Predictability of Equity Returns over Different Time Horizons: A Nonparametric Approach
(Preliminary Version)

Qingqing Chen
Department of Economics
Cornell University
Ithaca, NY14850

Yongmiao Hong
Department of Economics
Cornell University
Ithaca, NY14850

March 2009

Abstract

This paper aims to test an important hypothesis in financial economics: whether equity returns are predictable over various horizons? The conventional wisdom in the literature is that aggregate dividend yields strongly predict excess returns, and the predictability is stronger at longer horizons (Fama and French (1988), Campbell (1991), and Cochrane (1992)). In contrast, Ang and Bekaert (2007) find that dividend yields, together with the short rate, predict excess returns only at short horizons, and do not have any long-horizon predictive power. In this paper, we undertake an analysis of both in-sample and out-of-sample tests of equity return predictability to better understand the empirical evidence on return predictability over different time horizons. We first propose a nonparametric test to examine the predictability of equity returns, which can be interpreted as a signal-to-noise ratio test. Our empirical results show that the short rate, dividend yields and earnings yields have good predictability power for both short and long horizons, which is different from both the conventional wisdom and Ang and Bekaert (2007). Also, using our nonparametric test, a comprehensive in-sample and out-of-sample analysis documents that the predictor variables (dividend yields, earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio, investment to capital ratio, corporate issuing activity, and consumption, wealth, and income ratio) have predictability power on equity returns but this cannot be well captured by linear prediction models. In addition, we use the nonparametric test to compare the conventional long-horizon prediction regression models on predictor variables with the historical mean model, where there has exists a debate about which model has better forecasting power for equity returns (Campbell and Thompson (2007) and Goyal and Welch (2007)). We find that the prevailing prediction model has a better forecasting power than the historical mean model because the former has a lower neglected signal-to-noise ratio. Finally, our nonparametric predictive models have lower RMSE than the historical mean model at both short-horizon and long-horizon. Using our nonparametric methods, both combined and individual forecast outperform the historical average.

Key Words: Asset Return Predictability, bootstrap, Hypothesis Testing, Kernel, Nonlinearity, Signal-to-Noise Ratio, Time Horizons, Out-of-sample Inference, Pricing Error

1 Introduction

There is a long tradition in finance and economics to study the predictability of equity returns or equity premiums. Cochrane (1999) points out that one of financial economist's views about the investment world was that returns are unpredictable until the mid-1980. Towards the end of the last century, academic researchers came to take seriously the view that aggregate stock returns are predictable[1]. Fundamental economic forces are crucial determinants of equity premia in financial markets.[2] The vast literature has suggested that excess returns are predictable by such variables as dividend-price ratios, earnings-price ratios, dividend-earnings ratios, short rates, book-to-market ratio, and an assortment of other financial indicators.

The predictability of long-horizon returns has drawn great interests from researchers. Long-horizon asset returns are more informative than their shorter-horizon counterparts, so random walk models, and martingale models based on past asset returns are statistically weak to explain real data. It is more reasonable to study the price behavior using the models of asset returns in economics or finance. The most popular model which is used to predict asset returns is the discounted-cash-flow or present-value model explored by Rozeff (1984)[3], Campbell and Shiller (1987), Campbell and Shiller (1988a, 1988b)[4], and West (1988). This model relates the price of a stock to its expected future cash flows (i.e., its dividends) discounted to the present value using a constant or time-varying discount rate. The present value model assumes that the expected stock return is constant through time and makes no assumption about equity repurchases by firms which affect the time pattern of expected future dividends. While stock prices and dividends appear to grow exponentially over time rather than linearly, the linear model (even allowing for a unit root) is less appropriate than a nonlinear model which can better capture the properties of asset returns across time. Thereafter, researchers have proposed several nonlinear models to explain or predict asset returns. One is the dividend models with rational bubbles in which the bubble is a nonlinear function of the stock's dividends (Froot and Obstfeld (1991)). This nonlinear model with stochastic rational bubbles has its limitation in explaining the observed predictability of stock returns. Another nonlinear model is a loglinear present-value model (Campbell (1991), Ang and Bekaert (2007)), which suggests a nonlinear relation between equity returns and dividend ratio, interest rates, excess returns, or cash

[1] A new generation of empirical research in the late twenty century does substantially enlarge our view of "what activities provide rewards for holding risks, and they challenge our understanding of those risk premiums".

[2] Equity risk premia are closely related to economic conditions. Equity returns seem to be high at business cycle troughs and low at peaks. In line with the pioneering work by Ferson and Merrick (1987), Fama and French (1989), researchers suggest that predictors of excess returns should be correlated with economic conditions. Lettau and Ludvigson (2005) summarize the literature and point out that we should expect to find evidence from predictive regressions of excess returns on macroeconomic variables over business cycle horizons."

[3] Rozeff (1984) showed that dividend yields can forecast equity risk premia by a deterministic dividend discount model. For example, Under the Gordon growth model, $P_t = \sum_{i=1}^{\infty} \frac{D_t(1+g)^i}{(1+r)^i} = \frac{D_{t+1}}{r-g}$, where P is the stock price, D is the dividend, r is the discount rate, and g is the constant growth rate of dividend. In the certainty model, the discount rate is the expected return on the stock. If the stock price represents a claim to the future stream of dividends, the price can be exactly determined assuming constantly growing dividends and a known discount rate and the model suggests that dividend yields should capture variations in expected stock returns.

[4] Campbell and Shiller (1988a, 1988b) develop a stochastic approximation to the dividend discount model and estimate the model in a VAR framework.

flows. The loglinear model can capture the asset price behavior without imposing restrictions on the expected returns. These studies suggest that there exists strong nonlinearity in the models of predicting asset returns, and that expected asset returns and dividend ratios are highly persistent and time-varying. Thus, it is important to investigate the predictive relationship between asset returns and time horizons.

There are quite a few works which examine the predictive power of the dividend yield on excess stock returns over various time horizons. Fama and French (1988), Campbell and Shiller (1988a,b), and Nelson and Kim (1993) document evidence of predictability. However, empirical studies increasingly cast doubt on the forecasting power of price-based predictors of equity returns. There are two recent debates on the predictability of equity premiums in the literature. First, most of the theoretical and empirical work focus on the predictive prowess of the dividend yield,[5] especially at long horizons. The conventional wisdom in the literature is that aggregate dividend yields strongly predict excess returns, and the predictability is stronger at longer horizons (Fama and French (1988)[6], Campbell (1991), and Cochrane (1992)). The results at different horizons are reflections of a single underlying phenomenon. If daily returns are very slightly predictable by a slow-moving or persistent variable, then predictability adds up over long horizons. In contrast, Ang and Bekaert (2007) find that dividend yields, together with the short rate, predict excess returns only at short horizons and do not have any long-horizon predictive power. On the other hand, Goyal and Welch (2007) argue that the historical average excess stock return forecasts future excess stock returns better than regressions of excess returns on predictor variables. In response to their arguments, Campbell and Thompson (2007) show that many predictive regressions beat the historical average return by imposing restrictions on the signs of coefficients and return forecasts, or the coefficients relating valuation ratios to future returns based on steady-state models. The conclusions of the two debates are controversial.

Most of the existing empirical studies use linear regressions to forecast asset returns. There are a number of pitfalls applying those models to predict asset returns or evaluate the predictability power. First, several authors expressed concern that the apparent predictability of stock returns might be spurious given the fact that many predictor variables, such as valuation ratios, used are highly persistent. Nelson and Kim (1993) and Stambaugh (1999) pointed out that persistence leads to biased coefficients in predictive regressions if innovations in the predictor variable are correlated with returns (as is strongly the case for valuation ratios, although not for interest rates). Under the same conditions, the standard t-test for predictability has incorrect sizes in finite samples (Cavanagh et al., 1995). These problems become more serious if applied econometricians are data mining, considering large numbers of variables, and reporting only those results that are apparently statistically significant (Foster et al., 1997; Ferson, Sarkissian, and Simin, 2003). Sarkissian, and

[5] Fama and French (1988), Campbell and Shiller (1988a,b), Goetzmann and Jorion (1993, 1995), Hodrick (1992), Stambaugh (1999), Wolf (2000), Goyal and Welch (2003, 2007), Valkanov (2003), Lewellen (2004), Campbell and Yogo (2006), Campbell and Thompson(2007), and Ang and Bekaert (2007)

[6] Fama and French (1988) provide the strongest evidence in support of the dividend yield effect by using overlapping multiple-year horizon returns. They observe that the explanatory power of the dividend yield increases with the time horizon of the returns; over 4-year horizons, the R^2's reach an astonishing high value of 64%.

Simin (2003) explore spurious regressions and data mining in the presence of serially correlated explanatory variables and they conclude that many regressions based on individual predictor variables may result in spurious conclusions. Second, another problem is that the explanatory variable, the dividend yield, is not properly exogenous, but rather contains a price level that also appears in the regression (Stambaugh (1986)). Moreover, Fama and French (1988) point out an "errors-in-variables" problem due to the fact that yields contain forecasts of future returns and dividend growth. This may bias downward the regression coefficient in the dividend yield regression. Fama (1990) and Kothari and Shanken (1992) suggest that the errors-in-variables problem is a potentially major one, since a significant percentage of return variance may be explained by changes in the growth rate of future dividends.

Third, there exists serial correlation in the forecast error particularly when the time horizon h is large relative to the sample size. As a result, there exist some finite sample problems for reliable statistical inference (Hodrick (1992), Nelson and Kim (1993)). An active recent literature discusses alternative econometric methods or propose new statistical tests for correcting the bias and conducting valid inference on estimation of long-horizon predictive regression models with persistent variables and errors.[7] These studies have emphasized the bias toward rejection of the null hypothesis of no predictability. In particular, the usual corrections to the standard errors are only valid asymptotically, and there is some question as to whether "asymptotic" should be measured in terms of years, decades, or even centuries, especially for long-horizon forecasts. Hodrick (1992) examines the implications for hypothesis testing of using different specifications of the forecasting equation. Nelson and Kim (1993) analyze small-sample biases in simulations of a VAR system for returns and yields, under the null hypothesis of no predictability of returns. Using U.S. returns sampled annually, they report that the simulated distributions of t-statistics are displaced upward, and still find some spurious evidence of predictability at conventional significance levels. In the case of the dividend yield regression, however, price levels appear in both the regressor and the regressand. From the work of Dickey and Fuller (1976) and Stambaugh (1986), it is well-known that regressions on lagged dependent variables lead to biased coefficient estimates. Goetzmann and Jorion (1993) use the bootstrap methodology, as well as simulations, to examine the finite sample distribution of test statistics under the null hypothesis of no forecasting ability. These experiments are constructed so as to maintain the dynamics of regressions with lagged dependent variables over long horizons (up to four years). They find that the empirically observed statistics are well within the 95% bounds of their simulated distributions and overall there is no strong statistical evidence indicating that dividend yields can forecast excess equity returns. Wolf (2000) uses a new statistical method for finding reliable confidence intervals for regression parameters in the context of dependent and possibly heteroscedastic data, called subsampling and does not find convincing evidence for the predictability of stock returns. Ang and Bekaert (2007) find that excess return predictability by the dividend yield is not statistically significant at longer horizons or across countries and also uses the nonlinear present value model to examine the fit of regression-based

[7] See Cavanagh et al., 1995; Mark, 1995; Kilian, 1999; Lewellen, 2004; Campbell and Yogo, 2006; Polk et al., 2006; Ang and Bekaert, 2007; Valkanov, 2003

expected returns with true expected returns. Consistent with the data, they find that a univariate dividend yield regression provides a rather poor proxy for the true expected return. However, using both the short rate and dividend yield considerably improves the fit, especially at short horizons.

Fourth, while previous studies model returns and dividend yields using a finite-order VAR system (Hodrick (1992), Campbell and Shiller (1988a,b), Stambaugh (1999)), the VAR model cannot fully capture the nonlinear dynamics of dividend yields implied by the present value model. Indeed, for a linear predictive regression model, when an price-based estimator or regressor appeals to be statistically insignificant, one cannot conclude that the null hypothesis of no predictability holds, because there may exist neglected nonlinear predictibility. Fifth, a different critique[8] emphasizes that the most linear predictive regressions have often performed poorly out-of-sample (Goyal and Welch (2003, 2007); Campbell and Thompson (2007)). It is well-known that while in-sample diagnostic analysis is important and can reveal useful information on possible sources of model misspecification, it may cause overfitting and capture spurious predictability. Out-of-sample evaluation can capture the true predictability of a model or the data generating process.[9] The disparities between in-sample and out-of-sample test results of return predictability documented in the literature make an overall assessment of return predictability difficult. In particular, it is unclear whether the poor out-of-sample performance of linear prediction models is due to the limitation of linear models or due to the nonexistence of predictability of equity returns. Many of the earlier out-of-sample tests have focused on the dividend ratios. Fama and French (1988) interpret the out-of-sample performance of dividend ratios to have been a success. Bossaerts and Hillion (1999) interpret the out-of-sample performance of the dividend yield (not dividend price ratio) to be a failure. Torous and Valkanov (2000) find that a low signal-noise ratio of many predictive variables makes a spurious relation between returns and persistent predictive variables unlikely and would lead to no out-of-sample forecasting power. Rapach and Wohar (2006) explore out-of-sample performance for a number of variables and find that certain financial variables display significant in-sample and out-of-sample predictive ability with respect to stock returns. Goyal and Welch (2007) argue that the poor out-of-sample performance of predictive regressions is a systemic problem. They compare predictive regressions with historical average returns and find that historical average returns almost always generate superior return forecasts. They conclude that "the profession has yet to find some variable that has meaningful and robust empirical equity premium forecasting power." Campbell and Thompson (2007) find that most of these predictor variables perform better out-of-sample than the historical average return forecast, once weak restrictions are imposed on the signs of coefficients and return forecasts. The out-of-sample explanatory power is small, but nonetheless is economi-

[8] This critique had a particular force during the bull market of the late 1990s, when low valuation ratios predicted extraordinarily low stock returns that did not materialize until the early 2000s (Campbell and Shiller, 1998).

[9] Here are several important reasons why out-of-sample predictability check is important. First, the usual practice of extensive search for more complicated models using the same or similar data set may suffer from the so-called data snooping bias, as pointed out by Lo and MacKinlay (1989) and White (2000). A more complicated model may overfit idiosyncratic features of the data without capturing the true data generating process. Out-of-sample prediction evaluation will alleviate, if not eliminate completely, such data snooping bias. Second, a model that fits in-sample data well may not predict the future well because of unforeseen structural changes or regime shifts in the data generating process.

cally meaningful for investors. They also impose theoretical restrictions on the coefficients relating valuation ratios to future returns and theoretically restricted valuation models often outperform return forecasts based on the long-run historical mean of stock returns.

In this paper, we undertake an analysis of both in-sample and out-of-sample tests of stock return predictability in an effort to better understand the empirical evidence on return predictability. We are particularly interested in investigating the following problems:

- Does the predictability of valuation ratios such as dividend yields exist at various horizons?
- Do linear predictor variables-based regression models suffer from neglected nonlinear predictability? In particular, is the poor out-of-sample performance of most linear prediction models due to the limitation of linear models or due to the nonexistence of predictability of equaity returns?
- Does the predictor-based regression model beat the historical average excess stock return (historical mean model)?

For these purposes, we first develop a reliable out-of-sample nonparametric model-free predictability test, which has several appleaing features. First, as is well-known, the nonparametric method can capture a wide variety of linearities and nonlinearities without assuming any parametric model. Thus, it can assess directly the predictability of equity return data itself rather than the predictability of a specific model for equity return. Second, the nonparametric predictability test can be interpreted as a signal-to-noise ratio, because it is based on the average of the squared predictable components over the sample variance of pricing errors. Third, we propose to use a conditional bootstrap procedure which maintain the original dynamics of predictor variables and serial dependence structure of the multi-step-ahead forecast errors. Such a bootstrap procedure provides reliable statistical inference for sample sizes typically encountered in the literature. Simulation studies show that it has reasonable size and power in finite samples even when the regressors are highly persistent and the forecast horizon is relatively long.

We apply the proposed nonparametric test to examine whether there exists the predictability of equity returns at short or long horizons. Ang and Bekaert (2007) find that dividend yields, together with the short rate, predict excess returns only at short horizons. In contrast, we find that the short rate, dividend yields, and earnings yields have good predictability power at both short and long horizons. Second, the comprehensive in-sample and out-of-sample analysis suggests that such variables as dividend yields, earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio, investment to capital ratio, corporate issuing activity, and consumption, wealth, and income ratio have predictability power for equity returns, but this often cannot be captured by popular linear regression models. Third, we find that the prevailing prediction model beats the historical mean model because there is more neglected signal-to-noise ratio for the latter. Our conclusion is in contrast to Goyal and Welch (2007), and is consistent with Campbell and Thompson (2007), who find that predictor variables perform better out-of-sample than the historical average return forecasts, once weak restrictions are imposed on the signs of coefficients and return forecasts, or the coefficients relating valuation ratios to future returns based on steady-state models. In fact, the restrictions on coefficients is a form of nonlinearity.

In the literature, most papers focus on a set of predictors based on theoretical models. From an academic viewpoint, the use of model-based predictors facilitates an understanding of specific aspects of the economic mechanism. From an investor's viewpoint, however, these predetermined variables may not be enough to capture all information required in decision making. Forecast combination has recently received renewed attention in the forecasting literature; Stock and Watson (1999, 2003, 2004) with respect to forecasting inflation and real output growth. Rapach, Strauss, and Zhou (2009) propose a combination approach to improve the out-of-sample equity premium forecasting problem. In addition to the individual forecast, we also consider the combined forecast to improve equity premium forecasts, and examine the out-of-sample performance. On the other hand, previous studies suggest that there exists strong nonlinearity in the models of predicting asset returns, and that expected asset returns and dividend ratios are highly persistent and time-varying. The poor out-of-sample performance of most linear prediction models is due to the limitation of linear models. The lack of consistent out-of-sample evidence in Goyal and Welch (2008) indicates the need for improved forecasting methods to better establish the empirical reliability of equity premium predictability. In this paper, we propose nonparametric estimators to forecast the equity returns using both individual forecast and combined forecast.

Goyal and Welch (2007) argue that the historical average excess stock return forecasts future excess stock returns better than regressions of excess returns on predictor variables. With respect to the economic variables used to predict the equity premium, we use the 15 economic variables from Goyal and Welch (2008) to predict the individual predictive models. Common to all these papers is a focus on a small set of predictors based on theoretical models. From an academic viewpoint, the use of model-based predictors facilitates an understanding of specific aspects of the economic mechanism. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. We find that the combined forecast methods outperform the individual forecast methods. Fama and French (1989) and others show that these variables can detect changes in economic conditions that potentially signal fluctuations in the equity risk premium. But the dividend yield or term spread alone could capture different components of business conditions, and a given individual economic variable may give a number of "false signals" and/or imply an implausible equity risk premium during certain periods. Rapach, Strauss, and Zhou (2009) argue that if individual forecasts based on the predictors are weakly correlated, forecast combinatio should be less volatile and more reliably track movements in the equity risk premium. Our results are consistent with their argument that the combined forecast methods outperform the individual forecast methods. Combining forecast incorporates information from a host of economic variables while the historical average ignores economic variables. Combined forecasts have a substantially smaller bias than the historical average. Combining individual forecasts helps to reduce forecast variability.

The other important results we get are that our nonparametric predictive models have lower RMSE than the historical mean model at both short-horizon and long-horizon. Our nonparametric prediction can improve the out-of-sample performance without restrictions. Using our nonparametric methods, both combined and individual forecast outperform the historical average. One reasonable explanation is that the nonparametric predictive model can fit the equity return better based on the predictors. It is not restricted to the parametric forms. It can fit the data more better than simply the linear or nonlinear parametric model. Nonparametric prediction generates a forecast with a variance near that of the smooth real equity return data, thereby reducing the noise in the individual predictive regression model forecasts.

This paper is organized as follows. Section 2 introduces the proposed nonparametric predictability test. Section 3 describes the data. Section 4 presents and discusses the empirical results. Section 5 reports the out-of-sample performance of the individual and combined forecast and economic implication. Section 6 concludes the paper.

2 Nonparametric Test for Predictability

2.1 Hypotheses of Interests and Nonparametric Test

We are interested in whether the predictability of excess returns depends on time horizons. If future excess returns cannot be predicted by past dividend yield or other variables over any time horizon, then the null hypothesis holds.

Specifically, suppose $\{Y_t, X_t'\}'$ is a stationary time series process where Y_t is a scalar, and X_t is a d-dimensional vector. We are interested in testing the predictability of Y_{t+h} using X_t, where the integer h is the time horizon index for a multi-step ahead prediction. In our applications below, X_t is, for example, the dividend yield in period t, and Y_{t+h} is the asset return h periods ahead. Different h's will allow us to examine the relationship between asset return predictability and time horizons. The null hypothesis of interest can be written as

$$H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h}) \tag{3.1}$$

versus the alternative hypothesis

$$H_A : E(Y_{t+h}|X_t) \neq E(Y_{t+h}). \tag{3.2}$$

The null hypothesis H_0 is characterized by the horizon index h. It is possible that H_0 holds for a relatively long horizon but it does not hold for a relatively short horizon. This is one of our focuses in this paper, namely we will investigate the relationship between predictability of excess asset returns and the time horizon h, which has been a long-standing problem in empirical finance.

In empirical finance, often a linear predictive regression model

$$Y_{t+h} = X_t'\beta + \varepsilon_{t+h}, \tag{3.3}$$

is used to check predictability of excess asset returns. When an estimator for β is statistically insignificant, one does not find evidence for predictability power of X_t for Y_{t+h}. Strictly speaking, one cannot conclude that H_0 holds. This is because a zero parameter value for β is a necessary condition for H_0 but it is not a sufficient condition. A zero β implies that there is no linear predictive power of X_t for Y_{t+h}, but there may exist a nonlinear predictive power of X_t for Y_{t+h}. An example is that the true data generating process follows $Y_{t+h} = X_t^2 + \varepsilon_{t+h}$, where X_t is normally distributed with zero mean and the disturbance ε_{t+h} is independent of X_t. In this case, a linear regression coefficient β is exactly zero but $E(Y_{t+h}|X_t) = X_t^2$.

When an estimator for β is statistically significant, there exists evidence of the predictive power of X_t for Y_{t+h}. In this case, one may be interested in testing whether the linear regression model has the optimal predictive power for Y_{t+h}. Put it differently, one may be interested in testing whether there exists any nonlinear predictive power of X_t for Y_{t+h}, in addition to the documented linear predictability. In this case, the null hypothesis of interest

$$H_0 : E(\varepsilon_{t+h}|X_t) = 0 \tag{3.4}$$

versus the alternative hypothesis

$$H_A : E(\varepsilon_{t+h}|X_t) \neq 0, \tag{3.5}$$

where ε_{t+h} is the prediction error from the linear regression model in (3.3). The null hypothesis H_0 in (3.4) implies that the linear regression model in (3.3) has optimal predictive power. When H_A in (3.5) holds, there exists a nonlinear predictive relationship between X_t and Y_{t+h}, and a suitable nonlinear predictive model will outperform the linear regression model in (3.3). Because ε_{t+h} is unobservable, we need to use an estimated residual $\hat{\varepsilon}_{t+h} = Y_{t+h} - X_t'\hat{\beta}$, where $\hat{\beta}$ is an estimator for β. Note that when H_0 holds, $\{\varepsilon_{t+h}\}$ may not be a martingale difference sequence unless $h = 1$. In general, H_0 allows $\{\epsilon_{t+h}\}$ to follow a MA(h) dependence. This has an important implication on inference, particularly when h is relatively large.

In this section, we develop a unified nonparametric testing framework which is applicable to test hypotheses in (3.1) and (3.4). The basic idea is to use a nonparametric estimator for $E(Y_{t+h}|X_t)$ or $E(\varepsilon_{t+h}|X_t)$ and check if the estimator is close to constant or zero. As is well-known, the nonparametric method has an advantage that it does not require an ex ante model specification and can capture any predictive relationship no matter whether it is linear or nonlinear (c.f. Härdle (1993), Pagan and Ullah (1999)). Thus, it is quite suitable for our purpose here.

To avoid capturing spurious predictability due to in-sample overfitting, we consider out-of-sample predictability check. There are several important reasons why out-of-sample predictability check is important. First, the usual practice of extensive search for more complicated models using the same or similar data set may suffer from the so-called data snooping bias, as pointed out by Lo and MacKinlay (1989) and White (2000). A more complicated model may overfit some idiosyncratic features of the data without capturing the true data generating process. Out-of-sample prediction evaluation will alleviate, if not eliminate completely, such data snooping bias. Second, a model that fits in-sample data well may not predict the future well because of unforeseen structural changes or regime shifts in the data generating process. Therefore, in-sample analysis is not adequate and it is important to examine out-of-sample prediction. Third, out-of-sample prediction is more relevant to most economic applications in practice.

Specifically, suppose we have an observed sample $\{Y_t, X_t'\}_{t=1}^T$ of size T. We first split the sample into two parts: the first subsample contains R observations, and the second subsample contains $n = T - R$ observations. We will use the first subsample or a modification of it to estimate model parameter β and use the second subsample to check predictability. There are various methods to estimate parameter β. One simple method is to use the first subsample $\{Y_{t+h}, X_t'\}_{t=1}^R$ to estimate β. Another method is to use $\{Y_{t+h}, X_t\}_{t=i+1}^{R+i}$ to estimate β when predicting $Y_{R+h+1+i}$, for $0 \leq i \leq n - h - 1$. This is called the rolling estimation. One can also use the recursive estimation method, which uses the subsample $\{Y_{t+h}, X_t\}_{t=1}^{R+i}$ to estimate β when predicting $Y_{R+h+1+i}$. Generally, we use the notation $\hat{\beta}_t$ to denote an estimator for β when predicting Y_{t+h} in an out-of-sample context. The resulting estimated out-of-sample residual from a linear model (3.3) is

$$\hat{\varepsilon}_{t+h} = Y_{t+h} - X_t'\hat{\beta}_t, t = R + 1, \cdots T - h$$

To capture potentially neglected nonlinear predictable component in ε_{t+h}, we use a smoothed kernel method to estimate $E(\varepsilon_{t+h}|X_t)$. Put

$$\begin{aligned} \widehat{m}_h(x) &= \frac{1}{n-h}\sum_{s=R+1}^{T-h} \hat{\varepsilon}_{s+h} K_b(x - X_s), \\ \widehat{g}(x) &= \frac{1}{n-h}\sum_{s=R+1}^{T-h} K_b(x - X_s), \end{aligned}$$

where $x = (x_1, x_2, \cdots, x_d)'$, $y = (y_1, y_2, \cdots, y_d)'$, and $K_b(x - y) = \prod_{i=1}^{d} b^{-1}K[(x_i - y_i)/b]$. The kernel function $K(\cdot)$ is is a prespecified symmetric probability density function. Examples include a Gaussian kernel $K(u) = (2\pi)^{-1/2}\exp(-u^2/2)$ and a quatic kernel $K(u) = \frac{3}{4}(1 - u^2)\mathbf{1}(|u| \leq 1)$, where $\mathbf{1}(\cdot)$ is the indicator function, giving value 1 if $|u| \leq 1$ and value 0 otherwise. The bandwidth $b = b(n)$ vanishes to zero as the sample size $n \to \infty$, but at a slower rate. For simplicity, we use the same bandwidth for each components of X_t. In practice, one can first standardize each component of the vector X_t by its sample standard deviation. The regression estimator for $E(\varepsilon_{t+h}|X_t)$ is then defined as follows:

$$\widehat{r}_h(x) = \frac{\widehat{m}_h(x)}{\widehat{g}(x)}.$$

This is called the Nadaraya-Watson regression estimator. The estimator $\widehat{g}(x)$ in the denominator is a kernel estimator for the marginal density $g(x)$ of $\{X_t\}$. Under regularity conditions, $\widehat{r}_h(x) \to r_h(x) = E(\varepsilon_{t+h}|X_t = x)$ in probability as both $R, n \to \infty$.

Under H_0, $\widehat{r}_h(x)$ is close to zero for all x. Under the alternative hypothesis H_A, $\hat{r}_h(x)$ is not a zero function but is a nontrivial function of x subject to sampling variation. To measure the departure of $\hat{r}_h(x)$ from zero over all x, we use the following global measure

$$\widehat{Q}(h) = \frac{1}{n-h} \sum_{t=R+1}^{T-h} \widehat{r}_h^2(X_t) w(X_t),$$

where the positive weighting function $w(\cdot)$ can be chosen to trim the extreme observations where the estimation of $\hat{r}(x)$ is not reliable due to sparse observations (we allow the distribution of X_t has unbounded support). It can also be used to direct power of the proposed test to the region of interest, such as predictability when X_t is negative (in this case, we choose $w(x) = \mathbf{1}(x \leq 0)$. The statistic $\widehat{Q}(h)$ can be viewed as a measure of the magnitude of the "signal" that can be extracted to predict asset returns if (and only if) it contains no systematic predictable component in $E(\varepsilon_{t+h}|X_t)$, the estimator $\widehat{r}_h(X_t)$ and therefore $\widehat{Q}(h)$ will be close to zero.

Alternatively, we can directly use an integrated global measure

$$\widetilde{Q}(h) = \int \widehat{r}_h^2(x) \widehat{g}(x) w(x) dx$$

where the integral is over the support of $w(x)$, and it can be computed using either a numerical integration method (e.g. Gauss-Newton method) or a Monte Carlo simulation method[10].

