Department of Finance
Working Paper Series 1998

FIN-98-074

Time-Varying Sharpe Ratios and Market Timing

Robert F. Whitelaw

November 1997

This working paper series has been generously supported by a grant from

CDC Investment Management Corporation

TIME-VARYING SHARPE RATIOS AND MARKET TIMING

Robert F. Whitelaw*

NYU, Stern School of Business

November 19, 1997

*New York University, Stern School of Business (44 West 4th Street, Suite 9-190, New York, NY 10012, (212) 998-0338, FAX: (212) 995-4233, email: rwhitela@stern.nyu.edu). Thanks to Kobi Boudoukh, Michael Brennan, Ravi Jagannathan, and Matt Richardson for helpful comments.

Time-Varying Sharpe Ratios and Market Timing

Abstract

This paper documents predictable time-variation in stock market Sharpe ratios. Predetermined financial variables are used to estimate both the conditional mean and volatility of equity returns, and these moments are combined to estimate the conditional Sharpe ratio. In sample, estimated conditional Sharpe ratios show substantial time-variation that coincides with the variation in ex post Sharpe ratios and with the phases of the business cycle. Generally, Sharpe ratios are low at the peak of the cycle and high at the trough. In out-of-sample analysis, using 10-year rolling regressions, we can identify periods in which the ex post Sharpe ratio is approximately three times larger than its full-sample value. Moreover, relatively naive market-timing strategies that exploit this predictability can generate Sharpe ratios more than 70% larger than a buy-and-hold strategy.

1 Introduction

The empirical literature contains a wealth of evidence on predictable variation in the mean and volatility of equity returns.[1] Given the joint predictability of the mean and volatility, it is perhaps somewhat surprising that the literature has been relatively silent on predictable variation in equity market Sharpe ratios. Of course, predictable variation in the individual moments does not imply predictable variation in the Sharpe ratio. The key question is whether these moments move together, leading to Sharpe ratios which are more stable and potentially less predictable than the two components individually. The empirical evidence on this issue is somewhat mixed. Earlier work (e.g., French, Schwert and Stambaugh (1987)) suggests a weak positive relation between expected returns and volatility. However, several recent studies (e.g., Glosten, Jagannathan, and Runkle (1993), Whitelaw (1994), and Boudoukh, Richardson, and Whitelaw (1997)) appear to uncover a more complex relation. Specifically, on an unconditional basis, several empirical specifications, including a modified GARCH-M model and nonparametric kernel estimation, suggest a negative relation between the conditional mean and volatility of returns. This evidence would indicate the likelihood of substantial predictable variation in market Sharpe ratios.

This paper provides an empirical investigation of time-variation in equity market Sharpe ratios. The conditional mean and volatility of equity returns are modeled as linear functions of four predetermined financial variables using the specification in Whitelaw (1994). These estimated moments are then combined to form predictions of monthly Sharpe ratios, both on an in-sample and out-of-sample basis.

In sample, estimated, conditional Sharpe ratios exhibit substantial time-variation, with monthly values ranging from less than -0.3 to more than 1.0, relative to an unconditional Sharpe ratio of 0.14 over the full sample period. This variation in estimated Sharpe ratios closely matches variation in ex post Sharpe ratios measured over short horizons. Moreover, variation in both series appears to coincide with the phases of the business cycle. Sharpe ratios at business cycle peaks are, on average, 0.4 less than the ratios at business cycle troughs. Subsamples chosen on the basis of in-sample regressions have Sharpe ratios more than three times larger than Sharpe ratios over the full sample.

On an out-of-sample basis, using 10-year rolling regressions, subsample Sharpe ratios exhibit similar magnitudes. This result is robust across subperiods, even over data subsequent to the

[1] See, for example, Breen, Glosten, and Jagannathan (1989), Fama and French (1989), Kandel and Stambaugh (1990), Keim and Stambaugh (1986), and Schwert (1989).

estimation period in Whitelaw (1994). In addition, relatively naive market-timing strategies significantly outperform a buy-and-hold strategy. These active trading strategies involve switching between the market and the risk-free asset depending on the level of the estimated Sharpe ratio relative to a specific threshold. The Sharpe ratios of the resulting time series of returns are compared to that of the buy-and-hold strategy, and they exhibit improvements of as much as 75%.

There are two possible interpretations of these results. First, they could be a product of market inefficiency. If so, this paper provides a framework for analyzing and exploiting this inefficiency in order to generate superior performance. Second, the empirical results could be due to rational time-variation in the conditional moments of returns. For example, Whitelaw (1997) develops an equilibrium model in which the mean and volatility of market returns do not move together, implying substantial, rational time-variation in stock market Sharpe ratios. The intuition behind this result is simply that volatility, the denominator of the Sharpe ratio, is not providing a good proxy for priced risk. Regardless of the cause of predictable variation, the Sharpe ratio can be critical in asset allocation decisions. For example, in a partial equilibrium setting, the Sharpe ratio determines the fraction of wealth that an agent invests in the risky market portfolio.[2] In addition, these empirical results have implications for the use of Sharpe ratios in investment performance evaluation. If the market Sharpe ratio shows substantial predictable variation, then this variation needs to be accounted for when using the market as a performance benchmark.

The remainder of the paper is organized as follows. Section 2 provides a theoretical discussion and a setting in which to interpret time-varying Sharpe ratios. Section 3 introduces the estimation methodology and documents the economic and statistical significance of time-variation in stock market Sharpe ratios. In Section 4, the out-of-sample analysis is performed, and the performance of stylized market-timing strategies is examined. Section 5 concludes.

2 Theoretical Background

Harrison and Kreps (1979) show that the absence of arbitrage implies the existence of a pricing kernel, or stochastic discount factor (M_t) that prices all assets. Specifically, the expected value of the product of the pricing kernel and the gross asset return (R_t) must equal unity, i.e.,

$$\mathrm{E}_t[M_{t+1}R_{t+1}] = 1, \tag{1}$$

[2]See, for example, Kandel and Stambaugh (1996), who focus on the predictability of expected returns and its effect on asset allocation.

where E_t is the expectation based on information available at time t. Applying equation (1), the one-period risk-free, interest rate (R_{ft}) can be written as the inverse of the expectation of the pricing kernel:

$$R_{ft} = \mathrm{E}_t[M_{t+1}]^{-1}.$$

Equation (1) also implies that the expected risk premium on any asset is proportional to the conditional covariance of its return with the pricing kernel, i.e.,

$$\mathrm{E}_t[R_{t+1} - R_{ft}] = -R_{ft}\ \mathrm{Cov}_t[M_{t+1}, R_{t+1}],$$

where Cov_t is the covariance conditional on information available at time t. Consequently, the conditional Sharpe ratio of any asset, defined as the ratio of the conditional mean excess return to the conditional standard deviation of this return, can be written in terms of the volatility of the pricing kernel and the correlation between the pricing kernel and the return:

$$\begin{aligned} \frac{\mathrm{E}_t[R_{t+1} - R_{ft}]}{\sigma_t[R_{t+1} - R_{ft}]} &= \frac{-R_{ft}\mathrm{Cov}_t[M_{t+1}, R_{t+1}]}{\sigma_t[R_{t+1}]} \\ &= -R_{ft}\sigma_t[M_{t+1}]\mathrm{Corr}_t[M_{t+1}, R_{t+1}], \end{aligned} \tag{2}$$

where σ_t and Corr_t are the standard deviation and correlation, conditional on information at time t, respectively.

