A Recursive Modelling Approach to Predicting UK Stock Returns*

M. Hashem Pesaran
Trinity College, Cambridge

Allan Timmermann
London School of Economics

November 1996, Revised October 1997
This version September 1998

Abstract

This paper applies an extended and generalized version of the recursive modelling strategy developed in Pesaran and Timmermann (1995) to the UK stock market. The focus of the analysis is to simulate investors' search in 'real time' for a model that can forecast stock returns. It demonstrates the extent to which monthly stock returns in the UK were predictable over the period 1970-1993 after allowing for model specification uncertainty and possible shifts in the forecasting model. Due to a set of unique historical circumstances, UK stock returns were extremely volatile in 1974-1975, and we discuss how to design a modelling approach capable of accounting for this and similar low probability events. We find evidence of predictability in UK stock returns which could have been exploited by investors to improve on the risk-return trade-off offered by a passive strategy in the market portfolio. Alternative interpretations of this finding are briefly discussed.

JEL Classifications: G11, G12, E17, E44

Keywords: Stock returns, UK stock market, recursive modeling, switching portfolio, predicting stock prices.

*In preparing this version we have greatly benefited from the comments of three anonymous referees, an Associate Editor, the Editor and from discussions with Cliff Pratten. Partial financial support from the Isaac Newton Trust of Trinity College is gratefully acknowledged.

1 Introduction

Economists have long been fascinated by the sources of variations in the stock market. By the early 1970's a consensus had emerged among financial economists suggesting that stock prices could be well approximated by a random walk model and that changes in stock returns were basically unpredictable.[1] Historically, the 'random walk' theory of stock prices was preceded by theories relating movements in the financial markets to the business cycle. A prominent example is the interest shown by Keynes in the variation in stock returns over the business cycle. According to Skidelsky (1992) "Keynes initiated what was called an 'Active Investment Policy', which combined investing in real assets - at that time considered revolutionary - with constant switching between short-dated and long-dated securities, based on predictions of changes in the interest rate" (Skidelsky (1992, p. 26)).

Recently, a large number of studies in the finance literature have confirmed that stock returns can be predicted to some degree by means of interest rates, dividend yields and a variety of macroeconomic variables exhibiting clear business cycle variations. While the vast majority of these studies have looked at the US stock market, an emerging literature has also considered the UK stock market.[2] For example, Clare, Thomas and Wickens (1994) report that the gilt-equity yield ratio can forecast UK stock returns, Clare, Psaradakis and Thomas (1995) investigate the presence of deterministic seasonalities in UK stock returns, and Black and Fraser (1995) find that default- and term-premium variables have predictive power over UK stock returns. Common to all these studies is, however, that they condition their analysis on a given forecasting model, assumed to be in effect throughout the entire sample period.

Predictability of stock returns is of economic interest only in as far as it sheds light on investors' decision process and the functioning of financial markets. It is, for example, possible that researchers with access to data extending to the early nineties have found predictability of stock returns that could not have been detected by investors in earlier parts of the sample and hence does not reflect genuine *ex ante* predictability. For this reason it is important to model the behaviour in 'real time' of an investor who believes that stock returns can be forecast by means of business cycle factors, taking account of the uncertainty that surrounds the investor's choice of these factors and their relative importance in forecasting stock returns and avoiding, as far as possible, the benefits of hindsight. In the literature on recursive learning it is common to assume that agents know the 'correct' specification of the forecasting model, but not its true parameter values

[1] Fama (1970) provides an early, definitive statement of this position.

[2] References to studies conducted on US stock market data include Balvers, Cosimano and MacDonald (1990), Breen, Glosten and Jagannathan (1990), Campbell (1987), Fama and French (1989), Ferson and Harvey (1993), and Pesaran and Timmermann (1994, 1995). See Granger (1992) for a survey of the methods and results in the literature.

(see, for example, Bray and Savin (1986), Bulkley and Tonks (1989), Timmermann (1993)). This assumption does not appear to be plausible in the context of investors' portfolio decisions based on their predictions of stock returns. As witnessed by the many exploratory studies in the literature on predictability of stock returns, financial theory only provides very limited guidance for which state variables should have predictive power over stock returns. Finance theory does, however, suggest that in markets with risk averse agents stock returns would vary with the state of the business cycle, c.f. Lucas (1978) and Balvers, Cosimano and MacDonald (1990). Taken together, these statements suggest that a plausible analysis of investors' predictions of stock returns in 'real time' should be based on business cycle variables, taking appropriate account of the specification uncertainty that surrounds modelling stock returns at the end of each forecasting period.

In this paper we extend the literature, including our own earlier work (Pesaran and Timmermann (1995)), in a number of directions. We generalize the simulation of agents' 'real time' search over a set of factors/regressors by distinguishing between three sets of factors, referred to as 'focal', 'secondary focal', and 'potentially relevant' factors. In declining order of importance, these sets of regressors reflect agents' *a priori* beliefs that a given regressor should be included in the forecasting equation. Thus if a regressor is a focal variable, it is always included in the forecasting model. Secondary focal variables, on the other hand, are always considered but not necessarily included in the preferred forecasting model. Finally, the set of potentially relevant (or possible) variables are only considered in agents' modelling procedure after observing a residual which can be considered as an outlier. We expect these distinctions between regressors with differing degrees of *a priori* importance will become important in research using artificial intelligence to simulate agents' learning in situations where the true model is unknown. This classification of regressors helps in reducing the computational costs of the search strategy which could be substantial when it is required to search over a large number of regressors.

We also introduce a novel procedure for recursive selection of dummy variables whenever an extreme residual (more than three standard errors from zero) is obtained. The importance of such dummies in modelling UK stock returns can be exemplified by the set of special circumstances prevailing in the UK during the early 1970's. From the end of 1973 to the beginning of 1975 the UK stock market experienced a sequence of large, negative returns followed by extremely large positive returns in January and February of 1975. These dramatic movements were caused by particular events that took place in 1974, such as the quadrupling of oil prices, the miners' strikes, introduction of a three-day working week, increased political uncertainties (two elections were held in 1974), with some secondary banks facing runs on their deposits. This episode in the UK stock market presents real difficulties for any formal modelling approach, and one of the challenges of the present study is to develop an appropriate framework for dealing with special events that are unlikely to be repeated.

In recent papers Kandel and Stambaugh (1996) and Campbell and Viceira (1998) study the importance to optimal portfolio strategies of predictable components in stock returns. While these papers are a welcome contribution to the literature, they do not directly address our concern in the present paper which is the investor's uncertainty about the model specification. Our approach is explicitly designed to simulate the forecasting problem faced in real time by an investor who does not condition on a specific forecasting model and allows for the forecasting model to vary over time. For such an investor to characterize the optimal investment decisions is difficult and involves predicting future changes in the forecasting equation. Instead we establish a lower bound on the volatility of the stochastic discount factor derived from using the recursive forecasts in a simple, stylized investment strategy. This procedure is likely to produce a conservative estimate of the highest Sharpe ratio obtainable from using the forecast information. We find that the required movements in the stochastic discount factor are very large, even after accounting for transaction costs.

The rest of the paper is organized as follows: Section 2 describes the recursive modelling strategy and the various generalizations of the procedures already discussed in our earlier work. Section 3 deals with the choice of the factors to be used in the analysis and distinguishes between the 'focal', 'secondary focal' and 'potentially relevant' sets of factors. Section 4 reports the empirical results from a statistical view point, while section 5 provides an assessment of the economic importance of the predictions of stock returns based on trading results from spot and futures markets. Section 6 discusses the interpretation of our findings.

2 The Recursive Modelling Strategy

To simulate investors' search for a forecasting model we need to establish the set of regressors over which the search is to be conducted, the functional form of the estimated models and the criteria used to select a particular regression model. In Pesaran and Timmermann (1995) the forecasting equations were selected recursively from the *same* set of base regressors throughout the analysis. This procedure is subject to an important limitation: namely, it does not allow for the possibility of new variables being introduced into the forecasting exercise once the recursive strategy is set in motion. The regressors in the base set are decided at the start of the analysis and are then kept unchanged throughout. This is clearly restrictive, as it is very unlikely that all potentially relevant variables for forecasting stock returns could have been known *a priori*. Expanding the base set of regressors as a way of dealing with this problem is not computationally feasible either. For example, increasing the number of regressors in the base set from 10 to 16 increases the number of models to be estimated recursively each month from 1,024 to 65,536.[3] To deal with some of these difficulties, we propose the following generalization of our recursive modelling procedure.

We distinguish between three hierarchies of regressors. At the highest level are the set of 'focal variables', A_t, so called because they are believed to be important in forecasting stock returns, on *a priori* (or theoretical) grounds. These variables are always included in the forecasting equations. The second set of regressors, referred to as 'secondary focal' and denoted by B_t, are always considered in the forecasting exercise as potentially important for capturing short-term variations in risk premia due to business cycle fluctuations, although some or all of these regressors may be left out of the preferred forecasting model according to the model selection criterion in use. The combined set of regressors in A_t and B_t will be referred to as the "base set". Finally, a third set of regressors, C_t, are considered as potentially relevant, but utilized by the investors only if they discover clear evidence of the failure of the forecasting models obtainable from the regressors in the base set. This last set of variables are considered only occasionally since agents have weak reasons to believe that they should be included in the forecasting equation. Search across variables in the C_t set is triggered at time t when the most recent residual from the excess return equation using variables in the base set exceeds three (recursive) standard errors.[4]

Although in principle the variables in C_t could be combined with those in B_t, in practice, with a large number of regressors available, allowing some regressors only to be considered at relatively rare 'break points' is important on computational grounds. Once a search for regressors in the C_t set has been triggered, the variables in C_t that are chosen by a particular model selection criterion are then included in the B_t set in all subsequent periods. Hence the dimensions of B_t and C_t vary over time as indicated by the time subscript. In principle, one could also consider the problem of eliminating regressors from the base set. But in the absence of any compelling evidence suggesting that a regressor has ceased to be important over the sample period under investigation, we will not consider dropping regressors from the base set in the present study.

Extending the work in Pesaran and Timmermann (1995), we also allow for the possible inclusion of dummies in the estimated equations. Essentially we are asking whether, in the light of episodes such as January 1975 and October 1987, an investor would regard these events as truly exceptional, inexplicable by means of a linear regression model relating stock returns to a set of business cycle indicators. Although the magnitude of the stock market crash in October 1987 was very large seen in a historical perspective, and indeed proved to be an outlier even in the light of the subsequent events, it was not at all clear in the immediate

[3]Namely from 2^{10} to 2^{16}.

[4]For a broad class of distributions of stock returns, it seems reasonable to consider three standard errors from the mean as representing an extreme event. An alternative interpretation of this procedure in the context of the literature on Value at Risk is given by Duffie and Pan (1997).

aftermath of October 1987 whether the substantial fall in returns would be a one-off event or whether it could be repeated in the unsettled markets that followed. To deal with this problem once again we pursue a recursive modelling approach. We assume that investors consider inclusion of a dummy at any point in time if the residual in the forecasting equation lies more than three standard errors away from zero.[5] If such an extreme event occurs, the model selection criteria are again used to choose a model from the permutations on the expanded set of regressors that now includes the dummy variable. It is, however, important to note that the mere fact that a particular one-off dummy variable is considered in the search procedure does not necessarily mean that it will be selected as a factor in the forecasting model. Also even if a dummy variable (or for that matter any other variable) is selected in a particular period it does not mean that it will be automatically included in the forecasting model during subsequent periods. It could be de-selected at any point in time in the future.