The asymptotic behaviors of $\widehat{Q}(h)$ and $\widetilde{Q}(h)$ are similar. We now consider the decomposition

$$\begin{aligned} \widehat{Q}(h) &= \int \widehat{r}_h^2(x) \widehat{g}(x) w(x) dx \\ &= \int \frac{\widehat{m}_h^2(x)}{g(x)} w(x) dx + \int \widehat{m}_h^2(x) \left[\frac{1}{\widehat{g}(x)} - \frac{1}{g(x)} \right] w(x) dx \\ &= \int \widehat{m}_h^2(x) a(x) dx + O_p((Tb)^{-1} + (Tb)^{-\frac{1}{2}} + h^2), \end{aligned}$$

where $a(x) = w(x)/g(x)$, and the reminder term is dominated by the first (leading) term under suitable regularity conditions. Thus, we focus on the first term, which will determine the asymptotic distribution of the statistic $\widehat{Q}(h)$.

[10] The Monte Carlo method can be implemented as follows. Without loss of generality assume that $w(\cdot)$ is a prespecified probability density function.Then we can generate a large *i.i.d.* sample $\{X_i^*\}_{i=1}^N$ from the probability distribution $w(\cdot)$. Then the average $\widehat{Q}^*(h) = N^{-1} \sum_{i=1}^{N} \widehat{r}_h^2(X_i^*)$ will be arbitrarily close to $\widehat{Q}(h)$ if N is sufficiently large (much larger than the sample size n) by the law of large numbers.

For the first term, we have

$$\begin{aligned}
\int \widehat{m}_h^2(x)a(x)dx &= \int \left[\frac{1}{n-h}\sum_{s=R+1}^{T-h}\hat{\varepsilon}_{s+h}K_b(x-X_s)\right]^2 a(x)dx \\
&= \frac{1}{(n-h)^2}\sum_{|t-s|>h}\hat{\varepsilon}_{t+h}\hat{\varepsilon}_{s+h}\int K_b(x-X_t)K_b(x-X_s)a(x)dx \\
&\quad +\frac{1}{(n-h)^2}\sum_{|t-s|\le h}\hat{\varepsilon}_{t+h}\hat{\varepsilon}_{s+h}\int K_b(x-X_t)K_b(x-X_s)a(x)dx \\
&= \widehat{A}(h)+\widehat{B}(h),
\end{aligned}$$

where the term $\hat{A}(h)$ is a sum over (t,s) with $|t-s|>h$, and the term $\hat{B}(h)$ is a sum over (t,s) with $|t-s|\le h$. For the term $\hat{B}(h)$, we have

$$\begin{aligned}
\widehat{B}(h) &= \frac{1}{(n-h)^2}\sum_{|t-s|\le h}\hat{\varepsilon}_{t+h}\hat{\varepsilon}_{s+h}\int K_b(x-X_t)K_b(x-X_s)a(x)dx \\
&= \frac{1}{(n-h)^2}\sum_{t=R+1}^{T-h}\hat{\varepsilon}_{t+h}^2\int K_b^2(x-X_t)a(x)dx \\
&\quad +\frac{2}{(n-h)^2}\sum_{t=R+2}^{n-h}\sum_{s=\max(R+1,t-h)}^{t-1}\hat{\varepsilon}_{t+h}\hat{\varepsilon}_{s+h}\int K_b(x-X_t)K_b(x-X_s)a(x)dx \\
&= \frac{1}{(n-h)b}\sigma_\varepsilon^2\int w(x)dx\int K^2(u)du+\frac{2}{(n-h)}\sum_{j=1}^{h}\gamma(j)E[a(X_t)f_j(X_t,X_t)]+O_p((nb)^{-1}),
\end{aligned}$$

where $\sigma_\varepsilon^2 = var(\varepsilon_{t+h})$, $\gamma(j)=cov(\varepsilon_t,\varepsilon_{t-j})$ is the autocovariance function of $\{\varepsilon_t\}$, and $f_j(\cdot,\cdot)$ is the joint probability density of (X_t,X_{t-j}). Note that generally $\gamma(j)\neq 0$ for $0\le j\le h$ in a h-step ahead prediction model (3.3), even when H_0 holds. As noted earlier, $\{\varepsilon_{t+h}\}$ generally displays a MA$(h-1)$ structure under H_0.

Thus, $\hat{B}(h)$ depends on the serial dependence of $\{\varepsilon_{t+h}\}$ due to the existence of the second term. The effect of serial dependence in $\{\varepsilon_{t+h}\}$ on $\widehat{B}(h)$ is generally larger when the horizon parameter h is larger. In our construction of a test statistic, we could subtract the original form of $\hat{B}(h)$ directly from the global measure $\widehat{Q}(h)$, rather than use the asymptotic approximation of $\hat{B}(h)$. This will make the proposed test robust to the effect of serial dependence contained in $\hat{B}(h)$.

The term $\hat{A}(h)$ can be written as

$$\widehat{A}(h)=\frac{2}{(n-h)^2}\sum_{t=R+2}^{n-h}\sum_{s=R+1}^{t-h-1}\hat{\varepsilon}_{t+h}\hat{\varepsilon}_{s+h}\int K_b(x-X_t)K_b(x-X_s)a(x)dx.$$

Under H_0, $\widehat{A}(h)$ has an approximately zero mean. Its variance $var(\widehat{A}(h))$ depends on serial dependence in $\{\varepsilon_{t+h}\}$. However, when $\{\varepsilon_{t+h}\}$ has a MA$(h-1)$ structure where h is a fixed integer, the effect of serial dependence in $\{\varepsilon_t\}$ on $var[\widehat{A}(h)]$ is an asymptotically negligible higher order term, and it can be shown that the asymptotic variance of $b^{1/2}(n-h)\hat{A}(h)/\sigma_\varepsilon^2$ is given by

$$V = 8\int w^2(x)dx \int \left[\int K(u)K(u+v)du\right]^2 dv.$$

Using the central limit theorem for degenerate U-statistics, we can show $b^{\frac{1}{2}}(n-h)\widehat{A}(h)/\sigma_\varepsilon^2 \xrightarrow{d} N(0,V)$, as stated below:

Theorem 1 *Suppose Assumptions A.1–A.6 in the Appendix hold. Then*

(i) *under* H_0, *we have*

$$\hat{\mathbf{Q}}_h = \frac{\sqrt{b}(n-h)\hat{Q}(h)/\widehat{\sigma}_\varepsilon^2 - C/\sqrt{b}}{\sqrt{V}} \xrightarrow{d} N(0,1)$$

where $C = \int w(x)dx \int K^2(u)du$, $\widehat{\sigma}_\varepsilon^2 = (n-h)^{-1}\sum_{t=R+1}^{T-h} e_{t+h}^2$, *and* $e_{t+h} = \widehat{\varepsilon}_{t+h} - \widehat{r}_h(X_t)$.

(ii) *under* H_A,

$$\frac{\hat{\mathbf{Q}}_h}{\sqrt{b}(n-h)} \rightarrow \frac{V^{-1/2}\int r_h^2(x)g(x)w(x)dx}{\sigma_\varepsilon^2}.$$

The $\hat{\mathbf{Q}}_h$ test statistic has an appealing interpretation. Ignoring the centering and scaling factors, the $\hat{\mathbf{Q}}_h$ test statistic is essentially based on the ratio $\hat{Q}(h)/\widehat{\sigma}_\varepsilon^2$. Here, the denominator $\widehat{\sigma}_\varepsilon^2$ is the sample variance of pricing errors, and the numerator $\hat{Q}(h)$ is the average of the squared predictable components neglected by the linear regression model (3.3). Therefore, the ratio $\hat{Q}(h)/\widehat{\sigma}_\varepsilon^2$ can be viewed as an estimator for the neglected signal-to-noise ratio of the linear model. If the neglected pricing signal $\hat{Q}(h)$ is weak relative to the pricing noise $\widehat{\sigma}_\varepsilon^2$, the $\hat{\mathbf{Q}}_h$ test will not reject the null hypothesis H_0. If the neglected pricing signal $\hat{Q}(h)$ is sufficiently large relative to the pricing noise $\widehat{\sigma}_\varepsilon^2$, the $\hat{\mathbf{Q}}_h$ test will reject the null hypothesis H_0. How large the signal-to-noise ratio should be in order to be considered as sufficiently large is determined by the critical value of the test statistic.

Theorem 1(ii) shows that under H_A, the $\hat{\mathbf{Q}}_h$ statistic diverges to infinity at rate $\sqrt{b}(n-h)$. Thus, as long as $r_h(x)$ is not zero over the support of the weighting function $w(x)$ under H_A , the $\hat{\mathbf{Q}}_h$ test will be able to reject H_0 at any given significant level with probability approaching one as the sample sizes $R, n \rightarrow \infty$.

In computing the neglected pricing signal-to-noise ratio, we have used a nonparametric estimator for σ_ε^2. The variance estimator $\widehat{\sigma}_\varepsilon^2$ is based on the nonparametric residual e_{t+h} which is always consistent for the true pricing error $\varepsilon_t^o \equiv Y_{t+h} - E(Y_{t+h}|X_t)$ under both H_0 and H_A. One could also use the parametric variance estimator $\widetilde{\sigma}_\varepsilon^2 = \frac{1}{n-h}\sum_{t=R+1}^{T-h} \widehat{\varepsilon}_{t+h}^2$ using the estimated residuals from the linear regression model. This estimator is simpler than $\widehat{\sigma}_\varepsilon^2$, and may give better sizes in finite samples, because it is a better estimator for σ_ε^2 than $\widehat{\sigma}_\varepsilon^2$ under H_0. However, $\widetilde{\sigma}_\varepsilon^2$ is not consistent for the true error variance $Var(\varepsilon_t^o)$ under H_A, because it contains the neglected signals. Consequently, it may give a lower power in finite samples.

The test statistic $\hat{\mathbf{Q}}_h$ is constructed to check out-of-sample predictability of residual $\hat{\varepsilon}_{t+h}$ using X_t. It can also be used to test the null hypothesis H_0 in (3.1), namely the predictability of X_t for Y_{t+h}. This can be done by replacing the sample size n with T, and replacing the estimated residual $\hat{\varepsilon}_{t+h}$ with $Y_{t+h} - \bar{Y}$, where $\bar{Y} = (T-h)^{-1} \sum_{t=1}^{T-h} Y_{t+h}$ is the sample mean of $\{Y_{t+h}\}_{t=1}^{T-h}$. The resulting test statistic is still asymptotically N(0,1) under H_0 in (3.1).

Theorem 1(i) implies that approximately $\gamma(n-h)\hat{Q}(h)/\widehat{\sigma}_{\varepsilon}^2 \sim \chi^2_{\lambda_n}$ as $R, n \to \infty$ where the constant $\gamma = 2C/V$ and the degree of freedom $\lambda_n = 2C^2/bV$. Here, both constants γ and λ_n do not depend on any nuisance parameters or nuisance functions, such as the error distribution and the density function of X_t. In fact, they are independent of the data generating process. Therefore, the asymptotic null distribution of the scaled signal-to-noise ratio statistic $\gamma(n-h)\hat{Q}(h)/\widehat{\sigma}_{\varepsilon}^2$ is independent of nuisance parameters or nuisance functions, and approximately $\gamma(n-h)\hat{Q}(h)/\widehat{\sigma}_{\varepsilon}^2$ is distribulted as $N(\lambda_n, 2\lambda_n)$ where λ_n is known. This is the so-called Wilks' phenomena in statistics. One important implication of Wilks' phenomena is that one can simply simulate the null distributions by setting the nuisance parameters under the null hypothesis at reasonable values or estimates.

The asymptotic normality is quite convenient to use in practice. However, several reasons suggest that the asymptotic normal approximation may not work well in finite samples. First, the nonparametric estimator $\widehat{r}_h(x)$ converges slowly to the true function $r_h(x)$ particularly when the dimension d of X_t is relatively large. As it turns out, the neglected reminder terms in the asymptotic expansion of $\hat{Q}(h)/\widehat{\sigma}_{\varepsilon}^2$ are quite close to in order of magnitude to the dominating term which determines the asymptotic normal distribution of $\hat{\mathbf{Q}}_h$. Evidence in related literature shows that the size of nonparametric test statistics is generally very poor in finite samples. Second, in the present framework, $\{\varepsilon_{t+h}\}$ is not an i.i.d. or martingale difference sequence under the null hypothesis. Instead, it follows an MA$(h-1)$ structure in ε_{t+h} under the null hypothesis H_0 due to the h-step ahead prediction. Asymptotic analysis shows that the serial dependence in $\{\varepsilon_{t+h}\}$ has no impact on the asymptotic mean $C/\sqrt{b}$and the asymptotic variance V, but it may substantially affect the finite sample mean and variance of the test statistic $\hat{Q}(h)/\widehat{\sigma}_{\varepsilon}^2$, particularly when h is relatively large. Third, our asymptotic analysis shows that parameter estimation uncertainty in $\hat{\beta}_t$ has an asymptotically negligible impact on the asymptotic distribution of the proposed test, but the impact depends on the relative magnitude between two sample sizes R, n. When the ratio n/R is large (i.e., when n is large relative to R) , the impact of parameter estimation uncertainty of $\hat{\beta}_t$ may be substantial in finite samples.[11]

To obtain a reliable reference based on the proposed test in finite samples, we propose the following conditional bootstrap procedure which preserves the MA(h) structure in ε_{t+h} among other things:

Step 1: Use the first subsample $\{Y_{t+h}, X_t'\}_{t=1}^R$ to estimate the linear regression model

$$Y_{t+h} = X_t'\beta + \varepsilon_{t+h}, t = 1, ..., R.$$

[11] One implication of this result is that one should use a large R relative to n in practice to alleviate the impact of parameter estimation in $\widehat{\beta}$.

Obtain the parameter estimator $\hat{\beta}$. Alternatively, rolling estimation or recursive estimation could also be used.

Step 2: Use $\hat{\beta}$ to compute the out-of-sample estimated residual $\hat{\varepsilon}_{t+h} = Y_t - X_t'\hat{\beta}$ for $t = R+1, ..., T-h$.

Step 3: Compute the nonparametric estimates $\widehat{r}_h(X_t)$ and the nonparametric residual $\widehat{e}_{t+h} = \widehat{\varepsilon}_{t+h} - \widehat{r}_h(X_t)$ for $t = R+1, ..., T-h$.

Step 4: Compute the signal-to-noise ratio $\hat{Q}(h)/\widehat{\sigma}_\varepsilon^2$ using a prespecified kernel $k(\cdot)$ and bandwidth $b = (n-h)^{1/5}$. In practice, data-driven methods can be used to choose the bandwidth b.

Step 5: Estimate an MA$(h-1)$ model for the nonparametric residual

$$\widehat{e}_{t+h} = \sum_{j=1}^{h-1} \alpha_j v_{t+h-j} + v_{t+h}, t = R+1, ..., T-h.$$

This can be done via the conditional quasi-maximum likelihood estimation. Save the moving average parameter estimates $\{\hat{\alpha}_j\}_{j=1}^h$ and the estimated residual $\{\hat{v}_{t+h}\}_{t=R+1}^{T-h}$ in the MA$(h-1)$ model.

Step 6: Draw a bootstrap residual sample $\{\hat{v}_{t+h}^*\}_{t=R+1}^{T-h}$ from the centered empirical distribution of $\{\hat{v}_{t+h}\}_{t=R+1}^{T-h}$. Then obtain a bootstrap residual sample $\{\hat{\varepsilon}_{t+h}^*\}_{t=R+1}^{T-h}$ by the following MA$(h-1)$ model

$$\hat{\varepsilon}_{t+h}^* = \sum_{j=1}^{h-1} \hat{\alpha}_j \hat{v}_{t+h-j}^* + \hat{v}_{t+h}^*, t = R+1, ..., T-h,$$

where the parameter estimates $\{\hat{\alpha}_j\}_{j=1}^h$ are obtained in step 5. The bootstrap residual $\{\hat{\varepsilon}_{t+h}^*\}_{t=R+1}^{T-h}$ approximately preserves the MA$(h-1)$ structure of $\{\varepsilon_{t+h}\}$ under H_0.

Step 7: Use the bootstrap sample $\{\varepsilon_{t+h}^*, X_t\}_{t=R+1}^{T-h}$ to compute the bootstrap signal-to-noise ratio $\hat{Q}^*(h)/\widehat{\sigma}_\varepsilon^{*2}$ using the same kernel $k(\cdot)$ and bandwidth b as in Step 4.

Step 8: Repeat Steps 6 and 7 for a total of B times where B is a large number. Denote the obtained B bootstrap test statistics as $\{\hat{Q}_l^*(h)/\widehat{\sigma}_{\varepsilon l}^{*2}\}_{l=1}^B$.

Step 9: Compute the bootstrap p-value of the $\hat{\mathbf{Q}}_h$:

$$p^* = \frac{1}{B}\sum_{l=1}^{B} \mathbf{1}\left[\frac{\hat{Q}(h)}{\widehat{\sigma}_\varepsilon^2} < \frac{\hat{Q}_l^*(h)}{\widehat{\sigma}_{\varepsilon l}^{*2}}\right],$$

where $\mathbf{1}(\cdot)$ is the indicator function. Reject the null hypothesis H_0 at significance level α if and only if $p^* < \alpha$.

The above resampling approximation is a wild bootstrap. Here, one only need to calculate the signal-to-noise ratio $\hat{Q}(h)/\widehat{\sigma}_\varepsilon^2$ using the observed sample and bootstrap samples. There is no need to compute the original test statistic $\hat{\mathbf{Q}}_h$ which involves calculation of centering and scaling parameters. This follows because computing the bootstrap p-value involves ranking $\hat{\mathbf{Q}}_h$ and $\hat{\mathbf{Q}}_h^*$, which is equivalent to ranking the pricing signal-to-noise ratios $\hat{Q}(h)/\widehat{\sigma}_\varepsilon^2$ and $\hat{Q}^*(h)/\widehat{\sigma}_\varepsilon^{*2}$, given the fact that the centering and scaling factors do not depend on nuisance parameters and the data generating process. This greatly simplifies the computation of the test statistic.

When testing predictability of X_t for Y_{t+h} (i.e., testing H_0 in (3.1)), Steps 1 and 2 are not needed, the nonparametric residual in step 3 is replaced with $\widehat{e}_{t+h} = Y_{t+h} - \bar{Y}$, and the MA$(h-1)$ models in Steps 6 should be changed to the following:

$$Y^*_{t+h} = \overline{Y} + \sum_{j=1}^{h-1} \hat{\alpha}_j v^*_{t+h-j} + v^*_{t+h}, t = 1, ..., T-h.$$

respectively, where $\bar{Y}$ is the sample mean of $\{Y_{t+h}\}_{t=1}^{T-h}$.

We will examine the finite sample performance of the above conditional bootstrap procedure via simulation studies.

2.2 Simulation Design and Monte Carlo Evidence

It is well-known that for inference procedures based on linear prediction models, there exist two well-documented sources of size distortion that may arise in long-horizon regressions. First, many predictors, such as dividends and earning price ratios, interest rates, and forward premia, are highly persistent and only predetermined, rather than fully exogenous. Second, standard test-statistics based on prediction regressions do not have their usual limiting distribution (Cavanagh et al., 1995). The use of standard critical values is known to generate severe size distortion. These problems may carry over to the proposed nonparametric predictability test, particularly when h is large. In order to check the reliability of the proposed test, we investigate the finite performance (both size and power) of the proposed test using data-generating processes that could potentially be employed to capture the persistent behavior commonly observed in predictive regressors.

Table 0 summarizes the five data generating processes we use to investigate the empirical size of the tests for both linear and nonlinear predictability check.

Table 0 Summary of Simulation DGPs and Predictability Check

The data-generating processes are summarized in the table below, where $\varepsilon_{t+h} = \sum_{j=1}^{h} \alpha_j v_{t+h-j} + v_{t+h}$, $v_{t+h} \sim i.i.d.N(0,1)$, $u_t \sim i.i.d.N(0,1)$, and $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. Examine the case where $h = 1, 4, 12, 20$. Sample size $T = 250, 500, 1000$.

	DGP	Y_{t+h}	X_t	β_1, β_2, ρ
(1) Linear	$A.0(h)$	$Y_{t+h} = \alpha_0 + \varepsilon_{t+h}$	$X_t = \rho X_{t-1} + u_t$	$(0.1, 0.3, 0.5, 0.7, 0.9)$
	$A.1(h)$	$Y_{t+h} = \beta_0 + \beta_1 X_t + \varepsilon_{t+h}$	$X_t = \rho X_{t-1} + u_t$	$(0.1, 0.3, 0.5, 0.7, 0.9)$
	$A.2(h)$	$Y_{t+h} = \beta_0 + \beta_1 X_t^2 + \varepsilon_{t+h}$	$X_t = \rho X_{t-1} + u_t$	$(0.1, 0.3, 0.5, 0.7, 0.9)$
(2) Nonlinear	$B.0(h)$	$Y_{t+h} = \beta_0 + \beta_1 X_t + \varepsilon_{t+h}$	$X_t = \rho X_{t-1} + u_t$	$(0.1, 0.3, 0.5, 0.7, 0.9)$
	$B.1(h)$	$Y_{t+h} = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \varepsilon_{t+h}$	$X_t = \rho X_{t-1} + u_t$	$(0.1, 0.3, 0.5, 0.7, 0.9)$

Under $A.0(h)$, X_t has no predictive power for Y_{t+h}. This allows us to examine the size of the nonparametric test under $H_1 : E(Y_{t+h}|X_t) = E(Y_{t+h})$. Under $A.1(h)$ and $A.2(h)$, there exist linear and nonlinear (quadratic form) predictability of X_t for Y_{t+h}.This allows us to examine the power of the test under the alternatives. Next, under $B.0(h)$, there is no neglected nonlinear predictability of X_t for Y_{t+h}. This allows us to examine the size of the test for the null hypothesis $H_2 : E(\varepsilon_{t+h}|X_t) = 0$. Under $B.1(h)$, there exists neglected nonlinear predictability, and this allows us to examine the power of the test.

Tables 1a, 1b, 1c, and 1d report empirical rejection rates of the test at the 1%, 5%, and 10% nominal levels, for sample sizes of $T = 250, 500, 1000$ and horizons of $h = 1, 4, 12, 20$. The nonparametric test with the bootstrap procedure has quite reasonable sizes in finite samples under both the null hypotheses H_1 and H_2, which are robust to the length h of horizon and to the persistence of regressor X_t (as measured by the large value of the autoregressive coefficient ρ). Moreover, the proposed test has power under various alternatives to H_1 and H_2 respectively.

The reasonable and robust size and power performance of the proposed test is quite encouraging in view of the fact that due to both long-horizon returns and persistence of the regressors, there is an upward bias in the predictive coefficient on the regressors (Stambaugh 1999, Amihud and Hurvich 2004, Lewellen 2004), and existing long-horizon tests with robust Newey-West standard errors suffer from substantial overrejection.[12]

3 Data and Long-Horizon Predictability Regression

3.1 The long-horizon framework and Predictability Regression

Denote the gross return on equity by $G_{t+1} = (P_{t+1}+D_{t+1})/P_{t+1}$ and the continuously compounded return by $\widetilde{y}_{t+1} = log(G_{t+1})$. The long-horizon predictability regression considered is

$$Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h} \tag{4.1}$$

where $Y_{t+h} = (\tau/h)[(\widetilde{y}_{t+1} - r_t) + \cdots + (\ \widetilde{y}_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $\widetilde{y}_{t+1} - r_t$ is the one period excess return from time t to $t+1$. The constant τ is different, depending on the frequency of the data, i.e., $\tau = 1$ (annually), $\tau = 4$ (quarterly), and $\tau = 12$ (monthly). All returns are continuously compounded.

The error term $\varepsilon_{h,t+h}$ follows a $MA(h-1)$ process under the null hypothesis of no predictability $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$ and $H_0 : E(\varepsilon_{t+h}|X_t) = 0$. We will use different predictors as instruments in X_t, which are explained in details below. We estimate the regression (4.1) by OLS and compute standard errors of the parameters using the Newey and West(1987) and Hodrick (1992)

[12] Ang and Bekaert (2007) point out that the univariate dividend yield regression displays negligible size distortions in the shortest sample for the one-quarter horizon, but for the bivariate regressions, all tests slightly over-reject at asymptotic critical values with longer horizons.

standard error formula[13]. We use the test proposed in section 3 to check the predictability of different variables using the regression framework in (4.1).

3.2 Data

We will examine predictability for the equity returns by using data with different frequencies: annual, quarterly, and monthly. The most prominent X_t variables considered in the literature are dividend price ratio and dividend yield, earnings yield and dividend-earnings (payout) ratio, various interest rates and spreads, inflation rates, book-to-market ratio, investment-capital ratio, consumption, wealth, and income ratio (CAY), and aggregate net or equity issuing activity.

Stock Returns: Stock returns used are continuously compounded returns on the S&P 500 index, including dividends. Our quarterly data consist of price return (capital gain only), total returns (capital gain plus dividend), and dividends on the Standard & Poor's Composite Index from March 1936 to December 2001. This data is obtained from the Security Price Index Record, published by Standard & Poor's Statistical Service. For monthly data, we use S&P 500 index returns from January 1970 to December 2006 from CRSP's monthend values. Monthly dividends on the S&P 500 index are from Standard & Poor's Statistical Service. For yearly frequency, we get data from 1872 to 2005 provided in Robert Shiller's personal website.

Risk-free Rate: The risk-free rate from 1920 to 2005 is the T-bill rate. We follow the methods by Goyal and Welch (2007) to estimate T-bill rate prior to the 1920's.[14] For quarterly and monthly data, T-bill rates from 1934 to 2005 are the 3-Month Treasury Bill: the Secondary Market Rate from the economic research data base at the Federal Reserve Bank at St. Louis (FRED).

Dividend Yields, Earnings Yields, and Dividend Payout Ratio: Dividends and Earnings are the twelve-month moving sums of dividends and earnings paid on the S&P 500 index. The data from 1871 to 1970 are available from Robert Shiller's website. Quarterly dividends and earnings from 1936 to 2005 and monthly dividends and earnings from 1970 to 2006 are from the S&P Corporation. Dividends and Earnings are summed up over the past year. Monthly or quarterly frequency dividends and earnings are impossible to use because they are dominated by seasonal components. The dividend yield (d/y) is defined as $D_t^4/\ P_t$ with the superscript 4 to denote that it is constructed using dividends summed up over the past year (four quarters), where $D_t^4 = D_t + D_{t+1} + D_{t+2} + D_{t+3}$ represents dividends summed over the past year and P_t is the price level on S&P 500.[15] We also define the monthly dividend yield with a superscript of 12 to indicate that

[13] Using generalized method of moments, (GMM) has an asymptotic distribution $\sqrt{T}(\widehat{\theta} - \theta) \overset{a}{\sim} N(0, \Omega)$ where $\Omega = Z_0^{-1} S_0 Z_0^{-1}$, $Z_0 = E(x_t' x_t)$, and $x_t = (1\ z_t')'$. Hodrick(1992) sums $x_t' x_{t-j}$ into the past and estimates S_0 by $\widehat{S}_0 = \frac{1}{T} \sum_{t=h}^{T} wh_t wh_t'$, $wh_t = \varepsilon_{1,t+1} \sum_{i=0}^{h-1} x_{t-i}$.

[14] Commercial paper rates for New York City are from the NBER's Macrohistory data base. These are available from 1871 to 1970. We estimated a regression from 1920 to 1971, which yielded $T - billRate = -0.004 + 0.886 * CommercialPaperRate$, with an R^2 of 95.7% according to Goyal and Welch (2007). Therefore, we instrumented the risk-free rate from 1871 to 1919 with the predicted regression equation. The correlation for the period 1920 to 1971 between the equity premium computed using the actual T-bill rate and that computed using the predicted T-bill rate (using the commercial paper rate) is 99.8%.

[15] See, e.g., Ball (1978), Campbell (1987), Campbell and Shiller (1988a, 1988b), Campbell and Viceira (2002), Campbell and Yogo (2006), the survey in Cochrane (1997), Fama and French (1988), Hodrick (1992), Lewellen (2004), Menzly, Santos, and Veronesi (2004), and Ang and Bekaert(2007).

dividends have been summed over the past 12 months using the same method. We also denote log dividend yields as $dy_t^4 = \log(D_t^4/\ P_t)$ for quarterly data and $dy_t^{12} = \log(D_t^{12}/\ P_t)$ for monthly data. We use the similar definitions for log earnings yields for both quarterly and monthly. The Dividend Payout Ratio (d/e) is the difference between the log of dividends and the log of earnings.

Stock Variance (svar): Stock Variance is computed as sum of squared daily returns on the S&P 500. G. William Schwert provided daily returns from 1871 to 1926; data from 1926 to 2005 are from CRSP.

Book to Market Ratio: The Book to Market Ratio (b/m) is the ratio of book value to market value for the Dow Jones Industrial Average.[16] Book values from 1920 to 2005 are from Value Line's website, specifically their Long-Term Perspective Chart of the Dow Jones Industrial Average.

Corporate Issuing Activity: We follow the two measures of corporate issuing activity in Goyal and Welch (2007). Net Equity Expansion ($ntis$) is the ratio of twelve-month moving sums of net issues by S&P listed stocks divided by the total end-of-year market capitalization of S&P stocks. This dollar amount of net equity issuing activity (IPOs, SEOs, stock repurchases, less dividends) for NYSE listed stocks is computed from the CRSP data as $NetIssue_t = Mcap_t - Mcap_{t-1} \cdot (1 + vwretx_t)$, where $Mcap$ is the total market capitalization, and $vwretx$ is the value weighted return (excluding dividends) on the S&P 500 index. These data are available from 1926 to 2005. The second measure, Percent Equity Issuing ($eqis$), is the ratio of equity issuing activity as a fraction of total issuing activity. This is the variable proposed in Baker and Wurgler (2000).[17] The first equity issuing measure is relative to the aggregate market cap, while the second is relative to the aggregate corporate issuing.

Long Term Yield (lty): The data is from Goyal and Welch (2008). The long-term government bond yield data from 1919 to 1925 is the U.S. Yield On Long-Term United States Bonds series in the NBER's Macrohistory data base. Yields from 1926 to 2005 are from Ibbotson's Stocks, Bonds, Bills and Inflation Yearbook, the same source that provided the *Long Term Rate of Returns (ltr). The Term Spread (tms)* is the difference between the long term yield on government bonds and the T-bill. (See, e.g., Campbell (1987) and Fama and French (1989).)

