., J.R. Horst and M. Verbeek, 2002, Survival, look ahead bias and the performance of hedge funds, *working paper, Erasmus University Rotterdam.*

Baks, K., A. Metrick, and J. Wachter, 2001, Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation, *Journal of Finance 56, 45-85*

Baltagi, B.H., G. Bresson and A. Pirotte, 2001, Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators, *Document ERMES 01-13, University Paris II.*

Barber, B.M. and J. D. Lyon, 1997, Firm-size, book-to-market-ratio, and security returns: A holdout sample of financial firms, *Journal of Finance 52, 875-883.*

Bauer, R. and R. Molenaar, 2002, Is the value premium predictable in real time? *Working Paper 02-003, Limburg Institute of Financial Economics.*

Bauer, R., J. Derwall and R. Molenaar, 2003, The real-time predictability of the size and value premium in Japan, *Working Paper 03-011, Limburg Institute of Financial Economics.*

Becker. C., W. Ferson, D. Myers, M. Schill, 1999, Conditional market timing with benchmark investors, *Journal of Financial Economics 52, 119-148.*

Brown, G., P. Draper and E. McKenzie, 1997, The consistency of UK pension fund investment performance, *Journal of Business Finance and Accounting, Vol 24(2), 155-178.*

Blake, D. and A. Timmermann, 1998, Mutual fund performance: Evidence from the UK, *European Finance Review 2, 57–77.*

Black, F., 1993, Beta and return, *Journal of Portfolio Management 20, 8-18.*

Blake, D., N.L. Bruce, and A. Timmerman, 1999, Asset allocation dynamics and pension fund performance, *Journal of Business, Vol 72(4), pp.429-461.*

Bollen, N. and J.A. Busse, 2001, On the timing ability of mutual fund managers, *Journal of Finance, June, 1075-1094.*

Bollen, N. and J.A. Busse, 2005, Short-term persistence in mutual fund performance, *Review of Financial Studies, Vol 18, No.2, 569-597.*

Brown, J.S. and W.N. Goetzmann, 1996, Performance persistence, *Journal of Finance, 679-698*

Brown, K.C., W.V. Harlow, and L.T. Starks, 1996, Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry, *Journal of Finance 51, 85-110.*

Brown, S.J., and W.N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics 43, 373-399.*

Brunnermeier, M.K. and S Nagel, 2004, Hedge funds and the technology bubble, *Journal of Finance, vol. 59, issue 5, 2013-2040.*

Busse, J.A. and P.J. Irvine, 2002, Bayesian alphas and mutual fund persistence, *February working paper, Goizueta Business School, Emory University.*

Busse, J.A., 1999, Volatility timing in mutual funds: Evidence from daily returns, *Review of Financial Studies 12, 1009—1041.*

Busse, J.A., 2001, Another look at mutual fund tournaments, *Journal of Financial and Quantitative Analysis 36, 53—73.*

Capaul, Carlo, I. Rowley, and W.F. Sharpe, 1993, International value and growth stock returns, *Financial Analysts Journal, January/February, 27-36.*

Carhart M. Mark, 1997, On persistence in mutual fund performance, J*ournal of Finance, 57-81.*

Canova**,** F., G. de Nicolo, 2000, Stock returns, term structure, inflation and real activity: An international perspective, macroeconomic dynamics, *Vol. 4, issue 3, 343-372.*

Chan, L., H. Chen and J. Laknoishok, 2002, On mutual fund investment styles, *The Review of Financial Studies, Winter, Vol.15, No 5.*

Chauvet, M. and S. Potter, 2000, Coincident and leading indicators of the stock market *Journal of Empirical Finance 7, 87-111.*

Chen, Z. and P. Knez, 1996, Portfolio performance measurement: theory and applications, *The Review of Financial Studies, Summer, 511-555.*

Chen, Y., W. Ferson, and H. Peters, 2005, The timing ability of fixed-income mutual funds, *working paper, Boston College.*

Chevalier, J. and G. Ellison, 1999, Are some mutual fund managers better than others? cross-sectional patterns in behavior and performance, *Journal of Finance, 875-899.*

Chizawa, S., 2000, The stock market and investors waver between growth and value investment styles, *NLI Research Institute Financial Research Group, Aug.*

Christopherson, J., W.E. Ferson and D.A. Glassman, 1998, Conditioning manager alphas on economic information: Another look at the persistence of performance, *The Review of Financial Studies, Spring Vol.11, 111-142.*

Cochrane, J.H., 2001, Asset pricing, *Princeton University Press*

Connor, G. and R.A. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, *Journal of Financial Economics, Elsevier, vol. 15(3), pages 373-394, March*

Cumby R. and J. Glen, 1990, Evaluating the performance of international mutual funds, *Journal of Finance 45, 497-521.*

Dahlquist, M. and P. Soderlind, 1999, Evaluating portfolio performance with stochastic discount factors, *Journal of Business, Vol. 72, No. 3*

Daniel, Kent, M. Grinblatt, S. Titman, and R. Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance 52, 1035-1058.*