Denoting the conditional Sharpe ratio of the stock market at time t by S_t, equation (2) shows that this ratio is proportional to the correlation between the pricing kernel and the return on the market (R_{mt}):

$$S_t = -R_{ft}\sigma_t[M_{t+1}]\mathrm{Corr}_t[M_{t+1}, R_{mt+1}]. \tag{3}$$

Intuitively, if the Sharpe ratio varies substantially over time, then this variation is attributable mostly to variation in the conditional correlation. Note that R_{ft} is the gross, risk-free rate, which has varied between 1.00 and 1.02 for monthly, U.S. data. The conditional volatility of the pricing kernel is more difficult to pin down without imposing further structure, but, for the specifications discussed below, conditional heteroscedasticity is likely to be too small to account for large movements in the Sharpe ratio. Consequently, we focus on the correlation in equation (3) in the discussion that follows and when interpreting the empirical results.

The implications of equation (3) for time-variation in the Sharpe ratio depend critically on the modeling of the pricing kernel. One approach is to specify M_t as a function of asset returns.

For example, modeling the pricing kernel as a linear function of the market return produces the conditional CAPM. Risk aversion implies a negative coefficient on the market return; therefore, the correlation is -1 and the market Sharpe ratio is approximately constant over time. Alternatively, modeling the discount factor as a quadratic function of the market return gives the conditional three-moment CAPM, first proposed by Kraus and Litzenberger (1976).[3] This specification allows for some time-variation in market Sharpe ratios due to the pricing of skewness risk, but the correlation will still be pushed towards -1. Bansal and Viswanathan (1993) estimate the pricing kernel as a general, nonlinear function of the market return, but again time-variation in the correlation is limited by a specification which relies on variation in market returns to proxy for variation in the discount factor. A slightly different approach to generalizing the one-factor, conditional CAPM, without abandoning a linear specification, is to model the pricing kernel as a linear function of multiple asset returns. Based on explanatory and predictive power, a number of additional factors, including small firm returns and return spreads between long-term and short-term bonds, have been proposed and tested.[4] However, as above, correlations between the discount factor and the market return tend to be relatively stable, implying stability in the stock market Sharpe ratio.

A second branch of the literature, using results from a representative-agent, exchange economy (Lucas (1978)), models the pricing kernel as the marginal rate of substitution over consumption. The resulting consumption CAPM has been analyzed and tested in numerous contexts.[5] While the consumption CAPM literature is voluminous, there has been scant attention paid to the implications of the model for the risk/return relation. Intuitively, when the marginal rate of substitution depends on consumption growth and the stock market is modeled as a claim on aggregate consumption, one might expect the correlation and the Sharpe ratio to be relatively stable. In fact, Whitelaw (1997) shows that this result holds when consumption growth follows an autoregressive process. However, in a two-regime model, where mean consumption growth differs across the regimes and regime shifts are time-varying, this intuition is overturned. In this setting, the mean and volatility of market returns can be negatively correlated. Although the magnitude and variation of the market Sharpe ratio are not investigated, it is clear that the model implies economically significant time-variation. It is important to note that the regimes correspond loosely to the expansionary and contractionary

[3]Note that these models can also be derived by imposing restrictions on a representative agent's utility function. Harvey and Siddique (1996) provide an empirical investigation of these specifications and a more detailed discussion of the underlying theory.

[4]Among the numerous papers on this topic are Campbell (1987), Chen, Roll, and Ross (1986), and Fama and French (1993).

[5]See, for example, Hansen and Singleton (1983), Breeden, Gibbons, and Litzenberger (1989), and Ferson and Harvey (1992).

phases of the business cycle. Moreover, volatility and expected returns at any point in time depend critically on the probability of a regime shift. Consequently, this model predicts business cycle related variation in Sharpe ratios and large movements around transitions between the phases of the cycle.

3 Time-Variation in Market Sharpe Ratios

3.1 The Data

For this analysis, both the mean and volatility of stock market returns are estimated as functions of predetermined financial variables. The four variables – the Baa-Aaa spread, the commercial paper–Treasury spread, the one-year Treasury yield, and the dividend yield – are chosen based on their proven predictive power in earlier studies. The data for these four explanatory variables are from the Basic Economics database. The primary sources are Moody's Investor Service, "Bond Survey" and the U.S. Department of the Treasury, "Treasury Bulletin" for corporate bond yields; the Board of Governors of the Federal Reserve, "Selected Interest Rates and Bond Prices" for Treasury yields and commercial paper yields; and Standard and Poor's Corporation, "The Outlook" for dividend yield data. All data are monthly and cover the period April 1953 to November 1995. In addition to the four explanatory variables, the analysis uses monthly and daily returns on the value-weighted market portfolio from the CRSP data files. The monthly returns cover the period May 1953 to December 1995, and the daily returns start later in July 1962. Excess returns are calculated by subtracting the monthly yield on a 3-month T-bill from the corresponding stock return. The 3-month yield is used instead of the 1-month yield because of the well-documented idiosyncrasies in this latter time series (see Duffee (1996)).

3.2 Estimating Sharpe Ratios

The first step in the analysis is to determine if there is significant time-variation in estimated and ex post Sharpe ratios, and, further, to see if the variation in these two series coincides. Expected returns are estimated by regressing excess returns on a vector of predetermined variables. The absolute value of the residuals from this regression are then projected on predetermined variables to estimate the conditional volatility. Specifically, the conditional moments are modeled as follows:

$$\mathrm{E}_t[R_{t+1} - R_{ft}] = X_{1t}\beta_1$$

$$\sigma_t[R_{t+1}] = X_{2t}\beta_2,$$

and the corresponding regressions are specified as

$$R_{t+1} - R_{ft} = X_{1t}\beta_1 + \epsilon_{1t+1} \quad (4)$$

$$\sqrt{\frac{\pi}{2}}|\epsilon_{1t+1}| = X_{2t}\beta_2 + \epsilon_{2t+1}. \quad (5)$$

The specification and conditioning variables are chosen based on the results in Whitelaw (1994), and this system is estimated via GMM (Hansen (1982)). The Baa-Aaa spread, the dividend yield, and the one-year yield are used in the mean equation, and the one-year yield and the commercial paper–Treasury spread are used in the volatility equation. Whitelaw (1994) uses all four explanatory variables to estimate both conditional moments, but we drop the variables which are not significant in that study, both for ease of exposition and to reduce the dimension of the specification in the nonlinear analysis in Section 3.5.

Monthly regression results for the period 5/53-12/95 are presented in Table 1. Note that all the explanatory variables are significant at the 1% level and that the signs of the coefficients are consistent with those previously reported in the literature. The R^2 values of 6.6% for the mean equation and 11.9% for the volatility equation are also comparable to those obtained in the literature for similar specifications. The χ^2 statistics indicate that a null hypothesis of no predictable variation can be rejected for both the mean and volatility at all standard significance levels. Of course, these results must be considered in the context of the well known problem with data snooping (see, for example, Foster and Smith (1992)), although these concerns are mitigated somewhat by the out-of-sample exercise conducted in the Section 4.