Corporate Bond Returns: Long-term corporate bond returns from 1926 to 2005 are again from Ibbotson's Stocks, Bonds, Bills and Inflation Yearbook. Corporate Bond Yields on AAA and BAA-rated bonds from 1919 to 2005 are from FRED. The Default Yield Spread (dfy) is the difference between BAA and AAA-rated corporate bond yields. The Default Return Spread (dfr) is the difference between long-term corporate bond and long-term government bond returns. (See, e.g., Fama and French (1989) and Keim and Stambaugh (1986).)

Inflation ($infl$): Inflation is the Consumer Price Index (All Urban Consumers) from 1919 to 2005 from the Bureau of Labor Statistics.

Investment to Capital Ratio (i/k): The investment to capital ratio is the ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy.

Consumption, wealth, income ratio (cay): The variable cay is proposed by Lettau and Ludvigson

[16] See Kothari and Shanken (1997) and Ponti and Schall (1998).

[17] We get the data from http://pages.stern.nyu.edu/ jwurgler/

(2001)[18]. Data for *cay*'s construction at quarterly frequency from the second quarter of 1952 to the fourth quarter of 2005 are available from Martin Lettau's website. The annual data from 1948 to 2001 is available from Martin Lettau's website.

Table 2 summarizes the descriptive statistics of the predictors. Panels (a), (b), and (c) report the results for quarterly, monthly, and annual data respectively. Short rates, dividend and earnings yields, book-to-market ratio, and inflation are all highly persistent at different frequencies. Because the persistence of these instruments plays a crucial role in the finite sample performance of predictability test statistics, we report test statistics under the null of a unit root. Figure 1 plot excess returns, interest rate, dividend yields, and earnings yields from March 1936 to December 2001 quarterly. For annual data, Figure 2 and 3 plot dividend payout ratio, short rate, inflation, book-to-market ratio, investment to capital ratio(i/k), corporate issuing activity ($eqis$ and $ntis$), and consumption, wealth, and income ratio(cay) from 1872 to 2005 annually.

4 Is the Predictability There?

In this section, we first apply the nonparametric test to examine whether there exists the predictability of equity returns for both short and long horizons. Next, we will use the test to compare the conventional predictive regression models on predictor variables with the historical mean model according to Goyal and Welch (2007) and Campbell and Thompson(2007). Finally we provide a simulation study on the size and power of the proposed test to assess the reliability of the proposed test in finite samples.

4.1 Short-Horizon and Long-Horizon Predictability

The main regressions we consider are $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ in (4.1). We use quarterly, monthly, and annual data to check the predictability of equity returns. For quarterly and monthly data, we report results for four sample periods, from 1936 to 2001, from 1952 to 2001, from 1936 to 1990, and from 1952 to 1990, which are the same sample periods considered in Ang and Bekaert (2007).[19]

Table 3 summarizes the results on the excess return predictability for horizons of 1 quarter, 1 year, 3 years, and 5 years respectively. Table 3a focuses on the univariate regression with log dividend yields or log earnings yields as the regressor. The t-statistics in parentheses are computed using the Newey and West (1987) and Hodrick (1992) standard error formula respectively. The parameter estimates have similar patterns over the 4 periods, but the coefficient estimates are twice

[18] Lettau and Ludvigson (2001) estimate the following equation: $c_t = \alpha + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i} \Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i} \Delta y_{t-i} + \epsilon_t$, $t = k+1, ..., T-k$, where c is the aggregate consumption, a is the aggregate wealth, and y is the aggregate income. Using estimated coefficients from the above equation provides $cay = \widehat{cay}t = c_t - \alpha - \widehat{\beta}_a a_t - \widehat{\beta}_y y_t$, $t = 1, ..., T$.

[19] Interest rate data are hard to interpret before the 1951 Treasury Accord, as the Federal Reserve pegged interest rates during the 1930s and the 1940s. Hence, we examine the post-Accord period, starting in 1952. Second, the majority of studies establishing strong evidence of predictability use data before or up to the early 1990s. Studies by Lettau and Ludvigson (2001) and Goyal and Welch (2003) point out that predictability by the dividend yield is not robust to the addition of the 1990s decade. Hence, we separately consider the effect of adding the 1990s to the sample.

as large for the period omitting the 1990s from the sample. The Hodrick standard errors are smaller than the Newey-West standard errors. During the 1936-2001 and 1952-2001 periods, there is no evidence of predictability for dividend yields for both short and long horizons. For the 1936-1990 periods, there is strong predictability for dividend yields over the horizons of 1 quarter, 1 year, 3 years, and 5 years respectively. Yet for the 1952-1990 period, there exists only the predictability for short horizons of 1 quarter and 1 year.

Table 3a reports the bootstrap p-value for the predictability test under two hypotheses $H1$ and $H2$. Hypothesis $H1$ is $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$, namely that X_t has no predictive power for Y_{t+h}, and Hypothesis $H2$ is $H_0 : E(\varepsilon_{t+h}|X_t) = 0$, namely that X_t has no neglected nonlinear predictive power for Y_{t+h} beyond the linear model (4.1). As mentioned in Section 3, the $\hat{\mathbf{Q}}_h$ test has an appealing interpretation: it is essentially based on the ratio $\hat{Q}(h)/\widehat{\sigma}^2_\varepsilon$, where the denominator $\widehat{\sigma}^2_\varepsilon$ is the sample variance of pricing errors, and the numerator $\hat{Q}(h)$ is the average of the squared predictable components neglected by the linear regression model. Therefore, the ratio $\hat{Q}(h)/\widehat{\sigma}^2_\varepsilon$ can be viewed as an estimator for the neglected signal-to-noise ratio of the linear prediction model (4.1). If the neglected pricing signal $\hat{Q}(h)$ is weak relative to the pricing noise $\widehat{\sigma}^2_\varepsilon$, the $\hat{\mathbf{Q}}_h$ test will not reject the null hypothesis H_0. If the neglected pricing signal $\hat{Q}(h)$ is strong relative to the pricing noise $\widehat{\sigma}^2_\varepsilon$, the $\hat{\mathbf{Q}}_h$ test will reject the null hypothesis H_0. The results for testing $H1$ show that the $\hat{\mathbf{Q}}_h$ test strongly rejects the null hypothesis $H1$ for dividend yields over the 4 sample periods. This implies that dividend yield is a significant predictor of excess returns at all horizons, which is consistent with the prevailing result found by Campbell and Shiller (1988a,b).

Next, we examine whether there exists neglected nonlinear predictability of dividend yield. Table 3a show that the $\hat{\mathbf{Q}}_h$ test strongly rejects the null hypothesis $H2$ for all 4 sample periods. Thus, there exists a nonlinear predictive relationship between X_t and Y_{t+h}, and a suitable nonlinear predictive model is expected to outperform the linear regression model.

The right four columns of Table 3a also report a univariate regression with the earnings yield as the regressor. The results suggest that there is no strong evidence for linear predictability of earnings yields over 4 sample periods. The nonparametric tests for hypothesis $H1$ and $H2$, however, show that earnings yield is a good predictor for equity returns over all the horizons.

Table 3b summarizes the bivariate regression with log dividend yields and short rate together as regressors. The t-statistics in parentheses are computed using the Newey and West (1987) and Hodrick (1992) standard error formula. Horizons h are quarterly. Table 3b also reports the bootstrap p-value for the predictability test for six various hypotheses $H1 - H6$, where X_1 represents the short rate r and X_2 the dividend yield. The six Hypotheses are, respectively, $H1$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), $H2$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0$), $H3$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), $H4$ ($H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0$), $H5$ ($H_0 : E(Y_{t+h}|X_{1t}, X_{2t}) = E(Y_{t+h})$), and $H6$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}) = 0$). Hypotheses $H1$ and $H2$ are on predictability of the short rate or dividend yield separately and Hypothesis $H5$ is on the joint predictability of the short rate and dividend yield together. The results based the Newey-West standard errors suggest that the short rate has strong predictability over the 4 periods but the predictability only exists at short horizons when using the Hodrick (1992) standard errors. In the bivariate regression, there is evidence of predictability of dividend yields for equity returns when the sample period excludes the 1990s. The coefficient on the dividend yield is larger in the bivariate regression than in the univariate regression. This suggests that the univariate regression suffers from an omitted variable bias that lowers the marginal impact of dividend yields on expected excess returns.[20] The $\hat{\mathbf{Q}}_h$ test significantly rejects the hypotheses $H1 - H4$ for all 4 sample periods, indicating that that the short rate and dividend yield are two good predictors for equity returns which cannot be fully captured by linear prediction regressions.

The $\hat{\mathbf{Q}}_h$ test rejects the hypothesis $H5$ only at the horizon of 1 quarter and fails to reject at the horizons of 1 year, 3 year2, and 5 years for the 1936-2001 period. The $\hat{\mathbf{Q}}_h$ test rejects hypothesis $H5$ for the 1952-2001, 1936-1990, and 1952-1990 periods. These results suggest that there is evidence of joint predictability for the short rate and dividend yield together for the 3 sample periods (1952-2001, 1936-1990, and 1952-1990) and the short rate and dividend yield together have the predictability only at the short horizon of 1 quarter in the 1936-2001 period. The $\hat{\mathbf{Q}}_h$ test rejects hypothesis $H6$ for the 4 time periods with exception for the horizon of 5 years in the 1936-2001 and 1952-1990 periods. The bivariate linear regression does not have the optimal predictive power for equity returns over all the horizons in the 4 time periods. Nevertheless, it may have the long-horizon predictive power for the horizon of 5 years in the 1936-2001 and 1952-1990 periods since there is no strong evidence to reject hypothesis $H6$. Ang and Bekaert (2007) examine the predictive power of dividend yields for forecasting excess returns. They find that dividend yields predict excess returns only at short horizons together with the short rate and do not have any long-horizon predictive power. At short horizons, the short rate strongly negatively predicts returns.

To compare with Lamont (1998) and Ang and Bekaert (2007), we report a bivariate regression of excess returns on log dividend and log earnings yields. Lamont (1998) finds a positive coefficient on the dividend yield and a negative coefficient on the earnings yield. He argues that the predictive power of the dividend yield stems from the role of dividends in capturing permanent components of prices, whereas the negative coefficient on the earnings yield is due to earnings being a good measure of business conditions. Ang and Bekaert (2007) finds that dividend and earnings yields do not have a strong predictive power and only when the 1990s are excluded they find significant coefficients for dividend and earnings yields. Table 3c summarizes the bivariate regression with the log dividend yields and log earnings yields together as regressors. The dividend yields and earnings yields have a strong predictive power for equity returns over the 4 time periods when using the Newey-West (1987) standard errors. The results using the Hodrick (1992) standard errors are similar to Ang and Bekaert (2007). The $\hat{\mathbf{Q}}_h$ test rejects the six hypotheses over all the time horizons and for all 4 time periods. It supports Lamont (1998)'s arguments. Dividend yields and earnings yields have the predictability power for equity returns but the bivariate linear regression model cannot fully capture such predictability.

Table 3d summarize the trivariate regression with the short rate, log dividend yields, and log earnings yields together as regressors. When we add the short rate as a predictor in a trivariate regression of excess returns on risk-free rates, dividend and earnings yields, the coefficients on

[20]Engstrom (2003), Menzly, Santos, and Veronesi (2004), and Lettau and Ludvigson (2005) also note that a univariate dividend yield regression may understate the dividend yield's ability to forecast returns.

dividend and earnings yields remain insignificantly different from zero, and the sign on the earnings yield is fragile. For the post-1952 samples, the short rate, and dividend yields have predictive power in the presence of the earnings yield. The results for the $\hat{\mathbf{Q}}_h$ test show that the three variables short rate, dividend yields, and earnings yields do have the predictability power for the equity returns. The $\hat{\mathbf{Q}}_h$ test for the joint predictability of the three variables rejects the hypothesis $H7$ for most of the cases except the horizons of 1 year and 3 years in the 1936-2001 and 1936-1990 periods and the horizon of 5 year in the 1952-2001 period. And the trivariate regression does not capture the true equity returns and it needs a better nonlinear model to capture it.

We use the monthly data from January 1970 to December 2006 to test the predictability of the short rate, dividend yields, and earnings yields in univariate, bivariate, and trivariate regressions respectively. Table 4 summarizes the results for the regressions and predictability tests. We obtain similar results to those based on quarterly data. Using the Hodrick (1992) standard errors, our results suggest that the short rate has strong predictability. The nonparametric predictability tests show that the three variables are good candidates to predict equity returns but linear predive regression models cannot fully capture such predictability.

4.2 Does the prevailing models beat the historical mean?

Goyal and Welch (2007) reexamine the performance of variables that have been suggested by the academic literature to be good predictors of equity premiums. They find that those models have predicted poorly both in-sample and out-of-sample for thirty years and can not beat the historical mean model. We consider both In-Sample (IS) and Out-of-Sample (OOS) tests. Following Goyal and Welch (2007), the OOS forecasts use only the data available up to the time at which the forecast is made. Let e_N denote the vector of rolling OOS errors from the historical mean model and e_A denote the vector of rolling OOS errors from the OLS model. The OOS statistics are computed as $R^2 = 1 - \frac{MSE_A}{MSE_N}$, $\overline{R}^2 = R^2 - (1 - R^2) \cdot (\frac{T-k}{T-1})$, $\Delta RMSE = \sqrt{MSE_N} - \sqrt{MSE_A}$. It is important but difficult for OOS tests to choose the periods over which a regression model is estimated and subsequently evaluated. In this section we consider the annual prediction with similar data used in Goyal and Welch (2007). For the OOS test, we explore the time period specification which begins OOS forecasts twenty years after data are available.

We estimate regressions of form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ in (4.1), with X_t being log dividend yields, log earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio(b/m), investment to capital ratio(i/k), corporate issuing activity (Eqis and Ntis), and consumption, wealth, and income ratio(cay). The results are summarized in Table 5. The t-statistics in parentheses are computed using the Newey and West (1987) and Hodrick (1992) standard errors. We report the bootstrap p-value for the predictability test under two hypotheses $H1$ and $H2$. Hypothesis $H1$ is $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$ and Hypothesis $H2$ is $H_0 : E(\varepsilon_{t+h}|X_t) = 0$.Table 5 summarizes both in-sample and out-of-sample results. To compare the prevailing predictive models with the historical mean model, we introduce a criterion $\Delta(\frac{Q_h}{\sigma^2}) = \hat{Q}_N(h)/\widehat{\sigma}_\varepsilon^2 - \hat{Q}_A(h)/\widehat{\sigma}_\varepsilon^2$, where $\hat{Q}_N(h)/\widehat{\sigma}_\varepsilon^2$ and $\hat{Q}_A(h)/\widehat{\sigma}_\varepsilon^2$ are the signa-to-noise ratios of the historical mean model and the prevailing predictive regression model respectively. If $\Delta(\frac{Q_h}{\sigma^2}) > 0$, there is more neglected signal which cannot be explained by the historical mean model and thus the prevail predictive model performs better. If $\Delta(\frac{Q_h}{\sigma^2}) < 0$, there is more neglected signal which cannot be captured by the prevail predictive model and so the historical mean model performs better.

Table 5 shows that when a linear prediction model is used, all variables considered are insignificant and only several variables (dividend yield, short rate, eqis, and cay) are significant at the horizon of 1 year using the Newey-West standard errors. However, the results for both in-sample and out-of-sample nonparametric tests show that all variables considered (i.e., log dividend yields, log earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio(b/m), investment to capital ratio(i/k), corporate issuing activity (Eqis and Ntis), and consumption, wealth, and income ratio(cay)) have predictability power for equity returns. For all the cases considered, $\Delta(\frac{Q_h}{\sigma^2})$ is larger than zero for both in-sample and out-of-sample There exists more neglected signal which cannot be explained by the historical mean model and thus the prevail predictive model performs better. This conclusion differs from Goyal and Welch (2007) and supports Campbell and Thompson (2007).

5 Out-of-Sample Forecasting of Equity Returns

As mentioned in the previous sections, Goyal and Welch (2008) create enough of a controversy within the profession and argue that the historical average excess stock return forecasts future excess stock returns better than regressions of excess returns on predictor variables. Campbell and Thompson (2008) and Cochrane (2008) soon follow with opposing views. Campbell and Thompson argue that the empirical models can yield useful out-of-sample forecasts if one restricts their parameters in economically justified ways. In contrast, Cochrane (2008) argues that the types of out-of-sample tests performed by Goyal and Welch are relatively weak, and that in-sample tests provide far greater power and can be convincing on their own. The literature emphasizes that the most linear predictive regressions have often performed poorly out-of-sample (Goyal and Welch (2003, 2007); Campbell and Thompson (2007)). The lack of consistent out-of-sample evidence in Goyal and Welch (2008) indicates the need for improved forecasting methods to better establish the empirical reliability of equity premium predictability. Rapach, Strauss, and Zhou (2009) propose a combination approach to improve the out-of-sample equity premium forecasting problem. In this section, we propose a nonparametric estimator to forecast the equity returns.

5.1 Nonparametric forecast, linear predictive model, and Historical Mean Model

In the previous sections, our nonparametric test has proved that there exists the predictability of equity returns at short or long horizons. The predictors such as dividend yields, earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio, investment to capital ratio, corporate issuing activity, and consumption, wealth, and income ratio have predictability power for equity returns, but this often cannot be captured by popular linear regression models. We find that the poor out-of-sample performance of most linear prediction models is due to the limitation of linear models. We need to find the better fit of the equity returns.

Following the section 2, we use two nonlinear estimators to forecast the equity returns. The first estimator is to use a smoothed kernel method to estimate $E(\varepsilon_{t+h}|X_t)$ and capture potentially neglected nonlinear predictable component in ε_{t+h}. So the expected equity returns can be defined as The regression estimator for $E(\varepsilon_{t+h}|X_t)$ is then defined as follows:

$$\begin{aligned}
E(Y_{t+h}|X_t) &= X_t'\widehat{\beta} + E(\varepsilon_{t+h}|X_t) = X_t'\widehat{\beta} + \widehat{r}_h(x). \\
&= X_t'\widehat{\beta} + \frac{\widehat{m}_h(x)}{\widehat{g}(x)} \\
\widehat{m}_h(x) &= \frac{1}{n-h}\sum_{s=R+1}^{T-h} \hat{\varepsilon}_{s+h} K_b(x - X_s), \\
\widehat{g}(x) &= \frac{1}{n-h}\sum_{s=R+1}^{T-h} K_b(x - X_s)
\end{aligned}$$

where $x = (x_1, x_2, \cdots, x_d)'$, $y = (y_1, y_2, \cdots, y_d)'$, and $K_b(x-y) = \prod_{i=1}^{d} b^{-1}K[(x_i - y_i)/b]$. The kernel function $K(\cdot)$ is is a prespecified symmetric probability density function. The second estimator is to use a smoothed kernel method to estimate $E(Y_{t+h}|X_t)$ directly. We can predict the equity returns by

$$E(Y_{t+h}|X_t) = \frac{\frac{1}{n-h}\sum_{s=R+1}^{T-h} Y_{s+h} K_b(x - X_s)}{\widehat{g}(x)}$$

We want to compare the out-of-sample forecast results of four models: historical mean model, linear predictive model, and two nonlinear predictive models. The three measures we use are MSE (Mean squared error), MAE (Mean absolute error), and RMSE (Root mean squared error). The smaller the RMSE is and the model has a better fit.

$$\begin{aligned}
MSE &= \frac{1}{n-h}\sum_{s=R+1}^{T-h} (Y_{s+h} - \widehat{Y}_{s+h})^2 \\
MAE &= \frac{1}{n-h}\sum_{s=R+1}^{T-h} \left| Y_{s+h} - \widehat{Y}_{s+h} \right| \\
RMSE &= \sqrt{\frac{1}{n-h}\sum_{s=R+1}^{T-h} (Y_{s+h} - \widehat{Y}_{s+h})^2}
\end{aligned}$$

Table 6 show the out-of-sample results of the univariate linear predictive models. Table 6a summarize the MSE, MAE, and RMSE of linear predictive regression for dividend yield and earning yield during the period 1936-2001, and 1952-2001. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. We find that our second nonparametric predictive model has the lower RMSE than the historical mean model. The linear predictive model and the first nonparametric predictive model have higher RMSE than the historical mean model. The second nonparametric predicitve model can do better job than historical mean model in both short horizon and long horizon. Table 6b summarize the MSE, MAE, and RMSE of linear predictive regression for dividend yield and earning yield during the period 1936-1990, and 1952-1990. We find that both first and second nonparametric predictive models have the lower RMSE than the historical mean model. The linear predictive model has higher RMSE than the other three models. The two nonparametric predicitve models can do better job than historical mean model in both short horizon and long horizon.

Table 7 reports out-of-sample bivariate regression results with the short rate as an additional regressor. For the period of 1936-2001, 1936-1990, and 1952-1990, the second nonparametric predictive regression model has the smallest RMSE. For the post-Treasury Accord 1952–2001 sample, the linear predictive model and the first nonparametric predictive model have higher RMSE than the historical mean model. The second nonparametric predicitve model can do better job than historical mean model in both short horizon and long horizon. In the bivariate regression with earning yield and short rate, the second nonparametric predictive regression model is superior to the other three models during the period 1936-2001 and 1952-2001. Table 7b summarize the statistical results of bivariate linear predictive regression for dividend yield and earning yield by using the measure MSE, MAE, and RMSE during the period 1936-1990, and1952-1990. The linear predictive model has higher RMSE than the other three models. The two nonparametric predicitve models can do better job than historical mean model in both short horizon and long horizon. Ang and Bekaert (2007) find that dividend yields, together with the short rate, predict excess returns only at short horizons. In this section, we find that the nonparametric predictive model can capture the equity returns well and the short rate, dividend yields, and earnings yields have good predictability power at both short and long horizons.

Goyal and Welch (2007) argue that the historical average excess stock return forecasts future excess stock returns better than regressions of excess returns on predictor variables. With respect to the economic variables used to predict the equity premium, we consider the 15 variables from Goyal and Welch (2008) for which quarterly data are available for 1947:1–2007:4 and annual data are from 1872 to 2005. They are dividend-price ratio (D/P), dividend yield (D/Y), earnings-price ratio (E/P), dividend-payout ratio (D/E), stock variance ($SVAR$), book-to-market ratio (B/M), net equity expansion ($NTIS$), treasure bill rate (TBL), long-term yield (LTY), long-term return (LTR), term spread (TMS), default yield spread (DFY), default return yield (DFR), inflation ($INFL$), and investment-to-capital ratio (I/K). Common to all these papers is a focus on a small set of predictors based on theoretical models. From an academic viewpoint, the use of model-based predictors facilitates an understanding of specific aspects of the economic mechanism. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. Table 8 report the equity premium out-of-sample forecasting results using the annual data. Consistent with the previous results, the second non-parametric predicitve model can do better job than historical mean model and linear predictive model in both short horizon and long horizon. Table 10 report the equity premium out-of-sample forecasting results using the quarterly data from 1947:1–2007:4. We consider the out-of-sample forecast evaluation periods covering 1965:1–2007:4 consistent with Goyal and Welch (2008). The statistical results show that the second nonparametric predicitve model can do better job than historical mean model and linear predictive model in both short horizon and long horizon. For most predictors except dividend-price ratio (D/P), dividend yield (D/Y), earnings-price ratio (E/P), and book-to-market ratio (B/M), the two nonparametric models are superior to the historical mean model and linear regression model.

Figure 4 and 6 illustrate the out-of-sample performance for annual predictive regressions for individual methods. The black dotted line is the real data of the equity returns and the red dotted line is the unconditional historical average. The red and green solid line are the forecasted returns by the first and second nonparametric models respectively. A predictive regression model that always outperforms the historical average for any out-of-sample period will thus have a curve below the historical average curve. For individual predictor-based models, the second nonparametric prediction is mostly below the unconditional historical average line. Even for some periods, the nonparametric method is above the historical average yet on average it outperforms the historical average. Campbell and Thompson (2008) show that imposing theoretically motivated restrictions on individual predictive regression models can improve their out-of-sample performance. We find that our nonparametric prediction can improve the out-of-sample performance without restrictions. Figure 8, 10 and 12 illustrate the out-of-sample performance for quarterly predictive regressions for individual methods over 1-quarter, 1-year, and 3-year rolling windows. We find the similar results for the quarterly data.

5.2 Individual Forecast and Combined Forecast

In the literature, most papers focus on a set of predictors based on theoretical models. From an academic viewpoint, the use of model-based predictors facilitates an understanding of specific aspects of the economic mechanism. From an investor's viewpoint, however, these predetermined variables may not be enough to capture all information required in decision making. Forecast combination has recently received renewed attention in the forecasting literature; Stock and Watson (1999, 2003, 2004) with respect to forecasting inflation and real output growth. Rapach, Strauss, and Zhou (2009) propose a combination approach to improve the out-of-sample equity premium forecasting problem. In addition to the individual forecast, we also consider the combined forecast to improve equity premium forecasts, and examine the out-of-sample performance.

We follow the definition of the combined forecast by Rapach, Strauss, and Zhou (2009). The combination forecasts of Y_{t+1} made at time t are weighted averages of the M individual forecasts based on $\widehat{Y}_{c,t+h} = \sum_{i=1}^{M} \omega_{i,t} \widehat{Y}_{i,t+h}$ where $\{\omega_{i,t}\}_{i=1}^{M}$ are the ex ante combining weights formed at time t, and $\widehat{Y}_{i,t+h}$ is the out-of-sample forecast of the equity premium based on the individual predictive models[21]. For the individual predictors, we choose the 15 predictors used in the previous sections. We calculate five different combining methods based on the definition of the weights. The first three methods use simple averaging schemes: mean, median, and trimmed mean. The mean combination forecast sets $w_{i,t} = 1/M$ for $i = 1, \cdots, M$. The median combination forecast is the median of $\{\widehat{Y}_{i,t+h}\}_{i=1}^{M}$, and the trimmed mean combination forecast sets $w_{i,t} = 0$ for the individual forecasts with the smallest and largest values and $w_{i,t} = 1/(M-2)$ for the remaining individual forecasts. The other two combining methods are based on Stock and Watson (2004) and Rapach, Strauss, and Zhou (2009), where the combining weights formed at time t are functions of the historical forecasting performance of the individual models over the holdout out-of-sample period. Their discount mean square prediction error ($DMSPE$) combining method employs the following weights: $w_{i,t} = \phi_{i,t}^{-1}/\sum_{j=1}^{M}\phi_{j,t}^{-1}$, $\phi_{i,t} = \sum_{s=R}^{t-1}\theta^{t-1-s}(Y_{i,t+h} - \widehat{Y}_{i,t+h})^2$ and θ is a discount factor. The DMSPE method thus assigns greater weights to individual predictive regression model forecasts that have lower MSPE values (better forecasting performance) over the holdout out-of-sample period. We consider the two values of 1.0 and 0.9 for θ.

Table 9 report the equity premium out-of-sample combined forecasting results using the annual data. Consistent with the previous results, the two nonparametric predicitve models have lower $RMSE$ and can do better job than historical mean model and linear predictive model in both short horizon and long horizon. In addition, using combined method linear predictive model can outperform the historical mean model. Table 11 report the equity premium out-of-sample combined forecasting results using the quarterly data from 1947:1–2007:4. We consider the out-of-sample forecast evaluation periods covering 1965:1–2007:4 consistent with Goyal and Welch (2008). The statistical results show that the two nonparametric predicitve models can do better job than historical mean model and linear predictive model in both short horizon and long horizon. Rapach, Strauss, and Zhou (2009) find that forecast combination outperforms the historical mean model by statistically and economically meaningful margins for out-of-sample period. Our results are consistent with their conclusion. Using our nonparametric methods, both combined and individual forecast outperform the historical average. The combined forecast methods outperform the individual forecast methods.

Figure 5 and 7 illustrate the out-of-sample performance for annual predictive regressions for combined methods. The black dotted line is the real data of the equity returns and the red dotted line is the unconditional historical average. The red and green solid line are the forecasted returns by the first and second nonparametric models respectively. For combined predictor-based models, the two nonparametric prediction models are below the unconditional historical average line. We find that our nonparametric prediction can improve the out-of-sample performance without restrictions. Figure 9, 11 and 13 illustrate the out-of-sample performance for quarterly predictive regressions for combined methods over 1-quarter, 1-year, and 3-year rolling windows. We find the similar results for the quarterly data.

[21] $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$

5.3 Economic Implication

From the previous two sections, we get two important results: (1) our nonparametric predictive models have lower RMSE than the historical mean model at both short-horizon and long-horizon. Our nonparametric prediction can improve the out-of-sample performance without restrictions. (2) Using our nonparametric methods, both combined and individual forecast outperform the historical average. The combined forecast methods outperform the individual forecast methods. Why do our nonparametric predictive models do better to predict the equity returns?

We find the predictors have the predictability of the equity returns using our nonparametric test and linear predictive regression can not capture the nonlinear component of the true data. We use nonparametric model to predict the equity returns because it can capture the linear and nonlinear component without the model specification. It is not restricted to the parametric forms. It can fit the data more better than simply the linear or nonlinear parametric model.

Fama and French (1989) and others show that these variables can detect changes in economic conditions that potentially signal fluctuations in the equity risk premium. But the dividend yield or term spread alone could capture different components of business conditions, and a given individual economic variable may give a number of "false signals" and/or imply an implausible equity risk premium during certain periods. Rapach, Strauss, and Zhou (2009) argue that if individual forecasts based on the predictors are weakly correlated, forecast combinatio should be less volatile and more reliably track movements in the equity risk premium. This is one explanation why the combined forecast methods outperform the individual forecast methods.

