CORE Discussion Paper 2006/88

International Stock Return Predictability: Statistical Evidence and Economic Significance

Pierre Giot

Mikael Petitjean*

September 28, 2006

ABSTRACT

The predictability of stock returns in ten countries is assessed taking into account recently developed out-of-sample statistical tests and risk-adjusted metrics. Predictive variables include both valuation ratios and interest rate variables. Out-of-sample predictive power is found to be greatest for the short-term and long-term interest rate variables. Given the importance of trading profitability in assessing market efficiency, we show that such statistical predictive power is economically meaningless across countries and investment horizons. All in all, no common pattern of stock return predictability emerges across countries, be it on statistical or economic grounds.

*Pierre Giot is Professor of Finance at the Department of Business Administration & CeReFiM (FUNDP, University of Namur) as well as CORE Research Member at Université Catholique de Louvain, Belgium. Phone: +32 (0) 81 724887. Email: pierre.giot@fundp.ac.be. Mikael Petitjean is Assistant Professor of Finance at FUCAM (Catholic University of Mons), 151 Chaussée de Binche, 7000 Mons, Belgium. Phone: +32 (0)65 323381. Send all correspondence to mikael.petitjean@fucam.ac.be. We thank Chris Brooks, Amit Goyal, Georges Hübner, Sébastien Laurent, Jonathan Lewellen, Chris Neely, and Dave Rapach for their help. The usual disclaimer applies.

1 Introduction

Numerous studies have investigated the predictability of US stock returns using regression models. In these models, single- or multi-period real (excess) stock returns are regressed on variables deemed relevant. Early work such as Fama and French (1988, 1989) and Campbell and Shiller (1988) generally find strong evidence of stock return predictability. However, two serious econometric problems are ignored in these seminal papers: the small sample bias and the overlapping observations problem.[1] The overlapping observations problem occurs when multi-period returns are computed, which leads to serial correlation in the disturbance term. The small-sample bias (i.e. the bias in finite samples of the predictive variable's slope estimate) is due to the near persistence and/or endogeneity of the predictive variable.[2] Consequently, the distribution of the t-statistic for the predictive variable's slope estimate deviates from its usual form. Basing inferences on standard asymptotic results can therefore lead to considerable size distortions when testing the null hypothesis of no predictability.

While these econometric problems have been somewhat addressed, the vast majority of papers still rely exclusively on in-sample (IS) tests of stock return predictability. This raises concerns of data mining, also referred to as model overfitting or data snooping. Out-of-sample (OOS) tests generally address data mining issues, as the statistical models are tested using OOS observations (i.e. observations that are not used in the estimation of the model itself). The few recent studies that use OOS tests of return predictability provide contradictory evidence. While Rapach and Wohar (2006) and Campbell and Thompson (2005) find some OOS evidence of predictability, Goyal and Welch (2006) do not. All in all, it is challenging to argue for stock return predictability in the US.

There are surprisingly few studies of stock returns predictability that deal with non-US data. Besides, most of these rely on non-robust econometric methods. Harvey (1991,1995) and Ferson and Harvey (1993) consider various aspects of predictability in international stock returns. The datasets span a fair amount of countries, but the overall time span is limited to the 1969-1992 period and most results are thus based on regressions that use less than 20 years of data. In addi-

[1] See Mankiw and Shapiro (1986), Nelson and Kim (1993), Goetzman and Jorion (1993), Stambaugh (1999), Lewellen (2004), Campbell and Yogo (2006), and Moon, Rubia, and Valkanov (2006).

[2] This problem is acute for valuation ratios, like the dividend- and earnings-price ratios.

tion, no robust econometric methods are used. Ang and Bekaert (2004) analyze predictability in stock returns for four different countries in addition to the U.S. Their international sample dates back to 1975.[3] Polk, Thompson, and Vuolteenaho (2006) use an international sample consisting of 22 countries with observations dating back to 1975. However, they only analyze the predictive ability of their cross-sectional beta-premiums and do not consider traditional forecasting variables. Hjalmarsson (2004) and Rangvid, Rapach, and Wohar (2005) are the only up-to-date econometric studies that test the predictive ability of forecasting variables outside the US. Both papers find that interest rate variables have some ability in predicting international stock returns. However, Hjalmarsson (2004) only considers four forecasting variables, while Rangvid, Rapach, and Wohar (2005) exclusively focus on macro variables and do not consider valuation ratios. In addition, no study on international data explicitly addresses the issue of the small-sample bias in the OLS slope estimate by using either the Stambaugh (1999) or Lewellen (2004) correction approach.

This brief review shows that the extant literature is mainly concerned with testing for the existence of return predictability in population.[4] Another issue is whether a practitioner could have constructed a portfolio that earns above average, raw or risk-adjusted returns. A practitioner does not necessarily care about the difference in statistical metrics: he is mostly interested in selecting the model that generates the highest raw or risk-adjusted returns according to some pre-defined trading rule and profit-based metric.

The very few papers that pose the question of whether out-of-sample predictive power is economically meaningful suggest ‘thought experiments’ only. Campbell and Thompson (2005) use a certainty equivalence measure to evaluate the OOS predictive gains for a log-utility investor. By imposing a utility function, both an increase in the average return and a decrease in the average risk of a portfolio can bring welfare gain to a risk-averse investor. However, additional ex-ante restrictions are required (e.g. the risk-aversion parameter and the maximum degree of investment leverage). In particular, Campbell and Thompson (2005) calculate the welfare benefits for an investor with relative aversion of three ($\gamma = 3$). They also impose ‘realistic’ portfolio con-

[3]They advocate the use of Hodrick (1992) auto-correlation robust standard errors. However, these rely on the regressors being covariance stationary, which is usually a restrictive assumption for forecasting variables like interest rates or valuation ratios that are typically modeled as being nearly persistent processes.

[4]As emphasized by Fama and French (1989) and Fama (1991), return predictability in population does not imply that markets are inefficient, as time-varying returns may be an equilibrium phenomenon.

straints, preventing the investor from shorting stocks or taking more than 50% leverage, that is, confining the portfolio weight on stocks to lie between 0% and 150%. Goyal and Welch (2006) follow Campbell and Thompson's approach, but they also measure the economic gains of OOS predictive power through the root mean squared error metric.

In this paper, we re-examine the predictability of stock returns from an international perspective. Ten countries are considered: Australia, Canada, France, Germany, Japan, The Netherlands, South Africa, Sweden, the UK and the US. Our analysis employs a predictive regression framework, with samples of monthly data ending in late 2005 and beginning in the early 50's or late 60's. Five forecasting variables are considered, including both valuation ratios and interest rate variables. The two valuation ratios are the dividend-yield and the earnings-price ratios. They measure stock prices relative to fundamentals. Because each features a price in the denominator, the ratios may be positively related to expected returns. According to the mispricing view, the ratios are low when stocks are overpriced; they predict low future returns as prices return to fundamentals. The rational-pricing story claims, instead, that the ratios track the time-variation in discount rates: ratios are low when discount rates are low, and high when discount rates are high; they would predict returns by capturing information about the time-varying discount rate and, hence, risk premium.

The rational-pricing story also applies to the interest rate variables. According to this view, the level of short-term interest rates reveals the state of the business cycle. As discount rates vary with the business cycle, short-term interest rates capture information about time-varying expected returns. For instance, high short-term rates are associated with a tight monetary policy, low current investment opportunities, low output, and low expected returns. Interest rates on long maturity government bonds may also be affected by short-term interest rates. If long-term government loan rates influence corporate loan rates, real investment in plant and machinery is also affected. Hence, long-term interest rates may also influence real economic activity and may be negatively related to expected stock returns. Finally, the term spread (i.e. the slope of the yield curve) is believed to be one of the best variable for forecasting business cycles. In particular, it is believed to be pro-cyclical and positively related to expected returns. A widening term spread (i.e. steepening of the slope) indicates expansion and higher expected returns, while a tightening term spread (i.e. flattening of the slope) indicates contraction and lower expected returns.[5]

To guard against potential size distortions due to overlapping observations, we follow much of the recent predictability literature and base inferences concerning the predictive variable's slope coefficient on a bootstrap procedure similar to the procedures in Nelson and Kim (1993), Mark (1995), Kothari and Shanken (1997), Kilian (1999), and Rapach and Wohar (2006). In addition, we explicitly address the issue of the upward small-sample bias in the OLS slope estimate by applying the Stambaugh (1999) and Lewellen (2004) correction approaches.

Following Goyal and Welch's critique on the poor OOS performance of predictive variables, we keep a limited number of (100) observations for IS estimation and compute OOS forecasts of log returns using a recursive window. The largest OOS period covers 568 months for the UK and the USA, while the smallest OOS period includes 310 months for France. Following Rapach and Wohar (2006), we analyze the OOS forecasts using a pair of recently developed tests due to McCracken (2004) and Clark and McCracken (2001).[6]

To the best of our knowledge, the economic significance of predictive regression models has always been measured through thought experiments. In this respect, we follow Goyal and Welch (2006) in quantifying economic gains through the root mean squared error difference between the unconditional and conditional forecasts. However, we also aim at explicitly taking the practitioner's point of view. In particular, we design a simple investment strategy that relies upon the return forecasts of the predictive regression model. The 'predictive' trading strategy consists in buying (selling) stocks and selling (buying) the risk-free asset whenever the stock return forecast is greater (smaller) than the risk-free rate. By implementing this rule, we hope to measure the economic significance of statistical return predictability, net of transaction costs.