More precisely, suppose that the distribution of stock returns can be viewed as the outcome of a mixture of two underlying distributions, one representing relatively smooth, partially predictable business cycle variation, while the second component is a more volatile, essentially unpredictable news component related to major exogenous shocks. Then filtering out large extreme points, as our procedure does, would be a sensible modelling strategy.

In the second step of the search procedure we apply standard statistical criteria for model selection to select a preferred forecasting equation in each and every period. Let k_t^A, k_t^B, and k_t^C be the number of regressors in the A_t, B_t, and C_t set at time t, respectively. Then the total number of regression models to be evaluated at time t is either the number of possible permutations of regressors in B_t, or, if a wider search is triggered off, permutations across B_t and C_t. We refer to the number of regressors over which agents conduct their search as k_t. Then either $k_t = k_t^A + k_t^B$ or (at a trigger point) $k_t = k_t^A + k_t^B + k_t^C$. This gives a total of 2^{k_t} different sets of models, each of which is uniquely identified by a number, i, between 1 and 2^{k_t}. Since regressors in A_t are always included in the forecasting model, the number of models that need to be evaluated is somewhat smaller, namely $2^{k_t^B}$ or $2^{k_t^B+k_t^C}$ (at trigger points). Consider a $k_t \times 1$ column vector $\mathbf{v}_i$ with k_t^A ones followed by a string of ones or zeros, where a one in the j'th row means that the j'th regressor is included in the model whereas a zero in the j'th row means that this regressor is excluded from the model. Then each model at time t can be identified by the k_t-digit string of zeros and ones corresponding to the binary code of its number. Using a subscript i to indicate the model ($1 \leq i \leq 2^{k_t}$), we let $k_{t,i}$ be the number of regressors included in model i at time t. Then $k_{t,i} = \mathbf{e}_t^{'}\mathbf{v}_i$, where $\mathbf{e}_t$ is a k_t-vector of ones. Consider forecasting $\rho_{\tau+1}$, the excess return at time $\tau + 1$ by means of linear regressions

[5]The dummy variable would have the value of unity in the period in question and zeros elsewhere.

$$M_{t,i}: \ \rho_{\tau+1} = \boldsymbol{\beta}_i^{'} \mathbf{X}_{\tau,i} + \epsilon_{\tau+1,i}, \quad \tau = 1, 2, ..., t-1, \tag{1}$$

where $\mathbf{X}_{\tau,i}$ is a $k_{t,i} \times 1$ vector of regressors obtained as a subset of the regressors in contention, $\mathbf{X}_{\tau}$, and $M_{t,i}$ denotes the i'th regression model. Conditional on $M_{t,i}$, and given the observations $\rho_{\tau+1}$, $\mathbf{X}_{\tau,i}$, $\tau = 1, 2, ..., t-1$ (with $t \geq k_t + 2$), the parameters of model $M_{t,i}$ can be estimated by the OLS method. Denoting these estimates by $\widehat{\boldsymbol{\beta}}_{t,i}$, we have

$$\widehat{\boldsymbol{\beta}}_{t,i} = \left(\sum_{\tau=0}^{t-1} \mathbf{X}_{\tau,i} \mathbf{X}_{\tau,i}^{'} \right)^{-1} \sum_{\tau=0}^{t-1} \mathbf{X}_{\tau,i} \rho_{\tau+1}. \tag{2}$$

These OLS estimates are fairly simple to compute.

The particular choice of $\mathbf{X}_{\tau,i}$ to be used in forecasting of $\rho_{\tau+1}$ can be based on a number of model selection criteria suggested in the literature, such as the $\overline{R}^2$ (Theil (1958)), Akaike's Information Criterion (AIC) (Akaike (1974)) or Schwarz's Bayesian Information Criterion (BIC) (Schwarz (1978)). These criteria are likelihood based and assign different weights to the 'parsimony' and 'fit' of the models. The 'fit' is measured by the maximized value of the log-likelihood function (LL), and the 'parsimony' by the number of freely estimated coefficients. At time t, and under the linear regression model $M_{t,i}$, we have

$$\widehat{LL}_{t,i} = \frac{-t}{2} \left(1 + \ln(2\pi \widehat{\sigma}_{t,i}^2) \right), \tag{3}$$

where

$$\widehat{\sigma}_{t,i}^2 = \sum_{\tau=0}^{t-1} \left(\rho_{\tau+1} - \mathbf{X}_{\tau,i}^{'} \widehat{\boldsymbol{\beta}}_{t,i} \right)^2 / t. \tag{4}$$

The penalized log-likelihood model selection criteria choose the model which maximizes an expression of the form

$$g(\widehat{\beta}_{t,i}, k_{t,i}, t) = \widehat{LL}_{t,i} - f(k_{t,i}, t), \tag{5}$$

where $f(k_{t,i}, t) = (k_{t,i} + 1)$, $\frac{1}{2}(k_{t,i} + 1)\ln(t)$, and $\frac{t}{2} \ln \left(\frac{t}{t - k_{t,i} - 1} \right)$ for the AIC, BIC and $\overline{R}^2$ criteria, respectively.

3 On the Choice of Regressors

In simulating the historical process through which an investor may attempt to forecast stock returns it is important that the following are established:

(i) - The list of variables the investor is likely to consider in modelling stock returns, and their possible decomposition into different categories, such as A_t, B_t and C_t above.

(ii) - The criteria adopted to select a particular forecasting model.
(iii) - The estimation procedure applied.

In this section we discuss the choice of the variables which we assume will be considered by investors in forecasting stock returns. In the next section we explain the estimation and forecasting procedure in more detail.

In an attempt to minimize the effect that the "benefit of hindsight" might have on our analysis, we only include regressors that could be safely argued to have been considered ex ante by investors searching for a return forecasting specification. This points towards regressors that have been discussed in the early literature on variations in stock returns, and also seem a priori reasonable.

To motivate our specific choice of the variables to be included in the sets A_t, B_t and, C_t, consider the following decomposition of excess returns

$$\rho_{t+1} = e_t + u_{t+1},$$

where[6]

$$\rho_{t+1} = \Delta \ln(P_{t+1}) + \frac{D_{t+1}}{P_t} - rf_t, \tag{6}$$

and P_{t+1} is the end-of-period share price, D_{t+1} represents dividends per share paid during period $t+1$, rf_t is the "safe" rate of return known at time t for the period from the end of period t to the start of period $t+1$. e_t denotes the predictable part of excess returns, while u_{t+1} is a martingale difference process representing its unpredictable part. General asset pricing models such as Lucas (1978) imply that e_t need not be zero. The standard first-order Euler equation for a representative investor in a frictionless market is

$$E_t[M_{t+1}\rho_{t+1}] = 0,$$

where $E_t[.]$ is the conditional expectations operator with respect to the information available to the agent at time t, and M_{t+1} is the stochastic discount factor representing the investor's marginal rate of substitution between future consumption in period $t+1$ and current consumption in period t.[7] Rearranging this equation, we have

$$E_t[\rho_{t+1}] = e_t = \frac{-Cov_t(M_{t+1}, \rho_{t+1})}{E_t[M_{t+1}]}. \tag{7}$$

Hence, in an efficient market, expected excess returns may vary over time but only to the extent that it reflects a time-varying covariance between investors' marginal rate of substitution and excess returns relative to the variation in the conditional expectation of the marginal rate of substitution.

[6]In this specification we have used the approximation $\Delta \ln(P_{t+1}) \approx (P_{t+1} - P_t)/P_t$.

[7]See Campbell, Lo and MacKinlay (1997, Ch. 8) for a discussion of different models of the stochastic discount factor.

Hansen and Jagannathan (1991) develop a lower bound on the volatility of the stochastic discount factor required for it to be consistent with a given sample of asset returns. Using the definition of a conditional correlation and the fact that this is bounded between plus and minus one, it follows from (7) that

$$\frac{E_t[\rho_t]}{\sigma_t(\rho_{t+1})} \leq \frac{\sigma_t(M_{t+1})}{E_t[M_{t+1}]},$$

where $\sigma_t(\rho_{t+1})$ and $\sigma_t(M_{t+1})$ are the conditional standard errors of excess returns and the stochastic discount factor, respectivly. Thus the Sharpe ratio appearing on the left side of the above equation establishes a lower bound on the variation in the stochastic discount factor scaled by its conditional mean. The advantage of focusing on the Sharpe ratio when assessing the economic significance of our forecasts is that it does not require specifying a particular utility function. For a given mean of the stochastic discount factor this variance bound only depends on the Sharpe ratio. The equity premium puzzle shows that the Sharpe ratio of the passive market portfolio is already too high to be consistent with a wide range of utility functions. Findings of an even higher Sharpe ratio based on forecasting information would suggest either that there are indeed exploitable predictable components in stock returns or that the discount factor varies even more than previously thought.

There is a large literature that uses the dividend yield as a proxy for time variation in expected returns.[8] Similarly, there are strong a priori grounds and a long tradition for including some measure of inflation in the forecasting relation. For example, Fama (1981) argues that expected inflation is negatively correlated with shocks to future economic growth while stock returns are positively correlated with such shocks leading to a (non-causal) negative correlation between anticipated inflation and stock returns. Traditionally a short interest rate has been used as a proxy for short-term inflation expectations, but inflation is likely to have a substantial persistent component so we also include a lagged value of this variable as a regressor.

At the business cycle frequency, economic theory also provides some general guidance for the relevant class of forecasting variables that investors may consider *ex ante,* c.f. Balvers, Cosimano and McDonald (1990). Expected excess returns should be high (low) when future consumption is expected to be higher (lower) than current consumption, such as around troughs (peaks) of the business cycle. But notice that the theory does not provide much guidance as far as the selection of specific business cycle indicators is concerned. Even if the marginal rate of substitution is identified as some function of aggregate consumption, a variety of economic variables tracking the current and future state of the economy could be

[8]For example, in the context of a log-linearized present value model, Campell, Lo and MacKinlay (1997) show how the dividend-price ratio proxies for variations in expected future returns.

used to model business cycle variations in expected stock returns. Forecasting variables that qualify on this account include changes in interest rates, changes in industrial production and the rate of monetary growth.[9]

To be sure, such business cycle variables have long been linked to movements in stock returns. For instance, in his book "Investment for Appreciation. Forecasting Movements in Security Prices. Techniques of Trading in Shares for Profit" published in 1936, Angas writes that "the major determinant of price movements on the stock exchange is the business cycle.." (p. 15). Other examples of early studies emphasizing the systematic variation of stock returns over the business cycle include Prime (1946), Dowrie and Fuller (1950), Rose (1960), and Morgan and Thomas (1962). Variables suggested by these studies to be systematically linked with stock returns include changes in short-run and long-run interest rates, industrial production, company earnings, liquidity measures, and the inflation rate.

The first group of indicator variables, namely changes in interest rate, are widely considered to closely track the state of the business cycle. They are also frequently mentioned in the pre-1970 finance literature as important determinants of stock returns. Prime (1946, p.165) writes: "market quotations are influenced by three factors: (a) fundamental conditions .... Fundamental conditions include the state of business earnings, dividends, financial positions, management, and money rates". Similarly, in his study of predictability of stock returns, Angas (1936) writes that "falling interest rates tend, other things being equal (i.e. unless industrial earnings are falling even faster), to force ordinary shares up", and he goes on to distinguish between the impact on stock returns of changes in short and long interest rates. In view of these early conjectures we include changes in a short 3-month interest rate and changes in the consol yield as secondary focal variables, i.e. in the B_t set.