On the other hand, the nonparametric predictive model can fit the equity return better based on the predictors. First, n onparametric prediction generates a forecast with a variance near that of the smooth real equity return data, thereby reducing the noise in the individual predictive regression model forecasts. Second, combining forecast incorporates information from a host of economic variables while the historical average ignores economic variables. Combined forecasts have a substantially smaller bias than the historical average. Combining individual forecasts helps to reduce forecast variability.

6 Conclusion

The predictability of equity returns has been a long-standing problem in finance over decades. In this paper, we undertake an analysis of both in-sample and out-of-sample tests of stock return predictability in an effort to better understand the empirical evidence on return predictability. We use develop a reliable and powerful nonparametric predictability test and use it to examine whether there exists the predictability of equity returns for short and long horizons. We find that the prevailing variables, such as log dividend yields, log earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio, investment to capital ratio, corporate issuing activity, and consumption, wealth, and income ratio, have predictability power for equity returns at both short and long horizons. In contrast, the popular linear regression models cannot fully capture such predictability, apparently to due the neglected nonlinear predictable components. We also compare the conventional predictive regression models on predictor variables with the historical mean model according to Goyal and Welch (2007). We find that the prevailing predictive model outperforms the historical mean model in an out-of-sample content because it yields a smaller neglected signal-to-noise ratio.

We find that the poor out-of-sample performance of most linear prediction models is due to the limitation of linear models. We propose a nonparametric estimator to forecast the equity returns. Our nonparametric predictive models have lower RMSE than the historical mean model at both short-horizon and long-horizon. Our nonparametric prediction can improve the out-of-sample performance without restrictions. Using our nonparametric methods, both combined and individual forecast outperform the historical average. The combined forecast methods outperform the individual forecast methods.

Appendix

Assumptions

To prove theorem $1(i)$, we impose the following assumptions.

Assumption A.1: $\{Y_t, X_t\}$ is a stationary time series process with mixing condition. The marginal density function $g(x)$ of X_t is twice continuous differentiable with bounded second derivatives and $g(x)$ is strictly positive over the support of weighting function $w(\cdot)$ given in Assumption A.5. The dimension of X_t is d.

Assumption A.2: ε_t is a h-dependent process and ε_{t+h} is independent of $X_s, s \leq t$.

Assumption A.3: $\sqrt{R}(\hat{\beta} - \beta) = O_P(1)$, where $\beta = p\lim \hat{\beta}$.

Assumption A.4: $k(\cdot)$ is a symmetric probability density function.

Assumption A.5: $w(\cdot)$ is a positive continuous function over its support with $\int w(x)dx < \infty$ and $\int w^2(x)dx < \infty$.

Assumption A.6: $b = b(n) \to \infty, nb \to \infty$ as $n, R \to \infty$.

Assumption A.7: The function $r_h(x) = E(\varepsilon_{t+h}|X_t = x)$ is twice continuously differentiable with bounded second derivatives.

References

[1] Ang, Andrew and Geert Bekaert, 2007, Stock Return Predictability: Is It There?, Review of Financial Studies 20, 651-707.

[2] Andrew Ang, and Jun Liu, 2007, Risk, return, and dividends, Journal of Financial Economics, 85, 1- 38

[3] Baker, M., and J. Wurgler. 2000. The Equity Share in New Issues and Aggregate Stock Returns. Journal of Finance 55:2219–57.

[4] Campbell, J. Y., 1987, Stock returns and the Term Structure, Journal of Financial Economics, 18(2), 373- 399.

[5] Campbell, J. Y., 1991, A Variance Decomposition for Stock Returns, Economic Journal, 101, 157-179

[6] Campbell, J. Y., and R. J. Shiller, 1987, Cointegration and Tests of Present Value Models, Journal of Political Economy, 95, 1067-1087.

[7] Campbell, J. Y., and R. J. Shiller, 1988a, Stock Prices, Earnings, and Expected Dividends, Journal of Finance, 43(3), 661- 676.

[8] Campbell, J. Y., and R. J. Shiller,1988b, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, Review of Financial Studies, 1(3), 195- 227.

[9] Campbell, John Y., and S. Thompson, 2007, Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?, Forthcoming, Review of Financial Studies.

[10] Campbell, J. Y., and M. Yogo, 2006, Efficient Tests of Stock Return Predictability, Journal of Financial Economics, 81(1), 27- 60.

[11] Cavanagh, C. L., G. Elliott and J. H. Stock (1995), Inference in models with nearly integrated regressors. Econometric Theory 11, 1131–1147.

[12] Cochrane, J.H., 1999a, New Facts in Finance, Economic Perspectives 23, 36-58.

[13] Cochrane, J.H., 1999b. Portfolio Advice for a Multifactor World, Economic Perspectives 23, 59-78.

[14] Cochrane, J. H., 1992, Explaining the Variance of Price-Dividend Ratios, Review of Financial Studies, 5, 243- 280.

[15] Cochrane, John H., 2007, Financial Markets and the Real Economy, in John H. Cochrane, ed., Financial Markets and the Real Economy, Volume 18 of the International Library of Critical Writings in Financial Economics, London: Edward Elgar, p. xi-lxix.

[16] Cochrane, J.H., 2008, The Dog That Did Not Bark: A Defense of Return Predictability, Review of Financial Studies, 21, 1533–1575.

[17] Dickey, D., and Fuller, W. , 1979, Distribution of estimators for autoregressive time series with a unit root. Journal of American Statistical Association, 74, 427-31.

[18] Fama, E., 1990, Stock returns, expected returns, and real activity. Journal of Finance, 45 (September), 1089-1108.

[19] Fama, E., and F. French, 1988, Dividend Yields and Expected Stock Returns, Journal of Financial Economics, 22, 3- 26.

[20] Fama, Eugene F. and Kenneth R. French, 1989, Business Conditions and Expected Returns on Stocks and Bonds, Journal of Financial Economics, 25, 23- 49.

[21] Ferson,Wayne E. and Merrick John Jr., 1987, Non-stationarity and stage-of-the-businesscycle effects in consumption-based asset pricing relations, Journal of Financial Economics, 18(1), 127-146.

[22] Ferson, W. E., S. Sarkissian, and T. T. Simin. 2003. Spurious Regressions in Financial Economics? Journal of Finance 58:1393–413.

[23] Froot, K., and M. Obstfeld, 1991, Instrinsic Bubbles: The Case of Stock Prices, American Economic Review, 81, 1189-1217

[24] Foster, F. D., T. Smith, and R. E. Whaley. 1997. Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R2. Journal of Finance 52:591–607.

[25] Goetzmann, W. N., and P. Jorion, 1993, Testing the Predictive Power of Dividend Yields, Journal of Finance, 48(2), 663- 679.

[26] Goetzmann, W. N., and P. Jorion, 1995, A Longer Look at Dividend Yields, Journal of Business, 68, 483- 508.

[27] Goyal, A., and I.Welch, 2003, The Myth of Predictability: Does the Dividend Yield Forecast the Equity Premium?, Management Science, 49(5), 639- 654.

[28] Goyal, Amit and Ivo Welch, 2007, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, Forthcoming, Review of Financial Studies.

[29] Hodrick, R. J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, Review of Financial Studies, 5(3), 257- 286.

[30] Hong Y.and Lee T., 2003, Inference on via Genaralized Spectrum and Nonlinear Time Series Models, The Review of Economics and Statistics, November 2003, 85(4), 1048-1062

[31] Kothari, S., and J. Shanken, 1997, Book-to-market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis, Journal of Financial Economics, 44(2), 169- 203.

[32] Lewellen, J., 2004, Predicting Returns with Financial Ratios, Journal of Financial Economics, 74, 209- 235.

[33] Lamont, O., 1998, Earnings and Expected Returns, Journal of Finance, 53(5), 1563- 1587.

[34] Lettau, Martin, and Sydney Ludvigson, 2005, Measuring and Modeling Variation in the Risk-Return Tradeoff, Forthcoming in the Handbook of Financial Econometrics, edited by Yacine Ait-Shalia and Lars-Peter Hansen.

[35] Lettau, M. and S. Van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, Review of Financial Studies, 21, 1607–1652.

[36] Nelson, C., and Kim, M., 1993, Predictable stock returns: Reality or statistical illusion?, Journal of Finance , 48(June), 641-61.

[37] Polk, C., S. Thompson, and T. Vuolteenaho. 2006. Cross-Sectional Forecasts of the Equity Premium. Journal of Financial Economics 81:101–41.

[38] Ponti , J., and L. D. Schall, 1998, Book-to-Market Ratios as Predictors of Market Returns, Journal of Financial Economics, 49(2), 141-160.

[39] Rapach, D. E., and M. E. Wohar, 2006, In-Sample vs. Out-of-Sample Tests of Stock Return Predictability in the Context of Data Mining, Journal of Empirical Finance, 13(2), 231–247.

[40] Rapach, Strauss, and Zhou, 2009, Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, working paper

[41] Rozeff, M. ,1984, Dividend yields are equity risk premiums. Journal of Portfolio Management, 11 (Fall), 68-75.

[42] Stambaugh, R. ,1986, Biases in regressions with lagged stochastic regressors. Working Paper no. 156. Chicago: University of Chicago, Graduate School of Business.

[43] Stambaugh, R. F., 1999, Predictive Regressions, Journal of Financial Economics, 54, 375- 421.

[44] Stock, J.H. and M.W. Watson, 1999, Forecasting Inflation, Journal of Monetary Economics, 44, 293–335.

[45] Stock, J.H. and M.W.Watson, 2003, Forecasting Output and Inflation: The Role of Asset Prices, Journal of Economic Literature, 41, 788–829.

[46] Stock, J.H. and M.W.Watson, 2004, Combination Forecasts of Output Growth in a Seven-Country Data Set, Journal of Forecasting, 23, 405–430.

[47] Stock, J.H. and M.W.Watson, 2006, Forecasting with Many Predictors, In G. Elliott, C.W.J. Granger, and A. Timmermann (Eds.), Handbook of Economic Forecasting, Amsterdam: Elsevier, pp. 515–554.

[48] Torous, W., and R. Valkanov, 2000, Boundaries of Predictability: Noisy Predictive Regressions, Working Paper, UCLA.

[49] Valkanov, R., 2003, Long-Horizon Regressions: Theoretical Results and Applications, Journal of Financial Economics, 68(2), 201- 232.

[50] West, K., 1988, Dividend Innovations and Stock Price Volatility, Econometrica, 56, 37-61

[51] Wolf, M., 2000, Stock Returns And Dividend Yields Revisited: A New Way To Look At An Old Problem, Journal of Business and Economic Statistics, 18(1), 18- 30.

Tables and Graphics

Table 1a. Bootstrap Results for Predictability Check

Panel a: DGP $A.0(h)$ follows $Y_{t+h}= \alpha_0+\varepsilon_{t+h}$. ε_{t+h} has a MA$(h-1)$ property $\varepsilon_{t+h}= \sum_{j=1}^{h} \alpha_j v_{t+h-j}+v_{t+h}$. ρ is the autocoefficient of X_t. And $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. The table presents the rejection rate of bootstrap results for different h, and autocoefficient ρ under three cases (1) p-Value < 0.10, (2) p-Value < 0.05, (3) p-Value < 0.01 via the number of iterations $T = 250, 500, 1000$.

DGP				p-Value 10%			p-Value 5%			p-Value 1%		
$A.0$	h	β	ρ	$T = 250$	500	1000	$T = 250$	500	1000	$T = 250$	500	1000
	1	0	0.1	0.077	0.110	0.114	0.040	0.065	0.062	0.012	0.018	0.015
			0.3	0.089	0.124	0.098	0.050	0.071	0.045	0.015	0.024	0.010
			0.5	0.084	0.110	0.102	0.040	0.066	0.047	0.010	0.025	0.011
			0.7	0.085	0.091	0.098	0.051	0.059	0.043	0.012	0.015	0.014
			0.9	0.091	0.097	0.100	0.058	0.052	0.062	0.011	0.018	0.019

Table 1b. Bootstrap Results for Predictability Check

Panel b: DGP $A.2(h)$ follows $Y_{t+h}= \beta_0+\beta_1 X_t^2+\varepsilon_{t+h}$. ε_{t+h} has a MA$(h-1)$ property $\varepsilon_{t+h}= \sum_{j=1}^{h} \alpha_j v_{t+h-j}+v_{t+h}$. ρ is the autocoefficient of X_t. And $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. The table presents the rejection rate of bootstrap results for different h, and autocoefficient ρ under three cases (1) p-Value < 0.10, (2) p-Value < 0.05, (3) p-Value < 0.01 via the number of iterations $T = 250, 500, 1000$.

DGP				p-Value 10%			p-Value 5%			p-Value 1%		
$A.2$	h	β_1	ρ	$T = 250$	500	1000	$T = 250$	500	1000	$T = 250$	500	1000
	1	1	0.1	1.000	1.000	1.000	1.000	1.000	1.000	0.985	1.000	1.000
			0.3	1.000	1.000	1.000	1.000	1.000	1.000	0.989	1.000	1.000
			0.5	1.000	1.000	1.000	1.000	1.000	1.000	0.986	1.000	1.000
			0.7	1.000	1.000	1.000	1.000	1.000	1.000	0.989	1.000	1.000
			0.9	1.000	1.000	1.000	0.999	1.000	1.000	0.991	0.999	1.000

Table 1c. Bootstrap Results for Predictability Check

Panel c: DGP $A.1(h)$ follows $Y_{t+h} = \beta_0 + \beta_1 X_t + \varepsilon_{t+h}$. ε_{t+h} has a MA$(h-1)$ property $\varepsilon_{t+h} = \sum_{j=1}^{h} \alpha_j v_{t+h-j} + v_{t+h}$. ρ is the autocoefficient of X_t. And $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. The table presents the rejection rate of bootstrap results for different h, and autocoefficient ρ under three cases (1) p-Value < 0.10, (2) p-Value < 0.05, (3) p-Value < 0.01 via the number of iterations $T = 250, 500, 1000$.

(1) Test $H_0 : E(\varepsilon_{t+h}|X_t) = 0$, and $h = 1, 4, 12, 20$

DGP				p-Value 10%			p-Value 5%			p-Value 1%		
$A.1$	h	β	ρ	$T = 250$	500	1000	$T = 250$	500	1000	$T = 250$	500	1000
	1	0.5	0.1	0.091	0.108	0.097	0.051	0.060	0.055	0.014	0.015	0.011
			0.3	0.099	0.117	0.089	0.050	0.068	0.051	0.012	0.012	0.013
			0.5	0.101	0.110	0.100	0.046	0.064	0.058	0.014	0.013	0.011
			0.7	0.107	0.103	0.098	0.054	0.052	0.058	0.012	0.013	0.016
			0.9	0.111	0.107	0.086	0.069	0.059	0.046	0.019	0.015	0.016
	4	0.5	0.1	0.107	0.102	0.089	0.054	0.039	0.045	0.012	0.006	0.012
			0.3	0.093	0.092	0.095	0.054	0.052	0.050	0.015	0.015	0.008
			0.5	0.098	0.099	0.104	0.055	0.053	0.049	0.020	0.012	0.010
			0.7	0.097	0.113	0.100	0.051	0.065	0.061	0.012	0.023	0.014
			0.9	0.113	0.122	0.093	0.059	0.056	0.048	0.017	0.016	0.010
	12	0.5	0.1	0.081	0.090	0.091	0.045	0.052	0.052	0.012	0.009	0.013
			0.3	0.082	0.092	0.101	0.041	0.046	0.058	0.012	0.010	0.015
			0.5	0.091	0.085	0.103	0.042	0.046	0.061	0.011	0.014	0.013
			0.7	0.092	0.085	0.092	0.046	0.051	0.046	0.011	0.014	0.011
			0.9	0.088	0.108	0.088	0.040	0.062	0.049	0.009	0.017	0.014
	20	0.5	0.1	0.085	0.085	0.119	0.049	0.041	0.077	0.010	0.011	0.022
			0.5	0.082	0.092	0.097	0.041	0.052	0.049	0.009	0.017	0.013
			0.9	0.097	0.093	0.090	0.048	0.051	0.058	0.012	0.019	0.019

Table 1c. Bootstrap Results for Predictability Check

Panel c: DGP $A.1(h)$ follows $Y_{t+h} = \beta_0 + \beta_1 X_t + \varepsilon_{t+h}$. ε_{t+h} has a MA$(h-1)$ property $\varepsilon_{t+h} = \sum_{j=1}^{h} \alpha_j v_{t+h-j} + v_{t+h}$. ρ is the autocoefficient of X_t. And $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. The table presents the rejection rate of bootstrap results for different h, and autocoefficient ρ under three cases (1) p-Value < 0.10, (2) p-Value < 0.05, (3) p-Value < 0.01 via the number of iterations $T = 250, 500, 1000$.

(2) Test $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$, and $h = 1, 4, 12, 20$

DGP				p-Value 10%			p-Value 5%			p-Value 1%		
$A.1$	h	β	ρ	$T = 250$	500	1000	$T = 250$	500	1000	$T = 250$	500	1000
	1	0.1	0.1	0.107	0.183	0.310	0.062	0.121	0.218	0.010	0.034	0.083
		0.1	0.7	0.175	0.312	0.548	0.108	0.209	0.425	0.037	0.075	0.210
		0.3	0.1	0.649	0.955	0.998	0.526	0.910	0.997	0.249	0.729	0.981
		0.3	0.5	0.908	0.997	1.000	0.836	0.991	1.000	0.592	0.951	1.000
		0.5	0.5	0.993	1.000	1.000	0.984	1.000	1.000	0.913	0.999	1.000
		0.7	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.9	0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	4	0.1	0.1	0.104	0.175	0.282	0.054	0.099	0.193	0.007	0.028	0.060
		0.1	0.9	0.345	0.643	0.938	0.238	0.522	0.896	0.089	0.292	0.713
		0.3	0.1	0.620	0.917	0.999	0.498	0.869	0.997	0.217	0.642	0.974
		0.5	0.5	0.993	1.000	1.000	0.980	1.000	1.000	0.880	0.998	1.000
		0.7	0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		0.9	0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	12	0.1	0.1	0.098	0.174	0.294	0.052	0.112	0.194	0.012	0.038	0.075
		0.1	0.9	0.336	0.623	0.938	0.234	0.502	0.889	0.079	0.273	0.717
		0.3	0.1	0.587	0.910	1.000	0.472	0.855	0.999	0.221	0.628	0.972
		0.5	0.5	0.987	1.000	1.000	0.971	1.000	1.000	0.863	1.000	1.000
		0.7	0.5	0.999	1.000	1.000	0.999	1.000	1.000	0.996	1.000	1.000
		0.9	0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	20	0.1	0.1	0.094	0.167	0.303	0.056	0.074	0.217	0.020	0.020	0.077
		0.1	0.9	0.299	0.606	0.929	0.196	0.483	0.879	0.066	0.264	0.692
		0.3	0.1	0.559	0.889	0.998	0.424	0.808	0.994	0.194	0.572	0.961
		0.5	0.5	0.975	1.000	1.000	0.957	1.000	1.000	0.819	0.999	1.000
		0.7	0.5	1.000	1.000	1.000	1.000	1.000	1.000	0.984	1.000	1.000
		0.9	0.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 1d. Bootstrap Results for Predictability Check

Panel d: DGP $B.1(h)$ follows $Y_{t+h} = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \varepsilon_{t+h}$. ε_{t+h} has a MA$(h-1)$ property $\varepsilon_{t+h} = \sum_{j=1}^{h} \alpha_j v_{t+h-j} + v_{t+h}$. ρ is the autocoefficient of X_t. And $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. The table presents the rejection rate of bootstrap results for different h, and autocoefficient ρ under three cases (1) p-Value < 0.10, (2) p-Value < 0.05, (3) p-Value < 0.01 via the number of iterations $T = 250, 500, 1000$.

(1) Test $H_0 : E(\varepsilon_{t+h}|X_t) = 0$, and $h = 1, 4, 12, 20$

DGP					p-Value 10%			p-Value 5%			p-Value 1%		
$B.1$	h	β_1	β_2	ρ	$T = 250$	500	1000	$T = 250$	500	1000	$T = 250$	500	1000
	1	0.5	0.1	0.1	0.168	0.265	0.377	0.088	0.169	0.261	0.032	0.048	0.099
		0.5	0.1	0.5	0.215	0.377	0.590	0.132	0.274	0.475	0.039	0.118	0.241
		0.5	0.3	0.1	0.698	0.963	0.999	0.578	0.911	0.995	0.310	0.734	0.974
		0.5	0.3	0.5	0.874	0.996	1.000	0.808	0.989	1.000	0.585	0.930	1.000
		0.5	0.5	0.5	0.995	1.000	1.000	0.984	1.000	1.000	0.925	1.000	1.000
		0.5	0.7	0.7	1.000	1.000	1.000	1.000	1.000	1.000	0.990	1.000	1.000
		0.5	0.9	0.7	1.000	1.000	1.000	1.000	1.000	1.000	0.993	1.000	1.000
	4	0.5	0.1	0.1	0.178	0.229	0.394	0.103	0.155	0.283	0.024	0.040	0.101
		0.5	0.1	0.7	0.360	0.655	0.911	0.262	0.544	0.856	0.113	0.310	0.628
		0.5	0.3	0.1	0.656	0.947	1.000	0.537	0.898	0.999	0.298	0.708	0.977
		0.5	0.3	0.5	0.877	0.993	1.000	0.797	0.988	1.000	0.551	0.922	0.999
		0.5	0.5	0.5	0.997	1.000	1.000	0.991	1.000	1.000	0.921	0.999	1.000
		0.5	0.7	0.7	1.000	1.000	1.000	0.999	1.000	1.000	0.986	1.000	1.000
		0.5	0.9	0.7	1.000	1.000	1.000	1.000	1.000	1.000	0.988	1.000	1.000
	12	0.5	0.1	0.1	0.134	0.245	0.349	0.075	0.138	0.233	0.017	0.037	0.097
		0.5	0.1	0.5	0.188	0.335	0.586	0.105	0.224	0.460	0.028	0.076	0.220
		0.5	0.3	0.1	0.640	0.935	1.000	0.489	0.888	0.997	0.231	0.677	0.979
		0.5	0.3	0.5	0.845	0.994	1.000	0.754	0.985	1.000	0.496	0.918	0.999
		0.5	0.5	0.5	0.994	1.000	1.000	0.981	1.000	1.000	0.889	0.999	1.000
		0.5	0.7	0.5	0.999	1.000	1.000	0.998	1.000	1.000	0.963	1.000	1.000
		0.5	0.9	0.5	1.000	1.000	1.000	0.999	1.000	1.000	0.977	1.000	1.000
	20	0.5	0.1	0.1	0.140	0.235	0.410	0.079	0.146	0.286	0.016	0.035	0.110
		0.5	0.1	0.5	0.187	0.338	0.604	0.109	0.230	0.484	0.029	0.080	0.253
		0.5	0.3	0.1	0.610	0.934	1.000	0.466	0.882	0.999	0.206	0.670	0.981
		0.5	0.3	0.5	0.829	0.994	1.000	0.739	0.987	1.000	0.460	0.920	1.000
		0.5	0.5	0.5	0.993	1.000	1.000	0.975	1.000	1.000	0.867	0.997	1.000
		0.5	0.7	0.5	0.999	1.000	1.000	0.997	1.000	1.000	0.955	0.999	1.000
		0.5	0.9	0.5	1.000	1.000	1.000	0.999	1.000	1.000	0.973	0.999	1.000

Table 1d. Bootstrap Results for Predictability Check

Panel d: DGP $B.1(h)$ follows $Y_{t+h} = \beta_0+\beta_1 X_t + \beta_2 X_t^2 + \varepsilon_{t+h}$. ε_{t+h} has a MA$(h-1)$ property $\varepsilon_{t+h} = \sum_{j=1}^{h} \alpha_j v_{t+h-j} + v_{t+h}$. ρ is the autocoefficient of X_t. And $\{v_{t+h}\}$ and $\{u_t\}$ are mutually independent. The table presents the rejection rate of bootstrap results for different h, and autocoefficient ρ under three cases (1) p-Value < 0.10, (2) p-Value < 0.05, (3) p-Value < 0.01 via the number of iterations $T = 250, 500, 1000$.

(2) Test $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$, and $h = 1, 4, 12, 20$

DGP					p-Value 10%			p-Value 5%			p-Value 1%		
$B.1$	h	β_1	β_2	ρ	$T=250$	500	1000	$T=250$	500	1000	$T=250$	500	1000
	1	0.1	0.1	0.1	0.176	0.338	0.610	0.089	0.237	0.475	0.026	0.078	0.212
		0.1	0.5	0.1	0.975	1.000	1.000	0.936	0.999	0.999	0.759	0.990	0.994
		0.3	0.1	0.1	0.694	0.963	1.000	0.561	0.921	0.999	0.310	0.748	0.985
		0.3	0.5	0.1	0.995	1.000	1.000	0.978	1.000	1.000	0.868	0.997	1.000
		0.5	0.5	0.1	1.000	1.000	1.000	0.997	1.000	1.000	0.967	1.000	1.000
		0.7	0.7	0.1	1.000	1.000	1.000	1.000	1.000	1.000	0.992	1.000	1.000
		0.9	0.9	0.1	1.000	1.000	1.000	1.000	1.000	1.000	0.996	1.000	1.000
	4	0.5	0.1	0.1	0.978	1.000	1.000	0.947	1.000	1.000	0.804	0.995	1.000
		0.5	0.1	0.5	0.993	1.000	1.000	0.982	1.000	1.000	0.906	0.999	1.000
		0.5	0.3	0.1	0.995	1.000	1.000	0.981	1.000	1.000	0.900	1.000	1.000
		0.5	0.3	0.5	0.997	1.000	1.000	0.991	1.000	1.000	0.945	1.000	1.000
		0.5	0.7	0.5	1.000	1.000	1.000	1.000	1.000	1.000	0.983	1.000	1.000
		0.5	0.9	0.5	1.000	1.000	1.000	1.000	1.000	1.000	0.987	1.000	1.000
	12	0.5	0.1	0.1	0.970	1.000	1.000	0.944	1.000	1.000	0.765	0.994	1.000
		0.5	0.1	0.5	0.989	1.000	1.000	0.978	1.000	1.000	0.882	0.998	1.000
		0.5	0.3	0.1	0.987	1.000	1.000	0.973	1.000	1.000	0.865	0.997	1.000
		0.5	0.3	0.5	0.994	1.000	1.000	0.988	1.000	1.000	0.936	0.998	1.000
		0.5	0.5	0.5	0.999	1.000	1.000	0.996	1.000	1.000	0.964	1.000	1.000
		0.5	0.9	0.5	1.000	1.000	1.000	1.000	1.000	1.000	0.980	1.000	1.000
	20	0.5	0.1	0.1	0.956	1.000	1.000	0.909	1.000	1.000	0.713	0.990	1.000
		0.5	0.1	0.5	0.979	1.000	1.000	0.955	1.000	1.000	0.836	0.993	1.000
		0.5	0.3	0.1	0.978	1.000	1.000	0.953	1.000	1.000	0.813	0.997	1.000
		0.5	0.3	0.5	0.995	1.000	1.000	0.984	1.000	1.000	0.905	0.998	1.000
		0.5	0.5	0.5	0.997	1.000	1.000	0.994	1.000	1.000	0.951	0.999	1.000
		0.5	0.9	0.5	0.999	1.000	1.000	0.997	1.000	1.000	0.975	1.000	1.000

Table 2. Sample statistics

Panel (a) reports summary statistics of the data for S&P 500, all at a quarterly frequency. Panel (b) reports statistics for monthly frequency . Excess returns and short rates are continuously compounded. Sample means and standard deviations (SD) for excess returns, dividend, and earnings growth have been annualized by multiplying by $4(12)$ and $\sqrt{4}(\sqrt{12})$, respectively, for the case of quarterly (monthly) frequency data. Short rates are three-month T-bill rates. Dividend and earnings yields, and the corresponding dividend and earnings growth are computed using dividends or earnings summed up over the past year. Panel (c) reports statistics for annually frequency. The unit root test is the Phillips and Perron (1988) test for the estimated regression $x_t = \alpha + \rho x_{t-1} + u_t$ under the null $x_t = x_{t-1} + u_t$. The critical values corresponding to p-values of 0.01, 0.025, 0.05, and 0.10 are 3:46, 3:14, 2:88, and 2:57, respectively. $^{*}p < 0.05.^{**}p < 0.01$.