[5]For a more comprehensive economic motivation behind the choice of these variables, see Cochrane (1997) who surveys the dividend ratio prediction literature; Campbell and Shiller (1988),originally motivated by Graham and Dodd (1951), for the earnings price ratio; Keim and Stambaugh (1986) and Campbell (1987) for the interest rate variables.

[6]The McCracken (2004) test statistic is a variant of the Diebold and Mariano (1995) and West (1996) statistics designed to test for equal predictive ability, while the Clark and McCracken (2001) test statistic is a variant of the Harvey, Leybourne, and Newbold (1997) statistic designed to test for forecast encompassing. Importantly, Clark and McCracken (2001, 2005) find the variants to be considerably more powerful than the original statistics in extensive Monte simulations.

The predictive trading strategy is compared to three benchmark portfolios. The first benchmark is a passive, buy-and-hold portfolio that is fully and always invested in stocks. Two active strategies also serve as benchmarks. A naive strategy based on extremes consists in always investing the entire portfolio in stocks except when the value of the predictive variable is above the 90th percentile of its unconditional distribution. The last active strategy is based on a restricted version of the predictive regression model. In this restricted model, the predictive variable is excluded. If the restricted model delivers higher (risk-adjusted) returns than the unrestricted predictive model, the predictive variable does not help deliver better forecasts.

Both passive and active strategies are evaluated on the basis of raw returns (net of transaction costs). However, risk adjustment is essential when evaluating the usefulness of active strategies, as these spend time out of the market and may exhibit less volatile returns than buy-and-hold portfolios. (Kho, 1996; Brown, Goetzmann, and Kumar, 1998; Dowd, 2000). We therefore compute profit-based metrics that rely upon risk-adjusted returns like the Sharpe ratio (Sharpe, 1966), the X^* statistic (Sweeney, 1988), and the X_{eff} measure (Dacorogna, Gençay, Müller, and Pictet, 2001).

Our results can be summarized as follows. The short-term interest yield and, to a lesser extent, the long government bond yield are the best out-of-sample predictors of stock returns. However, the out-of-sample predictive power of these variables does not appear to be economically meaningful across countries and investment horizons. First, thought experiments that rely upon out-of-sample statistics show that forecasting gains are small, underscoring the notion from the extant empirical literature that the predictive component in stock returns is small. Second, investment strategies that rely upon the return forecasts of predictive regression models also fail to deliver meaningful economic gains across investment horizons and across countries. While risk is difficult to measure and any risk adjustment is subject to criticism, most profit-based metrics converge to the same conclusions: predictive regression strategies based on interest rate variables generate the most robust economic performance in the US, Canada, and France; they mostly fail in the other countries. Taking the evidence overall, no common pattern of stock return predictability emerges across countries, be it on statistical or economic grounds.

The remainder of the paper is structured as follows. We describe the econometric methodology in Section 2. In Section 3, we present the dataset. We discuss the empirical results in Section 4 and conclude in Section 5.

2 Methodology

Following Rapach and Wohar (2006) and much of the extant literature, we analyze stock return predictability using a predictive regression framework. The predictive regression model can be written as:

$$y_{t,t+k} = \alpha_k + \beta_k z_t + \gamma_k y_t + u_{t,t+k} \tag{1}$$

where y_t is the real total log stock return from period $t-1$ to period t, $y_{t,t+k} = y_{t+1} + ... + y_{t+k}$ is the real total log stock return from period t to $t+k$ (i.e. the k-period forward-looking return), z_t is the log of the predictive variable (i.e. the forecasting variable believed to potentially predict future real returns), and $u_{t,t+k}$ is a disturbance term. Under the null hypothesis, the z_t variable has no predictive power for future returns ($\beta_k = 0$); under the alternative hypothesis, z_t does have predictive power for future returns ($\beta_k \neq 0$).[7] Note that we include a lagged return in Equation (1) as a control variable when testing the predictive ability of z_t (Lettau and Ludvigson, 2001; Rangvid, Rapach, and Wohar, 2005).[8]

The in-sample estimation of the regression model is done as follows. Suppose we have observations for y_t and z_t for $t = 1,...,T$. This leaves us with $T-k$ usable observations with which to estimate the in-sample predictive regression model.[9] The predictive ability of z_t is assessed by examining the t-statistic for $\hat{\beta}_k$, the OLS estimate of β_k in Equation (1).

[7] Inoue and Kilian (2004) recommend using a one-sided alternative hypothesis if theory makes strong predictions about the sign of in Equation (1), as this increases the power of in-sample tests. For the variables consider here, theory always makes strong predictions as to the sign of β, hence we use a one-sided alternative hypothesis thereafter.

[8] Inspection of the partial autocorrelation function for real stock returns for each country indicates that a single real stock return lag is sufficient in Equation (1). This is not surprising, as stock returns are known to display only limited persistence.

[9] Following Baker and Wurgler (2000) and Rangvid, Rapach, and Wohar (2005), we first divide z_t by its standard deviation over the full sample when we analyze the forecasting variables in turn. This normalization has no effect on statistical inferences, but it makes it easier to compare the estimated coefficient in Equation (1) across forecasting variables, as the coefficient can be interpreted as the change in expected returns given a one-standard-deviation change in the forecasting variable.

2.1 Estimation pitfalls

While the predictive regression model is simple to estimate and understand, it nevertheless suffers from two well-documented econometric problems: the small-sample bias and the overlapping observations problem. We next detail these shortcomings and how the recent literature deals with them.

2.1.1 Overlapping observations

There are overlapping observations in returns when $k > 1$. For instance, when $k = 2$, the 2-period forward-looking returns are given by: $y_{1,3} = y_2 + y_3$ at $t = 1$, $y_{2,4} = y_3 + y_4$ at $t = 2$, etc. We see that the y_3 return observation is included into the 2-period forward-looking returns at both $t = 1$ and $t = 2$. In fact, the number of overlapping observations between two consecutive k-period forward looking returns are equal to $k - 1$.

A common procedure for dealing with overlapping observations is the use of Newey and West (1987) standard errors, as these are robust to heteroskedasticity and serial correlation in the disturbance term (Richardson and Stock, 1989).[10] However, even when robust standard errors are used to compute t-statistics, serious size distortions can occur when basing inferences on standard asymptotic distribution theory (Nelson and Kim, 1993; Goetzman and Jorion, 1993; Kirby, 1997). To address this issue, we follow much of the recent predictability literature and base inferences about β_k in Equation (1) on a bootstrap procedure similar to the procedures developed by Nelson and Kim (1993), Mark (1995), Kothari and Shanken (1997), Kilian (1999), and Rapach and Wohar (2006). The bootstrap procedure is described in detail below.

2.1.2 Small-sample bias

Assuming that the predictive variable follows a first-order autoregressive (AR1) process, Stambaugh (1999) derives the exact small-sample distribution of the slope estimate and works out an analytical expression for the small-sample bias. More recently, Lewellen (2004) shows that

[10] In the empirical analysis (Section 4), we use the Bartlett kernel and a lag truncation parameter of $f(1.5k)$, where $f()$ is the nearest integer function, when calculating Newey and West standard errors.

the Stambaugh's correction can substantially understate forecasting power when the predictive candidate is highly persistent.

Stambaugh's unconditional approach. Stambaugh (1999) shows that, in small samples, the predictive coefficients are biased if the independent variable is close to a random walk. When $k = 1$ in Equation (1),

$$y_{t,t+1} = \alpha_1 + \beta_1 z_t + \gamma_1 z_t + u_{t,t+1} \tag{2}$$

where $u_{t,t+1}$ is an independently and identically distributed disturbance term with mean zero. Stambaugh (1999) shows that the OLS estimate of β_1 is biased in finite samples when the DGP for $y_{t,t+1}$ is governed by Equation (2) and z_t is generated by the stationary first-order autoregressive process,

$$z_{t+1} = \phi + \rho z_t + \varepsilon_{t+1} \tag{3}$$

where ε_{t+1} is an independently and identically distributed disturbance term with mean zero and $0 < \rho < 1$ for the predictive variable. In particular, Stambaugh (1999) shows that:

$$E(\hat{\beta_1} - \beta_1) \approx (\sigma_{u\varepsilon}/\sigma_{\varepsilon}^2)E(\hat{\rho} - \rho) \tag{4}$$

where $\hat{\rho}$ is the OLS estimate of ρ in Equation (3); $\sigma_{u\varepsilon}$ is the covariance between the disturbances terms, u_t and ε_t; σ_{ε}^2 is the variance of ε_t. It is well known that $\hat{\rho}$ is a biased estimator of ρ, with the bias given by (Kendall, 1954; Shaman and Stine, 1988):

$$E(\hat{\rho} - \rho) \approx -(1 + 3\rho)/T \tag{5}$$

where T is the sample size. We see that $\hat{\rho}$ can substantially underestimate ρ in finite samples, with the bias in $\hat{\rho}$ increasing as ρ approaches unity (that is, as it becomes more persistent). The bias in $\hat{\rho}$ is fed into $\hat{\beta}_1$ via Equation (4) when $\sigma_{u\varepsilon} \neq 0$ (and σ_{ε}^2 is not 'too large'). Taken together, a persistent z_t series and highly correlated residuals across Equations (2) and (3) can result in a substantial bias in finite samples for $\hat{\beta}_1$.[11]