Fitted values from equations (4)-(5), using the parameter estimates from Table 1, can be used to compute conditional Sharpe ratios for each month. Specifically, based on information available at time t and parameter estimates from the estimation over the full sample, the estimated conditional Sharpe ratio is

$$\hat{S}_t = \frac{X_{1t}\widehat{\beta_1}}{X_{2t}\widehat{\beta_2}}. \quad (6)$$

The ex post, sample Sharpe ratio of the monthly returns can also be calculated as

$$S^* = \frac{\mu}{\sqrt{(1/(T-1))\sum_{t=1}^{T}[(R_{t+1} - R_{ft}) - \mu]^2}},$$

where

$$\mu = (1/T)\sum_{t=1}^{T}[R_{t+1} - R_{ft}].$$

The average conditional Sharpe ratio over the full sample is 0.175, compared to an ex post sample Sharpe ratio of 0.139. Note that these are the Sharpe ratios of monthly returns. The corresponding annualized ratio can be obtained by multiplying by $\sqrt{12}$. Also note that, in general, the average conditional Sharpe ratio will not equal the unconditional Sharpe ratio due to Jensen's inequality. Estimation error also accounts for some of the difference between the average conditional Sharpe ratio and the sample (unconditional) ratio.

3.3 In-Sample Forecasting

The estimated conditional Sharpe ratios can be compared to rolling Sharpe ratios computed over short horizons. The choice of horizon generates a tradeoff between two concerns: the shorter the horizon, the more likely it is that the rolling estimates will pick up time-variation; the longer the horizon, the lower the standard error of the estimate. As a compromise, we use a 19-month horizon. At this horizon, the volatility of the rolling estimates matches that of the estimated conditional Sharpe ratios, permitting easier comparison. The rolling estimate is compared to the conditional Sharpe ratio for the middle month of the 19-month sample. For example, the rolling estimate using data from May 1953 to November 1954 is compared to the conditional Sharpe ratio for February 1954.

Figure 1 graphs both the rolling estimates and the conditional Sharpe ratios over time. It is apparent that these measures tend to move together. The correlation between the two series is 0.456. To the extent that there is predictable variation in Sharpe ratios, both series appear to be detecting a similar pattern.

Of greater interest, perhaps, is the ability of the conditional Sharpe ratios to predict variations in the data. To address this issue, the data is divided into four equal subsamples using the estimated conditional Sharpe ratio (CSR). In other words, the CSR calculated at the beginning of the month is used to classify that month. The 128 months with the lowest estimated Sharpe ratios are put in the first group, the 128 months with the next lowest estimates are put in the second group, etc. Table 2 reports the minimum, maximum, and average CSRs, and the ex post Sharpe ratios for each subsample of returns. Note that the ordering of the ex post Sharpe ratios matches that of the estimated Sharpe ratios. The average CSRs and the ex post Sharpe ratios are not exactly equal, but, in each case, the ex post ratio lies between the minimum and maximum estimated CSR. Moreover, there is a dramatic difference in Sharpe ratios between group 1 (-0.181) and group 4 (0.481). For subsample 1, both the conditional and ex post Sharpe ratios are negative, i.e., the expected return on the market is less than the risk-free rate and so is the realized return. This result is somewhat puzzling, although negative risk premiums are not theoretically precluded in the framework of Section 2 (see Boudoukh, Richardson, and Whitelaw (1997)). Moreover, these results are consistent with the results in Kairys (1993), who uses commercial paper rates to predict negative risk premiums. For subsample 4, both the conditional and ex post Sharpe ratios are more than three times greater than the values over the full sample, indicating the economic significance of the results. Finally, a formal test for equality across subsamples ($\chi^2(3)$) yields a test statistic of 36.903, with a corresponding p-value of 0.000. Equality of Sharpe ratios across the subsamples can be rejected at all standard significance levels.

One potential issue with respect to the statistical test above is the small sample distribution of the ex post Sharpe ratios. The test is valid asymptotically, but it relies on the normality of the underlying estimators. To resolve this question, a bootstrapping exercise is performed. Specifically, 128 monthly returns are drawn at random from the full return series, and a Sharpe ratio is computed. This procedure is repeated 100,000 times and the distribution of the resulting Sharpe ratios is graphed in Figure 2. Each bar shows the probability of observing a Sharpe ratio in the specified range. The solid line gives the expected height of each bar under normality. Clearly, deviations from normality in the distribution of Sharpe ratios are not an issue. The mean Sharpe ratio is 0.143 and the standard deviation is 0.093. The former is close to the ex post ratio over the full sample, and the latter is marginally larger than the average standard error reported in Table 2.

3.4 Time-Varying Sharpe Ratios and Economic Activity

Given the significant time-variation in stock market Sharpe ratios documented above, a natural question is how movements in estimated Sharpe ratios correspond to fluctuations in economic activity.[6] Figure 1 shows not only the CSRs and the rolling Sharpe ratios, but also the NBER business cycle peaks and troughs (marked by vertical lines). There are only seven business cycles within the sample period; therefore, conclusions should be drawn with caution. Nevertheless, there appears to be striking cyclical variation in Sharpe ratios. Almost without exception, business cycle peaks correspond to low Sharpe ratios and business cycle troughs to high Sharpe ratios. The average rolling Sharpe ratio at the eight peaks is -0.106, while at the eight troughs it is 0.291. The

[6]Time-variations in both expected returns and volatility have been previously linked to the business cycle. See, for example, Fama and French (1989) and Schwert (1989).

time from peak to trough (i.e., the contraction) is short, averaging less than 12 months, but the average increase in rolling Sharpe ratio over the same period is 0.332.

The data indicate that the return/volatility tradeoff is significantly better entering expansions than it is leaving expansions. To some extent, this result is consistent with the theoretical results in Whitelaw (1997). At the end of expansions, when the probability of shifting to a contraction is high, the conditional volatility is also high. However, equity returns depend on whether a regime switch occurs, an event that is independent of the marginal rate of substitution. Consequently, the correlation in equation (3) is low and so is the expected return. Clearly, this combination will yield the low Sharpe ratios shown in Figure 1. The regime-shift model also generates a similar prediction for transitions from contractions to expansions, but a decrease in Sharpe ratios at this point in the cycle is not evident in the data. The one mitigating factor in the model in Whitelaw (1997) is that regime switch probabilities are much more stable within contractions. Consequently, the volatility and expected return effects are smaller.

The weakest relation between the business cycle and the rolling Sharpe ratios occurs around the contraction that ran from January 1980 to July 1980. Note, however, that this contraction preceded the shortest expansion in the sample period, lasting from July 1980 to July 1981. Given the 19-month horizon used in computing the Sharpe ratios, it is not surprising that this measure cannot pick up such quick cyclical changes. Interestingly, the conditional Sharpe ratio does seem to "identify" this rapid turnaround. There is a dramatic increase in this measure during the 1980 contraction.

The other apparent inconsistency between the rolling Sharpe ratios and the NBER business cycle turning points is the existence of Sharpe ratio cycles within the expansions of February 1961 to December 1969 and November 1982 to July 1990. Again, however, these extra cycles are less prominent in the conditional Sharpe ratio series, suggesting that they may result from estimation error in the rolling Sharpe ratios. An alternative explanation is that these cycles correspond to increased expectations of a cycle turning point but that no turning point actually occurred. Within the context of the model in Whitelaw (1997), an increase in the probability of a regime switch will increase ex ante volatility even if the switch does not occur ex post.

