The trading profitability of forecasts of the gilt–equity yield ratio

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

The Trading Profitability of Forecasts of the Gilt-Equity Yield Ratio

Chris Brooks and Gita Persand
ISMA Centre, Department of Economics, University of Reading
PO Box 242, Whiteknights, Reading, RG6 6BA.

Tel: (+44) 118 931 67 68
Fax: (+44) 118 931 47 41
E-mail: C.Brooks@reading.ac.uk

March 2000

Abstract

Research has highlighted the usefulness of the Gilt-Equity Yield Ratio (GEYR) as a predictor of UK stock returns. This paper extends recent studies by endogenising the threshold at which GEYR switches from being low to being high or vice versa, thus improving the arbitrary nature of the determination of the threshold employed in the extant literature. It is observed that a decision rule for investing in equities or bonds, based on the forecasts from a regime switching model, yields higher average returns with lower variability than a static portfolio containing any combinations of equities and bonds. A closer inspection of the results reveals that the model has power to forecast when investors should steer clear of equities, although the trading profits generated are insufficient to outweigh the associated transactions costs.

Keywords: GEYR; Markov switching; Regime model; Forecasting; Equity & bond returns; Trading rule.

JEL Classifications: C22, G11.

The Trading Profitability of Forecasts of the Gilt-Equity Yield Ratio

Chris Brooks is a Reader in Financial Econometrics at the ISMA Centre, University of Reading. His research interests are in the area of financial econometrics, principally time series modelling and forecasting and risk management, where he has published widely.

Gita Persand is Research Fellow, ISMA Centre, University of Reading. Her research interests are in the area of financial risk management and financial econometrics.

Abstract

Research has highlighted the usefulness of the Gilt-Equity Yield Ratio (GEYR) as a predictor of UK stock returns. This paper extends recent studies by endogenising the threshold at which GEYR switches from being low to being high or vice versa, thus improving the arbitrary nature of the determination of the threshold employed in the extant literature. It is observed that a decision rule for investing in equities or bonds, based on the forecasts from a regime switching model, yields higher average returns with lower variability than a static portfolio containing any combinations of equities and bonds. A closer inspection of the results reveals that the model has power to forecast when investors should steer clear of equities, although the trading profits generated are insufficient to outweigh the associated transactions costs.

1. Introduction

The Gilt-Equity Yield Ratio (GEYR) is defined as the ratio of the income yield on long-term government bonds to the dividend yield on equities. It has been suggested that the current value of the GEYR might be a useful tool for investment managers or market analysts in determining whether to invest in equities or whether to invest in gilts (see Clare *et al.*, 1994). Thus the GEYR is purported to contain information useful for determining the likely direction of future equity market trends. The GEYR is assumed to have a long-run equilibrium level, deviations from which are taken to signal that equity prices are at an unsustainable level. If the GEYR becomes high relative to its long-run level, equities are viewed as being expensive relative to bonds. The expectation, then, is that for given levels of bond yields, equity yields must rise which will occur via a fall in equity prices. Similarly, if the GEYR is well below its long-run level, bonds are considered expensive relative to stocks, and by the same analysis, the price of the latter is expected to increase. Thus, in its crudest form, an equity trading rule based on the GEYR would say, "if the GEYR is low, buy equities; if the GEYR is high, sell equities."

Two recent papers have sought to investigate the predictive power of the GEYR for changes in equity prices and whether the GEYR can be usefully employed in a trading rule to determine when investors should be in equities and when in gilts. Clare *et al.* (1994) evaluate a number of different trading rules incorporating the GEYR, and observe that the rules yield higher average returns and lower standard deviations than a buy-and-hold equities strategy. However, their trading rule analysis is based on just 11 quarterly out-of-sample observations which occur during the 1990's, at a time when GEYR was consistently above its historical average.

Levin and Wright (1998) extend the research of Clare *et al*. (1994) by allowing the threshold value of the GEYR, used to determine whether equities are cheap or expensive, to vary over time according to a set of underlying variables such as the expected inflation rate and the value of the equity risk premium. Their primary finding is that the GEYR provides a valuable asset allocation decision tool so long as allowances are made for time-variations in these factors.

The properties of the GEYR have also been investigated previously by Mills (1991, 1998) and the results also displayed in Mills (1993), where the GEYR is also known as a "confidence factor". Mills argues that a close monitoring of the GEYR could provide useful information concerning possible reversals in stock market prices. Mills (1991) also shows that the stationarity of the logarithm of the GEYR implies that prices, dividends and gilt yields must be cointegrated. The long-run performance of equity markets relative to fixed income is the subject of a book by Siegel (1994), where he shows that stocks substantially outperform bonds in the long run, both during times of high inflation and during eras of stable prices.

On the subject of the relationship between dividends and stock prices, there have been numerous studies spanning the past three decades. In a series of seminal papers, Fama and French (1992, 1995, for example) investigate the effect of size and market to book factors on earnings and stock returns. They find that, when allowing for the value of B/M, small stocks tend to be less profitable than large stocks, particularly in the 1980's. They also find a corresponding relationship between profits and earnings, and thus the market and size factors in earnings help to explain the market and size factors in returns. Further, Fama and French (1992) show that when size and book-to-market ratios are employed to model the cross-

sectional variation in stock prices, there is no role left for the CAPM beta measure in explaining risk for those stocks. Finally Shiller (1981) shows that stock price volatility over the long run is too high to be attributable to new information about expected future dividends.

This paper seeks to advance previous studies on this subject in a number of ways. First, the paper provides a motivation for, and a description of, the Markov switching regime approach, and explains how this might be usefully employed to model the GEYR. Thus the second innovation of the present paper is to provide a specification and forecasts using such a model for determining whether the GEYR is low or high, and therefore whether investors should be in stocks or bonds. This, at least in part, removes the somewhat arbitrary nature of the determination of the values of the thresholds employed by fund managers in their equity versus bond investment rules; for example, the Hoare Govett GEYR Rule states that investors should buy equities if GEYR < 2 and sell equities and invest in gilts if GEYR > 2.4 (see Levin and Wright, 1998). It is important that forecasts are evaluated using a relevant economic loss function as well as on pure statistical grounds, for Dacco and Satchell (1999) argue that traditional measures, such as mean squared error, are not useful methods for evaluating the forecasts of nonlinear models. They document that regime-shifting models will perform badly on mean squared error grounds due to the difficulty in forecasting the regime, even if the model is perfectly well specified within the regime. Dacco and Satchell therefore recommend evaluation methods which are appropriate for the problem, such as the profit-based measures employed in this study. Our focus is on the Gilt-Equity Yield Ratio for the UK, but we also conduct an empirical analysis of this ratio on US and German data over the same period to determine whether our results might be country-specific, or whether they are generally supported.

The remainder of the paper develops as follows. Section 2 outlines the data and its properties and motivates the econometric methodology employed, which is itself described in Section 3. The trading rules used to test the efficacy of the regime switching approach are also presented in Section 3. The results generated thereof are given in Section 4, while Section 5 offers some concluding comments.