Lewellen's conditional approach. Stambaugh's unconditional approach assumes that we have no information regarding $\hat{\rho} - \rho$. This approach is appropriate when $\hat{\rho}$ is small because the $\rho < 1$ constraint provides little information. In contrast, Lewellen (2004) shows that the unconditional approach can substantially misstate the predictive ability of z_t when the predictive variable's autocorrelation is close to one. In particular, he shows that $\hat{\beta}$ can be strongly correlated with the sample autocorrelation of the predictive variable. The bias-adjusted estimator is defined by:

$$\hat{\beta}_{adj} = \hat{\beta} - (\sigma_{u\varepsilon}/\sigma_{\varepsilon}^2)(\hat{\rho} - \rho). \quad (6)$$

Lewellen's test focuses on the distribution of $\hat{\beta}$ conditional upon knowledge of the autocorrelation of the predictive variable. As long as the predictive variable is stationary, the most conservative assumption for testing predictability is that $\rho \approx 1$. By assuming a lower bound on $\hat{\rho} - \rho$, Lewellen (2004) maximizes the bias $(\sigma_{u\varepsilon}/\sigma_{\varepsilon}^2)(\hat{\rho} - \rho)$ and minimizes the estimator $\hat{\beta}_{adj}$. If $\hat{\beta}_{adj}$ is significantly different from zero under this assumption, then it must be even more significant given the true value of ρ.

Joint significance test. The decision to use the Lewellen conditional test or the Stambaugh standard approach depends on the true value of ρ. From an ex-ante perspective, the conditional test has greater power when ρ is close to one, but the opposite is true once ρ drops below some level that depends on the other parameters. However, without prior information about ρ, it makes sense to rely on both tests and to calculate an overall, joint significance level that reflects the probability of rejecting the null hypothesis ($\beta = 0$) using either test. The modified Bonferroni upper bound provides a way to calculate the joint p-value (Lewellen, 2004). Under this approach, when the conditional and unconditional tests are both used, an overall p-value is given

[11] In the empirical analysis, the reported statistics on the uncorrected OLS take the persistence of the predictive variable into account because we bootstrapped for significance levels mimicking the IS autocorrelation of each predictive variable. However, we also apply the Stambaugh (1999) coefficient correction since it is a more powerful method in non-asymptotic samples.

by $min(2P, P+D)$, where P is the smaller of the two stand-alone p-values and D is the p-value for testing $\rho = 1$, based on the sampling distribution of $\hat{\rho}$.

2.2 Out-of-sample statistical analysis

As discussed in the introduction, we also perform OOS tests of return predictability. Following Rapach and Wohar (2006), a recursive scheme is used. We first divide the total sample of T observations into IS and OOS portions, where the IS portion spans the first R observations for y_t and z_t, and the OOS portion spans the last P observations for the two variables.

The first OOS forecast of the 'unrestricted' predictive regression model given in Equation (1) is generated in the following manner: estimate the unrestricted predictive regression model (Model 1) via OLS using data available up to period R; denote Model 1's OLS estimates of α_k, β_k and γ_k, using data available up to period R, as $\hat{\alpha}^1_{k,R}$, $\hat{\beta}^1_{k,R}$, and $\hat{\gamma}^1_{k,R}$. Using these OLS parameter estimates as well as y_R and z_R, construct a k-period return forecast, $y_{R,R+k}$, based on Model 1 using $\hat{y}^1_{R,R+k} = \hat{\alpha}^1_{k,R} + \hat{\beta}^1_{k,R} z_R + \hat{\gamma}^1_{k,R} y_R$. Denote the k-period return forecast error by $u^1_{R,R+k} = y_{R,R+k} - \hat{y}^1_{R,R+k}$.

OOS forecasts of the unrestricted predictive model are also generated after correcting for the bias in $\hat{\beta}$. Using Stambaugh's standard approach, construct a k-period return forecast such that $\hat{y}^{1S}_{R,R+k} = \hat{\alpha}^1_{k,R} + \hat{\beta}^{1S}_{k,R} z_R + \hat{\gamma}^1_{k,R} y_R$, where $\hat{\beta}^{1S}_{k,R}$ is Stambaugh's corrected beta coefficient estimated using data available up to period R. Correspondingly, construct a k-period return forecast such that $\hat{y}^{1L}_{R,R+k} = \hat{\alpha}^1_{k,R} + \hat{\beta}^{1L}_{k,R} z_R + \hat{\gamma}^1_{k,R} y_R$, where $\hat{\beta}^{1L}_{k,R}$ is Lewellen's corrected beta estimated using data available up to period R.

Campbell and Thompson (2005) make two interesting suggestions that we follow. First, they argue that a reasonable investor would not have used a model that has a $\hat{\beta}_k$ coefficient with the theoretically incorrect sign. They truncate such coefficients to zero. Second, they argue that a reasonable investor would not have used a model that forecasts a negative equity premium. Such predictions are also truncated to zero.

The initial k-period return forecast for the 'restricted' model is generated in a similar manner, except we set $\beta_k = 0$ in Equation (1). That is, we estimate the restricted regression model (Model

0) via OLS using data available up to period R in order to have the k-period return forecast $\hat{y}^0_{R,R+k} = \hat{\alpha}^0_{k,R} + \hat{\gamma}^0_{k,R} y_R$, where $\hat{\alpha}^0_{k,R}$ and $\hat{\gamma}^0_{k,R}$ are the OLS estimates of α_k and γ_k in Equation (1) with β_k restricted to zero using data available up to period R. Denote the k-period return forecast error corresponding to the restricted model as $u^0_{R,R+k} = y_{R,R+k} - \hat{y}^0_{R,R+k}$.

In order to generate a second set of forecasts, we update the above procedure by using data available up to period $1+R$. That is, we estimate the unrestricted and restricted predictive regression models using data available up to period $1+R$, and we use these parameter estimates and the observations for z_{R+1} and y_{R+1} in order to form unrestricted and restricted model k-period return forecasts for $y_{R+1,R+1+k}$ and their respective k-period return forecast errors, $u^1_{R+1,R+1+k}$ and $u^0_{R+1,R+1+k}$. We repeat this process through the end of the available sample, which yields two sets of $T-R-k-1$ recursive forecast errors, one each for the unrestricted and restricted regression models: $\{u^1_{t,t+k}\}_{t=R}^{T-k}$ and $\{u^0_{t,t+k}\}_{t=R}^{T-k}$.

2.2.1 Statistical metrics

The next step is to compare the OOS forecasts from the unrestricted and restricted predictive regression models. If the unrestricted model forecasts are superior to the restricted model forecasts, then the z_t variable improves the OOS forecasts of $y_{t,t+k}$ relative to the historical mean model which excludes z_t. A simple metric for comparing forecasts is Theil's U, the ratio of the unrestricted model forecast root mean squared error (RMSE) to the restricted model forecast RMSE. If the unrestricted model forecast RMSE is less than the restricted model forecast RMSE, then $U < 1$. In order to formally test whether the unrestricted regression model forecasts are significantly superior to the restricted model forecasts, we use three other OOS statistics: the McCracken (2004) MSE-F, the DRMSE statistic and the Clark and McCracken (2001) ENC-NEW statistics.

The MSE-F statistic is a variant of the Diebold and Mariano (1995) and West (1996) statistics designed to test for equal predictive ability. It is used to test the null hypothesis that the unrestricted model forecast mean-squared error (MSE) is equal to the restricted model forecast MSE against the one-sided (upper-tail) alternative hypothesis that the unrestricted model forecast MSE is less than the restricted model forecast MSE. Like the Diebold and Mariano

(1995) and West (1996) statistics, the MSE-F statistic is based on the loss differential, $\hat{d}_{t,t+k} = (\hat{u}^0_{t,t+k})^2 - (\hat{u}^1_{t,t+k})^2$. Let us define $\overline{d} = (T-R-k+1)^{-1}\sum_{t=R}^{T-k}\hat{d}_{t,t+k} = M\hat{S}E_0 - M\hat{S}E_1$, where $M\hat{S}E_i = (T-R-k+1)^{-1}\sum_{t=R}^{T-k}(\hat{u}^i_{t,t+k})^2$, i=0 or 1. The McCracken (2004) MSE-F statistic is given by:

$$\text{MSE-F} = (T-R-k+1)\overline{d}/M\hat{S}E_1. \tag{7}$$

A significant MSE-F statistic indicates that the unrestricted model forecasts are statistically superior to those of the restricted model. When comparing forecasts from nested models (as we do) and for $k = 1$, McCracken (2004) shows that the MSE-F statistic has a non-standard limiting distribution that is pivotal and a function of stochastic integrals of Brownian motion. Clark and McCracken (2005) demonstrate that the MSE-F statistic has a non-standard and non-pivotal limiting distribution in the case of nested models and $k > 1$. Given this last result, Clark and McCracken (2005) recommend basing inference on a bootstrap procedure along the lines of Kilian (1999). Following this recommendation, we base our inferences on the bootstrap procedure described below.