3.5 Nonlinearities

Although the linear model used in the analysis above yields good results, there is no *a priori* reason to believe that the relation between the conditional moments of market excess returns and predetermined financial variables should be linear. Boudoukh, Richardson, and Whitelaw (1997) document the existence of a nonlinear relation between the slope of the term structure of interest rates and both the mean and volatility of stock returns, although their analysis uses both a different explanatory variable and a different sample period. Brandt (1997) also detects significant nonlinearities within the context of a portfolio choice problem. Given the lack of strong priors about the functional form, one would like to use a nonparametric estimation technique such as kernel estimation. Unfortunately, the dimensionality of the dataset makes this infeasible. Instead, we employ a simple Taylor series expansion. We expand the regression to include higher order terms as follows:

$$\begin{aligned} R_{t+1} - R_{ft} &= X_{1t}\beta_1 + X_{1t}^2\gamma_1 + \epsilon_{1t+1} \\ \sqrt{\frac{\pi}{2}}|\epsilon_{1t+1}| &= X_{2t}\beta_2 + X_{2t}^2\gamma_2 + \epsilon_{2t+1} \end{aligned}$$

where X^2 denotes the vector of second order terms.

Rather than report the coefficients, which have little economic meaning, we report the test statistic for the significance of the higher order terms. There are six such terms in the mean equation (three squared terms and three cross-product terms) and three in the volatility equation (two squared terms and one cross-product term). For these two equations, the relevant χ^2 statistics are 4.274 and 2.594, with corresponding p-values of 0.640 and 0.459. From a slightly different perspective, adding the extra terms increases the R^2s by only 0.9% and 0.3% in the mean and volatility regressions, respectively. In neither case can we reject the linear specification, and the additional explanatory power is minimal. Consequently, we consider only the linear model in the analysis in Section 4.

3.6 Estimating Sharpe Ratios Directly

One disadvantage of the empirical methodology used above is that Sharpe ratios are not estimated directly. Instead, they are calculated using estimates of the conditional first and second moments. This two-stage procedure may not use the data efficiently. Unfortunately, it is not possible to estimate the monthly conditional Sharpe ratio directly using monthly returns. One feasible alternative is to employ more frequently sampled data (e.g., daily data) to estimate monthly volatility. Specifically, monthly ratios are computed as

$$S_{t+1} = \frac{R_{t+1} - R_{ft}}{\sigma_{t+1}},$$

$$\sigma_{t+1} = \sqrt{\sum_{n=1}^{N} R_{t+n}^2},$$

where n indexes the daily returns within month $t+1$. These ratios are then regressed on the conditioning variables as before:

$$S_{t+1}^* = X_t\beta + \epsilon_{t+1}. \tag{7}$$

It is an empirical question as to whether this specification produces better forecasting power than the specification in equations (4)-(6).

Table 3 presents the results from the estimation of equation (7). Note that the sample period is different from the one used in Table 1 because of the limited availability of daily data; therefore, the results are not directly comparable. Nevertheless, the signs of the coefficients are consistent with the earlier estimation. The default spread and the dividend yield, which are positively related to expected returns, again show up with positive coefficients. However, in this case, the default spread coefficient is less significant. In Table 1, expected returns are negatively related to the level of interest rates, and volatility is positively related to this variable; therefore, it is not surprising that the one-year rate shows up with a negative and strongly significant coefficient. Finally, the commercial paper–Treasury spread, which is positively related to volatility, is negatively but insignificantly related to the Sharpe ratio. A test for nonlinearity, analogous to the one used in Section 3.5, does not reject the linear specification.

The fitted values from the in-sample estimation of equation (7) look similar to those estimated previously. The two conditional Sharpe ratio series have a correlation of 0.759. As expected, subsamples sorted on this new estimate also exhibit significant variation in ex post Sharpe ratios. The out-of-sample performance of this estimate is considered below.

4 Exploiting Predictable Variation

4.1 Out-of-Sample Forecasting

The previous section documents statistically significant time-variation in conditional Sharpe ratios, and statistically and economically significant predictive power for estimates based on a simple linear model. From a practical perspective, however, the key issue is whether the empirical model has economically significant out-of-sample predictive power. Unfortunately, it is difficult to conduct a true out-of-sample test. The conditioning variables are chosen based on their correlation with returns and volatility in a sample that runs through April 1989. Consequently, only the most recent 6.5 years of data are truly out-of-sample. Nevertheless, it is worthwhile to consider the predictive power of out-of-sample regressions over the full sample.

There are two possible ways to specify the out-of-sample regressions for estimating the conditional Sharpe ratios. The first method is to choose a fixed sample size and to run rolling regressions. That is, a fixed number of observations are used to estimate each set of coefficients, and the estimation window is moved forward by one month at a time. The advantage of this approach is that if the coefficients vary over time, either because of misspecification of the empirical model or structural shifts, then the coefficients from the rolling regressions will "adapt" to these changes. The second method is to add the new monthly observation to the estimation dataset as we move through time. As a result, the number of observations increases through time, and the later coefficients of these cumulative regressions will be less subject to estimation error if the empirical specification is correct.

The most natural way to evaluate these alternatives is to examine their out-of-sample performance. For both the rolling and cumulative regressions, the initial estimation period is chosen to be 10 years. 120 monthly observations from May 1953 to April 1963 are used to estimate the first set of coefficients. These coefficients and the data on the explanatory variables from April 1963 are then used to estimate the conditional moments of stock market returns for May 1963. The estimation is then rolled forward one month, adding the most recent observation for both the rolling and cumulative regressions, but also dropping the oldest observation from the rolling regression. Both techniques generate a series of 392 out-of-sample conditional Sharpe ratios.

The first interesting question is how these out-of-sample estimates compare to the in-sample estimates from the previous section. The correlations between the out-of-sample and in-sample conditional Sharpe ratios are 0.812 and 0.915 for the rolling regressions and cumulative regressions, respectively. It is not surprising that the latter correlation is slightly higher, since the cumulative regression estimates towards the end of the sample are essentially identical to those from the in-sample regression because of the number of observations they have in common. Of greater interest is the fact that both correlations are high, suggesting that the out-of-sample regressions will also demonstrate significant predictive power.

To assess the predictive power of the two CSR series we conduct the same exercise as in Section 3.3. That is, the months are divided into four subsamples based on the magnitudes of the estimated Sharpe ratios. In this case, however, each subsample consists of 98 months due to the use of 120 observations in the initial regressions. Table 4 presents the results of this analysis. For each methodology and each subsample, the table gives the minimum, maximum, and average CSRs and the ex post Sharpe ratios. The standard errors for the latter ratios are in parentheses. Similar to the in-sample results in Table 2, the ex post Sharpe ratios exhibit substantial variation that coincides, for the most part, with the estimated ratios. The one minor exception is that the ratios for subsamples 1 and 2 from the cumulative regression have the wrong ordering, although the values themselves are not statistically distinguishable. In addition, the ex post Sharpe ratios do not always fall within the range of the minimum and maximum CSRs. As in Table 2, negative Sharpe ratios are both prevalent and predictable. The $\chi^2(3)$ test statistics for equality across the four subsamples are 10.403 and 20.170, with corresponding p-values of 0.015 and 0.000, respectively. Again, equality can be rejected for both out-of-sample regression techniques. There is little to choose between the rolling and cumulative regressions, although the estimated CSRs from the rolling regressions are more volatile due to the shorter estimation period. For ease of exposition, we focus on the rolling regressions in the remainder of this paper.