2. Data

The sample data used in this study comprises monthly stock index dividend yields and income yields on government bonds covering the period January 1975 until August 1997 for three countries – the UK, the US, and Germany – giving a total of 272 observations. This is a longer span of data than that employed in previous studies, and is of monthly frequency, in contrast to some previous work in the area which uses quarterly data; hence a sufficient amount of data is available to estimate model parameters validly and to conduct a credible out-of-sample forecasting framework. The GEYR is calculated by taking the ratio of the two quantities i.e. the gilt yield divided by the equity dividend yield. The equity variables employed in this study comprise the dividend yield and index values of the FTSE 100 (for the UK), the S&P 500 (for the US), the DAX (for Germany). The bond indices and redemption yields are based on the clean prices of UK government consols, and US/German 10-year government bonds (a consol series is not available for the whole of our sample, neither for the US nor Germany, and so in these cases we use the longest dated government bonds which are).

On plotting the value of the UK GEYR over time in Figure 1, we note that whilst the GEYR is fairly volatile, it seems to exhibit long swings.

[Insert figure 1]

4

For example, the GEYR rose from a low of 1.8 in late 1982 to a peak of 3.2 in October 1987, but then it fell back to 1.9 within three months. Thereafter until the end of the sample, it has remained in the historically narrow (and historically low) range of 1.8 to 2.4. The US and German series, displayed in Figures 2 and 3 respectively, show broadly similar features to the UK series.

[Insert figures 2 and 3]

However, the US series has a general upward trend over the sample period, with only one large downward correction in October 1987. The US series ended our sample period in August 1997 at the highest level (of 4.2) for over 20 years. The German GEYR series is considerably more volatile than the other two, in particular from the mid-1980's onwards, and it also ended the sample period on a historically high note, at around 4.

Summary statistics for the three GEYR series are presented in Table 1.

[Insert table 1]

The mean value of the UK GEYR series over the entire sample period is 2.19, and the range of values for the GEYR is 1.46 to 3.24. This range is smaller than that observed for either comparable GEYR series for the other countries or than that which existed in the 1960's or early 1970's (see Levin and Wright, 1998 for an illustration of the latter). (The results for a similar application of the principles applied here are available for other countries, but are not shown due to space constraints. They are available from the authors upon request). The overall average value of all three series is increased by a series of very high observations between August 1986 and October 1987. Many commentators have argued (with the benefit of hindsight, of course) that such a phenomenon is indicative that equities were over-valued and that a "market correction" was imminent. The three GEYR series are both significantly skewed and leptokurtic with the exception of the German GEYR, which is skewed but not

5

leptokurtic. Thus the Bera Jarque normality test statistics show a rejection of Gaussianity for all significance levels. Although the GEYR series show no evidence of non-stationarity, they are strongly autocorrelated, with the Ljung-Box Q* statistic rejecting the null of no autocorrelation convincingly. Similarly, an application of Engle's (1982) LM test for autoregressive conditional heteroscedasticity demonstrates that the variances of the GEYR are time-varying. These statistics provide motivation for the use of a model with some time-varying component, where the current estimates of both the mean and variance of the series are permitted to depend in some fashion upon their previous values.

In Figures 4 to 6, we also plot the unconditional distribution of the UK, US and German GEYR over the sample period in bold, together with the unconditional distribution of a set of normal variates with the same mean and variance.

[Insert figures 4 to 6]

It is evident that the GEYR series are not only far from normally distributed (being strongly skewed), confirming the inference from the summary statistics, but also have large appendages to one of the tails. There is an upper tail bulge for the UK and German GEYR, and a lower tail bulge for the US. The GEYR has previously been modelled in its logarithmic form (see, for example, Mills, 1991), but whilst the log transform can account for the skewness in the series, it would not adequately model the bulges in the tails. Thus it appears that we might be able usefully to separate the GEYR observations into two regimes: for example, in the UK case, one with a mean just below the unconditional mean of all the observations taken together, and one "high-GEYR" regime, whose observations are fewer in number but whose values are substantially higher than that of the whole series. This structure provides motivation for the use of a switching regime model with two regimes, corresponding to low and high values of the GEYR.

3. Methodology

3.1. Estimation of the Markov switching models

The Markov switching model provides a natural relaxation of the restrictive assumption that all the observations on a particular series are drawn from a normal distribution with constant mean and variance throughout the entire sample period. The model proposed by Hamilton (1989, 1990) assumes that there are K states of nature or regimes, each with mean μ_i and variance σ_i^2 ($i = 1, ..., K$). Restricting interest to the case where $K = 2$, the regime, denoted s_t, that contains the variable of interest, is an unobserved variable taking the value zero or one. This state variable is assumed to follow a first order Markov process with transition probabilities given by

$$p(\,s_t = 1 \mid s_{t\text{-}1} = 1\,) = p_{11}$$

$$p(\,s_t = 2 \mid s_{t\text{-}1} = 1\,) = 1 - p_{11}$$

$$p(\,s_t = 1 \mid s_{t\text{-}1} = 2\,) = p_{22}$$

$$p(\,s_t = 2 \mid s_{t\text{-}1} = 2\,) = 1 - p_{22} \tag{1}$$

Thus the parameters of the model which require estimation are $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p_{11}, p_{22})$. The parameter vector θ is estimated using Hamilton's EM algorithm for maximum likelihood estimation, a procedure which is claimed to be more robust to local function optima than, for example, the more familiar BFGS method (Broyden, 1965, 1967).

Denoting the observed data by y_t, $t = 1, ..., T$, we can express the conditional likelihood function assuming that $y_t \mid s_t \sim N(\mu_{s(t)}, \mathbf{\Omega}_{s(t)})$, $s = 1, 2$ as

$$p(y_t \mid s_t, s_{t-1}, ..., s_1, y_{t-1}, y_{t-2}, ..., y_1; \theta) = \frac{1}{(2\pi|\mathbf{\Omega}_{s(t)}|)^{1/2}} \exp\left[\frac{-\left[y_t - \mu_{s(t)}\right]' \mathbf{\Omega}_{s(t)}^{-1} \left[y_t - \mu_{s(t)}\right]}{2}\right] \tag{2}$$

(Note that the assumption made is that the observed variates y_t are normally distributed conditional upon being in a particular state. Maitland-Smith and Brooks (1999) show that this assumption is far less restrictive and thus more likely to be valid than the assumption that the unconditional distribution of all the data is Gaussian).

In fact, the sample likelihood is estimated as a biproduct of evaluation of the "smoothed" probabilities of the states $p(s_t | y_T, y_{T-1}, ..., y_1; \theta)$. These smoothed probabilities give estimates of the likelihood that y_t was in state 1 at any point in time. This is the quantity of most interest here since it determines whether GEYR is in a "high" or "low" regime. However, in order to ensure that the probability estimates only incorporate information that was available at the time to which the probability refers, the estimation procedure is carried out recursively, starting with the first five years of data and adding an observation at each stage. In each case, the last smoothed probability is recorded. The Markov switching approach is not a model in the usual context of the word; it is sometimes called a "filter" since all that is done effectively is to split the data into two sub-samples with corresponding probability estimates.

3.2. Specification testing

Ideally, one would like to test the null hypothesis that $\mu_1 = \mu_2$ and $\sigma_1^2 = \sigma_2^2$. The implication of finding this would be that the Markov switching model had been unable to separate out the data into two distinct regimes; in other words, that it is possible to restrict the data to be drawn from only one regime. Unfortunately, as Engel and Hamilton (1990) point out, the derivative of the likelihood function with respect to μ_1 and μ_2 is zero, and the information

8

matrix is singular under the null so that the usual regularity conditions for asymptotic validity of hypothesis tests do not hold.