The DRMSE statistic is a variant of Theil's U. It is given by:

$$\text{DRMSE} = (M\hat{S}E_0)^{0.5} - (M\hat{S}E_1)^{0.5} \tag{8}$$

where $(M\hat{S}E_0)^{0.5}$ is equal to the root mean squared error of the unconditional forecast (model 0) and $(M\hat{S}E_0)^{0.5}$ is the the root mean squared error of the conditional forecast (model 1). A positive number means superior OOS conditional forecast.

The ENC-NEW statistic is a variant of the Harvey, Leybourne, and Newbold (1998) statistic designed to test for forecast encompassing. Forecast encompassing is based on optimally constructed composite forecasts. Intuitively, if the forecasts from the restricted regression model encompass the unrestricted model forecasts, the forecasting variable included in the unrestricted model provides no useful additional information for predicting returns relative to the restricted model which excludes the forecasting variable; if the restricted model forecasts do not encompass the unrestricted model forecasts, then the forecasting variable does contain information useful for predicting returns beyond the information already contained in a model that excludes the forecasting variable. Tests for forecast encompassing are tantamount to testing whether the weight attached to the unrestricted model forecast is zero in an optimal composite forecast composed of the restricted and unrestricted model forecasts. The Clark and McCracken (2001) ENC-NEW statistic takes the form:

$$\text{ECN-NEW-F} = (T-R-k+1)\overline{c}/M\hat{S}E_1 \tag{9}$$

where $\hat{c}_{t,t+k} = \hat{u}^0_{t,t+k}(\hat{u}^0_{t,t+k} - \hat{u}^1_{t,t+k})$ and $\overline{c} = (T-R-k+1)^{-1}\sum_{t=R}^{T-k}\hat{c}_{t,t+k}$. Under the null hypothesis, the weight attached to the unrestricted model forecast in the optimal composite forecast is zero, and the restricted model forecasts encompass the unrestricted model forecasts. Under the one-sided (upper-tail) alternative hypothesis, the weight attached to the unrestricted model forecast in the optimal composite forecast is greater than zero, so that the restricted model forecasts do not encompass the unrestricted model forecasts. Similar to the MSE-F statistic, the limiting distribution of the ENC-NEW statistic is non-standard and pivotal for $k = 1$ (Clark and McCracken, 2001) and non-standard and non-pivotal for $k > 1$ (Clark and McCracken, 2005) when comparing forecasts from nested models. Again, Clark and McCracken (2005) recommend basing inference on a bootstrap procedure, given the non-pivotal limiting distribution. As indicated in the introduction, Clark and McCracken (2001, 2005) find that the MSE-F and ENC-NEW statistics have good size properties and are more powerful than the original statistics in extensive Monte Carlo simulations with nested models.

2.2.2 Bootstrap procedure

For the reasons discussed above, we base the IS and OOS test inferences on a bootstrap procedure similar to the procedures in Nelson and Kim (1993), Mark (1995), Kothari and Shanken (1997), Kilian (1999), and Rapach and Wohar (2006). We postulate that the data are generated by the following system under the null hypothesis of no predictability:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_{1,t} \tag{10}$$

$$z_t = b_0 + b_1 z_{t-1} + ... + b_p z_{t-p} + \varepsilon_{2,t} \quad (11)$$

where the disturbance vector $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})$' is independently and identically distributed with covariance matrix Σ. We first estimate Equations (10) and (11) via OLS, with the lag order (p) in Equation (11) selected using the AIC (considering a maximum lag order of twelve), and compute the OLS residuals $\{\hat{\varepsilon}_t = (\hat{\varepsilon_{1,t}}, \hat{\varepsilon_{2,t}})'\}_{t=1}^{T-p}$.[12] In order to generate a series of disturbances for our pseudo-sample, we randomly draw (with replacement) $T + 100$ times from the OLS residuals $\{\hat{\varepsilon}_t\}_{t=1}^{T-p}$, giving us a pseudo-series of disturbance terms $\{\hat{\varepsilon}_t\}_{t=1}^{T+100}$. Note that we draw from the OLS residuals in tandem, thus preserving the contemporaneous correlation between the disturbances in the original sample. Denote the OLS estimates of a_0 and a_1 in Equation (10) by $\hat{a}_0$ and $\hat{a}_1$, and the OLS estimate of $(b_0, b_1, ..., b_p)$ in Equation (11). Using $\{\hat{\varepsilon}_t^*\}_{t=1}^{T+100}$, $(\hat{a}_0, \hat{a}_1, \hat{b_0}, \hat{b_1}, ..., \hat{b_p})$ in Equations (10) and (11), and setting the initial observations for y_{t-1} and $(z_{t-1}, ..., z_{t-p})$ equal to zero in Equations (10) and (11), we can build up a pseudo-sample of $T + 100$ observations for y_t and z_t, $\{\hat{y}_t^*, \hat{z}_t^*\}_{t=1}^{T+100}$. We drop the first 100 transient start-up observations in order to randomize the initial y_{t-1} and $(z_{t-1}, ..., z_{t-p})$ observations, leaving us with pseudo-sample of T observations, matching the original sample. For the pseudo-sample, we calculate the t-statistic corresponding to β in the IS predictive regression model given in Equation (1), and the three OOS statistics given in Equations (7), (8) and (9). We repeat this process 1000 times, giving us an empirical distribution for the IS t-statistic and each of the OOS statistics. As both the OOS tests are one-sided and upper-tailed, the p-value is the proportion of the bootstrapped statistics that are greater than the statistic computed using the original sample. As the IS t-ratio test is one-sided and lower-tailed, the p-value is the proportion of the bootstrapped statistics that are lower than the statistic computed using the original sample.

2.3 Out-of-sample economic analysis

Switching from statistical OOS predictability to a useful economic strategy can be challenging and often weeds out many candidate predictors. Among the potential pitfalls, parameter insta-

[12]The lag order (p) has also been selected using the Schwarz's criterion that embodies a stiffer penalty for adding an extra term in Equation (11). Since no significant difference was observed when the bootstrap procedure was based on the Schwarz's criterion, we only report the results based on the AIC in the empirical section.

bility, transaction costs, and the fact that the strategy may expose the trader to idiosyncratic risks stand out. Besides, the algorithm may have not be stable.

To address these issues, we explore the economic significance of stock returns predictability by building the following 'predictive' investment strategy. Using data available up to period R and estimating the unrestricted model (Model 1), we construct the k-period return forecast such that $\hat{y}^1_{R,R+k} = \hat{\alpha}^1_{k,R} + \hat{\beta}^1_{k,R} z_R + \hat{\gamma}^1_{k,R} y_R$. Denote the k-period T-bill interest rate available at period R as $i_{R-k,R}$. If $\hat{y}^1_{R,R+k} > i_{R-k,R}$, the investor buys stocks and sell the risk-free instrument. If $\hat{y}^1_{R,R+k} > i_{R-k,R}$, the investor sells stocks and buys the risk-free instrument. In order to generate the next forecast $\hat{y}^1_{R+1,R+1+k}$, we update the above procedure one period by using data available through period $1+R$, and estimate Equation (1) using OLS. This process goes on until the end of the sample is reached. We follow the same methodology to measure the economic significance of forecasts based on Stambaugh's and Lewellen's corrected OLS slope coefficients.

2.3.1 Benchmark portfolios

The above 'predictive' investment strategies are compared to three benchmark portfolios. The first benchmark is a passive, buy-and-hold portfolio that is always fully invested in stocks (BH). The two other benchmarks are active strategies. The first one compares the current level of the predictive variable to its extreme historical values (EXT). It is a naive strategy in the sense that it does not make any forecast *per se*. It consists in always investing the entire portfolio in stocks except when the current value of the predictive variable is above the

90th percentile of its unconditional distribution.[13] This strategy does not predict the date of the turning point, but it points to a probable decline in stocks. For example, a fall in stock prices may be more likely than a further rise when the price-earnings ratio is above the 90th percentile of its unconditional distribution (Shen, 2003; Berge and Ziemba, 2006). The second active strategy (REST) is based on the restricted model discussed above (model 0). If $\hat{y}^0_{R,R+k} > i_{R-k,R}$, the investor buys stocks and sell the risk-free instrument. If $\hat{y}^0_{R,R+k} < i_{R-k,R}$, the investor sells stocks and buys the risk-free instrument. In order to generate the next forecast $\hat{y}^0_{R+1,R+1+k}$, we update the above procedure one period by using data available through period $1+R$, and

[13]The choice of the 90th percentile as the threshold implicitly assumes that the stock market moves 10% of the time far away from its fundamental value. This is consistent with the Black (1986) estimation.

estimate Equation (1) with β_k restricted to zero using OLS. This process goes on until the end of the sample is reached. If the restricted model delivers higher (risk-adjusted) returns than the unrestricted predictive model, the predictive variable may not bring any economic added value.