One potentially important issue that can be addressed using the rolling regressions is the instability of the coefficient estimates. Unstable coefficients indicate either structural shifts in the data or a misspecified model. In the former case, predictive power might be gained from shortening the estimation period further, although there is a clear tradeoff with estimation error as the number of observations decreases. In the latter case, alternative specifications, specifically more flexible functional forms, might prove useful. Figure 3 shows the rolling coefficient estimates for both the mean equation (top graph) and the volatility equation (bottom graph). Note that the date on the x-axis refers to the last date in the estimation period, e.g., the coefficient for December 1978 is actually based on the ten years of data from January 1969 to December 1978. The graphs show a good deal of instability. In the mean equation, the coefficient on the dividend yield ranges from approximately 0 to 6. The increase in the latter part of the sample is an adjustment to the low yields and high returns experienced in the 1990s. The coefficient on the default spread is as low as -3 and as high as 3. From an economic standpoint, the switching of the sign of the coefficient is especially worrisome, although the standard errors are high in both the early and late periods when the coefficients are negative. In the volatility equation, the coefficients on both the one-year rate and the commercial paper–Treasury spread switch signs, but there is significantly more stability in the latter half of the sample. Given these variations, the out-of-sample forecasting power of the rolling regression is perhaps even more surprising.

A final piece of information that the rolling regressions offer is the time series behavior of the estimated conditional Sharpe ratios. Given that all the explanatory variables are highly persistent, the estimated moments of excess returns are also persistent. The estimated coefficients change over time, but these estimates are also persistent because of the significant overlap in the estimation periods for nearby regressions. Rather than graph the conditional Sharpe ratios, which exhibit a pattern similar to that in Figure 1, Figure 4 shows the time series of group designations. That is, each month is designated by the subsample into which is falls, from lowest (subsample 1) to highest (subsample 4) CSR. These group designations exhibit the predicted persistence, but it is not the case that all the low or high estimates fall in a specific subperiod. Instead, the level of conditional Sharpe ratios tends to persist for 1 to 3 years. This periodicity is consistent with the business cycle interpretation of the results given in Section 3.4.

Although the group designations are dispersed throughout the sample period, it is not clear whether the predictive power is consistent across the dataset. To evaluate the subperiod performance of the rolling regressions, the sample is broken up into four equal subperiods, each of length 98 months. Due to the limited number of observations in each subperiod, the monthly returns are separated into two groups in each subperiod using the forecasted Sharpe ratios. Table 5 presents the subperiod results in the same format as used previously. Note first that the ordering of the ex post Sharpe ratios is consistent with the subsample designation for all four subperiods. However, the standard errors are relatively large (on the order of 0.14), therefore the differences are statistically significant at the 10% level for only two of the subperiods, 7/71-8/79 and 9/79-10/87.[7] The significant subperiods also exhibit negative conditional and ex post Sharpe ratios for subsample 1. To a large extent, the significant statistics coincide with periods of relatively stable coefficient estimates in both the mean and volatility equations (see Figure 3).

Of some interest, the fourth subperiod corresponds approximately to a true out-of-sample test. The choice of explanatory variables is based on a study that ends close to the beginning of this period. The results show that the subsample 2 Sharpe ratio is almost double the subsample 1 Sharpe ratio. While the test statistic is not large enough to reject equality due to the small sample size, the direction of the results is consistent with the posited effect.

The subperiod analysis is also an ideal setting in which to compare the methodology used above to the direct estimation of the Sharpe ratios described in Section 3.6. The available daily data are almost sufficient to duplicate the analysis for the last three subperiods using 10-year rolling regres-

[7]The p-values are conservative in the sense that we report the results for a 2-sided test. Using the alternative that the ex post Sharpe ratio for subsample 2 is larger than for subsample 1 yields a 1-sided test with lower p-values.

sions. The results from this exercise are mixed. The CSRs estimated directly perform marginally better in the initial subperiod, but perform significantly worse in the last two subperiods. The ex post Sharpe ratios exhibit the hypothesized relation, but the differences between the subsamples are not sufficiently large to generate significant test statistics. Consequently, the 2-step procedure is used in the market-timing strategies discussed below.

4.2 Trading Strategies

While the above results clearly demonstrate economically and statistically significant out-of-sample predictive power, they are not based on feasible trading strategies. At any point in time, it is not possible to know how the current month's conditional Sharpe ratio will rank relative to the estimates to come in the future. The possible set of trading strategies is huge; nevertheless, it is worthwhile to look at the performance of a few stylized strategies. Consider the strategy of estimating the conditional Sharpe ratio using the 10-year rolling regression and comparing this number to the ex post Sharpe ratio over the prior 10 years, or to a fixed threshold Sharpe ratio. If the estimated CSR is larger, then invest in the stock market; if it is smaller, then invest in the risk-free asset. The excess returns to any of these strategies will be a series of market excess returns interspersed with zero excess returns. This series can easily be compared to a buy-and-hold strategy that always holds the market.[8]

Table 6 reports the results from executing five strategies: a buy-and-hold strategy and four market-timing strategies where the threshold for investing in the market is the Sharpe ratio over the past 10 years, or three different pre-specified CSR levels (0.0, 0.1, and 0.2). The second column gives the number of months, out of a possible 392, in which the strategy is invested in the market. In the other months, the risk-free asset is held, yielding an excess return of zero. The table also shows the mean and volatility of monthly stock market returns for the months in which the market is held. For the buy-and-hold strategy, these are the sample moments for the full time period. All four market-timing strategies can identify months with mean returns more than twice the sample average. Of equal importance, these months also have lower volatilities. The reported Sharpe ratios include not only these returns, but also the returns from holding the risk-free asset. Consequently, both the mean and volatility are reduced relative to the numbers in columns three and four. The

[8]These market-timing strategies ignore both transaction costs and information in the magnitude of the CSRs relative to the threshold. A more sophisticated, and potentially better performing, strategy might involve time-varying market weights that depend on the prior position in the market and the relative magnitude of the CSR. Nevertheless, the stylized strategies are sufficient to illustrate the extent of predictable variation.

resulting Sharpe ratios exceed the Sharpe ratio of the buy-and-hold strategy by between 50% and 75%. The χ^2 test statistic for equality between the market-timing strategies and the buy-and-hold strategy is significant at the 10% level for only one strategy, but note that the p-values are for a 2-sided test (see footnote 7). Overall, the results confirm the earlier conclusion that there is economically significant, predictable variation in stock market Sharpe ratios.

5 Conclusion

This paper demonstrates the ability of a relatively straightforward linear specification of the conditional mean and volatility of equity returns to predict dramatic time-variation in monthly, stock market Sharpe ratios. This predictability is evident both in-sample and market-timing strategies outperform a buy-and-hold strategy in terms of ex post Sharpe ratios. This evidence provides further support for the contention that the mean and volatility of stock market returns do not move together.

Variations of the magnitude documented are inconsistent with the conditional CAPM and related models that imply a close to constant market Sharpe ratio. One possible explanation is that the results are due to market irrationality or inefficiency. However, the apparent relation between variation in Sharpe ratios and the business cycle suggests the possibility of an economic interpretation. Whitelaw (1997) provides a rational expectations, general equilibrium model that is broadly consistent with the empirical evidence. In this model, discrete shifts between expansions and contractions overturn the standard return/volatility relation. Further research in this area is warranted.