There is also a long tradition in finance for considering seasonalities in stock returns. The most famous seasonal regularity, the higher returns in January, has been known for long enough to be included in the B_t set. In a recent study Clare, Psaradakis and Thomas (1995) find evidence that UK stock returns tend to rise in January, April, and to a lesser extent in December, and fall in September. However, only the January dummy has a long track record and could thus reasonably have been expected to be used by investors during our sample period.

Finally, for the variables in the set of possible regressors (C_t) we consider various macroeconomic indicators capturing short-run business cycle variations in the economy. To account for the impact of liquidity on stock prices, we include the rate of change in the narrow monetary stock in this set. We also consider changes in industrial production as another potential forecasting variable. The

[9]Because aggregate consumption is notoriously difficult to measure and there is no widely accepted specification for M_t, we do not impose the condition that a variable predict stock returns only through its effect on the conditional covariance between agents' marginal rate of substitution and stock returns.

industrial production index is linked with company earnings, another variable mentioned in several of the early studies as being an important determinant of stock returns, and has the further advantage that observations on it are available on a monthly basis, whereas company earnings are typically reported on a half-yearly basis in the UK. Finally, the rate of change in the spot price of oil was included in C_t, to allow for the possible effect of oil price volatility, particularly during the early 1970's, on the stock market. One could reasonably argue that investors would have begun considering oil prices as an indicator for the state of the economy only after the large oil price shocks from 1973 and onwards and our experiment is set up to ensure that this regressor is not selected prior to this period.

Based on this review of the early literature and on our *a priori* reasoning concering the long-term determinants of the excess returns, we decided on the following three sets of regressors:

$$A_t = \{c, YALL_{t-1}, I3_{t-1}, \pi_{t-2}\},$$

$$B_t = \{DI3_{t-1}, DGILT_{t-1}, JAN_t\},$$

and

$$C_t = \{\Delta IP_{t-2}, \Delta M0_{t-2}, \Delta PSPOT_{t-1}\},$$

where $YALL$ is the dividend yield on the FT All Share Index, $I3$ is the 3-month T-bill rate, $DI3 = I3 - I3(-1)$, π is the rate of change of retail prices, $DGILT$ is the change in the yield on a 2.5 percent government consol, JAN is a January dummy (which takes the value of unity in January of each year and zeros elsewhere), ΔIP is the rate of change in the index of industrial production, $\Delta M0$ is the rate of change of the money supply (the narrow definition), and $\Delta PSPOT$ is the rate of change in the spot price of oil.[10] (See the Data Appendix for more details.). Only the most recently available observations were used in the regressions. This meant that the financial data were available with a one-month lag while the macroeconomic data were available with a two-month lag.

The above decomposition of the variables into categories A_t, B_t and C_t substantially reduces the number of models that need to be estimated. Under the procedure advanced in this paper, since the focal regressors (in A_t) are always included in the forecasting equations, and since three breakpoint dummies were identified in the sample, we had to estimate at most 512 models every period. However, treating the regressors in all the three categories symmetrically (except for the intercept term), would have required us estimating as many as 4096 models every period.

[10]Notice that changes in the interest rates were calculated on a month-on-month basis, while changes in the remaining macroeconomic variables were calculated as the rate of change of their twelve-monthly averages. These long moving averages were used in order to minimize the effects of data revisions on our forecasts.

4 Summary of the Statistical Results

Figure 1 displays the monthly movements of excess returns on the FT All Share Index over the period 1965-1993. Apart from the substantial changes in excess returns experienced during 1974-75 and 1987, the series seems reasonably well behaved and does not exhibit any obvious patterns. A closer inspection, however, reveals some weak evidence of serial correlation in excess returns. The estimated coefficients in a second order autoregressive process in excess returns estimated over the period 1965(3)-1993(12) are 0.148 (2.76) and −0.128 (−2.38) with the associated t-ratios in brackets. But even this weak evidence of auto-dependence in excess returns disappears once the autoregressive specification is augmented with a January dummy and the one-off dummy variables to take account of the special events in 1974/75 and for the October 1987 crash. The autoregressive coefficients in this augmented regression are estimated as 0.073 (1.42) and −0.069 (−1.48), respectively.

Absence of any discernible patterns does not, however, mean that excess returns cannot be (partly) predicted using information in other variables such as dividend yields, interest rates or business cycle indicators.

The extent to which excess returns in the UK stock market can be predicted using the various business cycle variables discussed in Section 3 can be seen from the regression results summarized in Table 1. This regression is estimated on monthly observations over the entire sample period 1965-1993 (348 observations) and contains all the variables discussed in the previous section (namely the variables in the sets A_t, B_t and C_t). In this regression we have also included dummy variables for January 1975, February 1975 and October 1987 which are identified by our recursive procedure as outliers in the sample. Out of the eight variables included in this regression (not counting the constant and the dummies) six turned out to be statistically significant at the 10 percent critical level, with the lagged dividend yield and change in oil prices being significant at a one percent level.

The signs of the coefficients of the various regressors in Table 1 are as to be expected from theory and follow findings for the US stock market. The dividend yield enters with a positive coefficient, possibly reflecting mean reversion in returns caused either by investor overreaction or persistent time-varying risk-premia. The rate of inflation and the growth in the monetary base both enter with negative coefficients in the excess return regression. Interest rates only appear to be weakly correlated with future excess returns, the most significant relationship being the negative coefficient on the change in the gilt yield. All these regressors measure different aspects of the economy's inflation rate. The change in the log of oil prices is negatively correlated with stock returns possibly because the large increases in oil prices in 1975 and 1979 contributed to the subsequent recession periods with associated low stock returns. The $\overline{R}^2$ of the excess return equation in Table 1 is 0.33, but due to the inclusion of the dummies, presents an exagger-

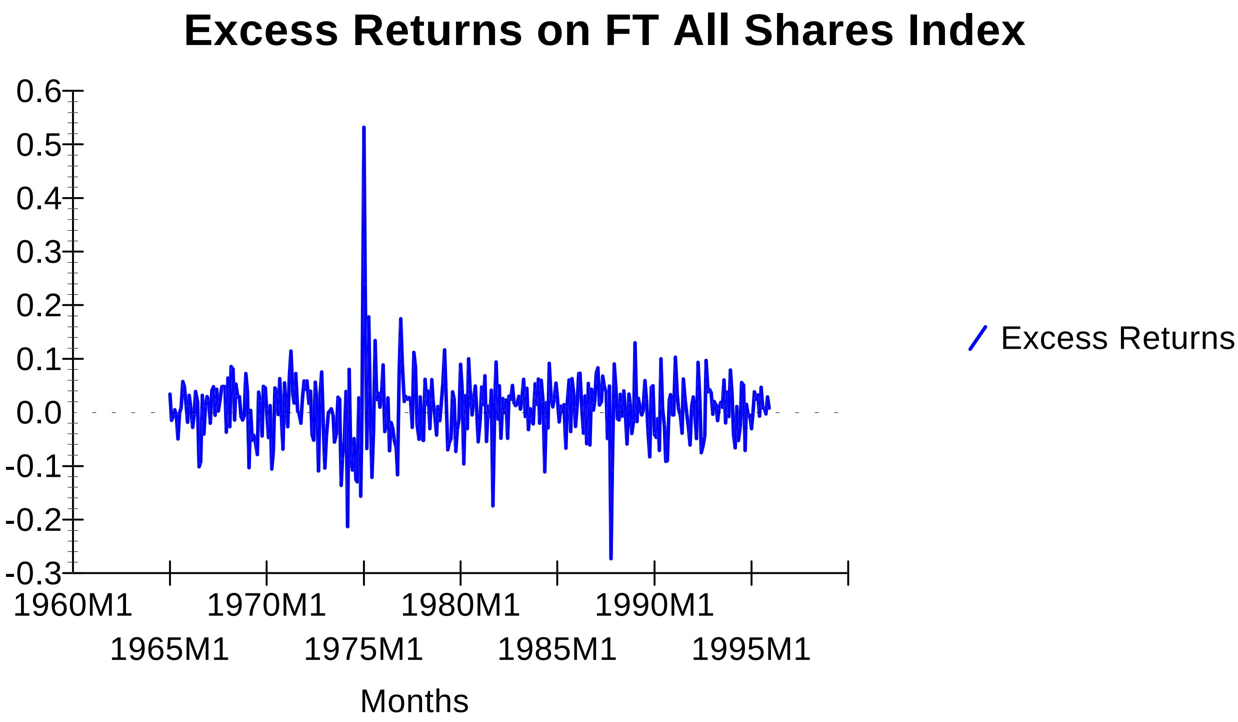

Figure 1: Monthly Excess Returns on Financial Times All Share Index

ated picture of the degree to which excess returns are predictable. The $\overline{R}^2$ for the excess return regression without the dummies is 0.12, which is in line with findings for the US stock market, c.f. the references cited in footnote 2. Only the diagnostic test for normality of the residuals is rejected for the specification in Table 1.

Predictions from the recursively selected and estimated forecasting models based on the $\overline{R}^2$, Akaike and Schwarz criteria, as well as on the model that includes all regressors in the A_t, B_t, and C_t sets are presented in the four windows of Figure 2. These forecasts are strongly serially correlated with first-order serial correlations ranging between 0.71 and 0.89, and there are periods where the predicted stock returns consistently stay negative (mainly during 1973-1975 and around 1981) or positive (late seventies, early eighties). This is as one would expect from the fact that our regressors are chosen to track business cycle variation, which is known to display a certain degree of persistence. The recursive forecasts from the model which includes all regressors are highly volatile in the early stages of the forecasting exercise.

As a simple, intuitive measure of the fit of the forecasting models we computed the (recursive) squared correlation between the forecasts from the recursively selected models and the realisations of monthly stock returns. These are presented in Figure 3. For the three model selection criteria, the squared correlation coefficient starts from a level of around 0.2, decreases to around 0.1 during 1974 and then jumps to 0.5 after January 1975, at which point a dummy gets included. After this the squared correlation coefficients gradually decline to a point around 0.4 at the end of 1993, only interrupted by an increase due to the inclusion of the October 1987 dummy.

Recursive standard errors for the estimated equations are provided in Figure 4. It is clear from this figure that the recursive standard errors build up from 1970 to 1982, and thereafter show a relatively smooth decline.

To study the importance of business cycle indicators and financial variables for forecasting of stock returns, we plot the inclusion frequencies of each regressor in Figures 5 and 6. The inclusion of a regressor by a particular model selection procedure is represented by a point in the relevant line of the graphs in these figures. The horizontal axis of these graphs also contain interesting information. When a variable is selected by all of the model selection procedures the corresponding point on the horizontal axis is left blank, otherwise it is shown as a solid point.

Consider first the inclusion frequency of the variables in the B_t set (namely $DI3_{t-1}$, $DGILT_{t-1}$, and JAN_t). Recall that the variables in the A_t set, namely a constant, the dividend yield, the 3-month T-bill rate and the rate of inflation are always included in the forecasting model. The change in the Gilt rate gets included by the Akaike and $\overline{R}^2$ criteria in all periods except for a spell between 1974 and 1980. This finding indicates that the finding in Clare, Thomas and Wickens (1994) on the importance of the gilt yield for forecasting of stock returns is quite robust with respect to the choice of model and sample period. Compared to the change in the gilt yield, the change in the 3-month T-bill rate is rarely selected by any of the model selection criteria. The January dummy is never selected by the Schwarz criterion, and is only rarely selected by the Akaike criterion. However, the $\overline{R}^2$ criterion selects the January dummy in around half of the time periods.