2.3.2 Risk-adjusted metrics

To compare the active and passive strategies, risk adjustment is important because active strategies are often out of the market and therefore may bear much less risk than the buy-and-hold strategies (Neely, 2003). Although there is no universally accepted method of adjusting returns for risk, we use four techniques: the Sharpe ratio (Sharpe, 1966), the X^* measure (Sweeney, 1988), Jensen α (Jensen, 1968) and the X_{eff} measure of Dacorogna, Gençay, Müller, and Pictet (2001).

The Sharpe ratio measures the expected excess return per unit of risk for a zero-investment strategy (Campbell, Lo, and MacKinlay, 1997). It is usually measured in annual terms as a portfolio's annual excess return over the riskless rate, net of transactions costs, divided by its annual standard deviation. Although the active strategies may exhibit lower returns than the buy-and-hold portfolios, the lower volatility may allow the portfolio to be leveraged and to exceed the buy-and-hold return with similar risk.

Sweeney (1988) developed another risk-adjustment statistic, the X^* measure.[14] He shows that, in the presence of a constant risk premium, an equilibrium k-period risk-adjusted return to a investment strategy would be given by:

$$X^* = \frac{1}{T_k}\sum_{t=0}^{T_k}[s_t y_{t+1}^k + (1-s_t)r_{t+1}^k] - \frac{n}{2T_k}\ln\left(\frac{1+c}{1-c}\right) - \left[\frac{p_1}{T_k}\sum_{t=0}^{T_k} y_{t+1}^k + \frac{p_2}{T_k}\sum_{t=0}^{T_k} r_{t+1}^k\right] \qquad (12)$$

where s_t is an indicator variable taking the value 1 if the strategy is in the market or 0 if the rule is in T-bills; y_{t+1}^k is the real total log return to holding stocks from period t to $t+k$; r_{t+1}^k is the real total log return to holding T-bills from period t to $t+k$; T_k is the number of k-period return observations; n is the number of one-way trades; c is the proportional transactions cost; p_1 is the proportion of the time spent in the market and p_2 is the proportion of the time spent in T-bills

[14] The X^* test statistic is virtually equivalent to the test statistic of the coefficient β_1 in the Cumby-Modest test of market timing, with the difference that transactions costs are omitted in the Cumby-Modest test.

$(p_1 + p_2 = 1)$.[15] Under the null of no market timing ability, there is no statistical difference between the actual return delivered by the active strategy and its expected return, so that X^* is not statistically different from zero. Statistically positive X^* statistics are interpreted as evidence of superior risk-adjusted returns.

According to Dacorogna, Gençay, Müller, and Pictet (2001), the performance of an active strategy should be viewed as favorable when: the total return is high; the total return curve increases linearly over time; and loss periods are not clustered. The Sharpe Ratio does not entirely satisfy these requirements. First, the definition of the Sharpe Ratio puts the variance of the return into the denominator which makes the ratio numerically unstable at extremely large values when the variance of the return is close to zero. Second, the Sharpe Ratio does not take into account the clustering of profits and losses. An even mixture of profit and loss trades is usually preferred to clusters of losses and clusters of profits, provided the total set of profit and loss trades is the same in both cases. Third, the Sharpe Ratio treats the variability of profitable returns (which are unimportant to investors) the same way as the variability of losses (which are an investor's major concern).

One of the risk-adjustment measures advocated by Dacorogna, Gençay, Müller, and Pictet (2001) is the X_{eff} measure. X_{eff} measures the utility that the strategy provides to a constant absolute risk averse individual over a weighted average of return horizons. According to these authors, X_{eff} exhibits fewer deficiencies than the Sharpe Ratio. First, X_{eff} is numerically stable. Second, the return curve (that reflects the real risk of a trading strategy) is assessed, through the sum of the accumulated total return and the current unrealized return of open positions. Accounting for unrealized returns avoids biases in favor of strategies with a low transaction frequency. X_{eff} reaches the value of the annualized total return if the return curve is a straight line as a function of time. For all nonlinear equity curves, $X_{eff}$is smaller than the annualized total return. The measure is:

[15]The sum of the last two terms estimates the expected return to a zero transactions-cost strategy that is randomly in the market on a fraction p_1 of the k month periods, earning the market premium, and in T-bills otherwise.

$$X_{eff} = \overbrace{\frac{(12/k)100}{T_k}\{\sum_{t=0}^{T_k}[s_t(y_{t+1}^k - r_{t+1}^k)] - \frac{n}{2}\ln(\frac{1+c}{1-c})\}}^{(1)} - \overbrace{\frac{\gamma}{2}\frac{\sum_{i=1}^{n}\tilde{w}_i\sigma_i^2[12/(k\Delta t_i)]}{\sum_{i=1}^{n}\tilde{w}_i}}^{(2)} \qquad (13)$$

where (1) is the annualized excess return to the strategy in percentage terms, net of transactions costs; (2) is the annualized cost of risk taking by the strategy; σ_i^2 is the variance of non-overlapping k-period returns over k-multiple time horizon Δt_i; $12/(k\Delta t_i)$ is the annualized factor, i.e. the number of k-period returns of length Δt_i in one year; γ is the risk aversion parameter. Dacorogna, Gençay, Müller, and Pictet (2001) recommend values of γ between 0.08 and 0.15; this paper follows Nelly (2003) in setting γ equal to 0.12. The weights $\tilde{w}_i$ are determined with a weighting function which allows the selection of the relative importance of the different horizons. The time horizons Δt_i are assumed to follow a geometric sequence (1,2,4,8,...months). The weighting function is:[16]

$$\tilde{w}_i = \frac{1}{2 + [\ln(\frac{\Delta t_i}{3/k})]^2} \qquad (14)$$

Finally, we consider the performance of active strategies according to Jensen (1968) α, the return in excess of the riskless rate that is uncorrelated with the excess return to the market:

$$s_t[y_{t+1}^k - r_{t+1}^k] - \frac{n}{2T_k}\ln(\frac{1+c}{1-c}) = \alpha_k + \beta_{M,k}[y_{t+1}^k - r_{t+1}^k] + \varepsilon_{t+1}^k \qquad (15)$$

where the variables are as defined in Equation (12). If the intercept ($\hat{\alpha}$) is positive and significant, then the trading rule delivers excess returns that cannot be explained by correlation with the market.

[16]This choice is motivated as follows. First, this weighting function encompasses a wide range of interval sizes Δt. Hence, drawdowns of many different sizes are captured. Second, the weighting function smoothly declines on both sides of a maximum at around three months as a function of $\ln\Delta t$. Extremely long intervals have a low weight, because the lifetime of the whole investment, as well as the period of available and relevant historical data, is limited to a few years. Regarding short intervals, there is also a limit: most investors do not regard short oscillations of their portfolio value as relevant for their investment. Third, the maximum of $\tilde{w}_i$ is at Δt = 3 months, which is a typical time horizon for many investors. This choice is reasonable, but special short-term or extremely long-term investors are free to shift the maximum weight to other interval sizes.

3 Data

The data is provided by Global Financial Data (GFD). All series are extracted on a monthly frequency. The total real log return to holding stocks from period $t-1$ to period t (RSR) are used to measure y_t. The total real log return to holding t-bills from period $t-1$ to period t (RTR) are used to measure i_t. Real prices are computed by dividing nominal prices (which include income gains in GFD) by the level of the CPI. Total real log returns from period $t-1$ to period t are computed in two steps: first, the real price at time t is divided by the real price at time $t-1$; then, the log of the ratio is computed.

The corresponding dividend-yield and price-earnings ratios in logs are used as predictive regressors, z_t. The dividend-price ratio (DPR) is defined as the sum of paid dividends during the past year, divided by the current price. The price-earnings ratio (PER) is defined in the same way, i.e. the current price divided by the latest 12 months of earnings available at the time. Both the dividend- and earnings-price ratios are thus expressed in annual units.

Following much of the extant literature, we measure interest rates as deviations from a moving average.[17] The relative short-term interest yield (STY) is computed as the difference between the short-term interest yield and a 12-month backward-looking moving average. The relative long-term government bond yield is defined in the same way, i.e. the difference between the long-term government bond yield and a 12-month backward-looking moving average. The short-term interest rate is the 3-month T-bill rate when available; if the latter is not available, the private discount rate or the interbank rate is used. The long rate is measured by the yield on long-term government bonds. When available, a 10-year bond is used; otherwise, the closest maturity to 10 years is relied upon. To be consistent with the computation of the one month stock returns, the short-term and long-term interest rates are also expressed on a monthly basis. Finally, the term spread is defined as the log difference between the long-term and the short-term rate.

When defining excess stock returns, two approaches are feasible: (i) the return on stocks, in the local currency, over the local short rate, or (ii) the return on stocks in dollar terms, over the

[17] As pointed out by Rangvid, Rapach, and Wohar (2005), measuring the interest rate as deviations from a backward-looking moving average may go some way toward making the nominal interest rate an effectively real interest rate (to the extent that the behavior of expected inflation is such that most of the fluctuations in the relative nominal interest rate reflect movements in the relative real component).