While the empirical evidence provides insights into the time series properties of equity returns and their underlying economics, it also has implications in other areas. In particular, substantial, predictable time-variation in market Sharpe ratios casts doubt on the ability of the volatility of even broadly diversified portfolios to proxy for priced risk. Consequently, standard measures of investment performance and traditional portfolio asset allocation rules may have to be re-thought.

Constant	DEF	DIV	1YR	CP	R^2	χ^2
	Explanatory Variables					
			Conditional Mean			
-1.292*	2.205**	0.634**	-0.439**		6.6%	35.573
(0.810)	(0.623)	(0.243)	(0.086)			[0.000]
			Conditional Volatility			
1.590**			0.127**	2.146**	11.9%	48.683
(0.327)			(0.044)	(0.413)		[0.000]

* Significant at the 10% level.
** Significant at the 1% level.

Table 1: Estimation of the Conditional First and Second Moments of Returns

Regressions of monthly, excess stock returns and volatilities for the CRSP value-weighted index on lagged explanatory variables for the period 5/53-12/95. The conditioning variables are the Baa-Aaa spread (DEF), the commercial paper–Treasury spread (CP), the one-year Treasury yield (1YR), and the dividend yield (DIV). The model is given in equations (4)-(5). Heteroscedasticity-consistent standard errors are in parentheses. The χ^2 statistic tests the hypothesis that all the coefficients except the constant are zero (p-value in brackets).

Subsample	Min. CSR	Max. CSR	Avg. CSR	Ex post SR
1	-0.372	-0.065	-0.160	-0.181 (0.080)
2	-0.062	0.152	0.034	0.140 (0.088)
3	0.153	0.339	0.244	0.190 (0.088)
4	0.347	1.049	0.583	0.481 (0.094)
$\chi^2(3)$				36.903 [0.000]

Table 2: In-Sample Conditional and Ex Post Sharpe Ratios

Minimum, maximum, and average monthly conditional Sharpe ratios and the ex post Sharpe ratios for four equal subsamples chosen by ranking based on conditional Sharpe ratios. The CSRs are calculated using equation (6) and the estimated coefficients from Table 1. Heteroscedasticity consistent standard errors for the ex post Sharpe ratios are in parentheses. The $\chi^2(3)$ statistic (p-value in brackets) tests for equality across the subsamples.

Explanatory Variables						
Constant	DEF	DIV	1YR	CP	R^2	$\chi^2(4)$
0.116	0.237*	0.258**	-0.148**	-0.032	4.452%	19.905
(0.264)	(0.168)	(0.113)	(0.034)	(0.132)		[0.001]

* Significant at the 10% level.

** Significant at the 1% level.

Table 3: Direct Estimation of Conditional Sharpe Ratios

Regressions of monthly, ex post Sharpe ratios on lagged explanatory variables for the period 8/62-12/95. The conditioning variables are the Baa-Aaa spread (DEF), the commercial paper–Treasury spread (CP), the one-year Treasury yield (1YR), and the dividend yield (DIV). The model is given in equation (7). Heteroscedasticity-consistent standard errors are in parentheses. The $\chi^2(4)$ statistic tests the hypothesis that all the coefficients except the constant are zero (p-value in brackets).

Subsample	Min. CSR	Max. CSR	Avg. CSR	Ex post SR
	Rolling Regression			
1	-0.868	-0.188	-0.378	-0.048 (0.101)
2	-0.188	-0.024	-0.087	-0.026 (0.099)
3	-0.020	0.239	0.086	0.207 (0.103)
4	0.239	1.658	0.612	0.318 (0.096)
$\chi^2(3)$				10.403 [0.015]
	Cumulative Regression			
1	-0.754	-0.202	-0.307	-0.035 (0.101)
2	-0.200	-0.027	-0.114	-0.135 (0.089)
3	-0.027	0.170	0.064	0.232 (0.095)
4	0.171	0.986	0.364	0.423 (0.100)
$\chi^2(3)$				20.170 [0.000]

Table 4: Out-of-Sample Conditional and Ex Post Sharpe Ratios

Minimum, maximum, and average monthly conditional Sharpe ratios and ex post Sharpe ratios for four equal subsamples chosen by ranking based on conditional Sharpe ratios from out-of-sample regressions. Rolling regressions use 10-year estimation periods, and cumulative regressions use the full sample prior to the prediction month. Heteroscedasticity consistent standard errors for the ex post Sharpe ratios are in parentheses. The χ^2 statistic (p-value in brackets) tests for equality across the subsamples.

Subsample	Min. CSR	Max. CSR	Avg. CSR	Ex post SR
		5/63-6/71		
1	-0.868	-0.062	-0.214	0.090 (0.144)
2	-0.055	0.489	0.067	0.099 (0.149)
$\chi^2(1)$				0.002 [0.962]
		7/71-8/79		
1	-0.670	0.059	-0.223	-0.133 (0.145)
2	0.060	1.658	0.582	0.211 (0.132)
$\chi^2(1)$				3.271 [0.071]
		9/79-10/87		
1	-0.739	0.012	-0.234	-0.098 (0.134)
2	0.016	1.090	0.292	0.372 (0.140)
$\chi^2(1)$				6.203 [0.013]
		11/87-12/95		
1	-0.615	0.043	-0.244	0.155 (0.151)
2	0.049	1.065	0.439	0.303 (0.141)
$\chi^2(1)$				0.534 [0.465]

Table 5: Subperiod Conditional and Ex Post Sharpe Ratios

Minimum, maximum, and average monthly conditional Sharpe ratios and ex post Sharpe ratios for two equal subsamples within each subperiod chosen by ranking based on conditional Sharpe ratios from out-of-sample rolling 10-year regressions. Heteroscedasticity consistent standard errors for the ex post Sharpe ratios are in parentheses. The $\chi^2(1)$ statistic (p-value in brackets) tests for equality across the subsamples.

Strategy	No. of Months	Mean	Vol.	Ex post SR	$\chi^2(1)$
Buy-and-Hold	392	0.441	4.332	0.102 (0.052)	
CSR > avg.	140	1.157	4.002	0.169 (0.045)	1.959 [0.162]
CSR > 0.0	180	1.070	4.043	0.177 (0.047)	2.833 [0.092]
CSR > 0.1	140	1.088	4.185	0.153 (0.047)	1.157 [0.282]
CSR > 0.2	106	1.252	4.072	0.155 (0.044)	1.127 [0.288]

Table 6: Ex Post Sharpe Ratios of Trading Strategies

Ex post monthly Sharpe ratios for five strategies based on 10-year rolling regressions. Strategy 1 is a buy-and-hold strategy. Strategy 2 holds the stock market when the estimated conditional Sharpe ratio exceeds the ratio from the past 10 years. Strategies 3, 4, and 5 hold the stock market when the conditional Sharpe ratio exceeds the given threshold. The column "No. of Months" indicates the number of months out of a possible 392 during which the strategy holds stocks. The mean and volatility are only for the months in which stocks are held. Heteroscedasticity consistent standard errors for the ex post Sharpe ratios are in parentheses. The χ^2 statistic (p-value in brackets) tests for the equality of the strategy Sharpe ratio versus the buy-and-hold strategy.