Panel (a) US S&P500 data, March 1936-December 2001 Quarterly

	Excess Return	Short Rate	Dividend Yield	earnings Yield	Dividend Growth	Earnings Growth
Mean	0.0803	0.0416	0.0408	0.0769	0.0566	0.0599
S.D.	0.3336	0.0320	0.0157	0.0299	0.1578	0.2303
Auto	0.092	0.9531	0.9221	0.9516	0.4306	0.6926
Test statistics						
H_0: unit root	−13.97*	−2.9290	−4.0337	−1.5765	−9.72**	−6.37**

Panel (b) US S&P500 data, January 1970-December 2006 monthly

	Excess Return	Short Rate	Dividend Yield	earnings Yield
Mean	0.0610	0.0597	0.0317	0.0694
S.D.	0.5259	0.0288	0.0129	0.0299
Auto	0.081	0.992	0.978	0.987
Test statistics				
H_0: unit root	−28.34*	−1.28	−1.51	−1.17

Panel (c) US S&P500 data, 1872-2005 annually

	Excess Return	Short Rate	Dividend Yield	earnings Yield	Dividend Payout Ratio	Inflation 1919−2005	b/m 1921−2005	i/k 1947−2005	ntis 1927−2005	eqis 1927−2005	cay 1945−2005
Mean	0.0171	0.0479	0.0451	0.0750	0.6213	0.0309	1.6924	0.0836	0.0229	0.1986	−0.0488
S.D.	0.170	0.0278	0.0163	0.0273	0.1987	0.0464	0.5338	0.0343	0.0772	0.1099	1.4925
Auto	0.077	0.831	0.762	0.737	0.615	0.566	0.918	0.912	0.436	0.493	0.564
Test statistics											
H_0: unit root	−10.53**	−1.05	−1.53	−1.54	−1.67	−4.32**	−0.90	−3.63**	−7.82**	−2.33*	−4.30**

Table 3a. Predictability of US Excess Returns(Quarterly)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (: y \; y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividend yields or log earnings yields. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are quarterly. The test column reports a p-value for the predictictability test under two hypothesises $H2$ and $H2$. Hypothesis $H2$ is $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$ and Hypothesis $H2$ is $H_0 : E(\varepsilon_{t+h}|X_t) = 0$. $^*p < 0.05$.$^{**}p < 0.01$. The time periods are from 1936 to 2001, from 1952 to 2001, from 1936 to 1990, and from 1952 to 1990.

	k	Univariate Regression $dy4$ (1)NW	(2)HK	Test (1)H1	(2)H2	$ey4$ (1)NW	(2)HK	Test (1)H1	(2)H2
1936	1	0.1441	0.1441	0.0565	0.0758	0.0630	0.0625	0.2526	0.0901
−2001		(2.28)*	(1.824)	(0.000)**	(0.000)**	(0.91)	(0.70)	(0.000)**	(0.000)**
	4	0.1618	0.1578	0.0525	0.1649	0.0918	0.0898	0.0379	0.0677
		(2.81)**	(2.030)	(0.000)**	(0.000)**	(1.42)	(1.30)	(0.015)*	(0.000)**
	12	0.1522	0.1455	0.0390	0.0454	0.0648	0.549	0.0377	0.0537
		(2.92)**	(1.568)	(0.000)**	(0.000)**	(1.05)	(0..90)	(0.016)*	(0.000)**
	20	0.1846	0.1028	0.0921	0.1552	0.0484	0.0348	0.0490	0.0549
		(3.84)**	(1.364)	(0.000)**	(0.000)**	(0.81)	(0.54)	(0.000)**	(0.000)**
1952	1	0.0417	0.0489	0.1048	0.1129	0.0467	0.0464	0.2952	0.0984
−2001		(0.51)	(0.771)	(0.000)**	(0.000)**	(0.59)	(0.46)	(0.000)**	(0.000)**
	4	0.0542	0.0530	0.1558	0.0910	0.0791	0.0798	0.1260	0.1178
		(0.69)	(0.773)	(0.000)**	(0.000)**	(1.41)	(1.218)	(0.000)**	(0.000)**
	12	0.0091	0.0234	0.0244	0.0454	0.0768	0.0756	0.0683	0.0887
		(0.12)	(0.344)	(0.024)*	(0.000)**	(1.25)	(1.03)	(0.000)**	(0.000)**
	20	0.0276	0.0287	0.0921	0.1552	0.0283	0.0223	0.1005	0.1065
		(0.35)	(0.477)	(0.000)**	(0.000)**	(1.80)	(0.320)	(0.000)**	(0.000)**
1936	1	0.2993	0.2203	0.0315	0.0515	0.0886	0.0896	0.3864	0.0387
−1990		(3.19)**	(2.416)*	(0.021)*	(0.000)**	(1.57)	(1.304)	(0.000)**	(0.015)*
	4	0.3228	0.2383	0.1560	0.0689	0.1417	0.1473	0.0333	0.0361
		(4.56)**	(3.097)**	(0.000)**	(0.000)**	(1.52)	(2.309)*	(0.019)*	(0.017)*
	12	0.2744	0.2380	0.1069	0.1016	0.0957	0.0945	0.0555	0.0979
		(5.08)**	(4.08)**	(0.000)**	(0.000)**	(1.29)	(1.02)	(0.000)**	(0.000)**
	20	0.2652	0.1787	0.1326	0.1042	0.1225	0.1078	0.0924	0.0882
		(5.97)**	(2.819)**	(0.000)**	(0.000)**	(1.38)	(1.392)	(0.000)**	(0.000)**
1952	1	0.1622	0.1482	0.0851	0.1006	0.1537	0.1028	0.2496	0.0223
−1990		(1.06)	(2.783)**	(0.000)**	(0.000)**	(0.60)	(1.168)	(0.000)**	(0.022)*
	4	0.1921	0.3070	0.0682	0.0639	0.1291	0.1287	0.0809	0.0685
		(1.40)	(2.50)*	(0.000)**	(0.000)**	(1.62)	(1.521)	(0.000)**	(0.000)**
	12	0.0988	0.1132	0.1849	0.1318	0.0576	0.0475	0.0732	0.1356
		(0.80)	(1.453)	(0.000)**	(0.000)**	(0.75)	(0.45)	(0.000)**	(0.000)**
	20	0.0622	0.0845	0.2669	0.1138	0.0771	0.0798	0.3169	0.1878
		(0.67)	(1.916)	(0.000)**	(0.000)**	(1.17)	(1.133)	(0.000)**	(0.000)**

Table 3b. Predictability of US Excess Returns(Quarterly)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividend yields and risk free rate together. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are quarterly. The test column reports a p-value for the predictictability test under six hypothesises $H1$–$H6$. X_1 represents the short rate r and X_2 represents the dividend yields. Hypothesis $H1$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), Hypothesis $H2$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0$),Hypothesis $H3$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), Hypothesis $H4$ ($H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0$), Hypothesis $H5$ ($H_0 : E(Y_{t+h}|X_{1t}, X_{2t}) = E(Y_{t+h})$), Hypothesis $H6$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}) = 0$). $^*p < 0.05$.$^{**}p < 0.01$. The time periods are from 1936 to 2001, from 1952 to 2001, from 1936 to 1990, and from 1952 to 1990.

	k	Bivariate Regression r (1)NW	(2)HK	$dy4$ (1)NW	(2)HK	Test (1)H1	(2)H2	(3)H3	(4)H4	(5)H5	(6)H6
1936	1	−0.9864	−1.0888	0.0849	0.0875	0.0409	0.0495	0.0475	0.0699	0.0088	0.0137
−2001		(−5.730)**	(−1.608)	(1.530)	(1.530)	(0.010)**	(0.008)**	(0.005)**	(0.000)**	(0.044)*	(0.032)*
	4	−0.8305	−0.5596	0.1068	0.1032	0.0694	0.0600	0.0525	0.0761	0.0064	0.0168
		(−5.520)**	(−0.827)	(2.200)*	(1.865)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.052)	(0.028)*
	12	−0.6993	−0.4321	0.0985	0.0980	0.0533	0.0469	0.0701	0.0892	0.0059	0.0072
		(−8.57)**	(−0.708)	(2.530)*	(1.265)	(0.000)**	(0.006)**	(0.000)**	(0.000)**	(0.056)	(0.046)*
	20	−0.6606	−0.6364	0.1246	0.1152	0.0634	0.1278	0.0913	0.2218	0.0040	0.0043
		(−8.83)**	(1.196)	(3.57)**	(1.255)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.062)	(0.064)
1952	1	−1.2958	−2.1623	0.1355	0.1362	0.1147	0.1236	(0.1593)	0.1151	0.0123	0.0095
−2001		(−6.74)**	(−2.912)**	(1.91)*	(2.152)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.034)*	(0.040)*
	4	−1.0484	−1.4433	0.1291	0.1313	0.1548	0.1316	0.1558	0.0909	0.0338	0.0472
		(−5.85)**	(−1.930)	(1.990)*	(1.921)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.017)*	(0.005)**
	12	−0.7141	−0.6272	0.0685	0.0874	0.0712	0.0488	0.0335	0.0823	0.0108	0.0153
		(−6.61)**	(−0.987)	(1.17)	(0.680)	(0.000)**	(0.006)**	(0.016)*	(0.000)**	(0.035)*	(0.031)*
	20	−0.5738	−0.4829	0.0596	0.0774	0.0534	0.0690	0.0759	0.1468	0.0198	0.0081
		(−5.99)**	(−0.745)	(1.00)	(0.600)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.026)*	(0.048)*
1936	1	−0.9709	−1.0380	0.1961	0.1917	0.0325	0.0321	0.0510	0.0503	0.0091	0.0090
−1990		(−5.49)**	(−1.543)	(2.44)*	(2.126)*	(0.021)*	(0.021)*	(0.000)**	(0.000)**	(0.044)*	(0.045)*
	4	−0.8131	−0.4865	0.2398	0.2254	0.0386	0.0390	0.0661	0.0698	0.0142	0.0144
		(−5.34)**	(−0.714)	(4.21)**	(3.006)**	(0.015)*	(0.012)*	(0.000)**	(0.000)**	(0.034)*	(0.032)*
	12	−0.7121	−0.6785	0.2114	0.1898	0.0827	0.0603	0.1069	0.0633	0.0126	0.0139
		(−8.26)**	(−0.879)	(5.65)**	(3.214)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.035)*	(0.033)*
	20	−0.6877	−0.6458	0.2097	0.1719	0.0791	0.0222	0.2107	0.0580	0.0335	0.0045
		(−8.58)**	(−1.289)	(7.43)**	(2.832)**	(0.000)**	(0.024)*	(0.000)**	(0.000)**	(0.017)*	(0.061)
1952	1	−1.4347	−2.7329	0.4102	0.4125	0.0947	0.1134	0.0851	0.0823	0.0183	0.0224
−1990		(−6.29)**	(3.504)**	(3.02)**	(3.672)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.029)*	(0.022)*
	4	−1.1908	−1.9840	0.4030	0.3935	0.0892	0.1049	0.0683	0.0607	0.0273	0.0193
		(−6.26)**	(−2.508)*	(3.78)**	(3.700)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.019)*	(0.025)*
	12	−0.8431	−0.7324	0.2623	0.2457	0.1210	0.0716	0.1849	0.1381	0.0628	0.0361
		(−9.02)**	(−1.278)	(4.28)**	(2.457)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.018)*
	20	−0.6679	−0.7120	0.2024	0.2057	0.3003	0.0541	0.2679	0.1058	0.1149	0.0029
		(−7.47)**	(−1.087)	(4.09)**	(2.387)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.112)

Table 3c. Predictability of US Excess Returns(Quarterly)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividends yields and log earnings yields together. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are quarterly. The test column reports a p-value for the predictictability test under six hypothesises $H1$–$H6$. X_1 represents the short rate r and X_2 represents the earnings yields. Hypothesis $H1$ $(H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h}))$, Hypothesis $H2$ $(H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0)$,Hypothesis $H3$ $(H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h}))$, Hypothesis $H4$ $(H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0)$, Hypothesis $H5$ $(H_0 : E(Y_{t+h}|X_{1t}, X_{2t}) = E(Y_{t+h}))$, Hypothesis $H6$ $(H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}) = 0)$. $^*p < 0.05$. $^{**}p < 0.01$.The time periods are from 1936 to 2001, from 1952 to 2001, from 1936 to 1990, and from 1952 to 1990.

	k	Bivariate Regression $dy4$ (1)NW	(2)HK	$ey4$ (1)NW	(2)HK	Test (1)H1	(2)H2	(3)H3	(4)H4	(5)H5	(6)H6
1936	1	0.3891	0.1842	−0.3003	−0.1019	0.0475	0.0771	0.0781	0.0887	0.0303	0.0482
−2001		(3.08)**	(1.330)	(−2.27)*	(−0.741)	(0.008)**	(0.000)**	(0.000)**	(0.000)**	(0.018)*	(0.006)**
	4	0.3485	0.1119	−0.2289	0.0012	0.0526	0.0781	0.0379	0.0680	0.0209	0.0211
		(3.11)**	(0.981)	(−1.97)*	(0.011)	(0.000)**	(0.000)**	(0.015)*	(0.000)**	(0.025)*	(0.024)*
	12	0.3503	0.1470	−0.2263	−0.1020	0.0701	0.0400	0.0306	0.0385	0.0148	0.0091
		(3.36)**	(1.40)	(−2.19)*	(−1.39)	(0.000)**	(0.011)*	(0.018)*	(0.014)*	(0.030)*	(0.044)*
	20	0.4752	0.2111	−0.3119	−0.0912	0.0913	0.0495	0.1493	0.1495	0.0121	0.0189
		(5.05)**	(1.55)	(−3.64)**	(−1.64)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.035)*	(0.027)*
1952	1	0.3857	0.1948	−0.3953	−0.1113	0.1593	0.0945	0.0941	0.0468	0.0898	0.0447
−2001		(2.46)*	(1.557)	(−2.61)**	(−0.873)	(0.000)**	(0.000)**	(0.000)**	(0.009)**	(0.000)**	(0.010)**
	4	0.3584	0.1615	−0.3485	−0.0636	0.1139	0.1161	0.1682	0.1255	0.0865	0.0810
		(2.57)*	(1.195)	(−2.83)**	(−0.450)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
	12	0.2843	0.1821	−0.2969	−0.0832	0.0336	0.0202	0.0458	0.0388	0.0087	0.0071
		(1.87)	(1.21)	(−2.65)**	(−0.76)	(0.016)*	(0.026)*	(0.009)**	(0.014)*	(0.045)*	(0.046)*
	20	0.3567	0.1458	−0.3402	−0.0765	0.0821	0.0986	0.0851	0.1001	0.0266	0.0525
		(2.23)*	(0.561)	(−3.32)**	(−0.520)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.022)*	(0.000)**
1936	1	0.7476	0.3848	−0.4719	−0.1800	0.0510	0.0497	0.0351	0.0388	0.0192	0.0221
−1990		(4.23)**	(1.849)	(−2.99)**	(1.062)	(0.000)**	(0.000)**	(0.019)*	(0.014)*	(0.027)*	(0.024)*
	4	0.6859	0.2869	−0.3823	−0.0532	0.0662	0.0666	0.0333	0.0381	0.0121	0.0163
		(4.58)**	(1.872)	(−2.76)**	(−0.411)	(0.000)**	(0.000)**	(0.020)*	(0.015)*	(0.035)*	(0.031)*
	12	0.6320	0.2987	−0.3677	−0.789	0.0706	0.0751	0.1741	0.1994	0.0501	0.0550
		(5.26)**	(1.98)	(−3.03)**	(−0.93)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
	20	0.6521	0.1916	−0.4005	−0.0139	0.1486	0.1245	0.0901	0.0834	0.1018	0.0824
		(6.77)**	(1.874)	(−4.25)**	(−0.181)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
1952	1	1.3055	0.9349	−0.94	−0.5240	0.0851	0.0907	0.0271	0.0219	0.0112	0.0089
−1990		(4.05)**	(3.71)**	(−3.90)**	(−2.618)**	(0.000)**	(0.000)**	(0.021)*	(0.024)*	(0.036)*	(0.045)*
	4	1.2350	0.8210	−0.8581	−0.4212	0.0683	0.0592	0.0809	0.0564	0.0427	0.0329
		(4.35)**	(3.234)**	(−4.21)**	(−2.080)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.009)**	(0.016)*
	12	1.0544	0.7824	−0.7689	−0.7320	0.0098	0.0648	0.0910	0.1074	0.0041	0.0311
		(6.89)**	(2.98)**	(−6.38)**	(−3.54)**	(0.043)*	(0.000)**	(0.000)**	(0.000)**	(0.061)	(0.017)*
	20	0.8561	0.4167	−0.6423	−0.1999	(0.0212)	0.0511	0.0458	0.0875	0.0824	0.0551
		(8.14)**	(2.685)**	(−7.43)**	(−1.570)	(0.025)*	(0.000)**	(0.006)**	(0.000)**	(0.000)**	(0.000)**

Table 3d. Predictability of US Excess Returns(Quarterly)

(a)We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being short rate, log dividends yields, and log earnings yields together. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are quarterly. The time periods are from 1936 to 2001, from 1952 to 2001, from 1936 to 1990, and from 1952 to 1990.

(b)The test column reports a p-value for the predictictability test under six hypothesises $H1-H8$. X_1 represents the short rate r , X_2 represents the dividends yields and X_3 represents the earnings yields. Hypothesis $H1$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), Hypothesis $H2$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0$),Hypothesis $H3$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), Hypothesis $H4$ ($H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0$), Hypothesis $H5$ ($H_0 : E(Y_{t+h}|X_{3t}) = E(Y_{t+h})$), Hypothesis $H6$ ($H_0 : E(\varepsilon_{t+h}|X_{3t}) = 0$),Hypothesis $H7$ ($H_0 : E(Y_{t+h}|X_{1t}, X_{2t}, X_{3t}) = E(Y_{t+h})$), Hypothesis $H8$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}, X_{3t}) = 0$). $^*p < 0.05$.$^{**}p < 0.01$.

TrivariateRegression

	k	r (1)NW	(2)HK	$dy4$ (1)NW	(2)HK	$ey4$ (1)NW	(2)HK
1936	1	−1.0703	−1.0334	0.1018	0.1000	−0.0132	−0.0168
−2001		(−5.40)**	(−1.296)	(0.80)	(0.605)	(−0.090)	(−0.103)
	4	−0.9325	−0.7759	0.0094	0.0472	0.1237	0.0654
		(−5.65)**	(−1.084)	(0.31)	(0.377)	(1.26)	(0.556)
	12	−0.7717	−0.6836	0.0175	0.0125	0.0862	0.0632
		(−8.49)**	(−1.923)	(0.20)	(0.042)	(1.10)	(0.70)
	20	−0.6231	−0.4190	0.1685	0.0666	−0.0435	0.0291
		(−6.63)**	(−0.682)	(1.79)	(0.486)	(−0.52)	(0.361)
1952	1	−1.4102	−2.6243	0.0327	0.0337	0.1276	0.1272
−2001		(−6.04)**	(−2.85)**	(0.21)	(0.239)	(0.80)	(0.815)
	4	−1.1074	−1.8146	0.0755	0.0479	0.0662	0.0629
		(−5.01)**	(−2.063)*	(0.56)	(0.687)	(0.52)	(0.487)
	12	−0.6863	−1.0213	0.0988	0.0453	−0.0337	−0.0134
		(−5.43)**	(−2.34)*	(0.81)	(0.43)	(−0.31)	(−0.23)
	20	−0.4535	−0.3548	0.2209	0.1199	−0.1591	−0.0418
		(−3.62)**	(−0.489)	(1.54)	(0.432)	(−1.36)	(−0.249)
1936	1	−0.9918	−0.6544	0.3706	0.3051	0.0245	0.0125
−1990		(−4.64)**	(−0.696)	(1.84)	(1.112)	(0.15)	(0.099)
	4	−0.852	−0.4708	0.1932	0.2300	0.0449	−0.0046
		(−4.98)**	(−0.031)	(1.33)	(1.336)	(0.38)	(−0.031)
	12	−0.6939	−0.4873	0.2341	0.2032	−0.0217	−0.0312
		(−8.11)**	(−6.46)**	(2.33)	(1.90)	(−0.24)	(−0.32)
	20	−0.6067	−0.4239	0.3085	0.1423	−0.0955	0.0297
		(−6.43)**	(−0.714)	(3.19)**	(1.597)	(−1.10)	(0.431)
1952	1	−1.2799	−2.0892	0.6690	0.6598	−0.2348	−0.2418
−1990		(−5.13)**	(−2.096)*	(1.84)	(1.78)	(−0.87)	(−0.949)
	4	−0.9802	−1.3320	0.7451	0.6565	−0.3121	−0.2392
		(−4.72)**	(−1.376)	(2.65)**	(2.303)*	(−1.58)	(−0.960)
	12	−0.5595	−0.4578	0.7738	0.6719	−0.4558	−0.3722
		(−6.62)**	(−2.23)*	(6.71)**	(2.13)*	(−4.14)**	(−1.80)
	20	−0.4047	−0.3068	0.6553	0.3793	−0.4111	−0.1570
		(−3.27)**	(−0.374)	(5.24)**	(2.188)*	(−3.23)**	(−0.991)

	k	Trivariate Regression Test (1)H1	(2)H2	(3)H3	(4)H4	(5)H5	(6)H6	(7)H7	(8)H8
1936 −2001	1	0.0409	0.0509	0.0475	0.0600	0.0781	0.0888	0.0089	0.0157
		(0.013)*	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.044)*	(0.026)*
	4	0.0694	0.0645	0.0526	0.0781	0.0379	0.0676	0.0058	0.0166
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.015)*	(0.000)**	(0.056)	(0.028)*
	12	0.0233	0.0466	0.0701	0.0901	0.0306	0.0434	0.0029	0.0043
		(0.022)*	(0.003)**	(0.000)**	(0.000)**	(0.018)*	(0.008)**	(0.089)	(0.060)
	20	0.0635	0.1266	0.0913	0.2269	0.1493	0.1108	0.0258	0.0352
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.022)*	(0.016)*
1952 −2001	1	0.1148	0.1079	0.1593	0.1033	0.0941	0.0518	0.0111	0.0026
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.035)*	(0.091)
	4	0.1270	0.1056	0.1558	0.0711	0.1260	0.1016	0.0195	0.0280
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.025)*	(0.020)*
	12	0.0448	0.0504	0.0861	0.1417	0.0676	0.1111	0.0245	0.0299
		(0.006)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.023)*	(0.019)*
	20	0.0901	0.0487	0.0751	0.0639	0.0841	0.0718	0.0047	0.0013
		(0.000)**	(0.002)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.058)	(0.110)
1936 −1990	1	0.0325	0.0268	0.0510	0.0427	0.0351	0.0287	0.0071	0.0049
		(0.019)*	(0.021)*	(0.000)**	(0.008)**	(0.018)*	(0.022)*	(0.046)*	(0.058)
	4	0.0087	0.0398	0.0662	0.0703	0.0333	0.0534	0.0040	0.0145
		(0.046)*	(0.013)*	(0.000)**	(0.000)**	(0.020)*	(0.000)**	(0.060)	(0.027)*
	12	0.0827	0.0599	0.0827	0.0639	0.0858	0.0535	0.0019	0.0130
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.101)	(0.035)*
	20	0.0817	0.0251	0.2098	0.0708	0.1070	0.0242	0.0349	0.0103
		(0.000)**	(0.022)*	(0.000)**	(0.000)**	(0.000)**	(0.023)*	(0.019)*	(0.037)*
1952 −1990	1	0.0947	0.1137	0.0851	0.0799	0.0271	0.0475	0.0062	0.0144
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.021)*	(0.000)**	(0.050)*	(0.028)*
	4	0.0892	0.0750	0.0683	0.0579	0.0809	0.0572	0.0233	0.0051
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.024)*	(0.058)
	12	0.1210	0.0679	0.1849	0.1499	0.1202	0.0942	0.0584	0.0266
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.022)*
	20	0.0711	0.1101	0.2362	0.0633	0.2116	0.0998	0.0957	0.0116
		(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.034)*

Table 4a. Predictability of US Excess Returns(monthly)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividend yields or log earnings yields. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are monthly. The test column reports a p-value for the predictictability test under two hypothesises $H2$ and $H2$. Hypothesis $H2$ is $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$ and Hypothesis $H2$ is $H_0 : E(\varepsilon_{t+h}|X_t) = 0$. $^*p < 0.05.^{**}p < 0.01$. The time periods are from Jan. 1970 to Dec. 2006.

	k	Univariate Regression $dy12$ (1)NW	(2)HK	Test (1)H1	(2)H2	$ey12$ (1)NW	(2)HK	Test (1)H1	(2)H2
197001	1	0.1810	0.1341	0.1699	0.0308	−0.4932	−0.4325	0.1699	0.0287
−200612		(2.25)*	(1.21)	(0.000)**	(0.019)*	(−4.02)**	(−1.01)	(0.000)**	(0.021)*
	12	0.1689	0.1256	0.0132	0.0328	−0.5032	−0.4132	0.0105	0.0473
		(2.34)*	(1.90)	(0.034)*	(0.017)*	(−3.47)**	(−1.74)	(0.036)*	(0.000)**
	60	−0.4365	−0.2874	0.0230	0.0641	−0.5025	−0.3389	0.0509	0.0263
		(4.16)**	(1.80)	(0.024)*	(0.000)**	(−5.53)**	(−2.42)	(0.000)**	(0.021)*

Table 4b. Predictability of US Excess Returns(Monthly)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividend yields and risk free rate together. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are monthly. The test column reports a p-value for the predictictability test under six hypothesises $H1$–$H6$. X_1 represents the short rate r and X_2 represents the dividend yields. Hypothesis $H1$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), Hypothesis $H2$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0$),Hypothesis $H3$ ($H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h})$), Hypothesis $H4$ ($H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0$), Hypothesis $H5$ ($H_0 : E(Y_{t+h}|X_{1t}, X_{2t}) = E(Y_{t+h})$), Hypothesis $H6$ ($H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}) = 0$). $^*p < 0.05.^{**}p < 0.01$. The time periods are from Jan. 1970 to Dec. 2006.

	k	BivariateRegression r (1)NW	(2)HK	$dy12$ (1)NW	(2)HK	Test (1)H1	(2)H2	(3)H3	(4)H4	(5)H5	(6)H6
197001	1	−14.084	−11.256	0.1868	0.1589	0.0224	0.0301	0.0359	0.0498	0.0143	0.0144
−200612		(−15.35)**	(−12.60)**	(2.53)*	(1.32)	(0.023)*	(0.019)*	(0.011)*	(0.000)**	(0.032)*	(0.032)*
	12	−0.8856	−0.6328	0.0513	0.0339	0.0448	0.0297	0.0249	0.0320	0.0124	0.0195
		(−10.98)**	(−8.03)**	(0.97)	(0.56)	(0.000)**	(0.020)*	(0.022)*	(0.017)*	(0.035)*	(0.028)*
	60	−0.1326	−0.0927	−0.3782	−0.2734	0.0487	0.0268	0.0230	0.0630	0.0150	0.0134
		(−1.71)	(−0.88)	(2.18)	(1.14)	(0.000)**	(0.022)*	(0.022)*	(0.000)**	(0.029)*	(0.034)*

Table 4c. Predictability of US Excess Returns(Monthly)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividends yields and log earnings yields together. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are monthly. The test column reports a p-value for the predictictability test under six hypothesises $H1-H6$. X_1 represents the short rate r and X_2 represents the earnings yields. Hypothesis $H1$ $(H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h}))$, Hypothesis $H2$ $(H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0)$,Hypothesis $H3$ $(H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h}))$, Hypothesis $H4$ $(H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0)$, Hypothesis $H5$ $(H_0 : E(Y_{t+h}|X_{1t}, X_{2t}) = E(Y_{t+h}))$, Hypothesis $H6$ $(H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}) = 0)$. $^*p < 0.05$.$^{**}p < 0.01$.The time periods are from Jan. 1970 to Dec. 2006.

	k	BivariateRegression $dy12$		$ey12$		Test					
		(1)NW	(2)HK	(1)NW	(2)HK	(1)H1	(2)H2	(3)H3	(4)H4	(5)H5	(6)H6
197001	1	−14.783	−11.215	0.2395	0.1945	0.0224	0.0250	0.0311	0.0352	0.0114	0.0897
−200612		(−13.63)**	(−16.34)**	(2.53)*	(1.45)	(0.023)*	(0.022)*	(0.020)*	(0.015)*	(0.034)*	(0.000)**
	12	−0.8311	−0.6834	−0.0099	−0.0043	0.0448	0.0450	0.0209	0.0475	0.0120	0.0241
		(−9.23)**	(−6.80)**	(0.15)	(0.00)	(0.000)**	(0.000)**	(0.025)*	(0.000)**	(0.032)*	(0.022)*
	60	0.0879	0.0632	−0.5493	−0.4222	0.0497	0.0267	0.0509	0.0306	0.0177	0.0137
		(1.07)	(0.54)	(1.74)	(0.35)	(0.000)**	(0.021)*	(0.000)**	(0.021)*	(0.030)*	(0.031)*

Table 4d. Predictability of US Excess Returns(Monthly)

(a)We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being short rate, log dividends yields, and log earnings yields together. T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are monthly. The time periods are from Jan. 1970 to Dec. 2006. (b)The test column reports a p-value for the predictictability test under six hypothesises $H1 - H8$. X_1 represents the short rate r , X_2 represents the dividends yields and X_3 represents the earnings yields. Hypothesis $H1$ $(H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h}))$, Hypothesis $H2$ $(H_0 : E(\varepsilon_{t+h}|X_{1t}) = 0)$,Hypothesis $H3$ $(H_0 : E(Y_{t+h}|X_{1t}) = E(Y_{t+h}))$, Hypothesis $H4$ $(H_0 : E(\varepsilon_{t+h}|X_{2t}) = 0)$, Hypothesis $H5$ $(H_0 : E(Y_{t+h}|X_{3t}) = E(Y_{t+h}))$, Hypothesis $H6$ $(H_0 : E(\varepsilon_{t+h}|X_{3t}) = 0)$,Hypothesis $H7$ $(H_0 : E(Y_{t+h}|X_{1t}, X_{2t}, X_{3t}) = E(Y_{t+h}))$, Hypothesis $H8$ $(H_0 : E(\varepsilon_{t+h}|X_{1t}, X_{2t}, X_{3t}) = 0)$. $^*p < 0.05.^{**}p < 0.01$.