U.S. short interest rate. In this paper, we use excess returns expressed in domestic currencies as it provides the international analogue of the typical forecasting regressions estimated for U.S. data. In addition, we avoid the pitfall that the predictability in exchange rates rather than in stock returns drives the results.

Transaction costs are evaluated following Sutcliffe (1997) estimates. As the active strategies can be easily replicated by using the corresponding T-bill and stock market futures contracts, we consider transaction costs on the futures markets. According to Sutcliffe (1997), an appropriate round-trip figure for the FTSE-100 futures is 0.116% (of the purchase and sale values). This figure is made up of bid/ask spread (0.083%) and commissions (0.033%).[18] To keep things simple, we assume identical costs for the two (stock and T-bill) future markets.

Table 1 reports the available sample for each country, as well as the mean, standard deviation, and coefficient of first-order autocorrelation for each variable. Regarding each of the six predictive variables, theory predicts negative β coefficients for the price-earnings ratio (PER), the short-term interest yield (STY), and the long government bond yield (LTY). Positive signs are expected for the dividend yield (DPR), and the term spread (SPR). For ease of computation, we always use a one-sided lower-tailed alternative hypothesis in Section 4. Therefore, DPR and SPR are constructed in such a way that a negative β coefficient is also expected (i.e. we take the negative of the log dividend-price ratio, the negative of the log dividend-earnings ratio, and the negative of the term spread as predictive variables). Of particular interest is the high persistence of all the predictive variables in each country, as revealed by the AC(1) coefficients. As a consequence, the OLS slope estimate is likely to be biased and a correction approach à la Lewellen (2004) may be appropriate. Table 2 gives the stock index in each country to which the total returns and dividend- and earnings-price ratios correspond.

4 Empirical analysis

Tables 3 to 12 report IS regression results for Equation (1) using data from the full sample for each country. Each forecasting variable is examined successively. The predictive ability

[18]Futures trading costs are not easy to gauge, but Goyal and Welch (2006) argue that a typical contract for a notional amount of $250,000 may cost around $10-$30. A 20% movement in the underlying index, which is about the annual volatility, would correspond to $50,000, which would amount to around 5 bp for a single transaction.

of the forecasting variables is estimated at the 1-month ($k = 1$), 3-month ($k = 3$), and 1-year ($k = 12$) horizons. IS statistical analysis is summarized by the β coefficient of Equation (1) and its level of significance. Theil's U, the MSE-F, DRMSE and ENC-NEW statistics for the OOS statistical analysis of forecasts are also reported. Tables 13 to 27 report OOS economic results (net of transaction costs) provided by the 'predictive' trading strategies and the benchmark portfolios. For each horizon, five profit-based metrics are reported: excess returns, Sharpe Ratio, X^*, X_{eff}, and Jensen α. For each country, we first describe the statistical results, focusing on the forecasting variables that are both IS and OOS significant.[19] Then, we investigate whether practitioners can exploit statistical superior OOS performance through the so-called 'predictive' investment strategy described in Section 2.3.

4.1 Australia

From Table 3, we see that statistical evidence of stock return predictability in Australia is weak. The long government bond yield (LTY) is the only forecasting variable that exhibits IS significance. The IS β slope estimate for LTY is significant at the 5% level for each of the horizons considered. IS significance persists even after explicitly correcting for the small-sample bias, like in Stambaugh (1999) or Lewellen (2004). The modified Bonferroni's upper bound also suggests joint IS significance at each of the horizons considered. Significant results also characterize the OOS tests on LTY, but ENC-NEW is the only statistic that points to superior OOS forecasts at the 3- and 12-month horizons.

The economic performance of the LTY predictive regression model is evaluated through the investment strategy described in Section 2.3. Since LTY is not OOS significant at the 1-month horizon, no economic investigation is carried out when $k = 1$. At the 3-month horizon, both STAM and LEW models generate negative raw excess returns relatively to the buy-and-hold equity portfolio (Table 18). While these models perform better than the restrictive model (REST), they cannot beat the other active benchmark portfolio based on extreme values (EXT). In terms of risk-adjusted returns, both STAM and LEW models display higher Sharpe ratios than the buy-and-hold portfolio (BH), but they fail to beat the EXT benchmark strategy (Table 19). Although

[19] There are a few models that show statistically superior OOS performance, but which we do not discuss and consider in the economic analysis because they are not IS significant, hence not robust.

STAM and LEW models generate positive and higher X^* metrics than both the EXT and BH portfolios (under the assumption of a constant risk premium), there is no statistical significance, even at the 10% level (Table 20). Furthermore, STAM and LEW models obtain negative X_{eff} metrics, indicating that riskless strategies provide constant absolute risk averse investors with greater utility over a weighted average of return horizons (Table 21). While STAM and LEW generate higher X_{eff} than EXT and BH, they cannot beat REST. Finally, STAM and LEW provide the two highest positive Jensen α metrics among all the portfolios considered (Table 22). However, these are not statistically and economically significant. At the 12-month horizon, the three versions (OLS,STAM, and LEW) of the predictive regression model are considered. Results are similar to those obtained at the 3-month horizon.

Overall evidence of stock return predictability in Australia is weak. The long-term government bond yield (LTY) is the only predictive variable that displays statistical significance. Although the predictive regression model based on LTY does not perform badly in comparison to the buy-and-hold portfolio, the investor should have known upfront that the LTY variable was the (only) good performer among all the candidates. In addition, the LTY predictive regression model does not seem to systematically outperform the naive benchmark strategy based on extremes.

4.2 Canada

The results for Canada in Table 4 show that there is IS and OOS evidence of predictability for each of the three interest rate variables, but none for the two valuation ratios. Among the three interest rate variables, IS and OOS significance is highest for the LTY variable. When economic gains of OOS predictive power are measured through the DRMSE metric (see Goyal and Welch, 2006), such gains appear to be limited. For example, the highest economic gain would be delivered by the simple OLS regression with LTY as predictive variable. It would amount to a mere 1.6 bp/month gain, which is likely to be offset by trading costs.

The OOS economic analysis of the predictive investment strategies at the 1-month horizon reveals that excess returns are systematically positive when LTY and STY are used as predictive variables (Table 13). STY predictive regression models are nevertheless the only models that deliver higher positive excess returns than the REST and EXT benchmark portfolios. Looking at Table 14, we see that each of the three interest rate variables generate higher Sharpe ratios than the buy-and-hold portfolio. However, only predictive models based on STY deliver higher Sharpe ratios than the two competing active strategies, EXT and REST. Results based on the X^* and X_{eff} metrics lead to similar conclusions: predictive models perform well relatively to the buy-and-hold portfolio, but none of them is able to beat the restrictive model, REST (Tables 15 and 16). Interestingly, all predictive models deliver positive and significant Jensen α's, but the only predictive models that beat both the EXT and REST competing models are the predictive regression models based on STY (Table 17).

At the 3-month horizon, LTY and TSP seem to offer economically meaningful performance. Excess returns delivered by the predictive models are positive and higher than those delivered by the REST model (Table 18). Predictive models based on LTY and TSP also generate both the highest Sharpe ratios and X^* metrics among all the active and passive portfolios (Tables 19 and 20). Results based on the X_{eff} metric lead to similar conclusions (Table 21). Finally, LTY and TSP predictive models obtain the highest Jensen α's, which are both positive and significant at the 5% level (Table 22).

STY and LTY predictive regression models demonstrate some ability in outperforming the other benchmark portfolios at the 12-month horizon also. Both STY and LTY predictive models generate positive excess returns, with STY predictive models displaying the highest positive excess returns among all competing models. The highest Sharpe ratios and X^* metrics are also delivered by both STY and LTY predictive regression models (Tables 24 and 25). While they cannot do better than the REST model with respect to the X_{eff} metric (Table 26), they generate the highest positive and significant Jensen α's.

There is, in summary, some statistical evidence of stock return predictability in Canada, provided interest rate variables are used as predictors. Although forecasting gains appear to be limited according to statistical criteria (such as Theil's U or DRMSE), they may be translated into meaningful economic gains through the use of predictive investment strategies based on LTY and STY variables. At the 1-month horizon, however, such gains do not appear to be significantly higher than those provided by the restrictive model.

4.3 France

From Table 5, we see that statistical evidence of stock return predictability in France is limited to interest rate variables. IS and OOS significance for the LTY variable is found at each horizon and for each of the three versions of the predictive model. In contrast, the OOS predictive power of the TSP variable is revealed by the ENC-NEW statistic only and is limited to the 3-month horizon. In any case, forecasting gains appear to be limited according to statistical criteria such as Theil's U or DRMSE.