Engel and Hamilton (1990) suggest testing this hypothesis indirectly by considering the following two slightly more general hypotheses:

$$H_0' : \begin{array}{l} p_{11} = 1 - p_{22} \\ \mu_1 \neq \mu_2 \\ \sigma_1 \neq \sigma_2 \end{array}$$

and

$$H_0'' : \begin{array}{l} \mu_1 = \mu_2 \\ \sigma_1 \neq \sigma_2 \end{array}$$

Wald test statistics for testing these hypotheses are given respectively by

$$W_1 = \frac{\left[\hat{p}_{11} - (1 - \hat{p}_{22})\right]^2}{\text{vâr}(\hat{p}_{11}) + \text{vâr}(\hat{p}_{22}) + 2\,\text{côv}(\hat{p}_{11}, \hat{p}_{22})} \tag{3}$$

and

$$W_2 = \frac{\left[\hat{\mu}_1 - \hat{\mu}_2\right]^2}{\text{vâr}(\mu_1) + \text{vâr}(\hat{\mu}_2) - 2\,\text{côv}(\hat{\mu}_1, \hat{\mu}_2)} \tag{4}$$

Both are distributed approximately as χ^2 variates with one degree of freedom under the null hypothesis. The first statistic determines whether there is stability in the regimes over time, by testing how frequently the process is estimated to shift from one regime to another, whilst allowing for differences in the means and variances under the null hypothesis. Intuitively, the higher the values of p_{11} and p_{22}, the more stable are the regimes and the more likely that the null will be rejected. The second null hypothesis tests whether the regimes have significantly different means, whilst allowing for differing variances under the null. Thus taken together, if these two hypotheses are rejected, this provides a strong suggestion that it is useful to allow for regime shifts in the data.

3.3. Alternative models for the GEYR

Although the focus of this study is the Markov regime switching approach, we estimate a number of other popular time series models for forecasting GEYR and for deriving trading rules, in order to benchmark the usefulness of the models. The alternative models employed are a self-exciting threshold autoregressive (SETAR) model, an MA(1)-GARCH(1,1) model, a random walk with drift (RWD), and an autoregressive model of order p where p is determined by in-sample minimisation of Akaike's (1974) information criterion. The MA-GARCH model is often used to summarise the first two moments of financial time series, while the RWD and AR models are standard linear techniques. A full description of these models and estimation issues is given in Brooks (1997) and Brooks and Hinich (1999); however, since the models are employed in this paper purely as benchmarks, and are in common usage in the financial econometrics literature, only the SETAR model is presented; the other models should be self-explanatory.

SETAR models

A general threshold autoregressive model (TAR) can be written

$$y_t = \sum_{j=1}^{J} I_t^{(j)} (\phi_0^{(j)} + \sum_{i=1}^{p_j} \phi_i^{(j)} y_{t-i} + \varepsilon_t^{(j)}), r_{j-1} \leq z_{t-d} \leq r_j \tag{5}$$

where $I_t^{(j)}$ is an indicator function for the jth regime taking the value one if the underlying variable is in state j and zero otherwise. z_{t-d} is an observed variable determining the switching point, r_j is a threshold value, and $\varepsilon_t^{(j)}$ is a zero-mean independently and identically distributed error process. If the regime changes are driven by own lags of the underlying variable, y_t (in this case, the value of the GEYR), i.e. $z_{t-d} = y_{t-d}$, then the model is a self-exciting TAR

(SETAR). The SETAR is associated primarily with Tong (Tong and Lim, 1980; Tong, 1983, 1990; Chan and Tong, 1986; Tsay, 1989), and in this study, the number of regimes, j, is 2, and the orders of the autoregressive components in each part of the model are determined using Akaike's information criterion.

The forecasts from the models are first evaluated using standard forecast error measures: Mean Squared Error (MSE), Mean Absolute Error (MAE), and the proportion of correct direction change predictions. Additionally, we employ the Pesaran–Timmerman (1992, 1994), hereafter PT, test for market timing ability. For this test, we work with changes in the value of the GEYR, so that a negative forecast, for example, implies that the value of the GEYR is predicted to fall during the next period. (We work with changes in the series for the PT test since both positive and negative forecasts are required for the test to be employed). The test is non-parametric, and is generalised from the Henriksson Merton (Henriksson and Merton, 1981; Merton, 1981) test for independence between forecast and realised values. We neither present nor derive the test statistic here, due to space constraints, but instead, we refer the reader to the original articles or Brooks (1997) for further discussion.

3.4. The trading rules

We use the regime switching model in the following way. Engel and Hamilton (1990) show that it is possible to give a forecast of the probability that a series y_t which follows a Markov switching process will be in a particular regime using the following equation

$$p^f_{1t+1|t} = \mu_2 + (\rho + (p_{11} + p_{22} - 1) \times (p_{1t} - \rho)) \times (\mu_1 - \mu_2) \tag{6}$$

where $\rho = \dfrac{1 - p_{22}}{(1 - p_{11}) + (1 - p_{22})}$, p_{1t} denotes the last observed probability of being in regime 1

for that observation and $p^f_{1t+1|t}$ denotes the forecasted probability of being in regime 1 at time

t+1 made at time t.

We use the first 60 observations (January 1975 – December 1979) for in-sample estimation of

the model parameters, $\theta = (\mu_1, \mu_2, \sigma^2_1, \sigma^2_2, p_{11}, p_{22})$. Then a 1-step ahead forecast is produced

of the probability that the GEYR will be in the high mean regime during the next period. If

the probability that the GEYR will be in the low regime during the next period is forecast to

be more that 0.5, then we forecast that the GEYR will be low and hence we buy or hold

equities; if the probability that the GEYR is in the low regime is forecast to be less than 0.5,

then we anticipate that the GEYR will be high and hence we invest in or hold gilts. The

model is then rolled forward one observation, with a new set of model parameters and

probability forecasts being constructed. This process continues until 212 such probabilities

are estimated with corresponding trading rules. The returns for each out-of-sample month for

the switching portfolio are calculated, and their characteristics compared with those of buy-

and-hold equities and buy-and-hold gilts strategies. Returns are calculated as continuously

compounded percentage returns on a stock (the FTSE in the UK, the S&P 500 in the US, the

DAX in Germany) or on a long-term government bond.

We also evaluate the trading performance of the other models, although in these cases,

determining whether we should be in equities or bonds is not so straightforward since the

forecasts from the alternative models are all of the value of the GEYR, and do not necessarily

pertain to a regime, and therefore no trading mechanism arises naturally from the predictions.

The following approach is therefore adopted. If the 1-step ahead forecasted value (using the

SETAR, RWD, AR, or MA-GARCH) of the GEYR is higher than its long-term mean, then we invest in bonds. If it is lower, then we invest in equities.

4. Results

4.1. The Markov switching models

Some summary statistics for the estimated Markov switching parameters are given in Table 2.

[Insert table 2]

The means and variances for the values of the GEYR for each of the two regimes are given in columns 1 to 4 of the table with standard errors associated with each parameter in parentheses. It is clear that the regimes switching model has split the data into two distinct samples – one with a high mean (of 2.43, 2.46 and 3.03 for the UK, US and Germany respectively) and one with a lower mean (of 2.07, 2.12, and 2.16), as was anticipated from the unconditional distribution of returns. Also apparent is the fact that the UK and German GEYR are more variable at times when it is in the high mean regime, evidenced by their higher variance (in fact, it is around four and twenty times higher than for the low GEYR state respectively). The number of observations for which the probability that the GEYR is in the high mean state exceeds 0.5 (and thus when the GEYR is actually deemed to be in this state) is 102 for the UK (37.5% of the total), while the figures for the US are 100 (36.8%) and for Germany 200 (73.5%). Thus, overall, the GEYR is more likely to be in the low mean regime for the UK and US, while it is likely to be high in Germany. The results suggest that the model is appropriate, given the "lump" in the upper tail of the unconditional distribution of the UK GEYR in Figure 4.