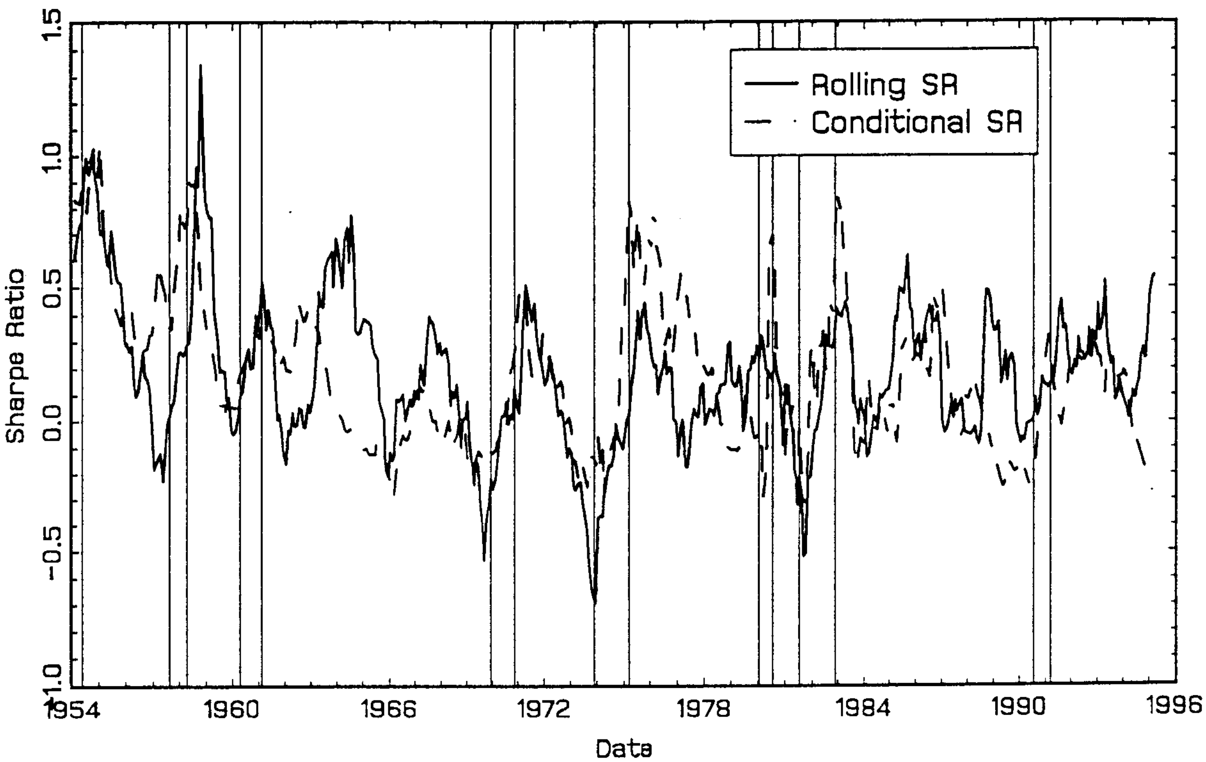


Figure 1: Expected Conditional Sharpe Ratios and Rolling Estimates

Monthly expected conditional Sharpe ratios computed using equation (6) and the estimated coefficients from Table 1 (dashed line) and rolling sample estimates based on 19 monthly observations (solid line) for the period February 1954 to March 1995. NBER business cycle peaks and troughs are marked by solid vertical lines.

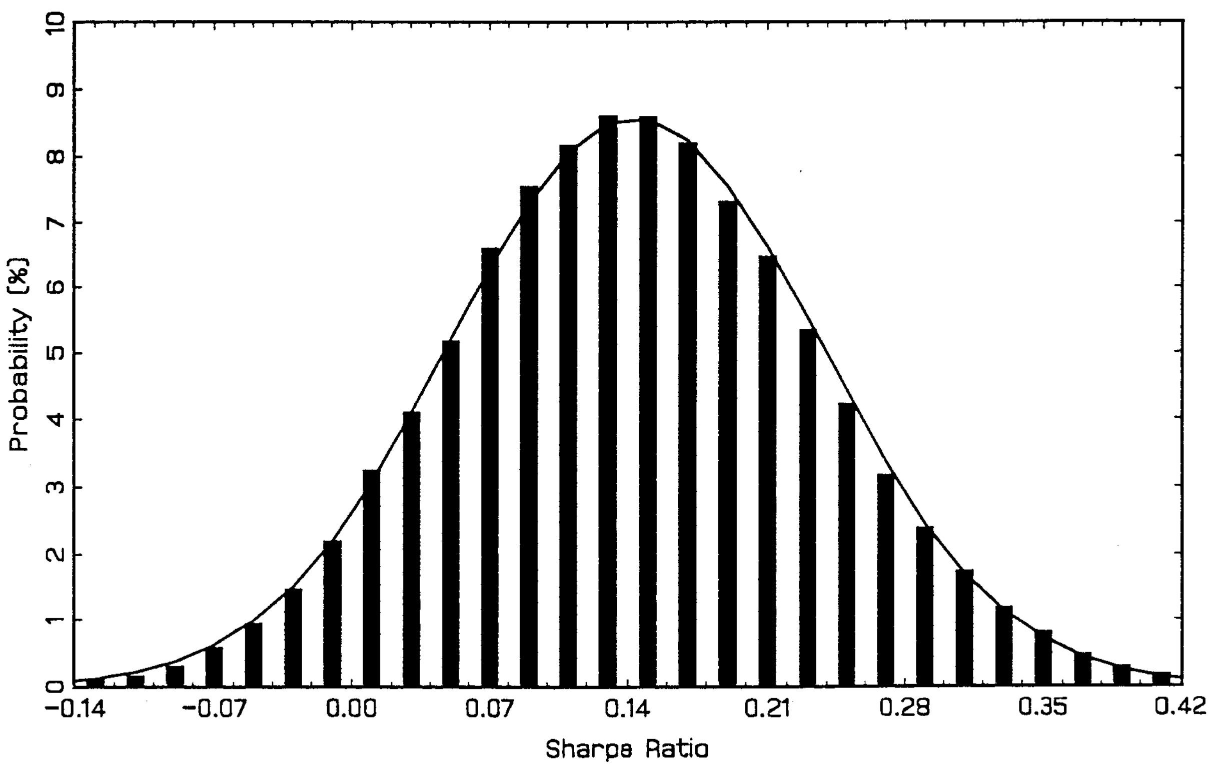
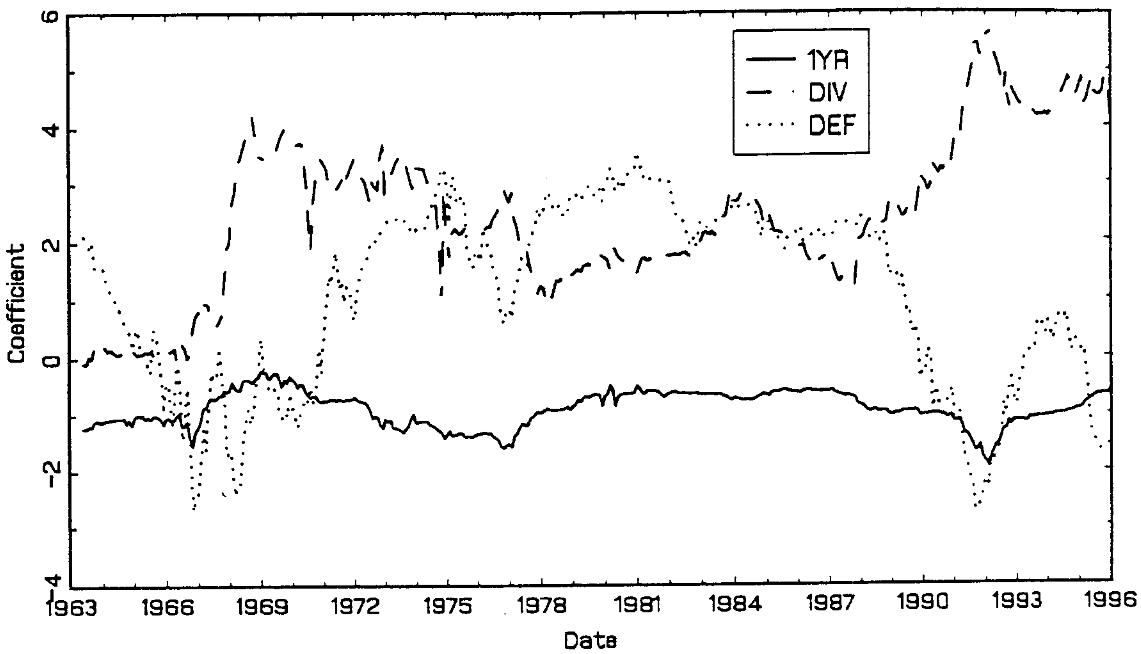