Looking at the 1-month economic performance (Tables 13 to 17), LTY predictive regression models perform best. They deliver higher excess returns, Sharpe ratios, X^* metric, and Jensen α's than the competing models. Among the three different versions of the predictive regression, the best performance comes from the LEW model. The only caveat is the inability of the LTY predictive regression models to beat the restrictive model with respect to the X_{eff} metric. Although predictive regression models based on STY seem to outperform the passive buy-and-hold portfolio on a risk-adjusted basis, they are unable to systematically beat the two other active strategies, EXT and REST. At the 3-month horizon, the TSP predictive regression models fail to deliver meaningful economic gains (Tables 18 to 22). In contrast, STY and LTY models perform well with respect to the five profit-based metrics. LTY is again the best predictive variable.

Similar results are found at the 12-month horizon (Tables 23 to 27). Profit-based metrics (except the X_eff statistic) point to superior economic performance of the STY and LTY predictive regression models relatively to the competing passive and active strategies. Although both STY and LTY predictive models generate meaningful economic gains, investment strategies based on STY seem to be the most successful ones.

Summarizing the results for France, evidence of stock return predictability is limited to the short- and long-term interest rates. Statistical criteria point to limited forecasting gains. Nevertheless, the use of predictive investment strategies based on LTY and STY variables may lead to meaningful economic gains, as revealed by the profit-based metrics.

4.4 Germany

Table 6 points to weak statistical evidence of stock return predictability in Germany. As in Australia, the long government bond yield (LTY) is the only forecasting variable that exhibits both IS and OOS significance, but at the 1- and 3-month horizons only. At the 1-month horizon, the predictive strategies based on LTY generate negative excess returns (Table 13). The other profit-based metrics point to superior economic performance of the LTY predictive regression models relatively to the other benchmark portfolios. However, both X^* and Jensen α metrics are not statistically different from zero (Tables 14 to 17). At the 3-month horizon, economic gains provided by the LTY predictive strategies seem to be very limited (Tables 18 to 22). Although these strategies generate positive and significant Sharpe ratios, they cannot do better than the buy-and-hold portfolio. In addition, no X^* metric or Jensen α, albeit positive, is significantly different from zero. Finally, the X_{eff} metric indicates that the restrictive model performs better than the unrestricted LTY predictive strategies. To conclude, there is little evidence of stock return predictability in Germany, be it on statistical or economic grounds.

4.5 Japan

In Japan, all the predictive variables (but TSP) exhibit IS evidence of predictability (Table 7). OOS statistical evidence of predictability is thin for STY but remains substantial for the LTY, DYR and PER variables. Statistical criteria point to higher economic gains that in previous countries, but such gains remain small in size. For example, one of the highest economic gains would be provided at the 12-month horizon by the simple OLS regression with PER as predictive variable. According to the DRMSE statistic, this would amount to a 90 bp gain per year. Although this may not be entirely offset by trading costs, any serious market practitioner is not likely to care about the remaining extra-normal returns.

Tables 13 to 17 show the economic value (net of transaction costs) of predictive regression strategies at the 1-month horizon. Predictive regression strategies based on valuation ratios underperform the benchmark strategies. LTY predictive regression models show some positive economic performance, but these results are generally not robust across profit-based metrics:

while Sharpe ratios are significant, X^* and Jensen α are not. Similar results are found at the 3-month and 12-month horizons.

Overall, IS evidence of stock return predictability is undeniable in Japan, but no substantial OOS forecasting gain is realized. In addition, predictive regression strategies based on valuation ratios fail to deliver extra-normal returns. While predictive regression models based on interest rate variables show some positive economic performance, these results are generally not robust across profit-based metrics.

4.6 Netherlands

The results for the Netherlands in Table 8 show that there is IS and OOS statistical evidence of predictability for two variables only: STY and LTY. Once again, statistical criteria point to small forecasting gains, except maybe for the LTY predictive regression models at the 3-month horizon.

The economic analysis of the predictive regression strategies at the 1-month horizon points to average performance. In particular, STY and (to a lesser extent) LTY predictive strategies are unable to consistently outperform the competing benchmark portfolios. First, they obtain lower Sharpe ratios than EXT models (Table 14). Second, they do not generate significantly superior X^* metrics than REST and EXT models (Table 15). The analysis of the X_{eff} and Jensen α metrics lead to similar conclusions (Tables 16 and 17). At the 3-month horizon, the LTY predictive regression strategies are the best performers, in agreement with the statistical analysis of OOS forecasts. First, each of the three approaches (OLS, STAM, and LEW) generate the highest Sharpe ratios and X^* metrics; both of them are statistically significant (Tables 19 and 20). Second, they provide the highest Jensen α's, which are positive and close to significance (Table 22). Third, they outperform the other benchmark portfolios with respect to the X_{eff} metric (Table 21). The outperformance of the LTY predictive regression models found at the 3-month horizon disappears at the 12-month horizon (Tables 23 to 27). First, they are outperformed by the buy-and-hold and extreme models with respect to the Sharpe ratio. Second, they provide negative Jensen α's and are outperformed by the EXT model. Third, they do worse than the REST model with respect to the X_{eff} metric.

To wrap it up, statistical evidence of stock return predictability in the Netherlands is concentrated on two variables only: STY and LTY. Economic significance is even further reduced: the only outperforming predictive regression strategies are based on the LTY variable and are limited to the 3-month horizon.

4.7 South Africa

IS evidence of predictability in South Africa is found for all variables, with the exception of TSP (Table 9). However, the OOS statistical analysis shows that such evidence mainly shrinks to the STY variable. Conditioning upon IS significance, OOS evidence of predictability is not even found at the 12-month horizon. As in the preceding countries, statistical criteria point to small forecasting gains.

Turning to the economic analysis, predictive regression strategies based on the STY variable show positive and significant results at the 1-month horizon (Tables 13 to 17). Besides providing the highest positive excess returns, they generate the highest Sharpe ratios, X^* metric, and Jensen α's, all of them being significant. Economic performance at the 3-month horizon is still positive but less convincing than at the 1-month horizon (Tables 18 to 22). Sharpe Ratios provided by the LTY predictive regression strategies are positive and significant, but still lower than the Sharpe ratio delivered by the EXT benchmark model. Regarding the STY predictive regression strategies, they generate the highest Sharpe ratios but these are not significant. Similar observations apply to the X^* metric. The X_{eff} metric points to underperformance of the predictive regression strategies with respect to the restrictive model. Since no single variable is both IS and OOS significant at the 12-month horizon, no economic investigation is carried out.

All in all, no robust OOS evidence of stock return predictability is found in South Africa. While the STY variable is the most informative predictor, it fails to deliver superior economic results across investment horizons.

4.8 Sweden

From Table 10, we see that there is almost no IS statistical evidence of stock return predictability in Sweden. The relative short-term yield (STY) is the only predictive variable that shows both IS and OOS significance, mainly at the 1-month horizon.

STY predictive regression strategies show some positive economic performance at the 1-month horizon (Tables 13 to 17). Excluding X_{eff}, all profit-based metrics point to superior performance. However, Sharpe ratios are the only metrics that show statistical significance. At the 3-month horizon, the STY predictive regression model à la Lewellen fails to deliver any meaningful economic gain (Tables 18 to 22). Since no single variable is both IS and OOS significant at the 12-month horizon, no economic investigation is carried out. In summary, no robust OOS statistical evidence and economic significance is delivered by predictive regression models in Sweden.

4.9 United Kingdom

Significant IS predictive ability in the UK is found for all the variables, with the exception of TSP (Table 11). Interestingly, IS evidence of predictability found by the OLS model persists after correcting for the small sample bias in $\hat{\beta}$. It is also worth noting that both IS and OOS evidence of predictability are more obvious for valuation ratios than for interest rate variables. In any case, forecasting gains appear to be small. As the DRMSE statistic shows, the maximum potential gain is delivered at the 12-month horizon and amounts to an average of 48 bp per year only.

Looking at the 1-month economic performance (net of transaction costs), predictive regression strategies fail to outperform the benchmark portfolios (Tables 13 to 17). First, they do not provide higher Sharpe ratios than EXT, the model based on extreme values of the predictive variable. In this respect, regression strategies based on valuation ratios perform worst as they are also unable to outperform BH and REST, the two other benchmark portfolios. Second, a quick inspection of the X^* metrics reveals that predictive regression strategies underperform the restrictive model. There is, in addition, no evidence of statistical significance. Similar results are found by looking at the Jensen α's. X_{eff} is the only statistic that points to some outperformance by the predictive regression strategies. However, all X_{eff} metrics are negative, suggesting that riskless strategies provide greater utility (over a weighted average of return horizons) to constant risk averse investors.

Although statistical significance remains an issue, more meaningful economic gains are provided by the predictive regression strategies at the 3-month horizon (Tables 18 to 22). The best results are obtained through the Sharpe ratios, which are all significant and higher than in the case of the benchmark portfolios. The X^* metric and Jensen α's also point to superior performance, but none of them is statistically significant. Finally, DYR and STY predictive regression strategies generate higher X_{eff} metrics than the benchmark portfolios, although they still all display negative values. In agreement with the results at the 1-month horizon, predictive regression strategies fail to outperform the benchmark portfolios at the 12-month horizon (Tables 23 to 27). A quick inspection of the Sharpe ratios, X^* metrics, and Jensen α reveal that no predictive regression strategy is able to outperform the EXT benchmark model. X_{eff} is the only statistic that points to some outperformance by the predictive regression strategies, but X_{eff} values are all negative.