Columns 6 and 7 of Table 2 give the values of p_{11} and p_{22} respectively, that is the probability of staying in state 1 given that the GEYR was in state 1 in the immediately preceeding month,

and the probability of staying in state 2 given that the GEYR was in state 2 previously, respectively. The high values of these parameters indicate that the regimes are highly stable with less than a 10% chance of moving from a low GEYR to a high GEYR regime and vice versa for all three series. This is also observed when we plot in Figure 7 the value of the GEYR for the UK (in bold) with the current probability that it is in the high GEYR regime (dotted line).

[Insert figure 7]

The regime switches occur only 16 times in the UK case, four times for the US, and six times for Germany during the out-of-sample period. After a long period of recommending that investors in the markets should choose equities (July 1992 – July 1997), during a time when the GEYR was low and stable and predicted to stay that way, our model suggests that equities were over-valued at the end of the sample, or at least that a period of instability is likely to occur. It is also worth noting that the probabilistic model is likely to be a more useful tool than simply looking at the current value of the GEYR since the former additionally uses past information and historical patterns in the value of the GEYR in order to make its predictions.

4.2. Results of specification tests for the Markov switching models

The W_1 and W_2 statistics of equations (3) and (4) are presented in the final two columns of Table 2. Since these figures are well in excess of their critical values (3.84 at the 5% level), it is concluded that the null hypotheses that there exist only one regime are very strongly rejected for all three GEYR series. This result provides further suggestion that regime switching models should capture important stylised features of the data.

4.3. Forecast evaluation using traditional statistical loss functions

The forecasting accuracy summaries of the various models are presented in Table 3.

[Insert table 3]

Panels A – E show the mean squared error, mean absolute error, percentage of correct direction predictions, and the value of the Pesaran–Timmerman statistic for the Markov switching model, the SETAR model, the RWD, the AR, and the MA-GARCH model respectively. It is evident that under standard prediction evaluation metrics, the Markov switching model is not a star performer. It never provides the smallest MSE or MAE; it is the second best of five models for the UK when measured in this way, the fourth best for the US, and the worst in the German case. On the other hand, the linear autoregressive model provides the lowest MSE and MAE for all three time series.

When evaluated according to the percentage of correct direction predictions, the Markov switching model is superior for the UK, predicting over 53% of next-day directions correctly (that is, correctly predicting GEYR to rise, or correctly predicting GEYR to fall). For the US and Germany, on the other hand, the switching model is again out-performed by the linear autoregression. The Hamilton Markov approach out-performs the explicit threshold model for the UK, as evidenced by its lower MSE and MAE, while the latter is slightly preferable for the US and Germany.

None of the models, for any of the three series, have market timing ability according to the Pesaran–Timmerman test. The best model is the Markov switching approach for the UK, which generates a PT statistic of 0.90, still a long way short of the corresponding standard normal critical values. Elsewhere in Table 3, we find a number of negative PT statistics, for example, the SETAR and MA-GARCH models for all three series, although none are significant. The interpretation of these is that the models have produced a lower proportion of correct direction changes than would have been the case if the forecasts and the realised

values were independent, or to put this another way, one would have had more chance of success in predicting directions over repeated trials by flipping a coin.

4.4. The profitability of the trading rules

The performance of the regime switching trading rule together with a strategy of buy-and-hold equities and buy-and-hold bonds are given in Table 4 for each of the three countries.

[Insert table 4]

The table demonstrates the superiority of the Markov switching model of the GEYR for making investment decision compared to a simple buy-and-hold equities strategy. In the UK context, the former yields higher average returns and lower standard deviations. The switching portfolio generates an average return of 0.69% per month, compared with 0.43% for the pure bond and 0.62% for the pure equity portfolios. It is also clear from Figure 3 that the use of the regime switching model has allowed the investor to move out of equities well before the stock market crash of October 1987, although we would have been in equities during the mid-late 1990's bear run. The improvements are not so clear-cut for the US and Germany.

It is not clear from the figures in Table 4, however, that the Markov switching model yields forecasts that represent a universal improvement over the buy-and-hold bonds or buy-and-hold equities portfolios, even in the UK, since the switching portfolio has higher standard deviation than the bond portfolio. Thus it might be that if investors invested (for the whole out-of-sample period) part of their wealth in equities, and part in bonds, then they may be able to achieve similar returns to the switching portfolio, but with lower standard deviation and therefore less risk. In order to investigate this possibility, Figure 8 shows the mean– variance

efficient frontier for all combinations of bond and equity portfolio weightings, together with the risk-return characteristics for the switching portfolio for the UK.

[Insert figure 8]

Again, the benefit from employing a regime switching approach is evident, for the switching portfolio risk-return combination lies well above the efficient frontier. In fact, for the given level of risk, the switching portfolio yields average monthly returns 0.15 percentage points higher than the static combination of bonds and equities. To state this in other terms, in order to achieve such a high return, one would have to short sell the risk-free asset by 35% of the total value of the portfolio, and invest the proceeds in equities. This would presumably imply an unacceptable level of risk to most investors. Therefore, it can be concluded that the switching portfolio dominates any static combination of stocks and gilts in terms of its risk-return characteristics.

In order to evaluate further the usefulness of the Markov switching model, and of its competitors, in generating profitable trading rules, we also calculate the Sharpe ratio, which measures the excess reward (over a risk-free rate) to risk ratio, and is defined as

$$S_i = \frac{R_i - RFR}{SD_i} \qquad (7)$$

where S_i denotes the value of the Sharpe ratio for strategy i, R_i denotes the average return over the period, RFR denotes the average value of a risk-free proxy over the period, and SD_i denotes the standard deviation of returns to strategy i. By definition, the buy-and-hold bonds strategy would have a zero Sharpe ratio, and the superior model is the one with the largest value of the ratio. (Strictly speaking, the returns to a short term government bond should be used as a risk-free proxy, rather than those of a long term bond. However, for comparability with the remainder of our analysis, which is focused on long term bonds, we use the return on

a long term government bond in the calculation of (9). This should be inconsequential as we treat all models identically, thus preserving the ordering of performance). The Sharpe ratio for the UK Markov switching portfolio is almost twice that of the buy-and-hold equities portfolio, suggesting that, after allowing for risk, the switching model provides a superior trading rule. The improvement in the Sharpe ratio for the other two models is, contrarily, only very modest.