Consider next the inclusion frequencies of the variables in the C_t set. Searches over the variables in this set were triggered by the Akaike and $\overline{R}^2$ criteria a total of three times during the sample, namely after 1975m1, 1975m2 and 1987m10.[11] According to the Akaike and $\overline{R}^2$ criteria, changes in the money supply appear to be important in forecasting excess returns after 1987. Schwarz's criterion, by contrast, does not select this regressor at all. Growth in industrial production is selected on an almost continuous basis by the Akaike and $\bar{R}^2$ criteria from 1975 onward, but only between 1975 and 1981 when the Schwarz criterion is used.

[11] Because of the difficulty of assessing the tails of the distribution of regression residuals, trigger points at which variables in the C_t set are evaluated were only considered after 100 observations from the start of the simulation experiment.

Changes in oil prices are included continuously after 1975 according to all three model selection criteria.

Inclusion of a dummy variable was only considered after the realization of the extremely high stock return in January 1975, February 1975 (not shown) and October 1987. Following these episodes, the Akaike, Schwarz and $\overline{R}^2$criteria all included these dummy variables continuously in all regressions.

The above summary clearly shows that whether a variable gets included in a forecasting equation critically depends on the choice of the model selection criteria. Among the three standard model selection criteria, Akaike and $\bar{R}^2$ tend to give similar outcomes and select a larger number of regressors in the forecasting equation as compared to the Schwarz criterion. This is as to be expected considering the much higher penality on inclusion of additional regressors in the forecasting equation imposed by the Schwarz criterion as compared to the other two criteria. The analysis also suggests that the best forecasting model is likely to change considerably over time. This is an important consideration in the discussion of optimal portfolio weights following in the next section.

5 The Economic Significance of Business Cycle Components in Stock Returns

5.1 Predictability of Returns and Optimal Portfolio Weights

Recent research investigates the implications of the evidence on predictability of asset returns on optimal investment and consumption decisions.[12] A particularly clear demonstration of the extent of time variation in optimal portfolio weights is provided by Kandel and Stambaugh (1996). These authors consider the optimal portfolio decisions of a Bayesian agent in the context of a single riskfree and a single risky asset whose returns can be predicted by means of a linear, time-invariant regression model. The coefficients of the regression model are unknown to the agent who maximizes expected utility over terminal wealth. Kandel and Stambaugh account for this parameter uncertainty in the agent's decision problem. They find that, even in cases where the agent has a diffuse prior over the parameter values and where predictability is extremely low according to standard statistical criteria such as the R^2 of the regression model, it is still optimal for agents with reasonable degrees of risk aversion to let their portfolio weights vary considerably over time.

Extensions of the investment and consumption decisions to the multiperiod case have been derived by Campbell and Viceira (1998) and Brandt (1998). Except for the special case with logarithmic utility which reduces to a sequence of

[12] An earlier asset pricing literature is concerned with estimating the preferences of a representative investor from the moment conditions corresponding to the investor's Euler equations.

single-period problems, the multi-period decision problem can yield solutions that are very different from the single-period problem. The reason is that the possibility of shifts in future investment opportunities generally creates an intertemporal hedging demand for the risky asset, reflecting a negative correlation between current realized and expected future asset returns. Campbell and Viceira use an approximate, log-linearized budget constraint and a Taylor expansion of the Euler equation corresponding to a power utility function with habit persistence to analytically derive an approximate optimal investment and consumption strategy for an infinitely-lived investor. Their analysis assumes that stock returns follow a stationary autoregressive process. They find that intertemporal hedging can more than double risk averse investors' demand for the risky asset. Like Kandel and Stambaugh, they report that the optimal investment strategy involves substantial timing of the stock market and that failing to vary the portfolio weights can result in substantial welfare losses.

Brandt (1998) considers power utility functions and estimates investors' optimal consumption and investment decisions from sample moments corresponding to their conditional Euler equations. Unlike Campbell and Viceira, Brandt does not obtain an analytical solution to the consumption and investment rules, and instead he characterizes these through moment conditions that are estimated nonparametrically. His results again reveal substantial dependency of the optimal portfolio weights with regard to changes in a set of standard forecasting variables.

Analytical solutions to investors' optimal investment decisions come at the cost of having to make strict parametric and distributional assumptions about the relationship between stock returns and the predictor variables. Common to the solutions derived in the literature is that they either do not account for the estimation risk facing an investor or, as in the case of Kandel and Stambaugh, only consider parameter uncertainty but disregard model uncertainty. Furthermore, these papers assume that the joint distribution of the forecasting variables and returns on the risky asset is stationary and does not account for conditional heteroskedasticity in asset returns, something which could significantly change the optimal solution if the same variables that forecast expected returns also predict the conditional volatility of returns. Without imposing considerable structure both on the functional form of the utility function and the data generating process of excess returns, it is not possible to characterize analytically optimal portfolio weights resulting from the evidence on predictable returns. Hence, in the present paper we confine our analysis to computing a conservative estimate of the economic value of the predictions that could arise from a simple, stylized trading rule.

Predictability of stock returns can also be the result of transaction costs. Luttmer (1997) enquires into how large transaction costs must be if we are to rule out the possibility of an investor being able to exploit the apparent predictability of returns. Luttmer computes the size of a fixed transaction cost facing investors every time they trade in the financial markets and large enough to ensure that it is not optimal for a representative investor to adjust portfolio weights in line with predicted variations in the investment opportunity set. Unsurprisingly this lower bound on the fixed cost strongly depends on the assumed utility function. For logarithmic utility, and in the case where the forecasting variables are observed ex ante, conservative estimates suggest that the representative consumer must face a fixed cost of at least three percent, and possibly higher than 10 percent, of monthly consumption to rule out expected utility gains from market timing. Consistent with this, He and Modest (1995) find that a combination of transaction costs and constraints on borrowing and short sales can reduce the volatility bound on the stochastic discount factor.

As pointed out by Campbell, Lo and MacKinlay (1997), these results are exploratory and may exaggerate the importance of transaction costs in a multi-period setting where investors can spread across several periods the transaction costs from switching into the highest paying asset. Optimal trading strategies under transaction costs are difficult to derive in the absence of assumptions about very specific time-invariant transaction cost technologies. We take the view that transaction costs can potentially be important and assess portfolio returns using a range of transaction costs.

5.2 Empirical Findings

As a first indication of the potential economic value of the predictability of stock returns we consider the non-parametric test of market timing skills proposed by Pesaran and Timmermann (1992). This statistic, which is asymptotically equivalent to the more familiar Henriksson-Merton (1981) test of market timing, tests the null that there is no information in the predictions of excess returns over the sign of subsequent realizations of excess returns. Leitch and Tanner (1991) found that the ranking of predictions according to sign tests is closely related to their ranking in terms of making money in simple trading strategies based on the predictions.

Consider the proportion of correctly predicted signs of excess returns over the whole sample period 1970-1993 used in the trading exercises. As can be seen from Table 2, the recursive forecasts based on models selected according to the Akaike, Schwarz, and $\overline{R}^2$ criteria generated values for the nonparametric sign test statistic which were significant at the 2.5 percent critical level.[13] The sign tests are, however, statistically less significant over the sub-sample periods as compared to the whole period. This reflects the lower power of the test when the sample size is relatively small. Finally, note that the proportion of correctly predicted signs is somewhat higher during the eighties than during the seventies.

The stylized investment strategy that we consider is a switching strategy instructing the investor to hold the asset for which the largest mean returns

[13]Notice that we have used a one-sided test which is appropriate here.

have been forecast. To avoid bankruptcy risks, we assume that investors are not allowed to use leverage or go short in either asset. While it may seem extreme to use an investment strategy that frequently shifts funds from a portfolio fully invested in stocks to a portfolio fully invested in riskfree T-bills, notice that this is also a feature of the optimal investment rule derived by Brennan, Schwartz, and Lagnado (1997). Furthermore, the high sensitivity of the switching portfolio weights with regard to predicted returns is a property shared by the optimal portfolio rules derived in the recent literature. For investors that are not very risk-averse, the switching portfolio is thus likely to capture some of the salient features of an optimal portfolio rule.

Tables 3-5 report various performance measures for a buy-and-hold strategy in the market index (the FT All Share Index), a portfolio consisting of rolling over T-bills, and switching portfolios based on forecasts from models recursively selected according to one of the three model selection criteria considered in this paper.

Table 3 reports the results for the zero transaction cost scenario. The market portfolio paid an annual arithmetic mean rate of return of 21.0 percent which was approximately 11 percent higher than the return on the T-bills (10.1 percent). As one would expect, the risk associated with holding equity is much higher than the risk of a portfolio of T-bills; the standard deviation of the annual returns on the market portfolio (36.6 percent) was almost 15 times higher than the standard deviation of the T-bill returns (2.6 percent). The switching strategy significantly improves the risk-return trade-off compared to the market portfolio as reflected in the high values of the Sharpe and Jensen indexes for the switching portfolios.[14] The Jensen index is statistically significant at the 1 percent critical level for the switching portfolios based on the three model selection criteria, as indicated by the t-values above 3.4 given in brackets following the Jensen measure.[15]

Performance results for the case with medium size transaction costs of 0.5 of a percent on dealings in shares and 0.1 of a percent on acquisitions of T-bills are reported in Table 4.[16] Payoffs on the buy-and-hold strategy are hardly affected by the introduction of transaction costs since this strategy only incurs

[14] Although our forecasting exercise focuses on first-moment predictability of stock returns, notice that we select the forecasting models so as to minimize standard statistical criteria. Hence we do not necessarily expect that the forecasts will generate a higher cumulated portfolio return which will depend on the full distribution of asset returns and is likely to be sensitive to outliers. However, it is also true that the forecasting information is only of economic value if it can be used to improve on standard economic performance measures.

[15] These are financial performance criteria which adjust the excess return on the portfolio under consideration for market risk (Jensen) or total risk (Sharpe). The Jensen measure is the intercept term in a regression of portfolio excess returns on a constant and excess returns on the market portfolio. The Sharpe measure is the ratio of mean excess portfolio returns to the standard deviation of portfolio returns.

[16] The particular formulas used to take account of transaction costs across different portfolios are described in Pesaran and Timmermann (1994).

such costs when dividends are reinvested. In contrast, mean returns on the switching portfolios decline by around 1 to 1.5 percent per annum when medium transaction costs are introduced. Despite this reduction in the returns generated by the switching portfolios, the portfolios based on the Akaike, Schwarz and $\overline{R}^2$ criteria still generate a much higher return-to-risk ratio compared to the market index and the values of the Jensen index remain significant at the five percent level.

Under the high transaction cost scenario where trades in shares cost one percent and acquisitions of T-bills cost 0.1 of a percent, the values of the Sharpe and Jensen indexes continue to indicate risk-adjusted outperformance of the switching portfolios based on the three model selection criteria.

We can conclude from the evidence in Tables 3-5 that increasing transaction costs has an important effect on the switching portfolios' returns: as compared to the mean returns under the zero transaction cost scenario, the mean returns of the switching portfolios in the high transaction cost scenario are between one and a half and three percent lower per annum. The reason for the importance of transaction costs for returns on the switching portfolios lies in the number of transactions between the stock portfolio and T-bills. For the period 1970-1993, there were between 17 and 42 switches, depending on the model selection criterion. Thus these portfolios would swap assets between once every 18 months and twice a year.