Panel a		BivariateRegression					
	k	r		$dy4$		$ey4$	
		(1)NW	(2)HK	(1)NW	(2)HK	(1)NW	(2)HK
197001	1	−14.799	−10.498	−0.0505	−0.0228	0.1930	0.1590
−200612		(−13.69)**	(−10.83)**	(0.49)	(0.12)	(1.42)	(0.98)
	12	−0.8328	−0.5639	0.1814	0.1434	−0.1779	−0.1034
		(−9.53)**	(−6.34)**	(2.42)*	(1.23)	(−1.85)	(−0.70)
	60	0.1111	0.7381	0.1467	0.1031	−0.6974	−0.4723
		(1.32)	(0.46)	(2.05)*	(1.37)	(−6.62)**	(−2.12)*

Panel b		Test							
		(1)H1	(2)H2	(3)H3	(4)H4	(5)H5	(6)H6	(7)H7	(8)H8
197001	1	0.0224	0.0290	0.0359	0.0376	0.0311	0.0620	0.0116	0.0093
−200612		(0.023)*	(0.019)*	(0.012)*	(0.008)**	(0.020)*	(0.000)**	(0.034)*	(0.044)*
	12	0.0448	0.0569	0.0249	0.0624	0.0209	0.0486	0.0087	0.0100
		(0.000)**	(0.000)**	(0.022)*	(0.000)**	(0.025)*	(0.000)**	(0.045)*	(0.042)*
	60	0.0487	0.0693	0.0230	0.0154	0.0509	0.0558	0.0020	0.0134
		(0.000)**	(0.000)**	(0.021)*	(0.028)*	(0.000)**	(0.000)**	(0.081)	(0.031)*

Table 5. Predictability of US Excess Returns(Annually)

We estimate regressions of the form $Y_{t+h} = \alpha_h + \beta_h' X_t + \varepsilon_{h,t+h}$ where $Y_{t+h} = (\tau/h)[(y_{t+1} - r_t) + \cdots + (\, y_{t+h} - r_{t+h-1})$ is the annualized h-period excess return for the aggregate stock market, r_t is the risk-free rate from t to $t+1$, and $y_{t+1} - r_t$ is the excess one period return from time t to $t+1$, with instruments z_t being log dividend yields, log earnings yields, dividend payout ratio, short rate, inflation, book-to-market ratio(b/m), investment to capital ratio(i/k), corporate issuing activity (Eqis and Ntis), and consumption, wealth, and income ratio(cay). T-statistics in parentheses are computed using Newey West(1987) and Hodrick (1992) standard errors. Horizons h are anuually. The test column reports a p-value for the predictictability test under two hypothesises $H2$ and $H2$. Hypothesis $H2$ is $H_0 : E(Y_{t+h}|X_t) = E(Y_{t+h})$ and Hypothesis $H2$ is $H_0 : E(\varepsilon_{t+h}|X_t) = 0$. $^*p < 0.05$.$^{**}p < 0.01$. In order to compare the prevailing predictive models with the historical mean model, we define a new variable $\Delta(\frac{Q_h}{\sigma^2}) = \hat{Q}_N(h)/\widehat{\sigma}_\varepsilon^2 - \hat{Q}_A(h)/\widehat{\sigma}_\varepsilon^2$, where $\hat{Q}_N(h)/\widehat{\sigma}_\varepsilon^2$ is computed by the historical mean model and $\hat{Q}_A(h)/\widehat{\sigma}_\varepsilon^2$ is computed by the prevailing predictive regression model. We also compute the IS $\overline{R}^2$, OOS $\overline{R}^2$, and $\Delta RMSE$ following the Goyal and Welch (2007)'s definition. The table summarizes the results for both in-sample and out-of-sample tests.

Variable	Data	k	Univariate Regression (1) NW	(2) HK	Test(IS) (1) H1	Test(IS) (2) H2	Test(OS) (1) H1	Test(OS) (2) H2	(1) $\Delta(\frac{Q_h}{\sigma^2})$	(2) $\Delta(\frac{Q_h}{\sigma^2})$	GW IS $\overline{R}^2$	GW OOS $\overline{R}^2$	$\Delta RMSE$ (%)
dy	1872	1	0.0795	0.0783	0.0674	0.0328	0.1613	0.0753	0.0367	0.0058	0.91	−1.93	−0.10
	−2005		$(2.07)^{*}$	(1.65)	$(0.000)^{**}$	$(0.019)^{*}$	$(0.000)^{**}$	$(0.000)^{**}$					
		5	0.0665	0.0634	0.0401	0.0430	0.0963	0.0921	0.0233	0.0183	6.04	−4.45	−0.68
			(1.60)	(1.04)	$(0.014)^{*}$	$(0.011)^{*}$	$(0.000)^{**}$	$(0.000)^{**}$					
ey	1872	1	0.0527	0.0478	0.0861	0.0494	0.1266	0.0238	0.0714	0.0515	1.08	−1.78	−0.08
	−2005		(1.14)	(0.78)	$(0.000)^{**}$	$(0.008)^{**}$	$(0.000)^{**}$	$(0.023)^{*}$					
		5	0.0344	0.0302	0.1055	0.0855	0.1030	0.1081	0.0300	0.0343	6.24	−1.04	−0.03
			(0.80)	(0.46)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
d/e	1872	1	0.0759	0.0696	0.0180	0.0451	0.0760	0.0988	0.0486	0.0235	−0.75	−4.33	−0.31
	−2005		(1.46)	(1.12)	$(0.028)^{*}$	$(0.009)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
		5	0.0698	0.0589	0.0298	0.0361	0.0938	0.0955	0.0208	0.0217	0.66	−4.87	−0.76
			(1.55)	(0.89)	$(0.021)^{*}$	$(0.018)^{*}$	$(0.000)^{**}$	$(0.000)^{**}$					
r	1872	1	−0.7299	−0.6545	0.0232	0.0409	0.1057	0.0706	0.0189	0.0011	0.34	−3.37	−0.14
	−2005		$(-1.95)^{*}$	(−0.98)	$(0.024)^{*}$	$(0.010)^{**}$	$(0.010)^{**}$	$(0.000)^{**}$					
		5	−0.5524	−0.4982	0.0340	0.0343	0.1145	0.1004	0.0415	0.0266	3.83	−17.66	−2.78
			(−1.02)	(−0.76)	$(0.019)^{*}$	$(0.019)^{*}$	$(0.010)^{**}$	$(0.000)^{**}$					
infl	1919	1	−0.3766	−0.3121	0.0409	0.1088	0.0064	0.0563	0.1182	0.0132	−1.00	−4.07	−0.20
	−2005		(−1.18)	(−0.87)	$(0.010)^{**}$	$(0.010)^{**}$	(0.052)	$(0.000)^{**}$					
		5	−0.4034	−0.3441	0.1145	0.1001	0.0877	0.0876	0.0147	0.0138	−1.21	−11.25	−1.70
			(−1.00)	(−0.49)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
b/m	1921	1	0.0641	0.0546	0.0352	0.1014	0.0629	0.1436	0.0617	0.0741	3.20	−1.72	−0.01
	−2005		(1.56)	(1.12)	$(0.018)^{*}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
		5	0.0182	0.0124	0.1317	0.1267	0.1413	0.1324	0.0683	0.0586	10.78	−13.06	−2. − 03
			(0.46)	(0.21)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
i/k	1947	1	0.0355	0.0312	0.1213	0.1035	0.3574	0.4443	0.2328	0.3748	6.63	−1.77	0.07
	−2005		(0.66)	(0.40)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
		5	0.0256	0.0212	0.0799	0.0821	0.1156	0.0732	0.0426	0.0006	33.99	12.99	3.39
			(0.40)	(0.21)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
ntis	1927	1	0.1833	0.1645	0.0177	0.0265	0.0034	0.0234	0.1212	0.0461	8.15	−5.07	−0.26
	−2005		(0.67)	(0.34)	$(0.028)^{*}$	$(0.021)^{*}$	(0.072)	$(0.023)^{*}$					
		5	0.1754	0.1437	0.0729	0.0726	0.0218	0.0060	0.0512	0.0678	6.59	−3.46	−0.32
			(0.65)	(0.45)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.025)^{*}$	$(0.046)^{*}$					
eqis	1927	1	−0.1125	−0.0937	0.0464	0.0426	0.1137	0.0670	0.0109	0.0025	9.15	2.04	0.30
	−2005		$(-2.98)^{*}$	(−1.60)	$(0.005)^{**}$	$(0.008)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
		5	−0.0515	−0.0332	0.1598	0.1599	0.1013	0.0813	0.0283	0.0075	9.50	−2.35	−0.11
			(−1.26)	(−0.90)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					
cay	1945	1	0.0529	0.0445	0.0128	0.1121	0.0066	0.1538	0.1194	0.0843	15.72	16.78	1.61
	−2005		$(4.08)^{**}$	$(3.68)^{**}$	$(0.036)^{*}$	$(0.000)^{**}$	$(0.043)^{*}$	$(0.000)^{**}$					
		5	0.0133	0.0092	0.1085	0.1055	0.1451	0.1647	0.0721	0.0909	36.05	30.35	7.50
			(1.11)	(0.86)	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$	$(0.000)^{**}$					

Table 6a. Equity Premium Out-of-Sample Forecasting Results(Univariate, Quarterly)

Table 6 show the out-of-sample results of the univariate linear predictive models. Table 6a summarize the MSE, MAE, and RMSE of linear predictive regression for dividend yield and earning yield during the period 1936-2001, and 1952-2001. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. Table 6a summarize the statistical results of univariate linear predictive regression for dividend yield and earning yield by using the measure MSE, MAE, and RMSE during the period 1936-2001, and1952-2001.

				HMM	LPM	NPM1	NPM2				HMM	LPM	NPM1	NPM2
1936 −2001	$dy4$	1	MSE	0.1389	0.1481	0.1444	0.1262	$ey4$	1	MSE	0.1389	0.1658	0.1469	0.1259
			MAE	0.2881	0.2803	0.2889	0.2728			MAE	0.2881	0.2972	0.2967	0.2718
			$RMSE$	0.3726	0.3848	0.3799	0.3553			$RMSE$	0.3726	0.4072	0.3833	0.3549
		4	MSE	0.0546	0.0629	0.0576	0.0388		4	MSE	0.0739	0.1005	0.0825	0.0593
			MAE	0.1733	0.1925	0.1897	0.1568			MAE	0.2061	0.2497	0.2243	0.1872
			$RMSE$	0.2336	0.2508	0.2399	0.1970			$RMSE$	0.2713	0.3170	0.2872	0.2436
		12	MSE	0.0297	0.0384	0.0281	0.0136		12	MSE	0.0563	0.0778	0.0593	0.0397
			MAE	0.1423	0.1642	0.1304	0.0939			MAE	0.1893	0.2324	0.1801	0.1406
			$RMSE$	0.1723	0.1961	0.1678	0.1168			$RMSE$	0.2372	0.2789	0.2435	0.1992
		20	MSE	0.0146	0.0380	0.0352	0.0115		20	MSE	0.0537	0.0744	0.0539	0.0377
			MAE	0.1042	0.1665	0.1605	0.0875			MAE	0.1861	0.2330	0.1697	0.1326
			$RMSE$	0.1208	0.1950	0.1876	0.1071			$RMSE$	0.2318	0.2727	0.2323	0.1942
1952 −2001	$dy4$	1	MSE	0.1260	0.1652	0.1613	0.1225	$ey4$	1	MSE	0.1260	0.1355	0.1341	0.1081
			MAE	0.2604	0.2996	0.3024	0.2679			MAE	0.2604	0.2772	0.2864	0.2531
			$RMSE$	0.3549	0.4065	0.4017	0.3500			$RMSE$	0.3549	0.3680	0.3662	0.3288
		4	MSE	0.0435	0.0738	0.0623	0.0358		4	MSE	0.0435	0.0695	0.0661	0.0344
			MAE	0.1589	0.2100	0.1945	0.1490			MAE	0.1589	0.2126	0.2117	0.1480
			$RMSE$	0.2085	0.2716	0.2496	0.1893			$RMSE$	0.2085	0.2637	0.2570	0.1854
		12	MSE	0.0193	0.0453	0.0308	0.0127		12	MSE	0.0193	0.0455	0.0404	0.0148
			MAE	0.1163	0.1783	0.1345	0.0898			MAE	0.1163	0.1741	0.1660	0.0983
			$RMSE$	0.1391	0.2129	0.1755	0.1126			$RMSE$	0.1391	0.2133	0.2010	0.1215
		20	MSE	0.0146	0.0322	0.0167	0.0091		20	MSE	0.0146	0.0208	0.0202	0.0112
			MAE	0.1042	0.1546	0.0989	0.0760			MAE	0.1042	0.1164	0.1140	0.0850
			$RMSE$	0.1208	0.1796	0.1291	0.0952			$RMSE$	0.1208	0.1443	0.1422	0.1057

Table 6b. Equity Premium Out-of-Sample Forecasting Results(Univariate, Quarterly)

Table 6 show the out-of-sample results of the univariate linear predictive models. Table 6b summarize the MSE, MAE, and RMSE of linear predictive regression for dividend yield and earning yield during the period 1936-1990, and 1952-1990. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. Table 6b summarize the statistical results of univariate linear predictive regression for dividend yield and earning yield by using the measure MSE, MAE, and RMSE during the period 1936-1990, and1952-1990.

				HMM	LPM	NPM1	NPM2				HMM	LPM	NPM1	NPM2
1936 −1990	$dy4$	1	MSE	0.1465	0.1515	|0.1256	0.1253	$ey4$	1	MSE	0.1465	0.1738	0.1364	0.1280
			MAE	0.2719	0.2780	0.2752	0.2725			MAE	0.2719	0.3026	0.2894	0.2761
			$RMSE$	0.3829	0.3892	0.3544	0.3539			$RMSE$	0.3829	0.4169	0.3693	0.3577
		4	MSE	0.0597	0.0649	0.0396	0.0362		4	MSE	0.0597	0.0872	0.0479	0.0370
			MAE	0.1873	0.1894	0.1592	0.1502			MAE	0.1873	0.2279	0.1743	0.1515
			$RMSE$	0.2443	0.2547	0.1990	0.1904			$RMSE$	0.2443	0.2952	0.2189	0.1924
		12	MSE	0.0361	0.0474	0.0142	0.0100		12	MSE	0.0361	0.0651	0.0211	0.0105
			MAE	0.1665	0.1862	0.0938	0.0805			MAE	0.1665	0.2176	0.1143	0.0827
			$RMSE$	0.1900	0.2176	0.1191	0.0999			$RMSE$	0.1900	0.2552	0.1452	0.1024
		20	MSE	0.0346	0.0401	0.0123	0.0093		20	MSE	0.0346	0.0639	0.0116	0.0059
			MAE	0.1724	0.1687	0.0885	0.0798			MAE	0.1724	0.2295	0.0835	0.0624
			$RMSE$	0.1861	0.2002	0.1108	0.0965			$RMSE$	0.1861	0.2528	0.1077	0.0768
1952 −1990	$dy4$	1	MSE	0.1520	0.2018	0.1481	0.1464	$ey4$	1	MSE	0.1520	0.1738	0.1318	0.1247
			MAE	0.2861	0.3357	0.3077	0.2997			MAE	0.2861	0.3129	0.2894	0.2786
			$RMSE$	0.3898	0.4492	0.3848	0.3826			$RMSE$	0.3898	0.4168	0.3630	0.3531
		4	MSE	0.0552	0.1054	0.0453	0.0411		4	MSE	0.0552	0.0923	00462	0.0357
			MAE	0.1798	0.2631	0.1664	0.1569			MAE	0.1798	0.2414	0.1665	0.1467
			$RMSE$	0.2349	0.3246	0.2128	0.2026			$RMSE$	0.2349	0.3039	0.2149	0.1890
		12	MSE	0.0234	0.0660	0.0185	0.0098		12	MSE	0.0234	0.0572	0.0204	0.0111
			MAE	0.1248	0.2204	0.1116	0.0811			MAE	0.1248	0.1950	0.1214	0.0876
			$RMSE$	0.1529	0.2569	0.1362	0.0989			$RMSE$	0.1529	0.2393	0.1427	0.1054
		20	MSE	0.0163	0.0489	0.0096	0.0058		20	MSE	0.0163	0.0340	0.0101	0.0078
			MAE	0.1084	0.1994	0.0844	0.0661			MAE	0.1084	0.1578	0.0911	0.0794
			$RMSE$	0.1278	0.2212	0.0978	0.0763			$RMSE$	0.1278	0.1843	0.1003	0.0883

Table 7a. Equity Premium Out-of-Sample Forecasting Results(Bivariate, Quarterly)

Table 7 show the out-of-sample results of the bivariate linear predictive models using quarterly data. Table 7a summarize the MSE, MAE, and RMSE of linear predictive regression for dividend yield and earning yield during the period 1936-2001, and 1952-2001. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. Table 7a summarize the statistical results of bivariate linear predictive regression for dividend yield and earning yield by using the measure MSE, MAE, and RMSE during the period 1936-2001, and1952-2001.

				HMM	LPM	NPM1	NPM2				HMM	LPM	NPM1	NPM2
1936 −2001	$dy4, r$	1	MSE	0.1389	0.1344	0.1310	0.1226	$ey4, r$	1	MSE	0.1389	0.1592	0.1443	0.1236
			MAE	0.2681	0.2705	0.2762	0.2701			MAE	0.2681	0.2917	0.2933	0.2705
			$RMSE$	0.3726	0.3667	0.3619	0.3501			$RMSE$	0.3726	0.3990	0.3799	0.3516
		4	MSE	0.0734	0.0877	0.0699	0.0375		4	MSE	0.0734	0.0917	0.0816	0.0580
			MAE	0.2048	0.2307	0.2098	0.1553			MAE	0.2048	0.2371	0.2210	0.1852
			$RMSE$	0.2710	0.2962	0.2643	0.1936			$RMSE$	0.2710	0.3028	0.2857	0.2407
		12	MSE	0.0540	0.0289	0.0227	0.0130		12	MSE	0.0540	0.0697	0.0598	0.0388
			MAE	0.1860	0.1397	0.1198	0.0915			MAE	0.1860	0.2182	0.1797	0.1386
			$RMSE$	0.2324	0.1700	0.1508	0.1141			$RMSE$	0.2324	0.2640	0.2445	0.1971
		20	MSE	0.0538	0.0145	0.0128	0.0093		20	MSE	0.0538	0.0632	0.0558	0.0370
			MAE	0.1893	0.0988	0.0915	0.0782			MAE	0.1893	0.2126	0.1742	0.1305
			$RMSE$	0.2320	0.1203	0.1130	0.0963			$RMSE$	0.2320	0.2514	0.2362	0.1923
1952 −2001	$dy4, r$	1	MSE	0.1076	0.1445	0.1153	0.1050	$ey4, r$	1	MSE	0.1076	0.2652	0.2466	0.1058
			MAE	0.2460	0.3003	0.2603	0.2508			MAE	0.2460	0.3864	0.3807	0.2511
			$RMSE$	0.3280	0.3802	0.3395	0.3240			$RMSE$	0.3280	0.5149	0.4966	0.3252
		4	MSE	0.0355	0.0690	0.0568	0.0332		4	MSE	0.0355	0.1473	0.1415	0.0335
			MAE	0.1471	0.2143	0.1897	0.1454			MAE	0.1471	0.2863	0.3030	0.1462
			$RMSE$	0.1885	0.2627	0.2383	0.1822			$RMSE$	0.1885	0.3837	0.3761	0.1830
		12	MSE	0.0179	0.0300	0.0283	0.0142		12	MSE	0.0179	0.0541	0.0550	0.0144
			MAE	0.1090	0.1404	0.1347	0.0964			MAE	0.1090	0.1804	0.1831	0.0969
			$RMSE$	0.1340	0.1733	0.1681	0.1190			$RMSE$	0.1340	0.2326	0.2345	0.1201
		20	MSE	0.0126	0.0196	0.0171	0.0110		20	MSE	0.0126	0.0279	0.0270	0.0111
			MAE	0.0930	0.1246	0.1101	0.0858			MAE	0.0930	0.1439	0.1395	0.0863
			$RMSE$	0.1124	0.1402	0.1309	0.1047			$RMSE$	0.1124	0.1670	0.1643	0.1055

Table 7b. Equity Premium Out-of-Sample Forecasting Results(Bivariate, Quarterly)

Table 7 show the out-of-sample results of the bivariate linear predictive models using quarterly data. Table 7a summarize the MSE, MAE, and RMSE of linear predictive regression for dividend yield and earning yield during the period 1936-2001, and 1952-2001. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models. Table 7b summarize the statistical results of bivariate linear predictive regression for dividend yield and earning yield by using the measure MSE, MAE, and RMSE during the period 1936-1990, and1952-1990.

				HMM	LPM	NPM1	NPM2				HMM	LPM	NPM1	NPM2
1936 −1990	$dy4,r$	1	MSE	0.1692	0.1378	0.1177	0.1210	$ey4,r$	1	MSE	0.1692	0.1733	0.1281	0.1217
			MAE	0.2999	0.2664	0.2669	0.2699			MAE	0.2999	0.3028	0.2796	0.2702
			$RMSE$	0.4114	0.3722	0.3431	0.3479			$RMSE$	0.4114	0.4163	0.3679	0.3488
		4	MSE	0.0710	0.1839	0.0698	0.0352		4	MSE	0.0710	0.0803	0.0432	0.0356
			MAE	0.2050	0.3450	0.2047	0.1483			MAE	02050	0.2156	0.1663	0.1491
			$RMSE$	0.2664	0.4288	0.2642	0.1877			$RMSE$	0.2664	0.2834	0.2079	0.1886
		12	MSE	0.0403	0.0885	0.0222	0.0097		12	MSE	0.0403	0.0599	0.0171	0.0098
			MAE	0.1758	0.2596	0.1120	0.0793			MAE	0.1758	0.2102	0.1027	0.0794
			$RMSE$	0.2006	0.2976	0.1491	0.0986			$RMSE$	0.2006	0.2449	0.1309	0.0988
		20	MSE	0.0384	0.0382	0.0078	0.0057		20	MSE	0.0384	0.0490	0.0098	0.0058
			MAE	0.1801	0.1747	0.0685	0.0621			MAE	0.1801	0.1987	0.0767	0.0622
			$RMSE$	0.1959	0.1954	0.0883	0.0758			$RMSE$	0.1959	0.2214	0.0991	0.0762
1952 −1990	$dy4,r$	1	MSE	0.1388	0.1270	0.1245	0.1358	$ey4,r$	1	MSE	0.1388	0.1421	0.1258	0.1198
			MAE	0.2572	0.2782	0.2713	0.2909			MAE	0.2572	0.2831	0.2886	0.2755
			$RMSE$	0.3726	0.3563	0.3528	0.3685			$RMSE$	0.3726	0.3769	0.3547	0.3462
		4	MSE	0.0552	0.0393	0.0328	0.0384		4	MSE	0.0552	0.0639	0.0441	0.0344
			MAE	0.1798	0.1583	0.1469	0.1519			MAE	0.1798	0.1999	0.1696	0.1436
			$RMSE$	0.2349	0.1981	0.1811	0.1960			$RMSE$	0.2349	0.2527	0.2101	0.1855
		12	MSE	0.0234	0.0218	0.0122	0.0093		12	MSE	0.0234	0.0357	0.0122	0.0095
			MAE	0.1248	0.1166	0.0908	0.0789			MAE	0.1248	0.1572	0.0918	0.0796
			$RMSE$	0.1529	0.1477	0.1106	0.0963			$RMSE$	0.1529	0.1889	0.1106	0.0973
		20	MSE	0.0163	0.0141	0.0135	0.0047		20	MSE	0.0163	0.0383	0.0080	0.0056
			MAE	0.1084	0.0995	0.0963	0.0576			MAE	0.1084	0.1744	0.0763	0.0654
			$RMSE$	0.1278	0.1189	0.1163	0.0688			$RMSE$	0.1278	0.1958	0.0896	0.0751

Table 8. Equity Premium Out-of-Sample Forecasting Results (Annually)

Table 8 summarize the out-of-sample results of univariate linear predictive regression for dividend yield and earning yield by using the measure MSE, MAE, and RMSE annually. The predictors are dividend yield (D/Y), earnings-price ratio (E/P), dividend-payout ratio (D/E), book-to-market ratio (B/M), net equity expansion ($NTIS$), treasure bill rate (TBL), Percent Equity Issuing ($eqis$), Consumption income wealth ratio (Cay), inflation ($INFL$), and investment-to-capital ratio (I/K) during the period 1972 and 2005. The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models.

				HMM	LPM	NPM1	NPM2					HMM	LPM	NPM1	NPM2
dy	1872	1	MSE	0.0261	0.0268	0.0261	0.0250	b/m	1921	1	MSE	0.0263	0.0385	0.0410	0.0243
	−2005		MAE	0.1238	0.1284	0.1271	0.1263		−2005		MAE	0.1230	0.1531	0.1568	0.1227
			$RMSE$	0.1615	0.1640	0.1617	0.1581				$RMSE$	0.1621	0.1962	0.2025	0.1559
		5	MSE	0.0249	0.0237	0.0239	0.0236			5	MSE	0.0267	0.0267	0.0279	0.0250
			MAE	0.1221	0.1220	0.1224	0.1148				MAE	0.1251	0.1291	0.1327	0.1275
			$RMSE$	0.1578	0.1540	0.1547	0.1535				$RMSE$	0.1635	0.1633	0.1672	0.1580
ey	1872	1	MSE	0.0261	0.0293	0.0285	0.0259	i/k	1947	1	MSE	0.0257	0.0332	0.0244	0.0227
	−2005		MAE	0.1238	0.1331	0.1355	0.1298		−2005		MAE	0.1213	0.1637	0.1252	0.1226
			$RMSE$	0.1615	0.1711	0.1687	0.1608				$RMSE$	0.1604	0.1822	0.1561	0.1507
		5	MSE	0.0249	0.0233	0.0216	0.0207			5	MSE	0.0266	0.0285	0.0264	0.0259
			MAE	0.1221	0.1199	0.1174	0.1101				MAE	0.1250	0.1476	0.1335	0.1328
			$RMSE$	0.1578	0.1528	0.1469	0.1439				$RMSE$	0.1630	0.1689	0.1625	0.1609
d/e	1872	1	MSE	0.0261	0.0264	0.0262	0.0243	ntis	1927	1	MSE	0.0268	0.0317	0.0261	0.0257
	−2005		MAE	0.1238	0.1255	0.1249	0.1186		−2005		MAE	0.1253	0.1578	0.1338	0.1337
			$RMSE$	0.1615	0.1624	0.1620	0.1558				$RMSE$	0.1637	0.1780	0.1615	0.1604
		5	MSE	0.0249	0.0233	0.0216	0.0207			5	MSE	0.0262	0.0297	0.0277	0.0277
			MAE	0.1221	0.1199	0.1174	0.1101				MAE	0.1239	0.1281	0.1330	0.1331
			$RMSE$	0.1578	0.1528	0.1469	0.1439				$RMSE$	0.1618	0.1722	0.1663	0.1665
r	1872	1	MSE	0.0261	0.0268	0.0261	0.0237	eqis	1927	1	MSE	0.0253	0.0262	0.0248	0.0248
	−2005		MAE	0.1238	0.1284	0.1267	0.1194		−2005		MAE	0.1205	0.1201	0.1262	0.1263
			$RMSE$	0.1615	0.1636	0.1614	0.1539				$RMSE$	0.1591	0.1620	0.1575	0.1577
		5	MSE	0.0249	0.0237	0.0256	0.0236			5	MSE	0.0262	0.0330	0.0262	0.0253
			MAE	0.1221	0.1216	0.1252	0.1182				MAE	0.1239	0.1377	0.1269	0.1249
			$RMSE$	0.1578	0.1539	0.1602	0.1535				$RMSE$	0.1618	0.1816	0.1618	0.1589
infl	1919	1	MSE	0.0249	0.0307	0.0251	0.0239	cay	1945	1	MSE	0.0249	0.0276	0.0231	0.0210
	−2005		MAE	0.1181	0.1308	0.1288	0.1247		−2005		MAE	0.1215	0.1296	0.1158	0.1166
			$RMSE$	0.1577	0.1753	0.1584	0.1547				$RMSE$	0.1578	0.1660	0.1521	0.1450
		5	MSE	0.0268	0.0368	0.0264	0.0257			5	MSE	0.0206	0.0257	0.0247	0.0240
			MAE	0.1242	0.1479	0.1337	0.1271				MAE	0.1191	0.1296	0.1265	0.1329
			$RMSE$	0.1638	0.1917	0.1625	0.1603				$RMSE$	0.1436	0.1604	0.1573	0.1549

Table 9. Equity Premium Out-of-Sample Forecasting Results for Combining Methods (Annually)

Table 9 report the equity premium out-of-sample combined forecasting results using the annual data. The combination forecasts of Y_{t+1} made at time t are weighted averages of the M individual forecasts based on $\widehat{Y}_{c,t+h} = \sum_{i=1}^{M} \omega_{i,t} \widehat{Y}_{i,t+h}$ where $\{\omega_{i,t}\}_{i=1}^{M}$ are the ex ante combining weights formed at time t, and $\widehat{Y}_{i,t+h}$ is the out-of-sample forecast of the equity premium based on the individual predictive models. The first three methods use simple averaging schemes: mean, median, and trimmed mean. Their discount mean square prediction error ($DMSPE$) combining method employs the following weights: $w_{i,t} = \phi_{i,t}^{-1} / \sum_{j=1}^{M} \phi_{j,t}^{-1}, \phi_{i,t} = \sum_{s=R}^{t-1} \theta^{t-1-s} (Y_{i,t+h} - \widehat{Y}_{i,t+h})^2$ and θ is a discount factor. We consider the two values of 1.0 and 0.9 for θ.