Figure 2: Bootstrapped Distribution of Subsample Sharpe Ratios

The distribution of ex post Sharpe ratios based on a bootstrapping experiment with 100,000 replications in which 128 monthly excess returns are picked at random from the data. The solid line shows the normal distribution with the same mean and variance.



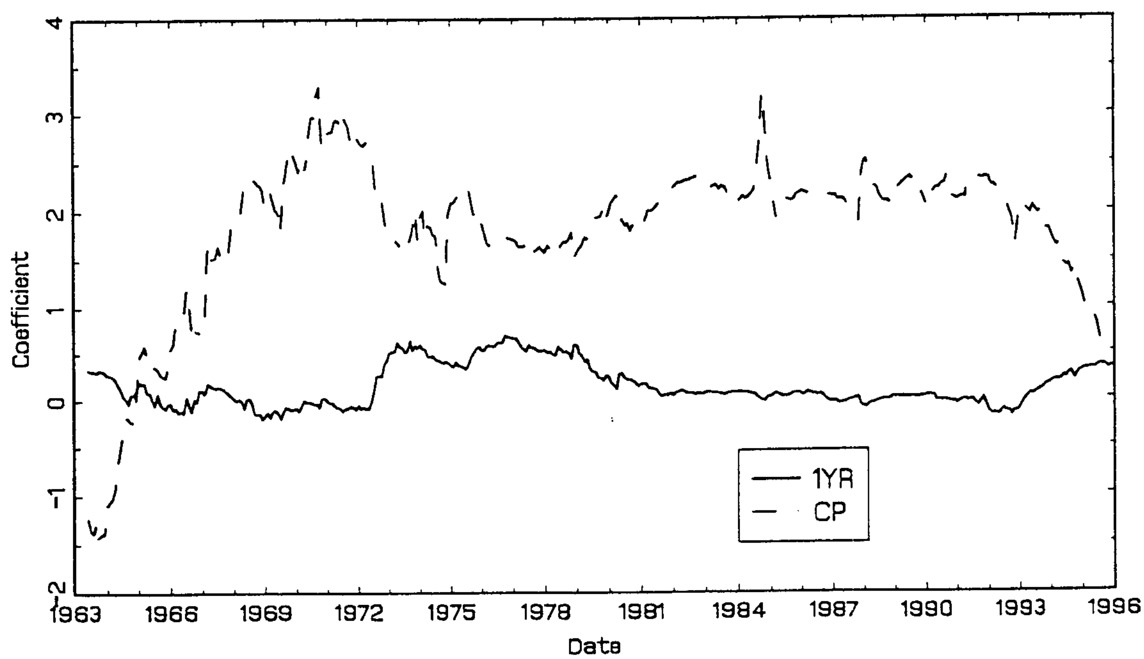


Figure 3: Estimated Rolling Coefficients

Estimated coefficients from 10-year rolling regressions for the mean equation (top) and the volatility equation (bottom). The model is given in equations (4)-(5).

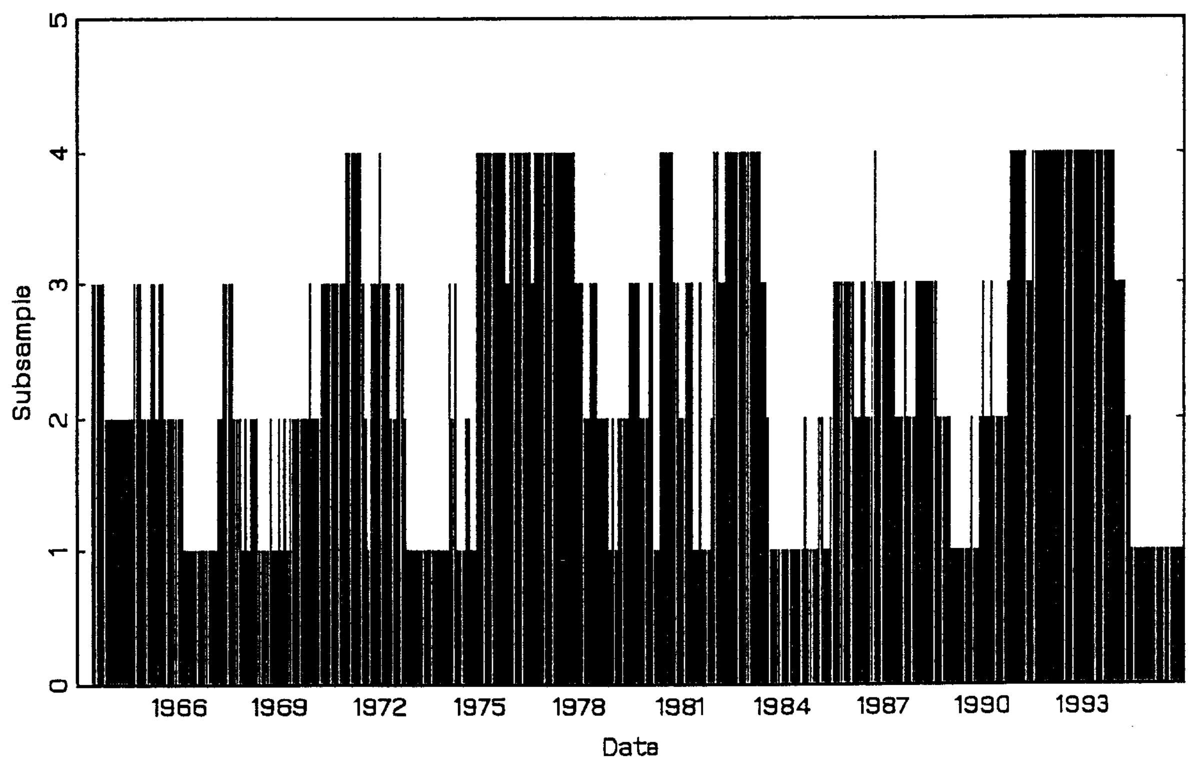


Figure 4: Predicted CSRs by Group

Month-by-month group designations of excess returns sorted by their conditional Sharpe ratios estimated using 10-year rolling regressions. Months are sorted from lowest (subsample 1) to highest (subsample 4) CSR.