However, a static portfolio (containing any combination of equities and bonds) would incur only two sets of charges: one when the position was opened at the start of the sample, and one when it is closed at the end. On the other hand, the switching model made a total of 16 switches, which will incur significantly higher transactions costs. Sutcliffe (1997) suggests that an appropriate "round-trip" figure for transacting in the FTSE-100 is 1.7% of the value of the purchase/sale per transaction for an investor. This figure is made up of bid/ask spread (0.8%), stamp duty (0.5%) and commission (0.4%). For simplicity, we assume identical costs for the other two markets. Calculating the average return for the switching portfolio taking this into account, gives a figure for the UK of 0.56%, compared with 0.69% previously. This would imply that the switching portfolio had lower average returns, but with lower risk, than a static equity portfolio. In fact, the UK switching portfolio would be very slightly below the mean–variance efficient frontier for the static portfolio, indicating that, net of transactions costs, the GEYR regime switching model is unable to beat a buy-and-hold strategy. Thus the Sharpe ratio is now slightly below that of the buy-and-hold equities portfolio. For the US and Germany, however, the Sharpe ratios are substantially below those of the buy-and-hold equities strategy, once transactions costs are allowed for, even though the number of trades is considerably smaller, at 4 and 6 respectively. Another feature worth noting is that, at least before transaction costs, the Markov model for the UK is able to generate useful trading

profits, in spite of the fact that the Pesaran–Timmerman statistic was not significant. This seems to indicate that the latter test is rather conservative, and perhaps even lacking in power to detect profitable trading predictions.

Table 5 presents the trading performance of the alternative forecasting models for comparison.

[Insert table 5]

For the US, the Markov switching model is superior to all of the alternatives, given its higher Sharpe ratio, although the story is different for the UK and Germany. In the latter case, the MA-GARCH model is the best, yielding a ratio of 0.17, larger than that of the buy-and-hold equities strategy. However, after allowing for transactions costs, the SETAR model is superior and has a very slightly higher Sharpe value due to the small number of switches that it makes. For the UK, both the SETAR and AR models have higher Sharpe ratios than the switching model, net of transactions costs.

5. Conclusions

This paper sought to employ a regime switching model for the GEYR in the UK. It has been shown that such a model yields forecasts which generate investment decisions with strongly superior risk-return characteristics compared with a buy-and-hold strategy for the UK, and slightly better for the US and German markets. The Markov switching model provides superior predictions to those of its competitors (SETAR, MA-GARCH, RWD, AR) for the UK when evaluated in this way, although the Markov approach is inferior on standard forecast error grounds. Our results broadly concur with existing studies in documenting the usefulness of the GEYR as a variable upon which to base investment decisions.

At first glance, such a conclusion may appear to be incompatible with the idea of efficient markets (see Fama, 1970 or 1991). Clare *et al.* (1994) suggest that the predictive power of GEYR may in part lie in the emphasis of pension fund managers on generating a continual stream of income. This group of large investors will thus focus their attention upon income flows derived from asset yields rather than the capital gains or losses that preoccupy the majority of academic studies. Levin and Wright (1998) point out that variations in the GEYR could signal either a profitable investment opportunity as a result of equity mis-pricing, which will disappear as the GEYR moves back to its equilibrium value, or that the equilibrium value of the GEYR itself has shifted. The implication of this is that the observed predictability of equity market returns may be a statistical anomaly, caused by time-varying equity risk premia, for example. In addition, if we take a more modern definition of market efficiency, such as that of Jensen (1978), then since the model employed here is unable to gain a profit net of transactions costs, the limited predictive power of the GEYR observed here is entirely consistent with financial market efficiency since the models are not able to make excess profits net of reasonable transactions costs. However, transactions costs for a market maker would be considerably lower (perhaps of the order of 0.5%), and additionally, transactions costs are always under pressure as market volumes increase and spreads narrow. Therefore, there do exist circumstances under which a regime switching model of the type employed here could make positive risk- and transactions cost-adjusted returns. This is at least indicative that further research is warranted on how the GEYR might best be modelled and forecast, and how it can be utilised to guide the investment decisions of investors. One possible avenue for further research would be to generate multi-step ahead predictions, or to use the actual value of the probability forecast in order to derive a filter rule – for example, "invest in equities iff the probability that the GEYR will be low during the next period exceeds 0.8" in order to minimise transactions costs. Additionally, other linear or nonlinear

time series or structural models might usefully be employed to capture the stylised features of the GEYR.

References

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control,* AC-19(6), 716-723.

Brooks, C. (1997). Linear and nonlinear (non-) forecastability of daily sterling exchange rates, *Journal of Forecasting,* 16, 125-145.

Brooks, C., & Hinich, M.J. (1999). Cross-correlations and cross-bicorrelations in sterling exchange rates, *Journal of Empirical Finance,* 6(4), 385-404.

Broyden, C.G. (1967). Quasi-Newton methods and their application to function minimisation, *Mathematics of Commutation,* 21, 368-381.

Broyden, C.G. (1965). A class of methods for solving nonlinear simultaneous equations, *Mathematics of Computation,* 19, 577-593.

Chan, W.S., & Tong, H. (1986). On estimating thresholds in autoregressive models, *Journal of Time Series Analysis*, 7(3), 179-190.

Clare, A.D., Thomas, S.H., & Wickens, M.R. (1994). Is the gilt-equity yield ratio useful for predicting UK stock returns?, *Economic Journal,* 104, 303-315.

Dacco, R., & Satchell, S. (1999). Why do regime-switching models forecast so badly?, *Journal of Forecasting,* 18, 1-16.

Engel, C. (1994). Can the Markov switching model forecast exchange rates?, *Journal of International Economics,* 36, 151-165.

Engel, C., & Hamilton, J.D. (1990). Long swings in the dollar: Are they in the data and do markets know it?, *American Economic Review,* 80, 689-713.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica,* 50, 987-1007.

Fama, E. (1991). Efficient capital markets: II, *Journal of Finance,* 46, 1575-1617.

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work, *Journal of Finance,* 25, 383-417.

Fama, E., & French, K.R. (1995). Size and book-to-market factors in earnings and returns, *Journal of Finance,* 50, 131-155.

Fama, E., & French, K.R. (1992). The cross-section of expected stock returns, *Journal of Finance,* 47, 427-465.

Hamilton, J.D. (1990). Analysis of time series subject to changes in regime, *Journal of Econometrics,* 45, 39-70.

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

Henriksson, R.D., & Merton, R.C. (1981). On market timing and investment performance: II. Statistical procedures for evaluating forecasting skills, *Journal of Business,* 54, 513-533.

Jensen, M.C. (1978). Some anomalous evidence regarding market efficiency, *Journal of Financial Economics,* 6, 95-101.

Levin E.J., & Wright, R.E. (1998). The information content of the gilt-equity yield ratio, *Manchester School Supplement,* 89-101.

Maitland-Smith, J.K., & Brooks, C. (1999). Threshold autoregressive and markov switching models: An application to commercial real estate, *Journal of Property Research*, 16(1), 1-19.

Merton, 1981

Mills, T.C. (1998). Recent developments in the modelling of nonstationary VARs, *Journal of Economic Surveys,* 12, 279-312.

Mills, T.C. (1993). *The Econometric Modelling of Financial Time Series,* Cambridge: University Press.

Mills, T.C. (1991). Equity prices, dividends and gilt yields in the UK: Cointegration, error correction, and confidence, *Scottish Journal of Political Economy,* 38(3), 242-255.

Pesaran, M.H., & Timmerman, A. (1994). A generalisation of the non-parametric Henriksson–Merton test of market timing, *Economics Letters,* 44, 1-7.

Pesaran, M.H., & Timmerman, A. (1992). A simple non-parametric test of predictive performance, *Journal of Business and Economic Statistics,* 10(4), 461-465.

Shiller, R.J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends, *American Economic Review,* 420-436.