In Figure 7 we present the fraction of total wealth invested in the stock market according to the switching portfolios following from the model selection criteria as well as from the predictions based on the full set of regressors. It is clear from that figure that the $\overline{R}^2$ criterion generates most switches between T-bills and stocks. The switching portfolios spend roughly 58 percent of the time in the stock market. Notice in particular the long period between 1973 and 1975 where these portfolios were out of the stock market. This is a key to understanding the risk-return trade-off for the switching portfolios. Being out of the stock market between 1973 and 1975 means that the switching portfolios avoided the sequence of large negative returns during 1973 and 1974, but also that these portfolios did not profit from the very large positive returns during early 1975.

We also considered the hyper-selection criterion introduced in Pesaran and Timmermann (1995). Based on the cumulated wealth accrued by trading on the predictions generated by models chosen according to the various selection criteria, this hyper-selection criterion establishes a way to recursively choose the model selection criterion itself. In principle, no clear-cut conclusion about the possibility of outperforming the market index can be drawn unless a recursive procedure for choosing the model selection criterion is established. Since maximizing expected terminal wealth is similar to maximizing expected utility for an investor with a logarithmic utility function, our procedure can also be interpreted as choosing the model selection criterion based on expected utility optimization. Using our utility- or wealth based hyper-selection criterion, it is clear from Tables 3-5 that, on a risk-adjusted basis, it would have been possible for investors to choose model selection criteria recursively, use forecasts from the models that these criteria choose, and outperform the market index.

We finally studied the sub-sample trading results for the 1970's and 1980's, two decades with many different features. Consider the zero transaction cost scenario. During the 1970s, arithmetic mean returns on the market portfolio and the switching portfolios chosen by the Akaike, Schwarz and $\overline{R}^2$ model selection criteria were 20.46, 14.21, 14.84 and 13.98 percent, respectively. Their standard deviations were 56.48, 8.90, 8.78, and 9.24 percent per year, resulting in Sharpe ratios of 0.20, 0.59, 0.67 and 0.55. During the 1980s, these four portfolios paid mean returns of 24.30, 23.96, 28.34 and 24.20 percent per year. Their annual standard deviations were 10.58, 8.65, 10.05 and 8.58 percent, and the Sharpe ratios were 1.22, 1.46, 1.69 and 1.50, for the market portfolio and the three switching portfolios, respectively. These findings suggest that although the largest return-to-risk ratio improvements occur in the volatile 1970s, the switching portfolios were also capable of generating a higher return-to-risk trade-off during the much calmer markets of the 1980s. Perhaps these findings are not all that surprising: Our regressors were chosen to track variations in stock returns related to the business cycle, variation in which was much more pronounced during the 1970s than during the 1980s.

5.3 Trading Results from the Futures Market

A potential criticism of the trading results in the previous section is that it would have been difficult - and perhaps even more costly than assumed - for an investor to switch in and out of the broad FT All Share portfolio. We do not think this is a major concern because in practice investors could have held a far smaller number of liquid stocks which, if carefully selected, would be very strongly correlated with returns on the FT All Share portfolio.

Nevertheless, to address these issues we used our recursive predictions in a trading strategy based on the FTSE-100 Futures contract. Transaction costs in the futures markets are very small (close to zero in percentage terms) and it is straightforward to execute trades in portfolios of shares. A futures contract on the FTSE-100 stock market index has been traded in the UK since May 1984. We run our forecasting experiment from January 1985 to December 1993, giving a total of 108 monthly observations. To get a continuous time series, each month the futures price was based on the settlement value of the contract nearest to maturity, except for the month of delivery where the contract whose expiration date is second nearest to maturity was used. This procedure was also used in Buckle, Clare and Thomas (1994) and the associated return sequence resembles the payoffs to a long position in the nearest-to-maturity contract.[17]

Based on a mean return of 12.8 percent and a standard deviation of 13.4 percent over the period 1985-1993, the long position in the futures contract resulted in a Sharpe ratio of 0.17. Switching portfolios using the predictions generated according to the Akaike, Schwarz and $\overline{R}^2$ criteria produced Sharpe ratios of 0.32, 0.60, and 0.34, respectively.[18] The mean returns of these switching portfolios were 14.0, 15.8 and 14.2 percent, respectively, while their standard deviations were 10.9, 8.7, and 10.8 percent. It is likely that the returns from these switching strategies could be increased by modifying the trading rule to allow for short selling. Most importantly, these results suggest that the earlier findings about the possibility of using our recursive predictions to improve on the market portfolio's risk-return trade-off do not depend on the tradeability of the FT All Share Index.

6 Conclusion

The main difference between the recursive modelling approach applied in this paper and the standard recursive estimation techniques lies in the treatment of model uncertainty. This is particularly important in the analysis of stock returns where there is very little guidance from theory as to the precise nature of factors that could be used in such an analysis. In this paper by applying an extended and modified version of the recursive modelling procedure originally put forward in Pesaran and Timmermann (1995) to the UK stock market, we have been able to identify a number of genuine *ex ante* predictors of excess returns and evaluate their economic significance in trading contexts.

While there are some variations across the performances of the different switching portfolios examined in this paper, overall we show that it would have been possible for investors to choose model selection criteria recursively, use forecasts from the models that these criteria choose, and improve on the risk-return trade-off offered by the market portfolio. This finding is robust to the design of the simulation experiment.[19] Cumulated wealth from the switching portfolios can be sensitive to the presence of outliers. The difference in cumulated wealth between getting the sign of excess returns in January 1975 right or wrong is 50 percent. In contrast, the risk-adjusted performance measures are more robust to such events. A switching rule that was long in stocks in January 1975 would have a higher mean but also a much higher variance than a short position in

[17] We are grateful to Andrew Clare for providing us with the futures price data.

[18] Since the earlier forecasting equations refer to excess returns while payoffs on the futures contract are based on capital gains only, we added to our previous predictions the T-bill rate and subtracted the lagged dividend yield (at monthly rates). This ensures that our predictions are comparable to returns on the FTSE-100 futures contract.

[19] For example, an earlier version of this paper utilized a slightly different set of forecasting variables but arrived at a very similar conclusion.

stocks for that month. These effects would partially cancel out in the Sharpe ratio. In this context it is noteworthy that all the switching portfolios considered in the paper were out of the stock market during the period 1973-1975, which meant that these portfolios managed to avoid the sequence of large negative returns experienced during 1973 and 1974, but also failed to take advantage of the sharp, unprecedented rises in the stock market that took place during January and February of 1975, overall doing better than the market portfolio.

After accounting for investors' "real time" search for a forecasting model and transaction costs in executing their investment decisions it still seems possible to improve the risk-return trade-off of the market portfolio. Two interpretations of this finding are possible. The efficient market interpretation regards variations in expected stock returns as changing risk premia and suggests that negative risk premia are compatible with stocks providing a hedge against states with a high marginal utility of consumption. Alternatively, predictability of stock returns could reflect an inefficient stock market dominated by investors who do not use publicly available information efficiently, possibly because they systematically overreact to news (see, e.g., Bulkley and Harris (1997), De Bondt and Thaler (1990)).

Since neither agents' expectations nor their preferences are directly observable, it is not possible to conclusively establish which interpretation is correct. We will make two observations, however. First, the Sharpe ratio varies substantially across sub-samples so any attempt at interpreting expected excess returns in terms of time-varying risk premia requires a model displaying far more variability in agents' intertemporal marginal rates of substitution than standard consumption-based asset pricing models are capable of delivering. Second, negative expected excess returns can only be consistent with a general equilibrium model of risk averse, rational investors if stocks provide a hedge against states with high marginal utility of consumption. Our analysis identified two periods (1973-1975 and around 1981) with persistently negative expected excess returns. If it cannot be reasonably argued that investors would use stocks to hedge against adverse economic states during these periods, then it would be difficult to interpret variations in expected returns as changes in risk premia. Even if this interpretation of our results is acceptable, we are still left with the task of attributing changes in expected returns to factors other than risk premia, such as "noise-trading" or the importance of "psychological" factors in market transactions. However, these issues are beyond the scope of the present paper.

Appendix: Optimal Portfolio Weights for Investors With Logarithmic Utility

To better understand the properties of the switching portfolio used in Section 5, this appendix shows that the switching strategy closely resembles the optimal trading strategy of an agent maximizing expected cumulated wealth subject to a no-short-selling constraint. Consider the set of two-fund portfolios mixing risk-free T-bills paying rf_t between the end of time t and time $t+1$, and risky stocks paying a return of

$$r_{s,t+1} = \rho_{t+1} + rf_t,$$

where ρ_{t+1} is the excess return on stocks. Disallowing short-sale of assets, suppose that the fraction $\alpha \in [0,1]$ of the available funds is invested in risky stocks such that the return on the switching portfolio becomes

$$r_{t+1} = \alpha(\rho_{t+1} + rf_t) + (1-\alpha)rf_t = rf_t + \alpha\rho_{t+1}$$

For an investor with logarithmic utility over terminal wealth, maximizing expected utility under zero transaction costs is equivalent to maximizing the geometric rate of return. This is because

$$W_T = W_0 \prod_{t=1}^{T}(1+r_t)$$

so

$$\ln(W_T/W_0) = \sum_{t=1}^{T} \ln(1+r_t)$$

where W_T is wealth at the terminal date, T, and W_0 is initial wealth. Thus maximization of the expected value of $\ln(W_T/W_0)$ is identical to maximizing the expected value of $\ln(1+r_{t+1})$ with respect to the available information at time t, period by period. Clearly this involves maximizing a concave function in the portfolio returns, r_{t+1}, so, by Jensen's inequality, the optimal portfolio may not maximize the arithmetic mean return, which is what the switching portfolio aims to do. Suppose excess returns are generated by the following linear model closely related to the recursive procedure analyzed earlier in the paper

$$\rho_{t+1} = \boldsymbol{\beta}_i' \mathbf{X}_{t,i} + \epsilon_{t+1,i}\ , \quad \epsilon_{t+1,i} \sim\ IID(0, \sigma_i)$$

Here $\mathbf{X}_{t,i}$ is a set of factors known at time t, so that $\epsilon_{t+1,i}$ and $\boldsymbol{\beta}_i$ are the only unknown factors. Assuming that $\epsilon_{t+1,i}$ is drawn from a density function $h_t(.)$, the expected period $t+1$ rate of return on the portfolio with weight α_t on stocks becomes

$$E_t[\ln(1+r_{t+1})] = \int \ln(1+rf_t+\alpha_t(\boldsymbol{\beta}_i'\mathbf{X}_{t,i}+\epsilon_{t+1,i}))h_t(\epsilon_{t+1,i})d\epsilon_{t+1,i} \qquad (8)$$

Because the residuals from the excess return equation are non-normally distributed we used recursive non-parametric boot-strapping methods based on the empirical density function to draw residuals $(\widehat{\epsilon}_{1,i}, \widehat{\epsilon}_{2,i}, ..., \widehat{\epsilon}_{t,i})$ from the excess return equation chosen by a particular model selection criterion. Based on these draws, we choose the value of α_t within the interval [0;1] which maximizes equation (8).[20] Because of the concavity of the utility function, whenever a negative excess return has been predicted $(\widehat{\boldsymbol{\beta}}_{t,i}'\mathbf{X}_{t,i} < 0)$, the investor prefers to go short in shares. However, since the choice of the portfolio weight is constrained to lie in the interval [0;1], the best constrained solution will be for the investor to place all funds in T-bills and nothing in stocks. Similarly, if a large positive excess return has been predicted, the unconstrained solution for α_t may exceed 1, and the best constrained solution is to invest all funds in shares.