Draper, P., 1989, The investment trust industry in the UK, *Avebury, Aldershot.*

Dybvig, P. and S.A. Ross, 1985, Differential information and performance measurement using a security market line, Journal of Finance, *American Finance Association, vol. 40(2), pages 383-99, June*

Elton, E., M.J. Gruber and C.R. Blake, 1996, The persistence of risk-adjusted mutual fund performance, *Journal of Business, 133-157.*

Falkenstein, G., 1996, Preferences for stock characteristics as revealed by mutual fund portfolio holdings, *Journal of Finance, 111-135.*

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

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

Fama, E.F. K. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance 51, 55-84.*

Ferson, W. and R. Korajczyk. 1995, Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business 68, 309-349.*

Ferson, W. and R.W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance, 425-461.*

Ferson, W., H. Farnsworth, D. Jackson and S. Todd, 2002, Performance evaluation with stochastic discount factors, *working paper, NBER working paper No.8791.*

Ferson, W., and V. Warther, 1996, Evaluating fund performance in a dynamic market, *Financial Analysts Journal 52, 20-28.*

Ferson,W. and K.Khang, 2002, Conditional performance measurement using portfolio weights: Evidence for Pension Funds, *February, NBER Working Paper No. W8790.*

Ferson,W. and A.Siegel, 2001, Testing portfolio efficiency with conditioning information, *April.*

Fletcher, J., 1995, An examination of the selectivity and market timing performance of UK unit trusts, *Journal of Business Finance and Accounting, Vol 22(1), pp. 143-156.*

Fletcher, J., 1997, An examination of UK unit trust performance within the arbitrage pricing framework, *Review of Quantitative Finance and Accounting, Vol 8, pp.91-107.*

Fletcher, J., 1999, The evaluation of the performance of UK American unit trusts, *International Review of Economics and Finance, Vol 8, pp.455-466.*

Fletcher, J. and D. Forbes, 2002, An exploration of the persistence of UK unit trust performance, *Journal of Empirical Finance, 2002, 9, 475-494*.

Fletcher, J. and D. Forbes, 2004, Performance evaluation of UK unit trusts within the stochastic discount framework, *Journal of Financial Research, 27, pp. 289-306.*

Fletcher, J. and J. Kihanda, 2005, An examination of alternative CAPM-based models in U.K. stock returns, *Journal of Banking and Finance, Vol29, Issue 12,2995-3014*

Fletcher, J., P. Ntozi and A. Byrne, 2006, An exploration of the conditional market timing performance of U.K. unit trusts, *Journal of Business Finance and Accounting, 33, 816-838.*

Gemmill G. and H. Salmon, 2005, Performance measurement with Loss aversion, *CEPR Discussion Papers with No.5173*

Girardin E. and Z. Liu, 2003, The Chinese stock market: a casino with 'buffer zones'? *Journal of Chinese Economic and Business Studies, Vol.1, No.1, pp.57-70*

Giles T., T. Wilsdon and T. Worboys, 2002, Performance persistence in UK equity funds--- A literature review, *Published in Investment Management Association website, Jan.*

Giles T., T. Wilsdon and T. Worboys, 2002, Performance persistence in UK equity funds— An empirical analysis, *published in Investment Management Association Website, Oct.*

Grinblatt M. and S. Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business, 393-416.*

Grinblatt M. and S.Titman, 1992, The persistence of mutual fund performance, *Journal of Finance, 1977-1984.*

Gruber, M.J., 1996, Another Puzzle: The growth in actively managed mutual funds, *Journal of Finance, 51, 783-811.*

Haigh, M. and J. List, 2005, Do professional traders exhibit myopic loss aversion? An experimental analysis, *Journal of Finance, Feb Vol 60, Issue 1, 523-534.*

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

Hansen, L. P. and K. J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica, 1269-1286.*

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica, 1029-1054.*

Hendricks, D., J. Patel and R. Zeckhauser, 1993, Hot hands in mutual funds: short-run persistence of relative performance, 1974-1988 , *Journal of Finance, 93-129.*

Henriksson, R.D., 1984, Market timing and mutual fund performance: An empirical investigation, *Journal of Business, 73-96.*

Henriksson, R.D., and R.C. Merton, 1981, On market timing ad investment performance II Statistical procedures for evaluating forecasting skills, *Journals of Business 54, 513-533.*

Henriksson, R.D., 1984, Market timing and mutual fund performance: An empirical investigation*, Journal of Business 57, 73-96.*

Ippolito, R.A., 1989*,* Efficiency with costly information: A study of mutual fund performance, *Quarterly Journal of Economics, 104, 1-23.*

Jagannathan, R., and R.A. Korajczyk, 1996, Assessing the market timing performance of managed portfolio, *Journal of Business 59, 217-235.*

Jensen, G.R., R. Johnson and J. Mercer, 1996, Business conditions, monetary policy and expected security returns, *Journal of Financial Economics 40*

Jensen, G.R., R. Johnson and J. Mercer, 1997, New evidence on size and price-to-book effect in stock returns, *Financial Analysts Journal.*

Jiang, W., 2003, A nonparametric test of market timing, *Journal of Empirical Finance 10, 399-425.*