Siegel, J.J. (1994). *Stocks for the Long Run,* Chicago: Irwin.

Sutcliffe, S. (1997). *Stock Index Futures: Theories and International Evidence,* Second Edition, London: Thomson Business Press.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical Systems Approach,* Oxford: University Press.

Tong, H. (1983). *Threshold Models in Nonlinear Time Series Analysis*, N.Y.: Springer-Verlag.
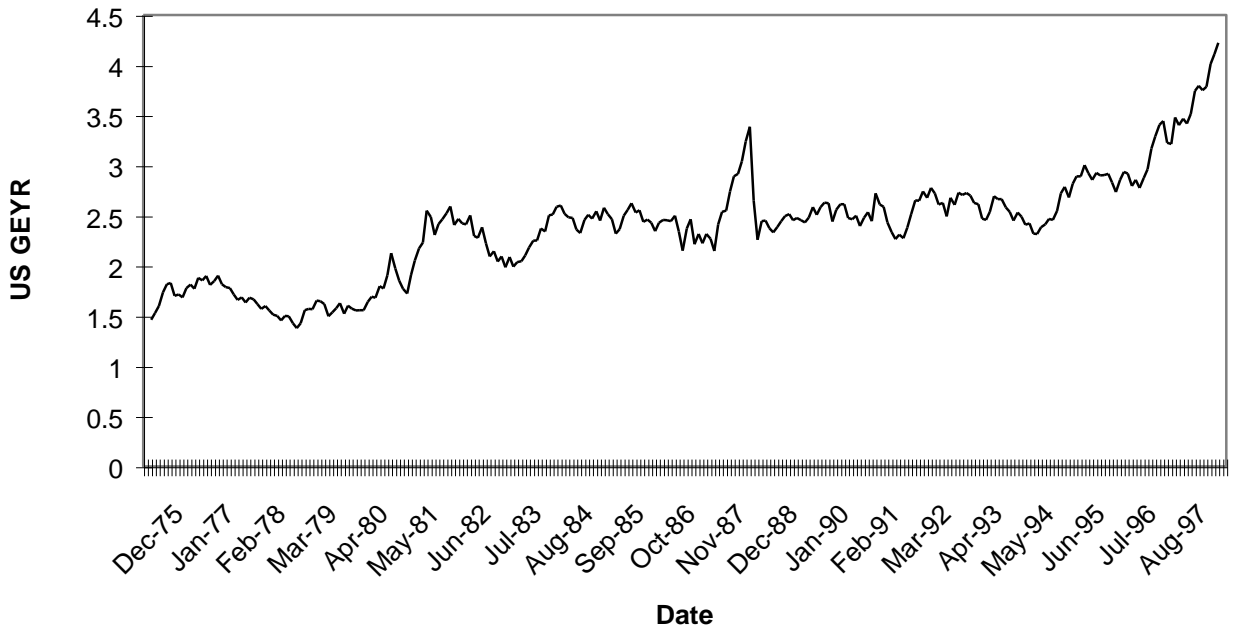
Tong, H., & Lim, K.S. (1980). Threshold autoregression, limit cycles and cyclical data, *Journal of the Royal Statistical Society B,* 42(3), 245-292.

Tsay, R.S. (1989). Testing and modelling threshold autoregressive processes, *Journal of the American Statistical Association,* 84, 231-240.
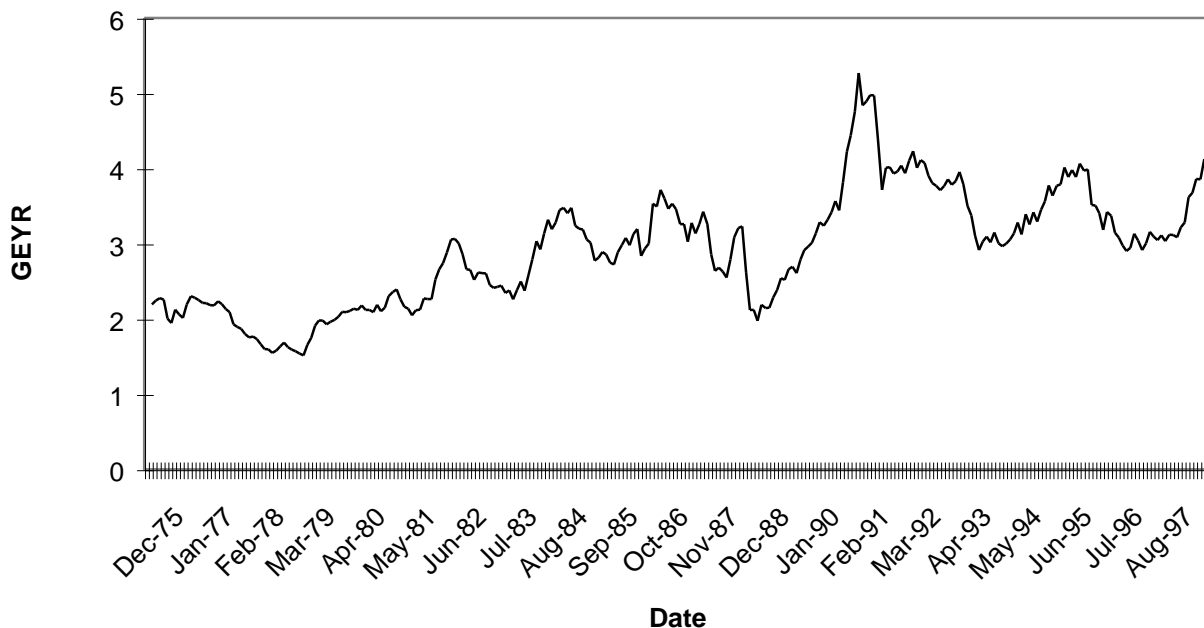
Figure 1: The value of UK GEYR 1975-97
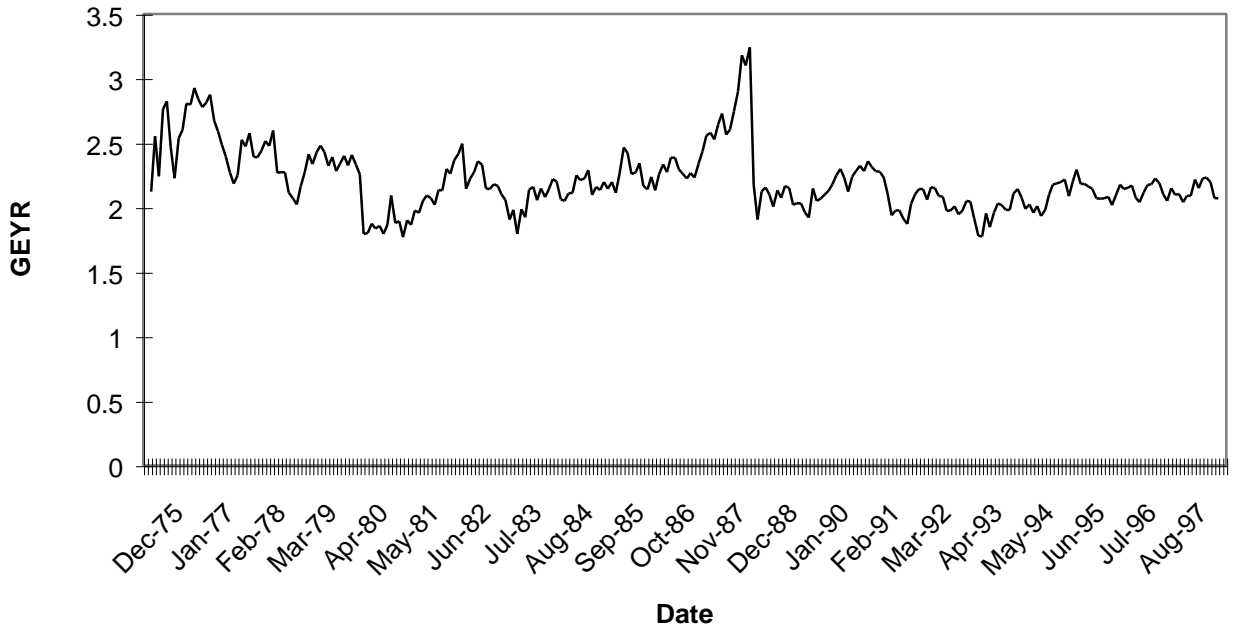
Figure 2: The value of US GEYR 1975-97