Using predictions based on the models selected by Akaike's criterion, Figure 8 presents the sequence of portfolio weights that maximize expected utility under logarithmic preferences. Only in 11 out of 288 forecasting periods, or in about seven percent of all months, did the optimal portfolio weights lie in the interior of [0;1]. In this perspective it is not surprising that the performance of the portfolio based on the weights plotted in Figure 8 was very similar to the performance of the switching portfolio based on the Akaike criterion. For example, in the case of zero transaction costs, the mean and standard deviation of returns based on Akaike's criterion were 17.7 and 9.7 percent, respectively, very similar to the values taken by the corresponding switching portfolio in Table 3. These findings suggest that, subject to the constraint that investors cannot use leverage or go short, the investment weights of the switching portfolios in the previous section are approximately optimal for an expected utility optimizing investor with logarithmic preferences over terminal wealth.

[20]The true value of $\beta_{t,i}$ is unknown, so we condition on the estimate, $\widehat{\beta}_{t,i}$, in the bootstrap experiment. Our procedure also implicitly assumes that the residuals are homoskedastic.

Data Appendix

We used Data Stream as the source for our data which included the following variables

P	:	Value of the Financial Times All-Share (FTALL) Price Index at the end of the month.
DIV	:	12-month moving average of dividends paid by firms included in the FTALL index.
$YALL$	:	Dividend yield computed as DIV_t/P_t.
$GILT$	:	yield on a 2.5 percent government consol measured at the end of the month.
$I3$	:	3-month Treasury Bill rate
	:	
π	:	rate of inflation computed as $\ln(RPI12/RPI12(-12))$, where $RPI12$ is a twelve-month moving average of the Retail Price Index
ΔIP	:	year-on-year rate of change in industrial production of the manufacturing sector.
$\Delta M0$	:	year-on-year rate of change in the narrow money stock $M0$. $M0$ only becomes available in 69m6. Before this period we used notes and coins outstanding (total in million pounds) provided in the CSO publication 'Financial Statistics'. The two measures are very similar, the only difference being Bankers' Operational Deposits with the banking department at the Bank of England.
$\Delta PSPOT$	:	year-on-year change in the spot price of oil (in logs).
Nominal Stock Returns	:	$(P_t + DIV_t - P_{t-1})/P_{t-1}$.
Excess Returns (ERALL)	:	Nominal Stock Returns minus the three month T-bill rate converted to a monthly rate.

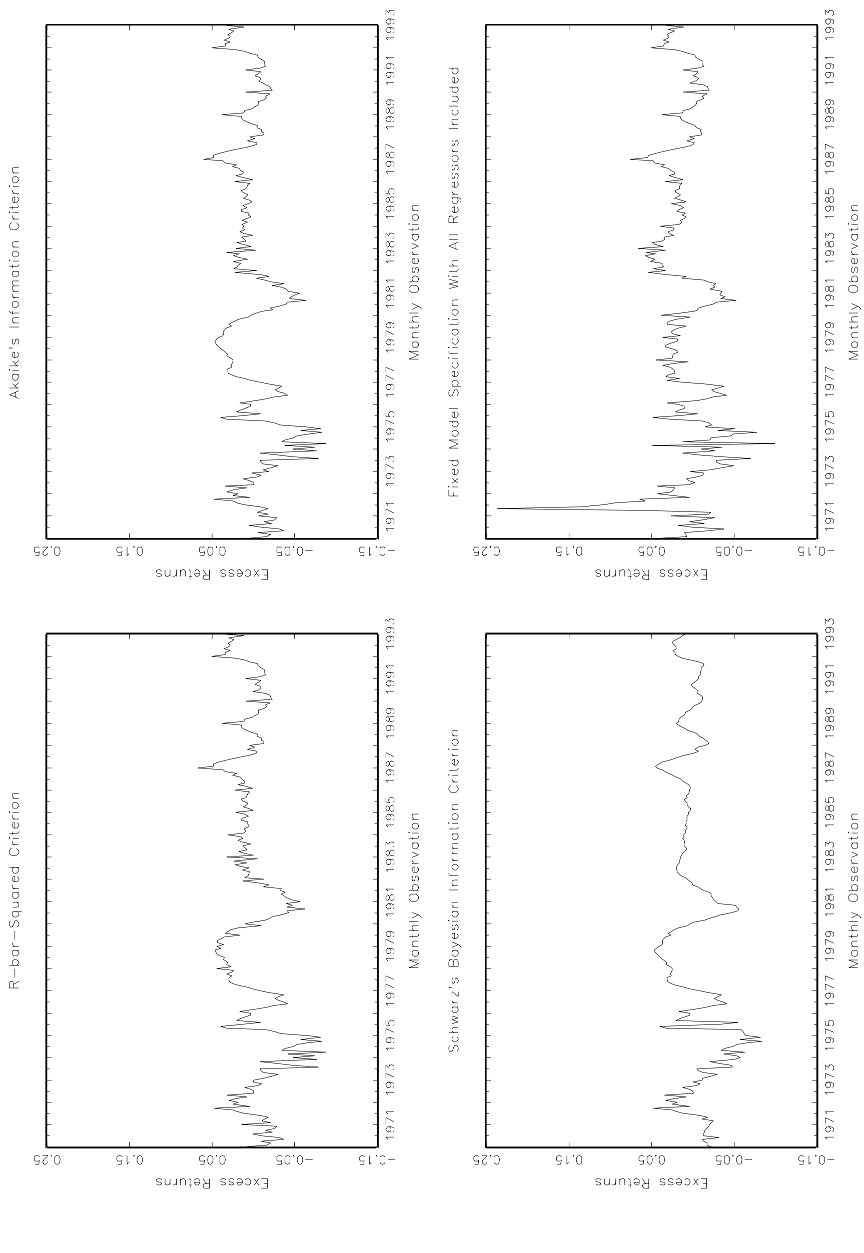

Figure 2: Forecasts of excess returns based on alternative model selection strategies

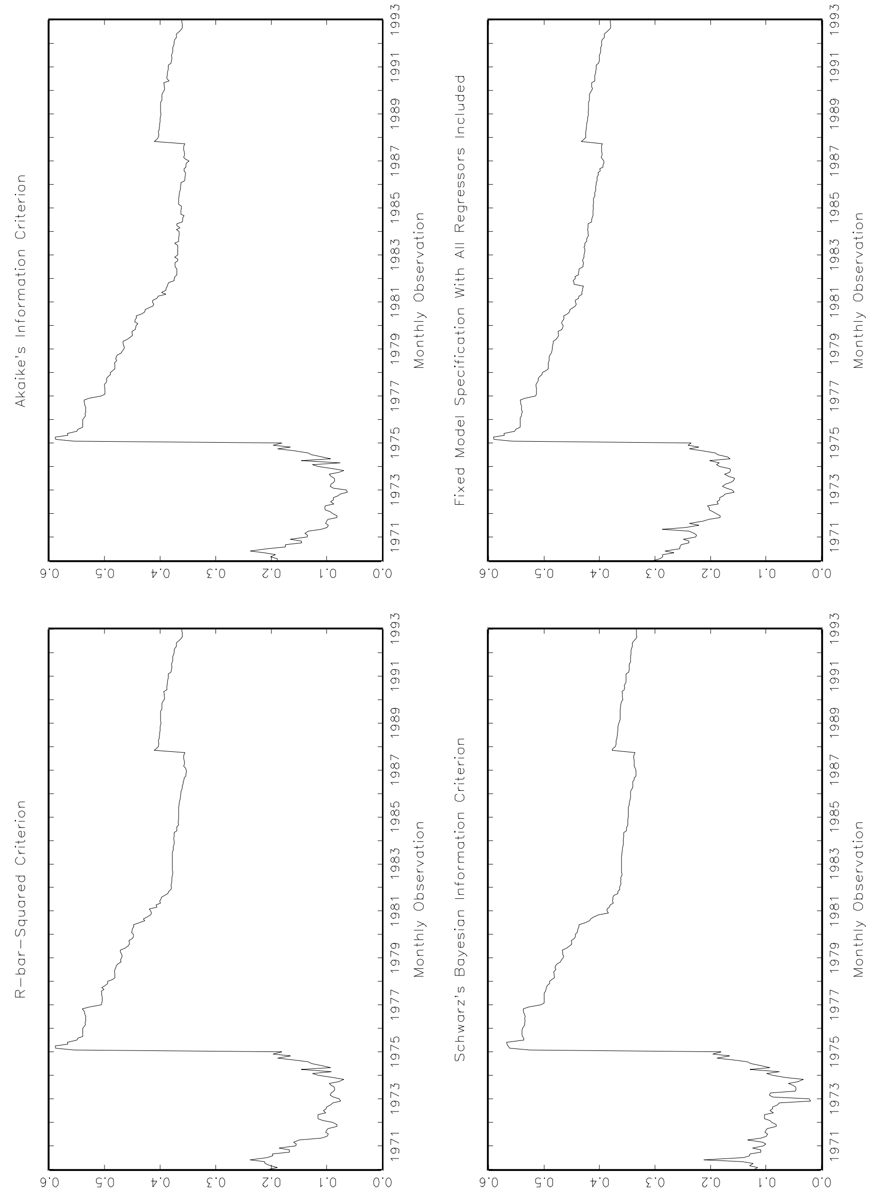

Figure 3: The squared correlation coefficient between the recursively computed forecasts and the realisations of excess returns under alternative model selection strategies