Kosowski, R., A. Timmermann, R.Wermers, and H. White, 2001, Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis, *Discussion Paper, University of California San Diego*

La Porta, R., 1996, Expectations and the cross-section of stock returns, *Journal of Finance 51, No.5, 1715-1742.*

Lakonishok, J., A. Shleifer and R.W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance 49, 1541-1578.*

Levis, M. and M. Liodakis, 1999, The profitability of style-rotation strategies in the United Kingdom depends on temporal volatility of underlying return spread between style, *The Journal of Portfolio Management, Fall.*

Lewellen, J., 1999, The time-series relations among expected return, risk and book-to-market, *Journal of Financial Economics 54, 5-43.*

Liu, L., H. Liu and S. Varshney, 2009, Conditional performance evaluation: evidence from UK unit trusts, *Journal of International Finance & Economics*

Lo, A. and A. MacKinlay, 1990, When are contrarian profits due to stock market overreaction? *Review of Financial Studies 3, 175-205.*

Lucas A., R. van Dijk and T. Kloek, 2001, Stock Selection, style Rotation, and risk, *Tinbergen Institute Discussion Paper Number 01-021/2.*

Lunde A. and A.Timmerman, 2000, Duration dependence in stock prices: an analysis of bull and bear markets, *Working paper, Department of Economics, university of California, San Diego, August.*

Macedo, R., 1995, Value, relative strength, and volatility in global equity country selection. *Financial Analysts Journal, March/April, 70-78.*

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

Maddala, G.S., 1991, To Pool or Not to Pool: That is the question, *Journal of Quantitative Economics 7, 255-264.*

Maddala, G.S., R.P. Trost, H. Li and F. Joutz, 1997, Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators, *Journal of Business and Economic Statistics 15, 90-100.*

Moreno, D. and R. Rodríguez, 2006, Performance evaluation considering the coskewness: A stochastic discount factor framework, *Managerial Finance, Vol. 32 Iss: 4, pp.375 – 392*

Ragsdale, E.K., G.R. Rao, and L. Fochtman, 1995, Small versus large cap stocks: quantifying the fundamental reasons behind relative market performance. In R. Klein and J. Lederman, eds., *Small Cap Stocks: Investment and Portfolio Strategies for the Institutional Investor. Chicago: Probus Publishing Company.*

.

Rhodes, M., 2000, Past imperfect? The performance of UK equity managed funds, *FSA Occasional Paper 9, 1-58.*

Richardson, M., 1999, Current Topics, *presented to the Faculty of Actuaries Students Society.*

Pagan, A. R. and K.A. Sossounov, 2003, A simple framework for analyzing bull and bear markets, *Journal of Applied Econometrics 18, 23-46.*

Pastor, L. and R. Stambaugh, 2001a, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics*

Pastor, L. and R. Stambaugh, 2002b, Investing in equity mutual funds, *Journal of Financial Economics*

Quigley, G. and R.A. Sinquefield, 2000, Performance of UK asset management *Journal of Asset Management, Vol 1, 72-92.*

Schwert, G. W., 1990, Stock returns and real activity: A century of evidence, *Journal of Finance 45, 1237-1257.*

Schwob, R., 1999, Global Style Advisor, paper produced by Style Investment Research *Associates Limited, January*

Sharpe, W., Winter, 1992, Asset allocation: management style and performance measurement: An asset class factor model can help make order out of chaos, *The Journal of Portfolio Management, 7-19.*

Sirri, E.R. and P. Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance, 53, 1589-1622.*

Sorensen, E. and C. Lazzara, 1995, Equity style management: the case of growth and value, *chapter 4 from Equity Style Management, Irwin Publishing.*

Swinkels, L. and P.J. van der Sluis, 2001, Return-based style analysis with time-varying exposures, *working paper, Tilburg University.*

Swinkels, L, P.J. van der Sluis and M. Verbeek, 2002, Mutual fund manager ability when styles are changing, *working paper, Erasmus University Rotterdam.*

Teo, M., and S. Woo, 2001, Persistence in style-adjusted mutual fund returns, *working paper, Harvard University.*

Treynor, L., and K. Mazuy, 1966, Can mutual funds outguess the market? *Harvard Business Review 44, 131-136*

Tripathy, N.P., 2005, An empirical evaluation of market timing abilities of Indian fund managers on equity linked saving scheme, *Delhi Business Review Vol.6, No.2, July-December*

W.M. Company, 1999, *Company Report for Virgin Direct.*

W.M. Company, 2002, A comparison of active and passive management of unit trusts- 2002, *Produced for Virgin Money Personal Financial Service, Edinburgh, l-20.*

Wang, K., 2005, Multifactor evaluation of style rotation, *Journal of Financial and Quantitative Analysis, Vol. 40, 349-372.*

Warther, V.A., 1995, Aggregate mutual fund flows and security returns, *Journal of Financial Economics 39, 209-235.*

Wermers, R., 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expense, *Journal of Finance, 1655-1703.*

Zheng, L., 1999, Is money smart? A study of mutual fund investors' fund selection ability, *Journal of Finance, 901-933.*