	OS			HMM	LPM	NPM1	NPM2		OS			HMM	LPM	NPM1	NPM2
Mean	1965	1	MSE	0.0261	0.0244	0.0251	0.0239	DMSPE	1965	1	MSE	0.0261	0.0246	0.0256	0.0240
	−2005		MAE	0.1238	0.1205	0.1217	0.1206	$\theta = 1.0$	−2005		MAE	0.1238	0.1214	0.1239	0.1212
			$RMSE$	0.1615	0.1563	0.1585	0.1545				$RMSE$	0.1615	0.1568	0.1599	0.1549
		5	MSE	0.0249	0.0240	0.0219	0.0215			5	MSE	0.0249	0.0266	0.0252	0.0214
			MAE	0.1221	0.1168	0.1114	0.1174				MAE	0.1221	0.1275	0.1247	0.1164
			$RMSE$	0.1578	0.1548	0.1479	0.1466				$RMSE$	0.1578	0.1632	0.1589	0.1463
Median	1965	1	MSE	0.0261	0.0233	0.0249	0.0230	DMSPE	1965	1	MSE	0.0261	0.0248	0.0260	0.0240
	−2005		MAE	0.1238	0.1169	0.1190	0.1170	$\theta = 0.9$	−2005		MAE	0.1238	0.1220	0.1248	0.1211
			$RMSE$	0.1615	0.1528	0.1578	0.1517				$RMSE$	0.1615	0.1575	0.1613	0.1549
		5	MSE	0.0249	0.0256	0.0213	0.0213			5	MSE	0.0249	0.0303	0.0288	0.0214
			MAE	0.1221	0.1195	0.1109	0.1178				MAE	0.1221	0.1355	0.1332	0.1162
			$RMSE$	0.1578	0.1600	0.1458	0.1460				$RMSE$	0.1578	0.1740	0.1698	0.1462
Trimmed	1965	1	MSE	0.0261	0.0243	0.0248	0.0233								
Mean	−2005		MAE	0.1238	0.1200	0.1200	0.1187								
			$RMSE$	0.1615	0.1557	0.1575	0.1526								
		5	MSE	0.0249	0.0244	0.0217	0.0213								
			MAE	0.1221	0.1171	0.1117	0.1161								
			$RMSE$	0.1578	0.1562	0.1471	0.1461								

Table 10. Equity Premium Out-of-Sample Forecasting Results (Quarterly)

Table 9 report the equity premium out-of-sample forecasting results using the quarterly data from 1947:1–2007:4. The out-of-sample forecast evaluation periods are 1965:1–2007:4 consistent with Goyal and Welch (2008). Table 8 summarize the MSE, MAE, and RMSE of linear predictive regression for dividend-price ratio (D/P), dividend yield (D/Y), earnings-price ratio (E/P), dividend-payout ratio (D/E), stock variance ($SVAR$), book-to-market ratio (B/M), net equity expansion ($NTIS$), treasure bill rate (TBL), long-term yield (LTY), long-term return (LTR), term spread (TMS), default yield spread (DFY), default return yield (DFR), inflation ($INFL$), and investment-to-capital ratio (I/K). The benchmark model is historical average equity returns. The alternative models are linear predictive model and two nonparametric predictive models.

			HMM	LPM	NPM1	NPM2				HMM	LPM	NPM1	NPM2
1965:1	−2007:4												
D/P	1	MSE	0.0402	0.0548	0.0466	0.0331	LTY	1	MSE	0.0402	0.0475	0.0318	0.0315
		MAE	0.1469	0.1720	0.1659	0.1397			MAE	0.1469	0.1791	0.1405	0.1366
		$RMSE$	0.2004	0.2340	0.2159	0.1819			$RMSE$	0.2004	0.2181	0.1782	0.1776
	4	MSE	0.0235	0.0380	0.0301	0.0157		4	MSE	0.0235	0.0157	0.0119	0.0115
		MAE	0.1137	0.1499	0.1376	0.0984			MAE	0.1137	0.0985	0.0862	0.0851
		$RMSE$	0.1534	0.1950	0.1734	0.1255			$RMSE$	0.1534	0.1254	0.1090	0.1075
	12	MSE	0.0180	0.0318	0.0230	0.0093		12	MSE	0.0180	0.0094	0.0072	0.0070
		MAE	0.1099	0.1443	0.1192	0.0753			MAE	0.1099	0.0756	0.0706	0.0707
		$RMSE$	0.1343	0.1783	0.1516	0.0967			$RMSE$	0.1343	0.0969	0.0847	0.0837
D/Y	1	MSE	0.0402	0.0546	0.0474	0.0331	LTR	1	MSE	0.0402	0.0601	0.0320	0.0310
		MAE	0.1469	0.1708	0.1665	0.1397			MAE	0.1469	0.1906	0.1375	0.1353
		$RMSE$	0.2004	0.2337	0.2176	0.1820			$RMSE$	0.2004	0.2451	0.1789	0.1761
	4	MSE	0.0235	0.0378	0.0298	0.0157		4	MSE	0.0235	0.0488	0.0171	0.0161
		MAE	0.1137	0.1491	0.1372	0.0985			MAE	0.1137	0.1856	0.1005	0.0989
		$RMSE$	0.1534	0.1944	0.1728	0.1255			$RMSE$	0.1534	0.2210	0.1308	0.1269
	12	MSE	0.0180	0.0315	0.0221	0.0093		12	MSE	0.0180	0.0353	0.0097	0.0095
		MAE	0.1099	0.1423	0.1174	0.0753			MAE	0.1099	0.1632	0.0764	0.0758
		$RMSE$	0.1343	0.1775	0.1487	0.0965			$RMSE$	0.1343	0.1879	0.0983	0.0976
E/P	1	MSE	0.0402	0.0594	0.0432	0.0337	TMS	1	MSE	0.0402	0.0879	0.0390	0.0336
		MAE	0.1469	0.1789	0.1591	0.1406			MAE	0.1469	0.2434	0.1544	0.1398
		$RMSE$	0.2004	0.2438	0.2079	0.1835			$RMSE$	0.2004	0.2965	0.1974	0.1833
	4	MSE	0.0235	0.0428	0.0269	0.0159		4	MSE	0.0235	0.0515	0.0159	0.0135
		MAE	0.1137	0.1544	0.1295	0.0994			MAE	0.1137	0.1984	0.0985	0.0933
		$RMSE$	0.1534	0.2069	0.1641	0.1262			$RMSE$	0.1534	0.2269	0.1263	0.1162
	12	MSE	0.0180	0.0361	0.0206	0.0094		12	MSE	0.0180	0.0389	0.0095	0.0086
		MAE	0.1099	0.1482	0.1108	0.0755			MAE	0.1099	0.1785	0.0754	0.0748
		$RMSE$	0.1343	0.1899	0.1435	0.0969			$RMSE$	0.1343	0.1973	0.0974	0.0925
D/E	1	MSE	0.0402	0.0632	0.0340	0.0335	DFY	1	MSE	0.0402	0.0741	0.0364	0.0339
		MAE	0.1469	0.1947	0.1418	0.1409			MAE	0.1469	0.2130	0.1464	0.1414
		$RMSE$	0.2004	0.2515	0.1844	0.1831			$RMSE$	0.2004	0.2723	0.1908	0.1840
	4	MSE	0.0235	0.0526	0.0179	0.0159		4	MSE	0.0235	0.0581	0.0189	0.0159
		MAE	0.1137	0.1923	0.1040	0.0987			MAE	0.1137	0.2015	0.1078	0.0993
		$RMSE$	0.1534	0.2292	0.1338	0.1262			$RMSE$	0.1534	0.2410	0.1375	0.1262
	12	MSE	0.0180	0.0434	0.0114	0.0095		12	MSE	0.0180	0.0417	0.0106	0.0094
		MAE	0.1099	0.1819	0.0815	0.0756			MAE	0.1099	0.1774	0.0814	0.0756
		$RMSE$	0.1343	0.2083	0.1069	0.0972			$RMSE$	0.1343	0.2042	0.1030	0.0968

			HMM	LPM	NPM1	NPM2				HMM	LPM	NPM1	NPM2
1965:1	−2007:4												
SVAR	1	MSE	0.0402	0.0672	0.0338	0.0337	DFR	1	MSE	0.0402	0.0591	0.0340	0.0338
		MAE	0.1469	0.1984	0.1438	0.1417			MAE	0.1469	0.1870	0.1419	0.1419
		$RMSE$	0.2004	0.2592	0.1840	0.1837			$RMSE$	0.2004	0.2431	0.1845	0.1840
	4	MSE	0.0235	0.0482	0.0163	0.0161		4	MSE	0.0235	0.0425	0.0161	0.0161
		MAE	0.1137	0.1820	0.1010	0.0994			MAE	0.1137	0.1699	0.0994	0.0994
		$RMSE$	0.1534	0.2194	0.1277	0.1267			$RMSE$	0.1534	0.2063	0.1269	0.1269
	12	MSE	0.0180	0.0388	0.0095	0.0095		12	MSE	0.0180	0.0349	0.0096	0.0095
		MAE	0.1099	0.1723	0.0767	0.0762			MAE	0.1099	0.1619	0.0767	0.0769
		$RMSE$	0.1343	0.1969	0.0976	0.0976			$RMSE$	0.1343	0.1869	0.0978	0.0977
B/M	1	MSE	0.0402	0.0845	0.0815	0.0331	INFL	1	MSE	0.0402	0.0572	0.0338	0.0329
		MAE	0.1469	0.2254	0.2258	0.1397			MAE	0.1469	0.1831	0.1415	0.1355
		$RMSE$	0.2004	0.2907	0.2855	0.1820			$RMSE$	0.2004	0.2392	0.1840	0.1814
	4	MSE	0.0235	0.0718	0.0686	0.0157		4	MSE	0.0235	0.0411	0.0161	0.0156
		MAE	0.1137	0.2243	0.2147	0.0984			MAE	0.1137	0.1667	0.0992	0.0978
		$RMSE$	0.1534	0.2679	0.2619	0.1254			$RMSE$	0.1534	0.2026	0.1271	0.1251
	12	MSE	0.0180	0.0619	0.0572	0.0093		12	MSE	0.0180	0.0361	0.0100	0.0095
		MAE	0.1099	0.2134	0.1984	0.0751			MAE	0.1099	0.1643	0.0769	0.0754
		$RMSE$	0.1343	0.2488	0.2392	0.0964			$RMSE$	0.1343	0.1901	0.0999	0.0975
NTIS	1	MSE	0.0402	0.0580	0.0336	0.0334	I/K	1	MSE	0.0402	0.0457	0.0333	0.0315
		MAE	0.1469	0.1835	0.1412	0.1408			MAE	0.1469	0.1603	0.1402	0.1338
		$RMSE$	0.2004	0.2408	0.1832	0.1827			$RMSE$	0.2004	0.2138	0.1825	0.1774
	4	MSE	0.0235	0.0392	0.0164	0.0159		4	MSE	0.0235	0.0358	0.0158	0.0149
		MAE	0.1137	0.1584	0.0998	0.0990			MAE	0.1137	0.1515	0.0983	0.0961
		$RMSE$	0.1534	0.1979	0.1280	0.1259			$RMSE$	0.1534	0.1891	0.1255	0.1219
	12	MSE	0.0180	0.0291	0.0108	0.0094		12	MSE	0.0180	0.0313	0.0094	0.0090
		MAE	0.1099	0.1399	0.0833	0.0766			MAE	0.1099	0.1515	0.0756	0.0749
		$RMSE$	0.1343	0.1707	0.1040	0.0971			$RMSE$	0.1343	0.1768	0.0969	0.0949
TBL	1	MSE	0.0402	0.0341	0.0318	00268							
		MAE	0.1469	0.1457	0.1362	0.1244							
		$RMSE$	0.2004	0.1847	0.1783	0.1636							
	4	MSE	0.0235	0.0155	0.0159	0.0103							
		MAE	0.1137	0.0914	0.0984	0.0801							
		$RMSE$	0.1534	0.1245	0.1260	0.1015							
	12	MSE	0.0180	0.0127	0.0095	0.0059							
		MAE	0.1099	0.0885	0.0749	0.0608							
		$RMSE$	0.1343	0.1129	0.0973	0.0767							

Table 11. Equity Premium Out-of-Sample Forecasting Results for Combining Methods (Quarterly)

Table 11 report the equity premium out-of-sample combined forecasting results using the quarterly data. The combination forecasts of Y_{t+1} made at time t are weighted averages of the M individual forecasts based on $\widehat{Y}_{c,t+h} = \sum_{i=1}^{M} \omega_{i,t} \widehat{Y}_{i,t+h}$ where $\{\omega_{i,t}\}_{i=1}^{M}$ are the ex ante combining weights formed at time t, and $\widehat{Y}_{i,t+h}$ is the out-of-sample forecast of the equity premium based on the individual predictive models. The first three methods use simple averaging schemes: mean, median, and trimmed mean. Their discount mean square prediction error ($DMSPE$) combining method employs the following weights: $w_{i,t} = \phi_{i,t}^{-1} / \sum_{j=1}^{M} \phi_{j,t}^{-1}, \phi_{i,t} = \sum_{s=R}^{t-1} \theta^{t-1-s} (Y_{i,t+h} - \widehat{Y}_{i,t+h})^2$ and θ is a discount factor. We consider the two values of 1.0 and 0.9 for θ.

	OS			HMM	LPM	NPM1	NPM2		OS			HMM	LPM	NPM1	NPM2
Mean	1965:1	1	MSE	0.0402	0.0290	0.0312	0.0233	DMSPE	1965:1	1	MSE	0.0402	0.0260	0.0303	0.0223
	−2007:4		MAE	0.1469	0.1411	0.1462	0.1175	$\theta = 1.0$	−2007:4		MAE	0.1469	0.1315	0.1442	0.1174
			$RMSE$	0.2004	0.1703	0.1766	0.1494				$RMSE$	0.2004	0.1611	0.1740	0.1494
		4	MSE	0.0235	0.0250	0.0273	0.0228			4	MSE	0.0235	0.0240	0.0249	0.0214
			MAE	0.1137	0.1283	0.1354	0.1190				MAE	0.1137	0.1248	0.1280	0.1137
			$RMSE$	0.1534	0.1580	0.1652	0.1511				$RMSE$	0.1534	0.1548	0.1578	0.1464
		12	MSE	0.0180	0.0126	0.0139	0.0036			12	MSE	0.0180	0.0079	0.0086	0.0036
			MAE	0.1099	0.1055	0.1115	0.0522				MAE	0.1099	0.0808	0.0840	0.0504
			$RMSE$	0.1343	0.1122	0.1181	0.0601				$RMSE$	0.1343	0.0888	0.0925	0.0596
Median	1965:1	1	MSE	0.0402	0.0254	0.0275	0.0224	DMSPE	1965:1	1	MSE	0.0402	0.0248	0.0260	0.0221
	−2007:4		MAE	0.1469	0.1305	0.1356	0.1178	$\theta = 0.9$	−2007:4		MAE	0.1469	0.1280	0.1306	0.1165
			$RMSE$	0.2004	0.1594	0.1658	0.1497				$RMSE$	0.2004	0.1575	0.1611	0.1487
		4	MSE	0.0235	0.0233	0.0245	0.0208			4	MSE	0.0235	0.0229	0.0239	0.0212
			MAE	0.1137	0.1225	0.1266	0.1115				MAE	0.1137	0.1208	0.1246	0.1126
			$RMSE$	0.1534	0.1527	0.1564	0.1444				$RMSE$	0.1534	0.1513	0.1544	0.1455
		12	MSE	0.0180	0.0080	0.0095	0.0034			12	MSE	0.0180	0.0063	0.0074	0.0036
			MAE	0.1099	0.0819	0.0910	0.0506				MAE	0.1099	0.0715	0.0778	0.0519
			$RMSE$	0.1343	0.0893	0.0977	0.0586				$RMSE$	0.1343	0.0795	0.0860	0.0598
Trimmed	1965:1	1	MSE	0.0402	0.0282	0.0303	0.0224								
Mean	−2007:4		MAE	0.1469	0.1393	0.1442	0.1177								
			$RMSE$	0.2004	0.1680	0.1740	0.1496								
		4	MSE	0.0235	0.0245	0.0259	0.0218								
			MAE	0.1137	0.1269	0.1316	0.1153								
			$RMSE$	0.1534	0.1565	0.1611	0.1477								
		12	MSE	0.0180	0.0111	0.0123	0.0036								
			MAE	0.1099	0.0980	0.1040	0.0517								
			$RMSE$	0.1343	0.1052	0.1108	0.0596								

Figure 1 US Excess Returns, Interest Rates, Dividend Yield, and Earnings Yield

We plot excess returns, interest rate, dividend yields, and earnings yields from March 1936 to December 2001 quarterly. (March 1936-December 2001 Quarterly)

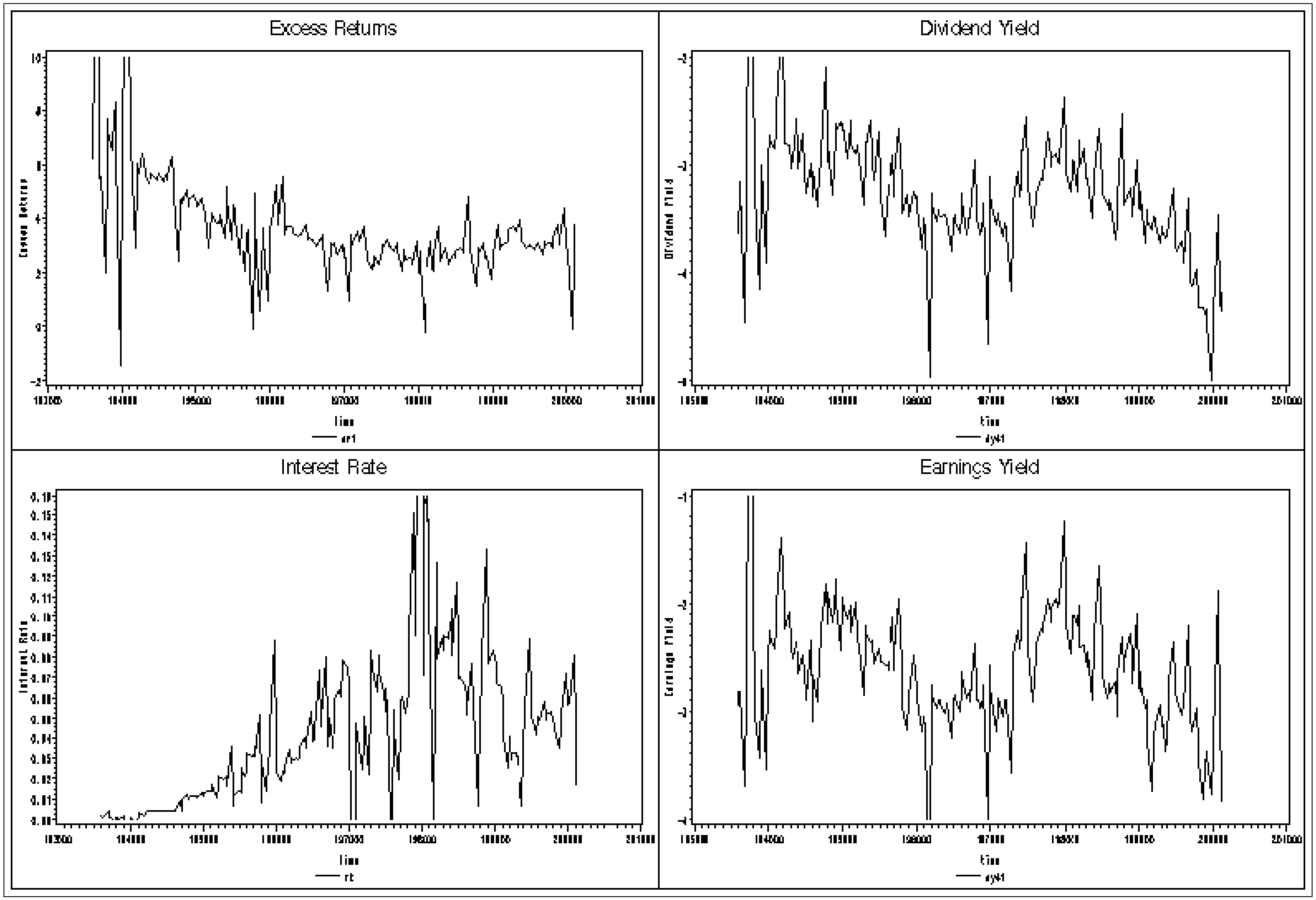

Figure 2 US Dividend Payout Raio, Inflation, Short Rate, and Book-to-Market Ratio (Annually)

We plot dividend payout ratio, short rate, inflation, and book-to-market ratio from 1872 to 2005 annually.

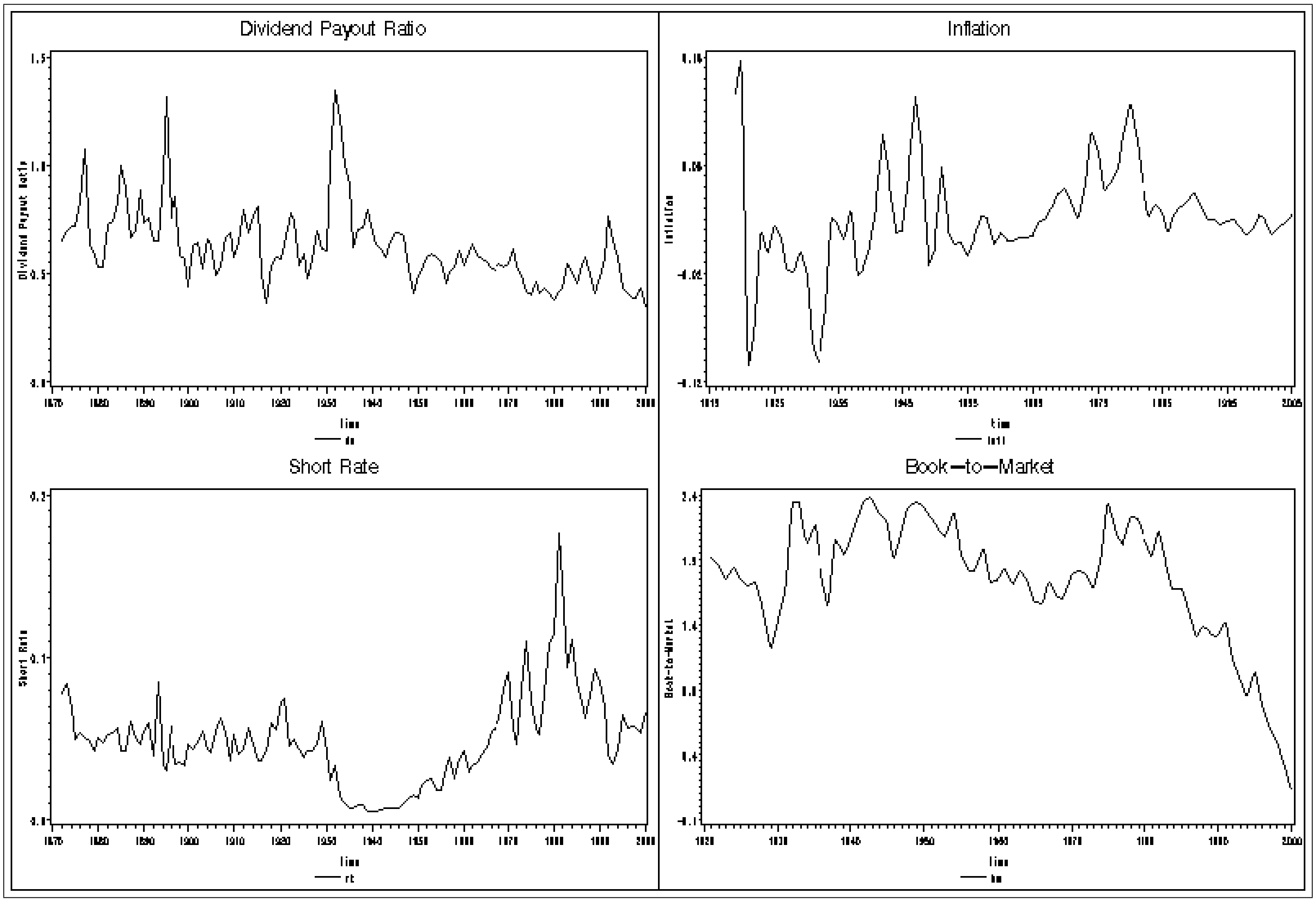

Figure 3 US Investment to Capital Raio, Corporate Issuing Activity, and Consumtion Income Ratio (Annually)

We plot investment to capital ratio(i/k), corporate issuing activity (Eqis and Ntis), and consumption, wealth, and income ratio(cay) annually.

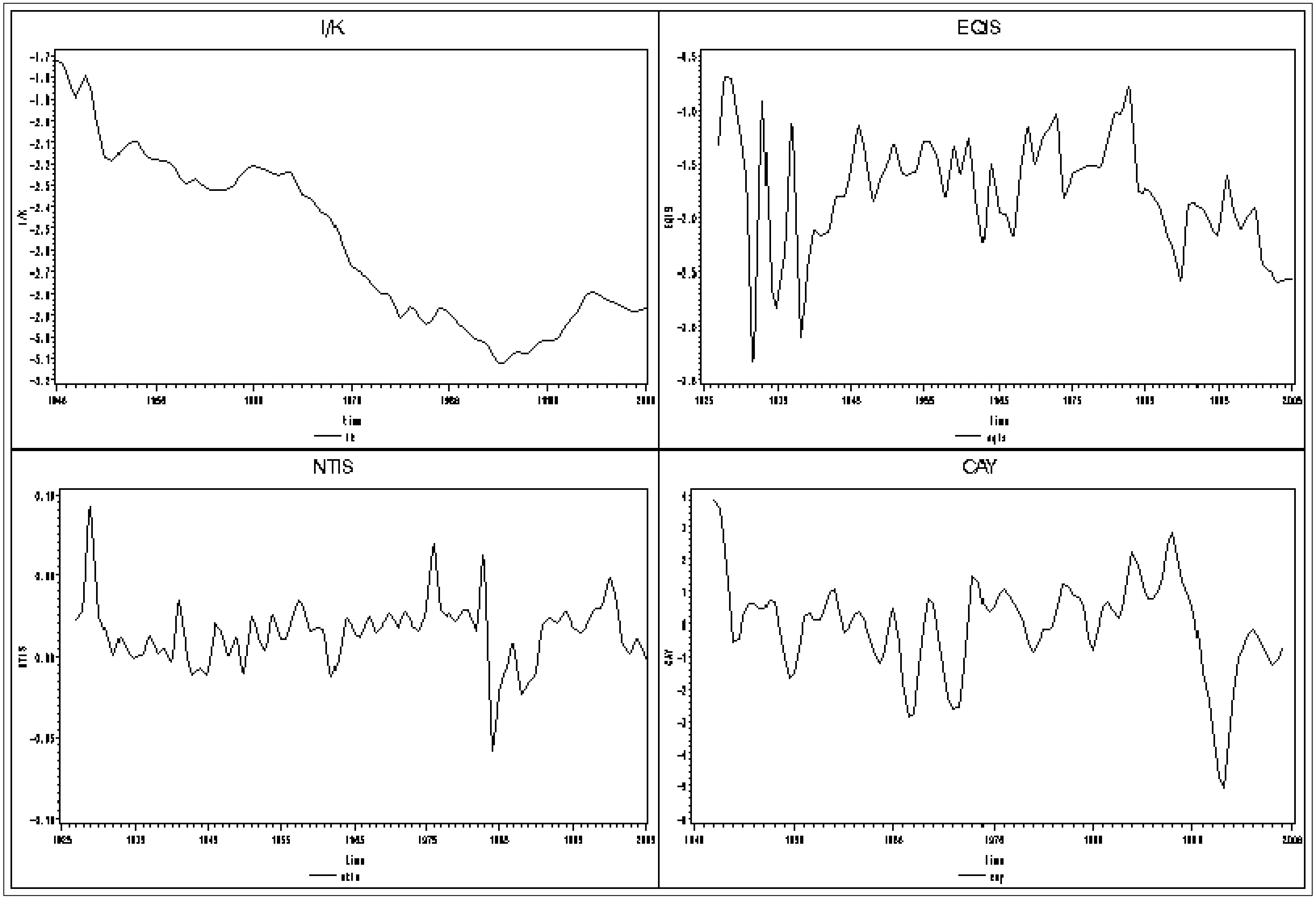

Figure 4 Equity Premium Out-of-Sample Forecasting Results for Individual Methods (Annually)
Figure 4 illustrate the out-of-sample performance for annual predictive regressions for individual methods using annual data.

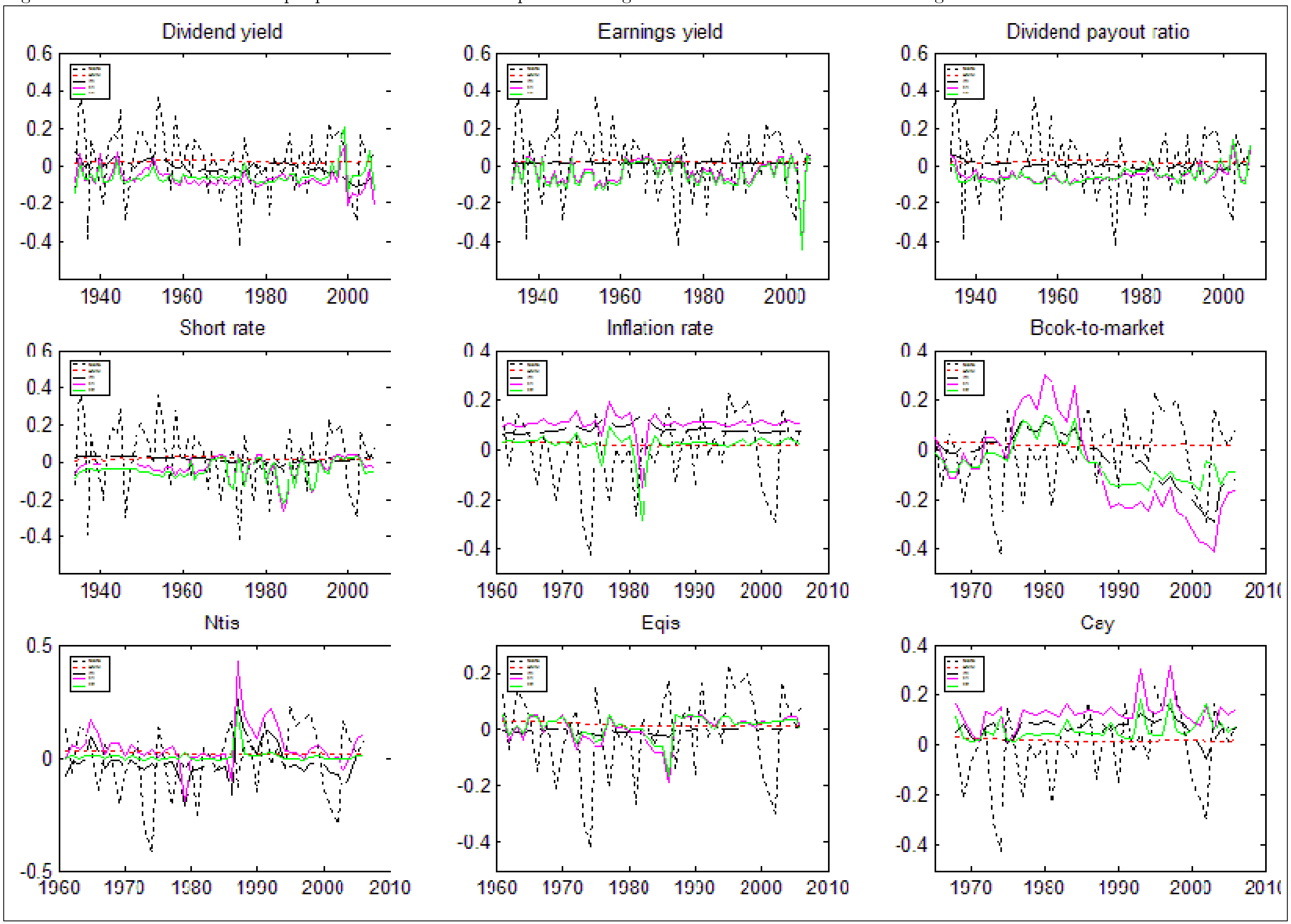

Figure 5 Equity Premium Out-of-Sample Forecasting Results for Combined Methods (Annually)

Figure 5 illustrate the out-of-sample combined performance for annual predictive regressions for combined methods using annual data.