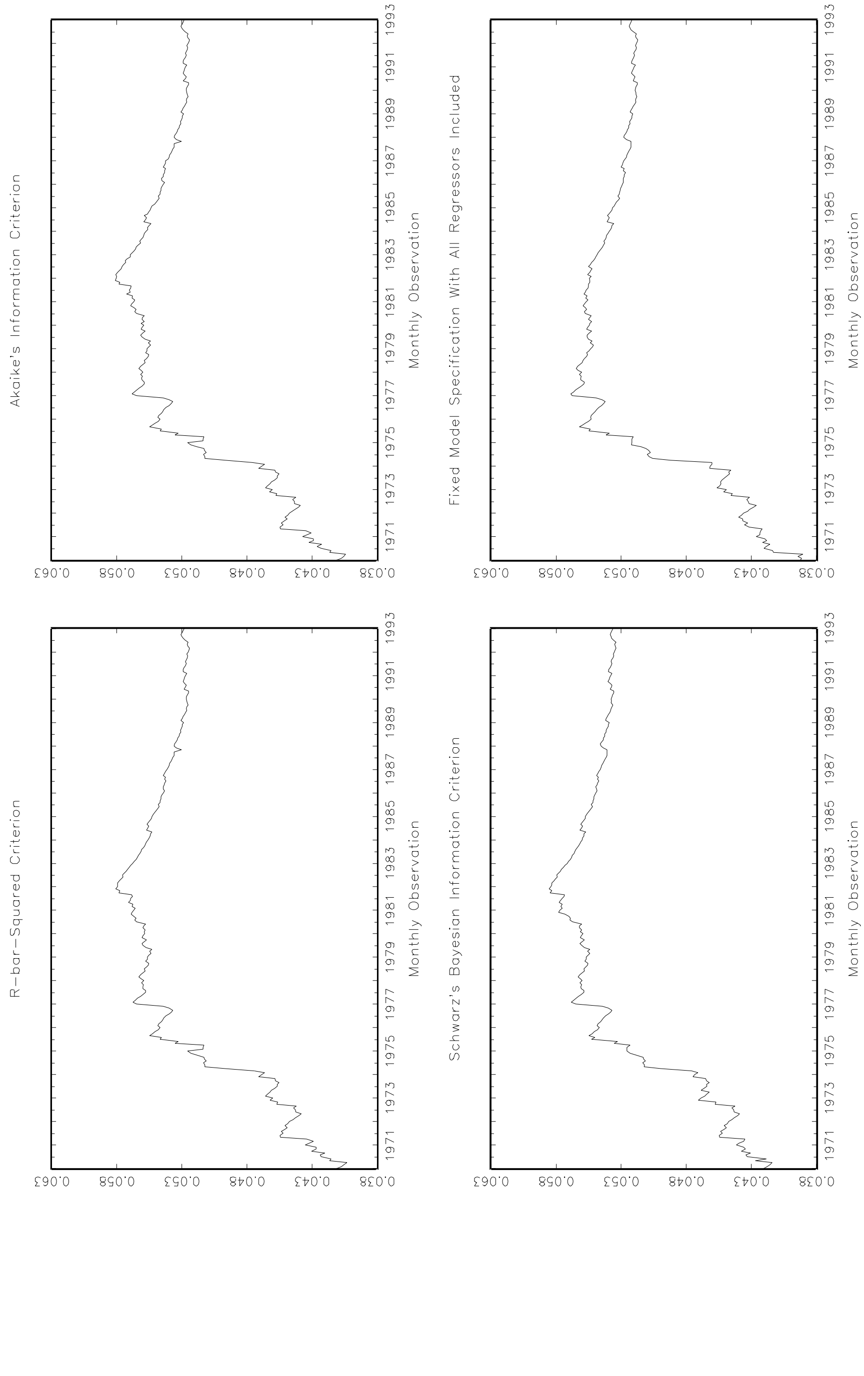

Figure 4: Recursively computed standard errors of excess return regressions under alternative model selection strategies

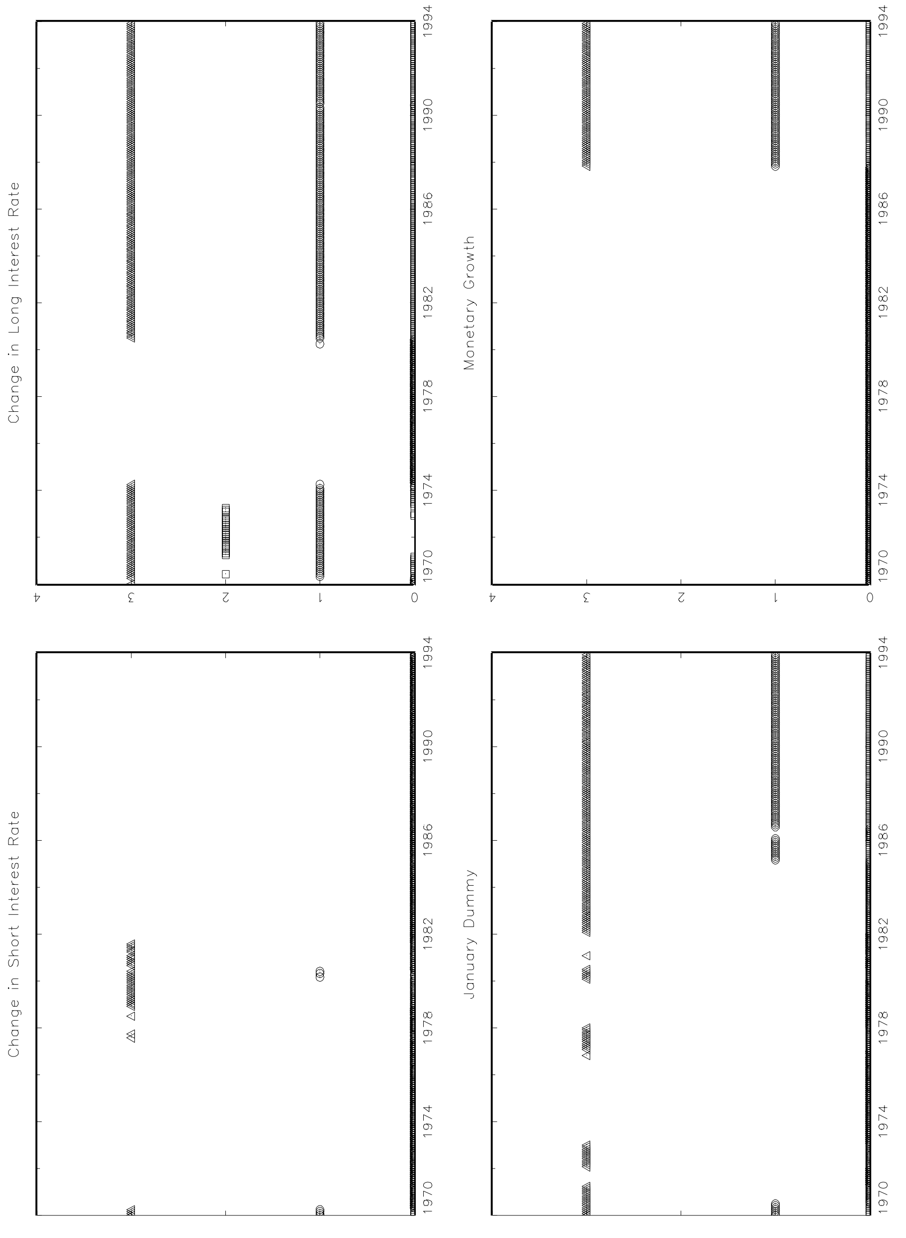

Figure 5: Inclusion frequencies of the variables in the sets B_t and C_t under alternative model selection strategies - labels 1, 2 and 3 represent AIC, Schwarz and $\overline{R}^2$ model selection criteria, respectively.