## Figure 3: The value of German GEYR 1975-97

Date

Table 1: Summary statistics for UK, US and German GEYR
December 1975 - August 1997

Statistic	UK	US	Germany
Number of observations	272	272	272
Mean	2.190	2.314	2.828
Variance	0.066	0.200	0.569
Skewness	0.898** [0.000]	0.991** [0.000]	0.552** [0.000]
Excess Kurtosis	2.362** [0.000]	3.172** [0.000]	-0.081 [0.789]
Minimum	1.463	1.448	1.561
Maximum	3.241	4.222	5.166
Bera-Jarque Normality Test	106.019** [0.000]	1578350** [0.000]	24.908** [0.000]
Dickey Fuller Test	-4.961**	-2.977*	-4.758** [0.000]
Augmented Dickey Fuller Test (12 lags)	-3.722**	-10.740**	-7.865** [0.000]
Ljung-Box Q*-statistic	778.555** [0.000]	1752.697** [0.000]	4229.389** [0.000]
ARCH(4)	214.304** [0.000]	245.714** [0.000]	252.078** [0.000]

Notes: Marginal significance levels are shown in square brackets where appropriate. The 5% critical values for the Dickey Fuller and augmented Dickey Fuller tests are -2.8726; ARCH(4) denotes Engle's LM test for autoregressive conditional heteroscedasticity of up to fourth order. The test statistic follows a $\chi^2(4)$ under the null of no ARCH.

Figure 4: The unconditional distribution of the UK GEYR
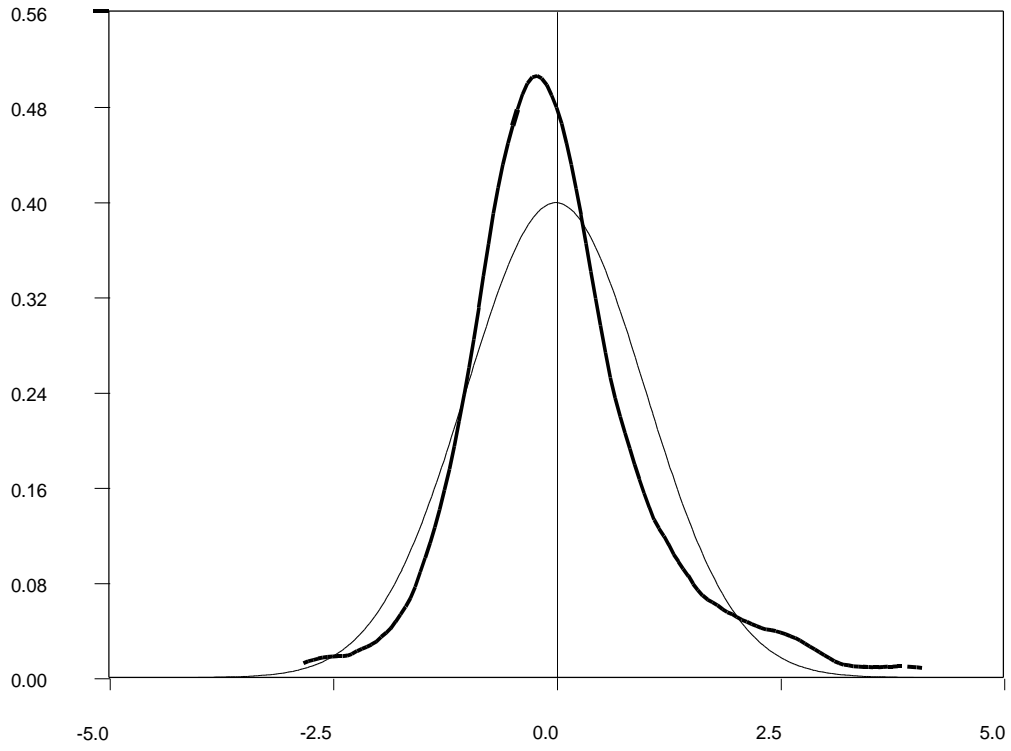
Figure 5: The unconditional distribution of US GEYR
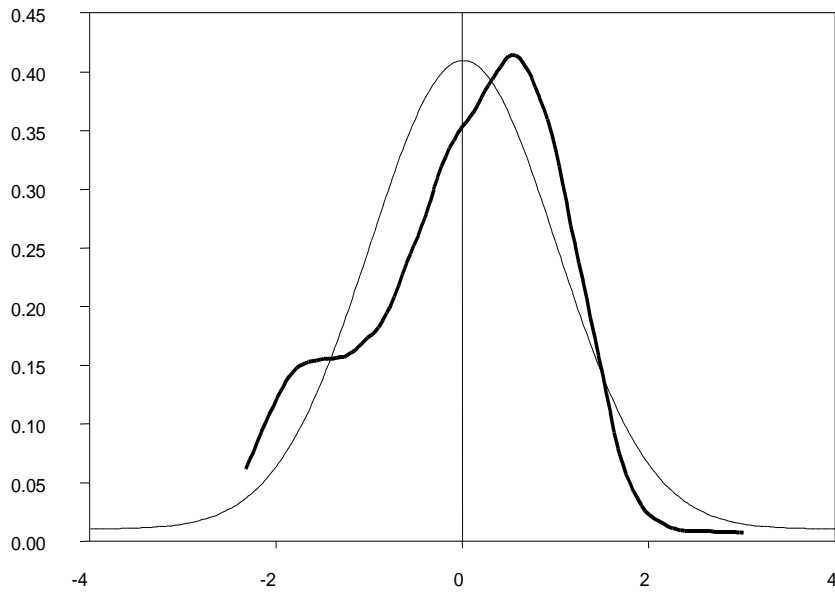
Figure 6:The unconditional distribution of German GEYR
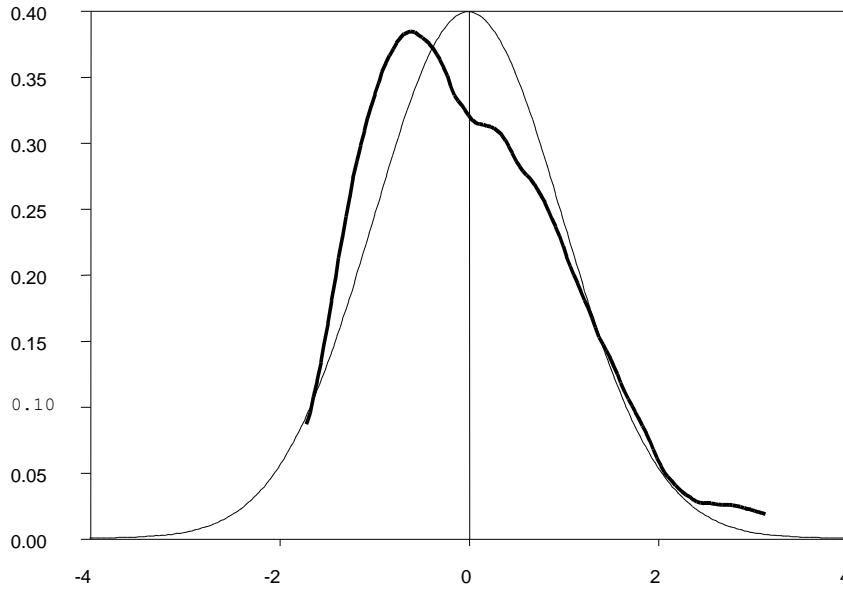