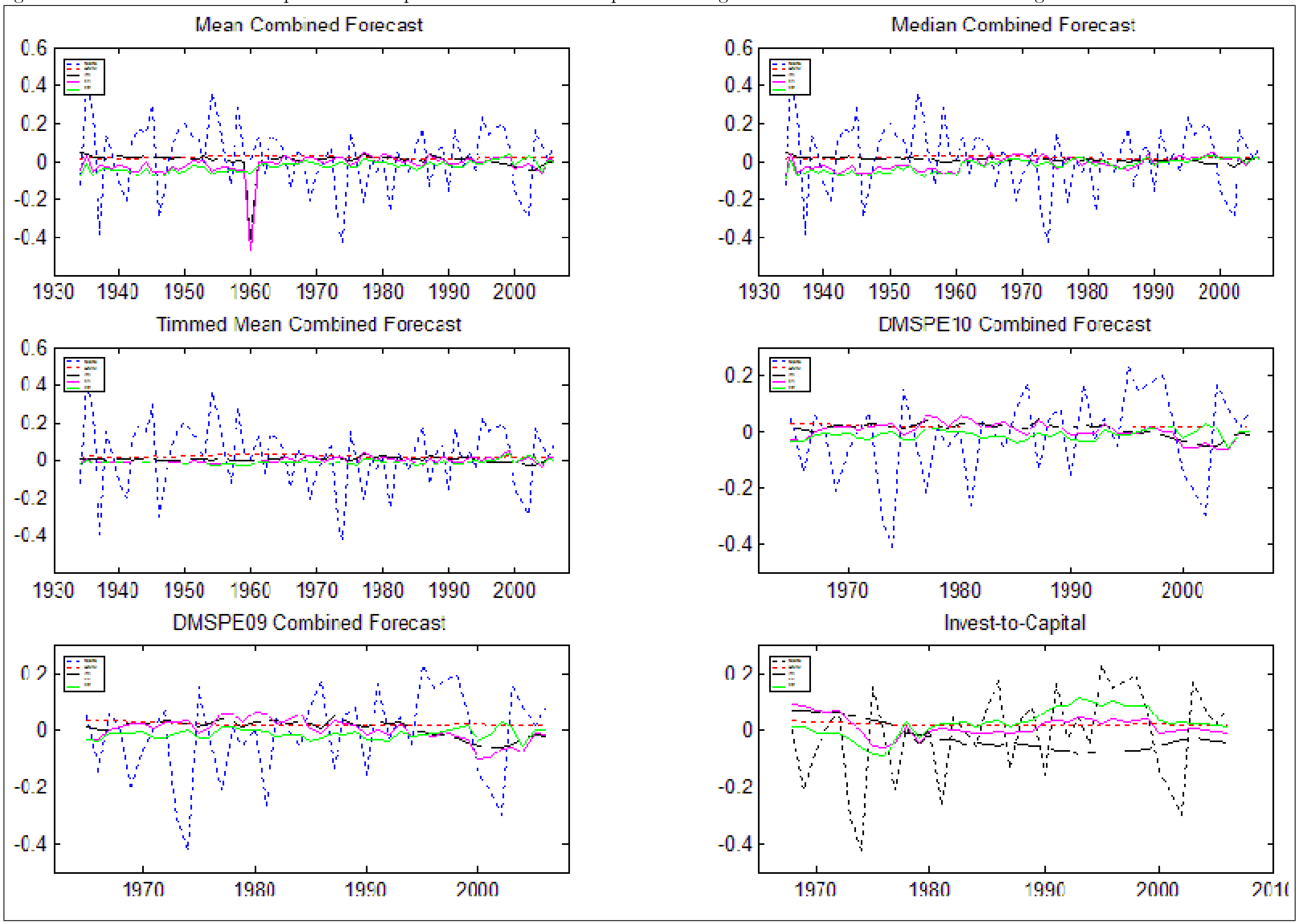

Figure 6 Equity Premium Out-of-Sample Forecasting Results for Individual Methods (Annually, 5-year ahead)
Figure 6 illustrate the out-of-sample performance for annual predictive regressions for individual methods using annual data over 5-year rolling window.

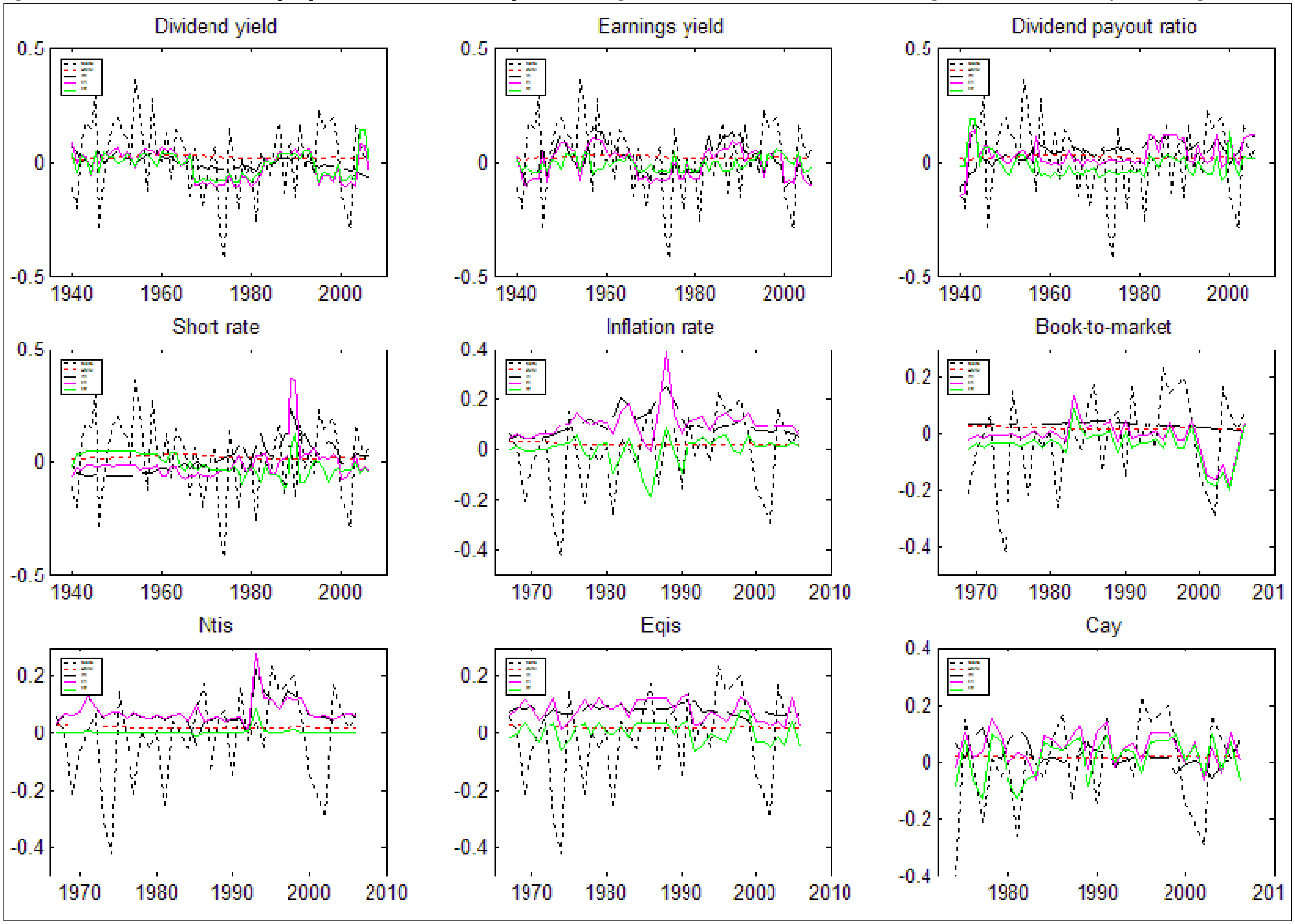

Figure 7 Equity Premium Out-of-Sample Forecasting Results for Combined Methods (Annually, 5-year Ahead)

Figure 7 illustrate the out-of-sample performance for annual predictive regressions for combined methods using annual data over 5-year rolling window.

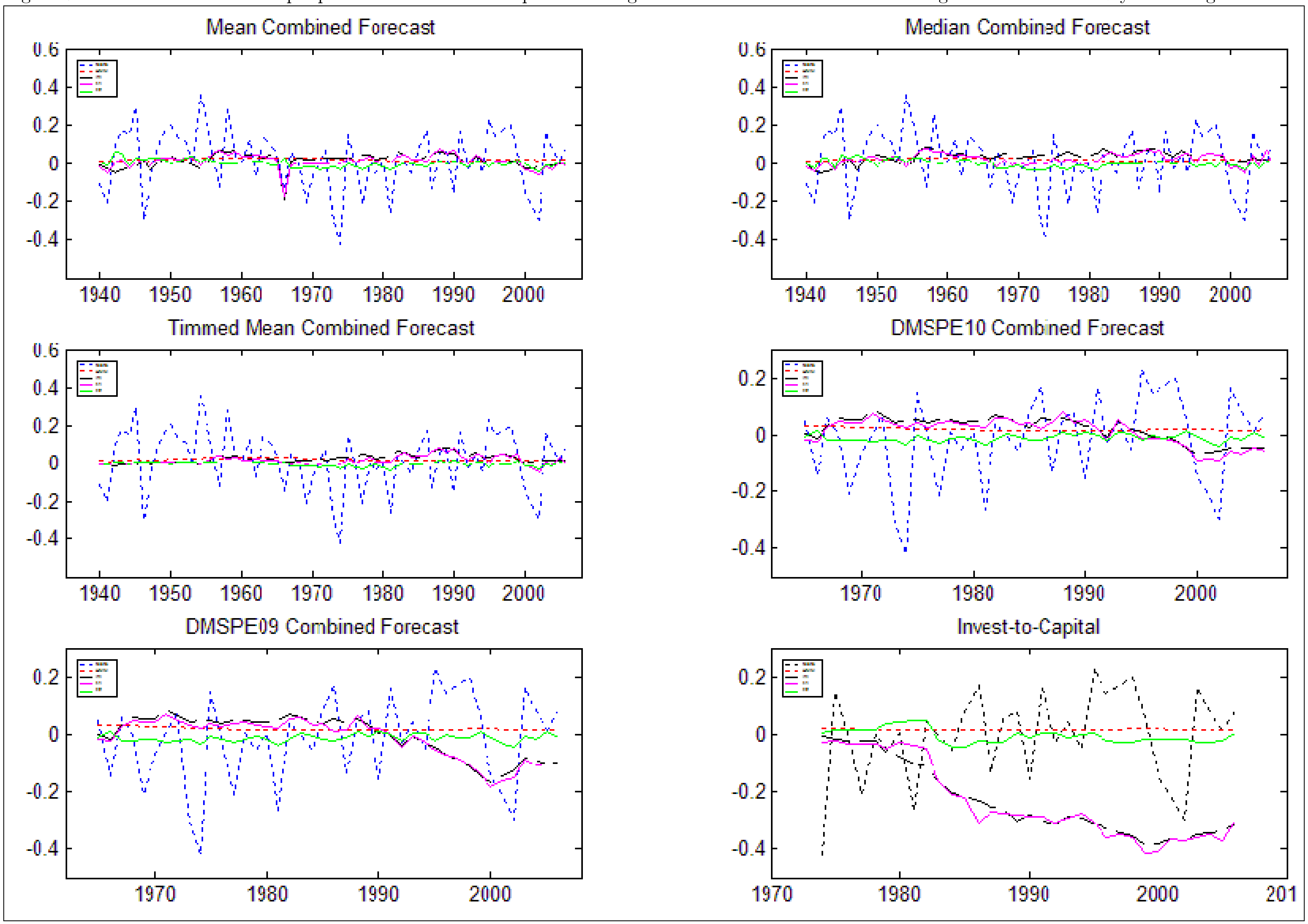

Figure 8 Equity Premium Out-of-Sample Forecasting Results for Individual Methods (Quarterly, 1-period ahead)

Figure 8 illustrate the out-of-sample performance for quarterly predictive regressions for individual methods over 1-quarter rolling window.

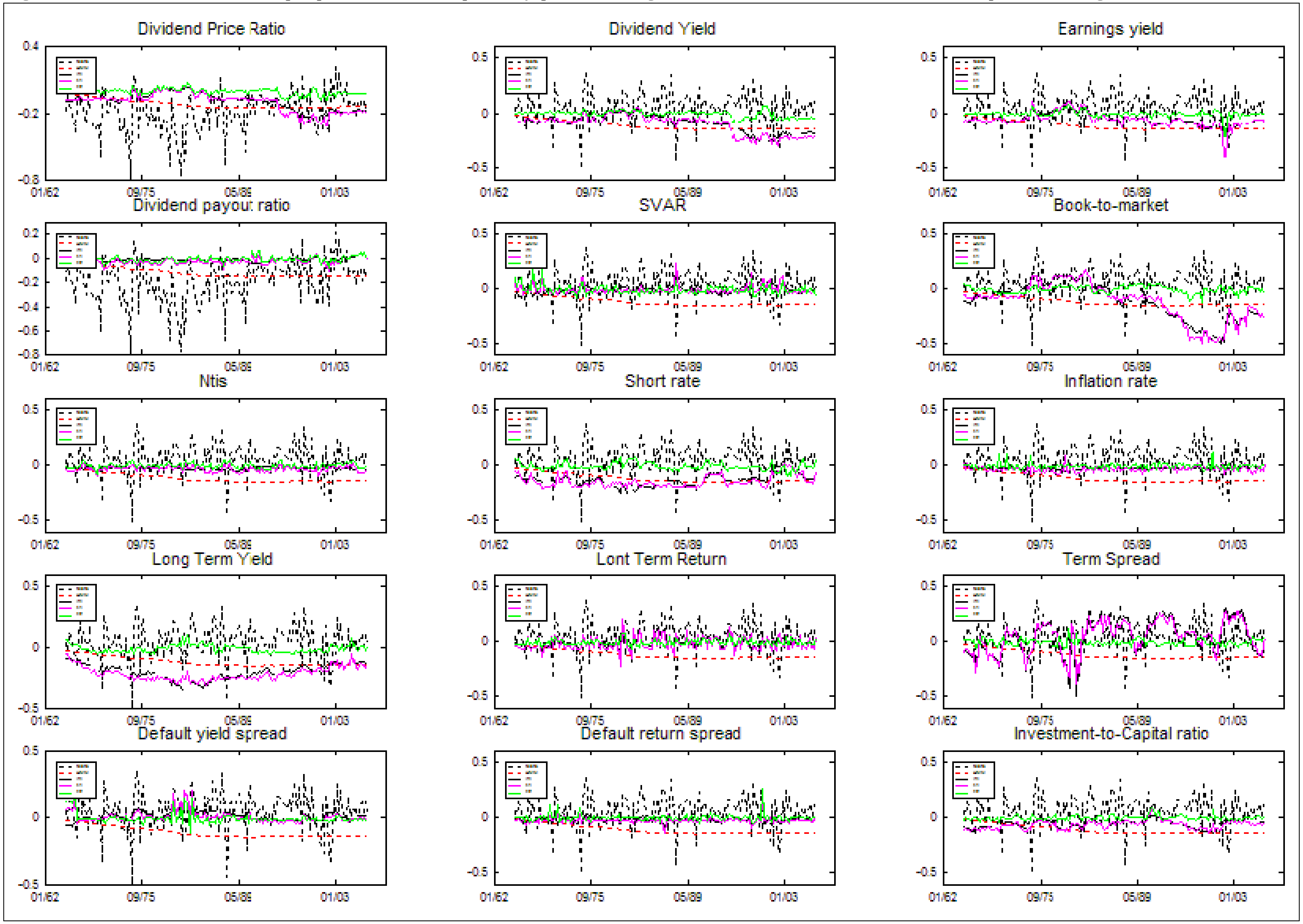

Figure 9 Equity Premium Out-of-Sample Forecasting Results for Combined Methods (Quarterly, 1-Period ahead)

Figure 9 illustrate the out-of-sample performance for quarterly predictive regressions for combined methods over 1-quarter rolling window.

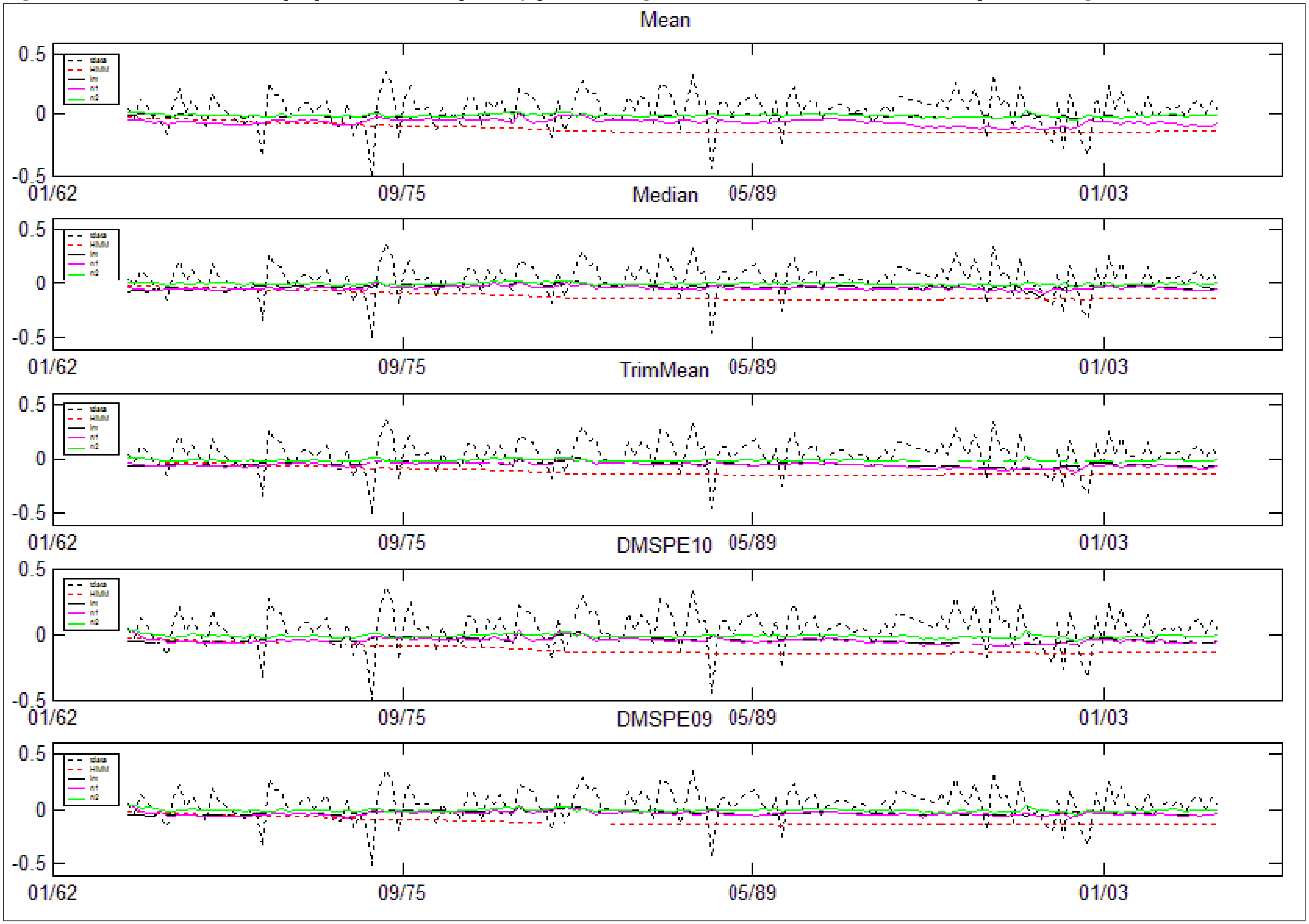

Figure 10 Equity Premium Out-of-Sample Forecasting Results for Individual Methods (Quarterly, 4-period ahead)

Figure 10 illustrate the out-of-sample performance for quarterly predictive regressions for individual methods over 4-quarter rolling window.

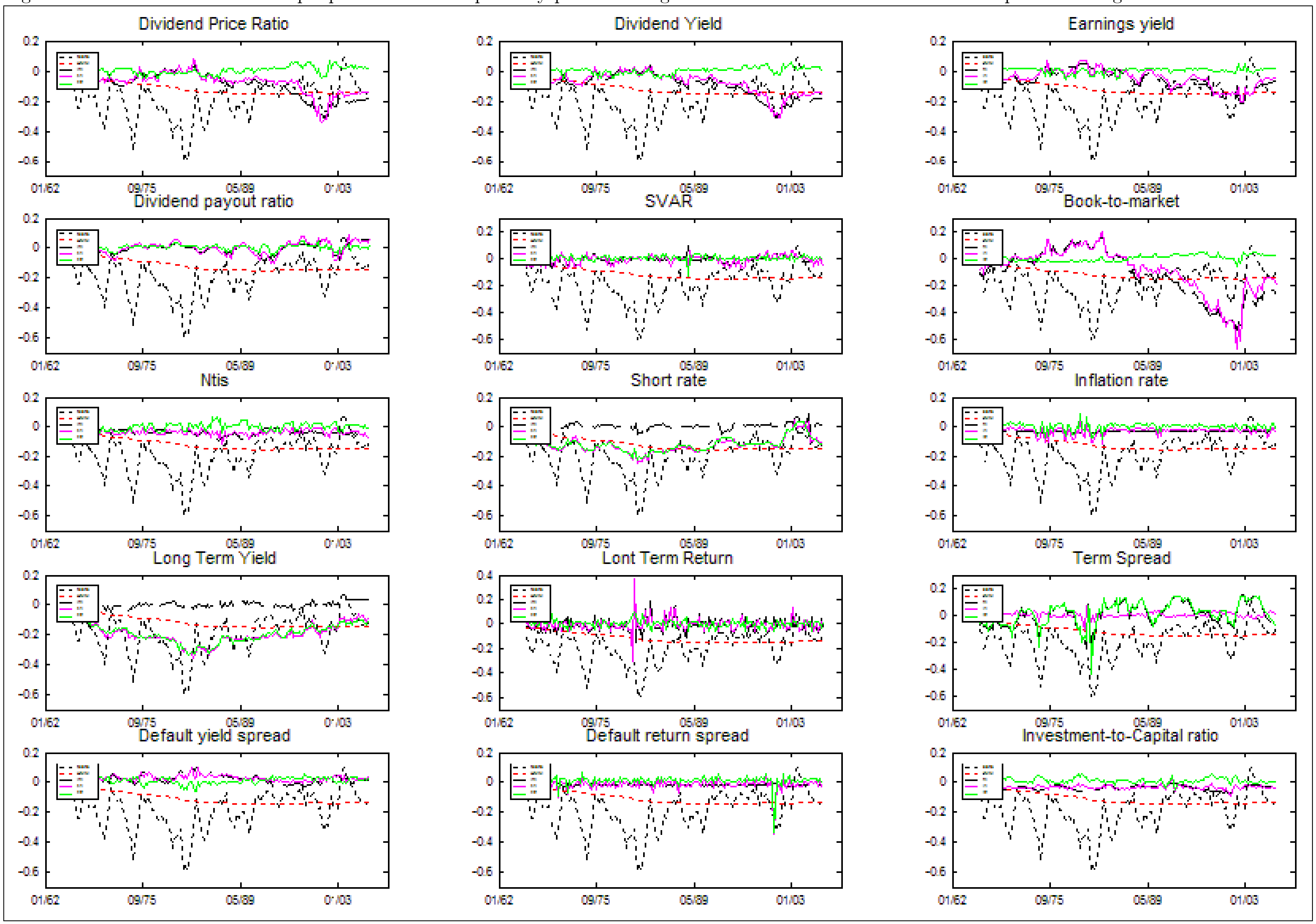

Figure 11 Equity Premium Out-of-Sample Forecasting Results for Combined Methods (Quarterly, 4-Period ahead)

Figure 11 illustrate the out-of-sample performance for quarterly predictive regressions for combined methods over 4-quarter rolling window.

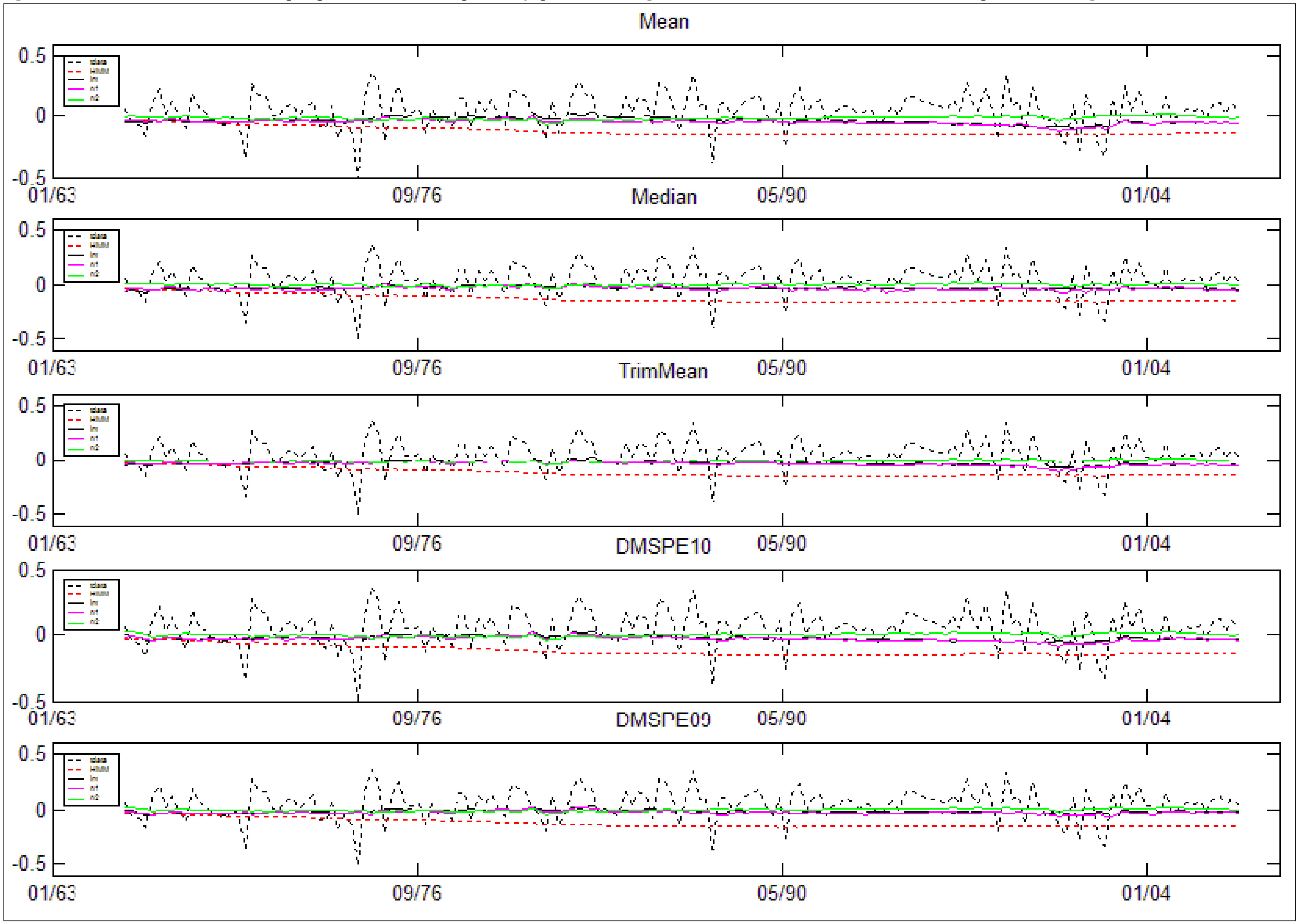

Figure 12 Equity Premium Out-of-Sample Forecasting Results for Individual Methods (Quarterly, 12-period ahead)
Figure 12 illustrate the out-of-sample performance for quarterly predictive regressions for individual methods over 12-quarter rolling window.

Figure 13 Equity Premium Out-of-Sample Forecasting Results for Combined Methods (Quarterly, 12-Period ahead)

Figure 13 illustrate the out-of-sample performance for quarterly predictive regressions for combined methods over 12-quarter rolling window.