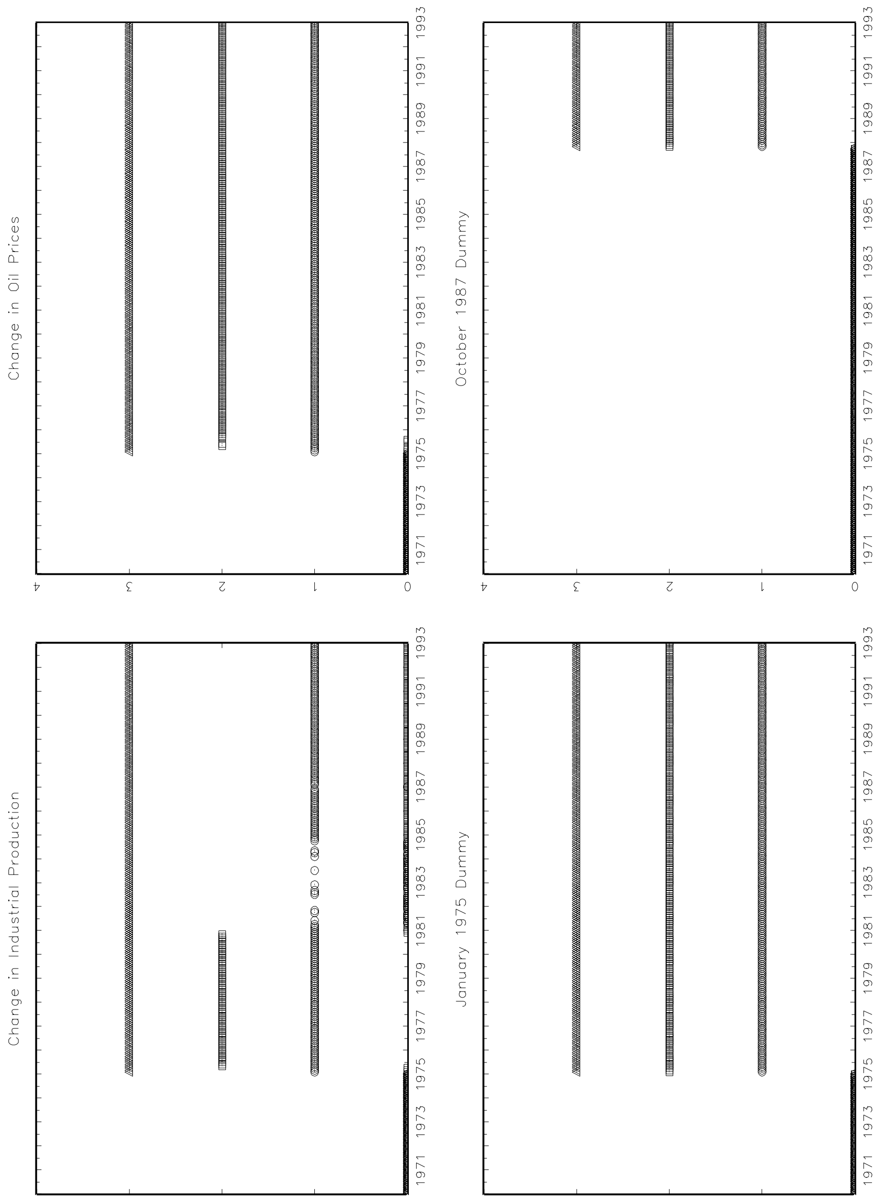

Figure 6: Continuation of Figure 5

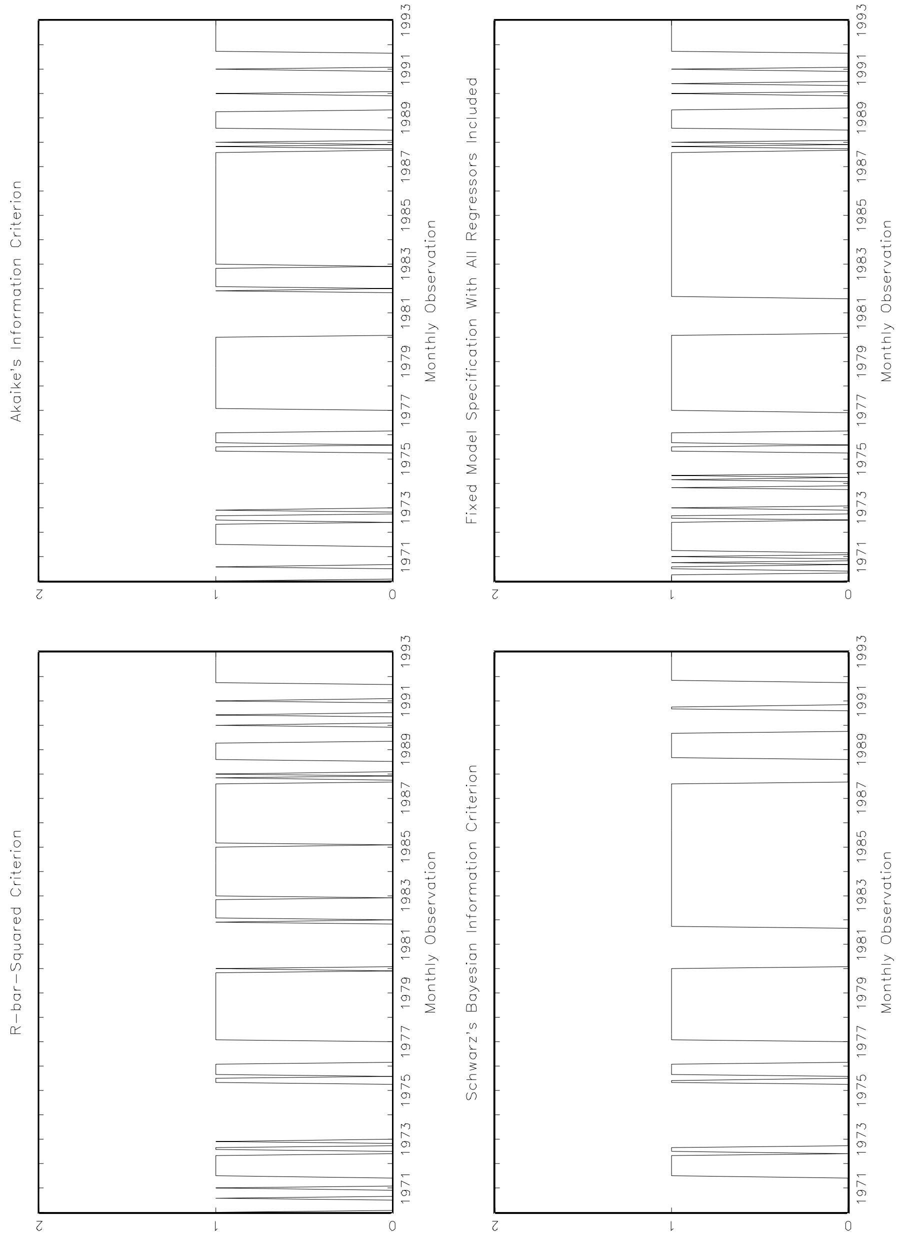

Figure 7: Investment weights (0 for T-bills and 1 for stocks) of the switching portfolios based on alternative model selection strategies

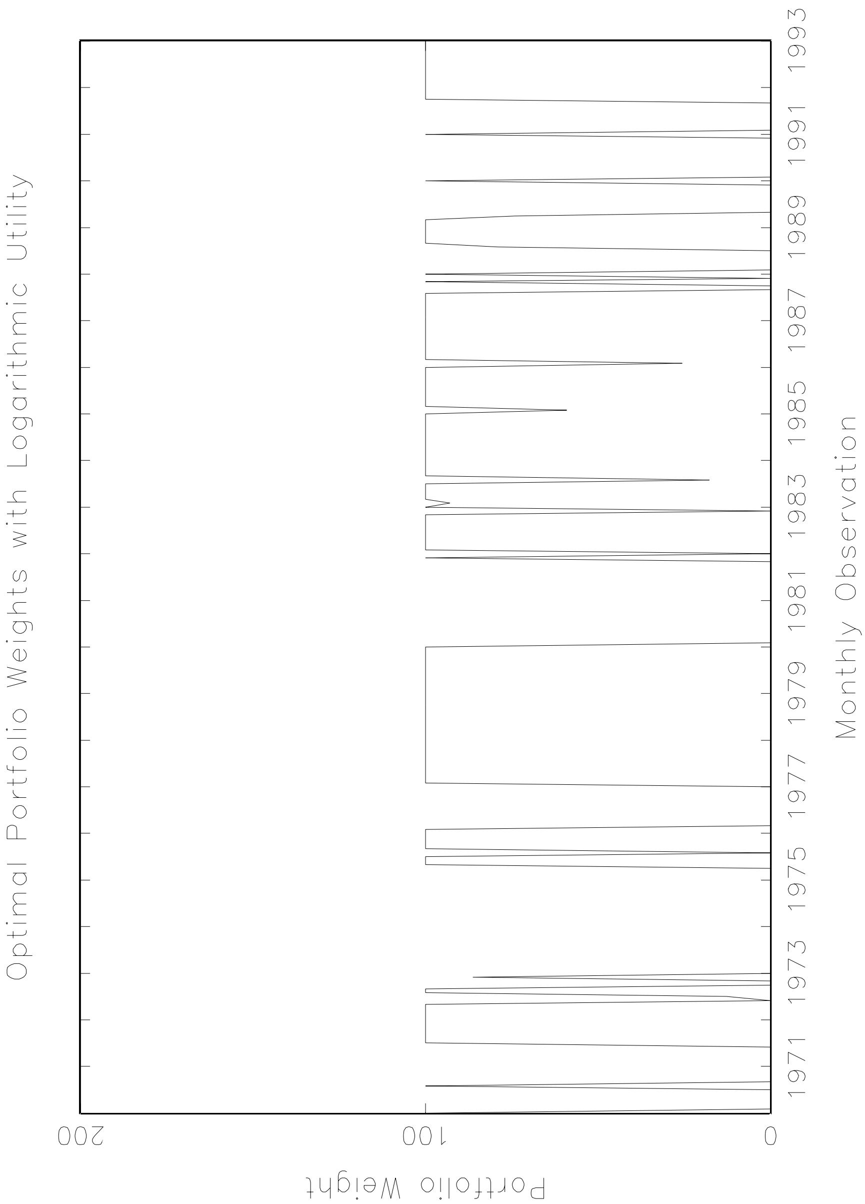

Figure 8: Optimal portfolio weights under logarithmic utility function (o for T-bills and 1 for stocks)

REFERENCES

Akaike, H. (1973) "Information Theory and an Extension of the Maximum Likelihood Principle". In B.N. Petrov and F. Csaki (eds), *Second International Symposium on Information Theory*, pp 267-281. Budapest: Akademiai Kiado.

Angas, L.L.B. (1936) *Investment for Appreciation. Forecasting Movements in Security Prices. Technique of Trading in Shares for Profit.* Macmillan, London.

Balvers, R.J., Cosimano, T.F. and MacDonald, B. (1990) "Predicting Stock Returns in an Efficient Market". *Journal of Finance*, 45, 1109-28.

Black, A. and Fraser, P. (1995) "UK Stock Returns: Predictability and Business Conditions". *The Manchester School Supplement*, 85-102.

Brandt, M.W. (1998) "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach". Working paper, The Wharton School, University of Pennsylvania.

Bray, M. and Savin, N.E. (1986) "Rational Expectations Equilibria, Learning and Model Specification". *Econometrica*, 54, 1129-60.

Breen, W., L.R. Glosten, and R. Jagannathan (1990) "Predictable Variations in Stock Index Returns". *Journal of Finance*, 44, 1177-1189.

Brennan, M.J., Schwartz, E.S., and Lagnado, R. (1997) "Strategic Asset Allocation". Journal of Economic Dynamics and Control, 21, 1377-1403.

Buckle, M.G., Clare, A.D., and Thomas, S.H. (1994) "Predicting the Returns from Stock Index Futures". *Discussion Paper 94-04*, Brunel.

Bulkley, G. and Harris, R.D.F. (1997) "Irrational Analysts' Expectations as a Cause of Excess Volatility in Stock Prices". Forthcoming in *Economic Journal.*

Bulkley, G. and Tonks, I. (1989) "Are UK Stock Prices Excessively Volatile? Trading Rules and Variance Bounds Tests". *The Economic Journal*, vol. 99, 1083-98.

Campbell, J.Y. (1987) "Stock Returns and the Term Structure". *Journal of Financial Economics*, 373- 99.

Campbell, J.Y., and Viceira, L.M. (1998) "Consumption and Portfolio Decisions when Expected Returns are Time Varying". Working paper, Department of Economics, Harvard University.

Campbell, J.Y., Lo, A. W., and MacKinlay, A.C. (1997) "The Econometrics of Financial Markets". Princeton University Press, Princeton, New Jersey.

Clare, A.D., Thomas, S.H., and Wickens, M.R. (1994) "Is the Gilt- Equity Yield Ratio Useful for Predicting UK Stock Return?". *Economic Journal*, 104, 303-15.

Clare, A.D., Psaradakis, Z., and Thomas, S.H. (1995) "An Analysis of Seasonality in the UK Equity Market". *Economic Journal*, 105, 398-409.

De Bondt, W. and Thaler, R. (1990) "Do Security Analysts Overreact?". *American Economic Review*, 80, 52-57.

Dowrie, G.W., and Fuller, D.R. (1950) *Investments.* Second Edition, John Wiley, New York.

Duffie, D. and Pan, J. (1997) "An Overview of Value at Risk". *Journal of Derivatives*, Spring 1997, 7 - 49.

Fama, E.F. (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work". *Journal of Finance*, 25, 383-417.

Fama, E.F. (1981) "Stock Returns, Real Activity, Inflation, and Money". *American Economic Review*, 71, 545-565.

Fama, E.F., and French, K.R. (1989) "Business Conditions and Expected Returns on Stocks and Bonds". *Journal of Financial* Economics, 25, 23-49.

Ferson, W.E., and Harvey, C.R. (1993) "The Risk and Predictability of International Equity Returns". *Review of Financial Studies*, 6, 527-566.

Granger, C.W.J. (1992) "Forecasting Stock Market Prices: Lessons for Forecasters". *International Journal of Forecasting*, 8, 3-13.

Hansen, L.P., and Jagannathan, R. (1991) "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns". Journal of Political Economy, 99, 225-262.

He, H. and Modest, D. (1995) "Market Frictions and Consumption-Based Asset Pricing". Journal of Political Economy, 103, 94-117.

Henriksson, R.D., and Merton, R.C. (1981) "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills". *Journal of Business*, 54, 513-533.

Kandel, S. and Stambaugh, R.F. (1996) "On the Predictability of Stock Returns: An Asset-Allocation Perspective". Journal of Finance, 51, 385-424.

Leitch, G. and Tanner, J.E. (1991) "Economic Forecast Evaluation: Profits Versus the Conventional Error Measures". *American Economic Review*, 81, 580-90.

Lucas, R.E. Jr. (1978) "Asset Prices in an Exchange Economy". *Econometrica*, 46, 1426-1445.

Luttmer, E.G.J. (1997) "What Level of Fixed Costs can Reconcile Asset Returns and Consumption Choices". Working paper, London School of Economics.

Morgan, E.V., and Thomas, W.A. (1962) *The Stock Exchange. Its History and Functions.* Third Edition, Harper Collins Publishers, New York.

Pesaran, M.H. and Pesarann, B. (1997) *Working with Microfit 4.0: Interactive Econometric Analysis.* Oxford University Press, Oxford.

Pesaran, M.H. and Timmermann, A. (1992) "A Simple Nonparametric Test of Predictive Performance". *Journal of Business and Economic Statistics*, 10, 461-65.

Pesaran, M.H. and Timmermann, A. (1994) "Forecasting Stock Returns. An Examination of Stock Market Trading in the Presence of Transaction Costs". *Journal of Forecasting*, 13, 330-365.

Pesaran, M.H. and Timmermann, A. (1995) "The Robustness and Economic Significance of Predictability of Stock Returns". *Journal of Finance*, 50, 1201-1228.

Prime, J.H. (1946) *Investment Analysis.* Prentice Hall, New York.

Rose, H.B. (1960) *The Economic Background to Investment.* Cambridge University Press.

Schwarz, G. (1978) "Estimating the Dimension of a Model". *Annals of Statistics*, 6, 461-464.

Skidelsky, R. (1992) *John Maynard Keynes. The Economist as Savior.* Allen Lane, The Penguin Press.

Theil, H. (1958) *Economic Forecasts and Policy.* Amsterdam, North Holland.

Timmermann, A. (1993) "How Learning in Financial Markets Generates Excess Volatility and Predictability of Excess Returns". *Quarterly Journal of Economics*, 108, 1135-1145.