Table 2: Estimated parameters for the Markov switching models (whole sample)

Statistic	μ_1	μ_2	σ_1^2	σ_2^2	p_{11}	p_{22}	N_1	N_2	W_1	W_2
UK	2.429	2.075	0.062	0.014	0.955	0.972	102	170	1038.230	61.280
	(0.030)	(0.037)	(0.009)	(0.002)	(0.073)	(0.013)				
US	2.455	2.122	0.029	0.040	0.972	0.982	100	172	583.220	26.530
	(0.181)	(0.062)	(0.060)	(0.004)	(0.017)	(0.011)				
Germany	3.025	2.156	0.551	0.013	0.982	0.933	200	72	734.790	78.340
	(0.054)	(0.015)	(0.057)	(0.002)	(0.011)	(0.032)				

Notes: Standard errors in parentheses; N_1 and N_2 denote the number of observations deemed to be in regimes 1 and 2 respectively. W_1 and W_2 are asymptotically distributed as a $\chi^2(1)$ with critical value 3.84 at the 5% level under the null hypothesis.

Table 3: Forecast summary statistics

	UK	US	Germany
Panel A: Markov switching model			
MSE	0.057	0.221	1.189
MAE	0.144	0.313	0.877
% correct signs	53.365	47.596	0.486
PT statistic	0.905	-0.940	-0.492
Panel B: SETAR model			
MSE	0.064	0.165	0.366
MAE	0.187	0.262	0.466
% correct signs	44.286	48.571	46.667
PT statistic	-1.255	-0.881	-0.883
Panel C: Random walk with drift			
MSE	0.066	0.163	0.358
MAE	0.189	0.272	0.470
% correct signs	-	-	-
PT statistic	-	-	-
Panel D: AR(p) model			
MSE	0.016	0.018	0.033
MAE	0.086	0.094	0.131
% correct signs	42.381	52.381	51.905
PT statistic	-1.469	0.373	0.074
Panel E: MA(1)-GARCH model			
MSE	0.258	0.878	0.646
MAE	0.424	0.773	0.845
% correct signs	49.048	46.667	50.952
PT statistic	-0.035	-1.649	-0.768

Table 4: Average returns and volatility of returns for

a bond, an equity, and the Markov switching portfolios

|	Buy-and-hold bonds portfolio	Buy-and-hold equities portfolio	Switching portfolio		Number of switches
Panel A: UK					
Average return	0.430	0.622	0.689	(0.561)	16
Standard deviation of returns	0.898	2.145	1.582		
Sharpe ratio	0.000	0.090	0.164	(0.083)	
Panel B: US					
Average return	0.051	0.466	0.326	(0.294)	4
Standard deviation of returns	1.097	1.725	1.072		
Sharpe ratio	0.000	0.241	0.257	(0.227)	
Panel C: Germany					
Average return	0.024	0.376	0.166	(0.1129)	6
Standard deviation of returns	0.739	2.111	0.845		
Sharpe ratio	0.000	0.167	0.168	(0.1049)	

Notes: Figures are given in percent per month; entries in parentheses represent returns and Sharpe ratios net of transactions costs.

Table 5: Average returns and volatility of returns for alternative models

	SETAR switching portfolio		MA-GARCH switching portfolio		AR(p) switching portfolio	
Panel A: UK						
Average return	1.519	(2.311)	0.941	(0.448)	1.542	(1.218)
Standard deviation of returns	5.515		2.715		5.444	
Sharpe ratio	0.161	(0.160)	0.151	(0.008)	0.149	(0.145)
Number of switches	1		49		3	
Panel B: US						
Average return	0.944	(0.928)	0.293	(-0.148)	0.683	(0.372)
Standard deviation of returns	4.393		2.947		4.262	
Sharpe ratio	0.203	(0.200)	0.082	(-0.067)	0.148	(0.075)
Number of switches	2		55		32	
Panel C: Germany						
Average return	0.570	(0.554)	0.533	(0.212)	0.268	(0.139)
Standard deviation of returns	3.132		2.428		2.769	
Sharpe ratio	0.174	(0.169)	0.210	(0.077)	0.090	(0.081)
Number of switches	2		40		16	

Notes: Figures are given in percent per month; entries in parentheses represent returns and Sharpe ratios net of transactions costs.

Figure 7: Value of GEYR and probability that we are in the high GEYR regime for the UK
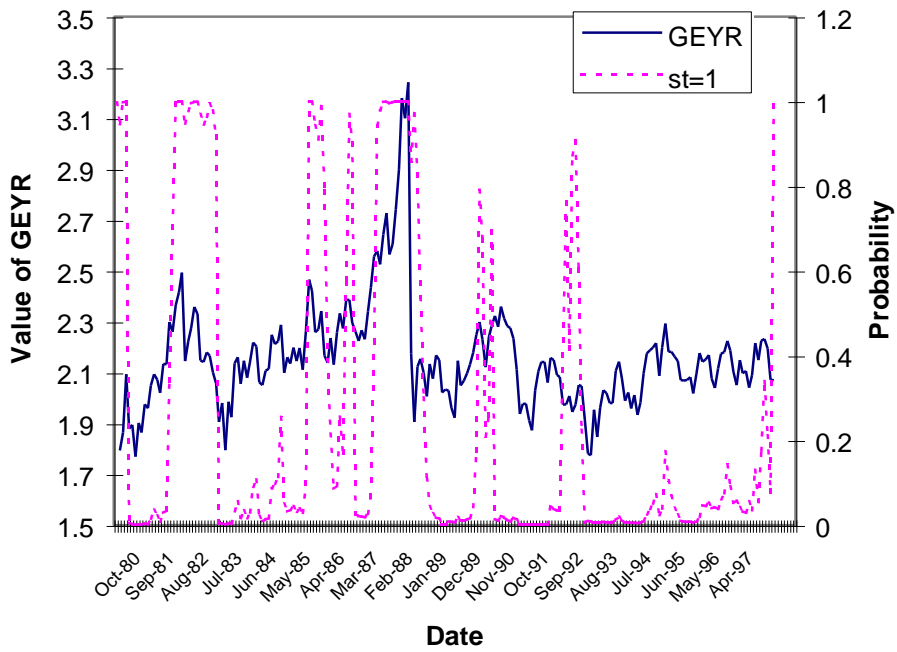
Figure 8: Mean-variance efficient frontier for the UK