Summarizing the results for the UK, we find IS predictive ability for four (out of five) variables. Although valuation ratios exhibit undeniable OOS evidence of predictability, forecasting gains appear to be tiny. This is in general agreement with the economic performance of the predictive regression strategies. Although they sometimes generate higher economic gains than the buy-and-hold portfolio, none of them systematically outperforms the alternative benchmark strategies.

4.10 United States

All the variables exhibit some evidence of in-sample predictive ability in the US (Table 12). However, no evidence of OOS predictive ability is found for the price-earnings ratio. OOS evidence of predictability is also limited for the term spread. In any case, forecasting gains are small as indicated by the DRMSE and Theil's U statistics. This holds true whatever the predictive variable or investment horizon.

Nevertheless, the OOS economic analysis of the predictive investment strategies points to meaningful economic gains at the 1-month horizon (Tables 13 to 17). In particular, predictive investment strategies based on LTY and STY outperform the benchmark portfolios with respect to all risk-adjusted metrics. Evidence of superior performance, albeit weaker, is also provided by the TSP predictive regression strategies and the OLS regression strategy based on DYR. Interestingly, predictive regression strategies generate positive X_{eff} metrics, suggesting that they provide risk averse investors with greater utility than riskless strategies.

Results at the 3-month horizons are in agreement with the 1-month observations (Tables 18 to 22). The only significant difference comes from TSP, which is not OOS significant anymore. At the 12-month horizon, overall economic performance is still positive but less convincing than before (Tables 23 to 27). First, no valuation ratio is investigated from an economic perspective, as they do not show both IS and OOS statistical evidence of predictability at the 12-month horizon. Second, the STY predictive regression strategies are the only ones that generate both higher Sharpe ratios and X_{eff} metrics than each of the benchmark strategies. In this respect, the EXT benchmark performs better than the TSP and LTY regression strategies. Finally, both STY and TSP regression strategies do best with respect to the X^* metric and Jensen α.

Taking the results together for the US, stock returns appear to be predictable in-sample. While such evidence does not completely vanish out-of-sample, forecasting gains appear to be very limited. The economic analysis of predictive regression strategies generally confirms these findings. Nevertheless, meaningful outcomes are generated by the STY variable across each of the investment horizons.

5 Conclusion

We measure and assess the statistical and economic predictability of stock returns in 10 countries over 1-month, 3-month, and 1-year horizons. Five forecasting variables are taken into account, including both valuation ratios and interest rate variables: the dividend-yield, the price-earnings ratio, the short-term interest rate, the long-term interest rate, and the term spread.

In earlier studies, the economic significance of predictive regression models has always been measured through thought experiments. In this paper, we also measure the economic gains of statistical predictability by designing a simple investment strategy that relies upon the return forecasts of the predictive regression model. As the return forecasts depend on the slope estimate of the predictive variable, we also implement the strategy taking into account the Stambaugh (1999) or Lewellen (2004) bias correction method.

The predictive trading strategy then competes against three benchmark portfolios. The first benchmark is a passive, buy-and-hold portfolio that is fully and always invested in stocks. Two active strategies also serve as benchmarks. A naive strategy based on extremes consists in always investing the entire portfolio in stocks except when the value of the predictive variable is above the 90th percentile of its unconditional distribution. The last active strategy is based on a restricted version of the predictive regression model. In this restricted model, the predictive variable is excluded. Both passive and active strategies are evaluated on the basis of raw and risk-adjusted metrics, including the Sharpe ratio (Sharpe, 1966), the X^* statistic (Sweeney, 1988), and the X_{eff} measure (Dacorogna, Gençay, Müller, and Pictet, 2001).

Our out-of-sample statistical analysis shows that the short-term interest yield and, to a lesser extent, the long government bond yield are the most informative out-of-sample predictors of stock returns. However, the out-of-sample predictive power of these variables does not appear to be economically meaningful across countries and investment horizons.

First, thought experiments that rely upon out-of-sample statistics show that forecasting gains are small, underscoring the notion from the extant empirical literature that the predictive component in stock returns is small. According to Theil's U, the reduction in forecasting errors is never greater than 5%. Correspondingly, the DRMSE statistic shows that the highest performance across countries and investment horizons never goes above 100 bp per year, before transaction costs. Furthermore, there are cases where the only out-of-sample significant statistic is the Clark and McCracken (2001) ENC-NEW statistic designed to test for forecast encompassing. This indicates that a given predictive variable contains information that is useful in forecasting stock returns, even though no forecasting gain is realized according to mean squared error metrics.

Second, investment strategies that rely upon the return forecasts of predictive regression models also fail to deliver meaningful economic gains across investment horizons and across countries. While any risk adjustment is subject to criticism, most profit-based metrics yield the same conclusions: predictive regression strategies based on interest rate variables generate the most robust economic performance in the US, Canada, and France; they mostly fail in the other countries. Interestingly, the average performance of predictive regression strategies is similar to the one provided by the 'extreme' strategy, which consists in comparing the current value of the predictive variable to the 90th percentile of its unconditional distribution.

All in all, we find no common pattern of stock return predictability across countries, be it on statistical or economic grounds. The ability of predictive regression models to predict international stock returns appears to be very limited.

According to Goyal and Welch (2006), model instability is probably the reason why predictive regression models generally perform poorly. Model instability can be reduced through the use of structural change models (by better modelling time-changing correlations, for example). However, specification search is clearly an issue as some of these models are bound to work both in-sample or out-of-sample by pure luck. In this respect, the approach of Lettau and Van Nieuwerburgh (2005) seems promising in that it seeks to model structural change not based on the forecasting regression, but based on mean shifts in the dependent variables.

References

Ang, A., and G. Bekaert, 2004, Stock return predictability: Is it there?, Manuscript, Columbia University.

Baker, M., and J. Wurgler, 2000, The equity share in new issues and aggregate stock returns, *Journal of Finance* 55, 2219–2257.

Berge, K., and W.T. Ziemba, 2006, *The predictive ability of the bond stock earnings yield differential in US and foreign markets*, in R. Mehra (ed.), Handbook of Investments: The Equity Risk Premium, Forthcoming.

Brown, S.J., W.N. Goetzmann, and A. Kumar, 1998, The Dow theory: William Peter Hamilton's track record reconsidered, *Journal of Finance* 53, 1311–1333.

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

Campbell, J., and R. Shiller, 1988, Stock prices, earnings and expected dividends, *Journal of Finance* 43, 661–676.

Campbell, J., and S.B. Thompson, 2005, Predicting the equity premium out of sample: Can anything beat the historical average?, *NBER Working Paper Series* , No. 11468.

Campbell, J.Y., and M. Yogo, 2006, Efficient tests of stock return predictability, *Journal of Financial Economics,* Forthcoming.

Clark, T.E., and M.W. McCracken, 2001, Tests of equal forecast accuracy and encompassing for nested models, *Journal of Econometrics* 105, 85–110.

Clark, T.E., and M.W. McCracken, 2005, Evaluating direct multi-step forecasts, *Econometric Reviews* 24, 369–404.

Cochrane, J.H., 1997, Where is the Market Going? Uncertain Facts and Novel Theories, *Federal Reserve Bank of Chicago - Economic Perspectives* 21, 3–37.

Dacorogna, M.M., R. Gençay, U.A. Müller, and O.V. Pictet, 2001, Effective return, risk aversion and drawdowns, *Physica A* 289, 229–248.

Diebold, F., and R. Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* 13, 253–263.

Dowd, K., 2000, Adjusting for risk: An improved Sharpe ratio, *International Review of Economics and Finance* 9, 209–222.

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

Fama, E., and K. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Fama, E., and K. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Ferson, W.E., and C.R. Harvey, 1993, The risk and predictability of international equity returns, *Review of Financial Studies* 6, 527–566.

Goetzman, W.N., and P. Jorion, 1993, Testing the predictive power of dividend yields, *Journal of Finance* 48, 663–679.

Goyal, A., and I. Welch, 2006, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies,* Forthcoming.

Graham, B., and D.L. Dodd, 1951, *Security analysis*. (3rd edition, McGraw-Hill, New York).

Harvey, C.R., 1991, The world price of covariance risk, *Journal of Finance* 46, 111–157.

Harvey, C.R., 1995, Predictable risk and returns in emerging markets, *Review of Financial Studies* 8, 773–816.

Harvey, D., S. Leybourne, and P. Newbold, 1997, Testing the equality of prediction mean squared errors, *International Journal of Forecasting* 13, 281–291.

Hjalmarsson, E., 2004, On the predictability of global stock returns, Manuscript, Department of Economics, Yale University.

Inoue, A., and L. Kilian, 2004, In-sample or out-of-sample tests of predictability: Which one should we use?, *Econometric Reviews* 23, 371–402.

Jensen, M.C., 1968, Problems in selection of security portfolios: The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 389–416.

Keim, D.B., and R.F. Stambaugh, 1986, Predicting Returns in the Stock and Bond Markets, *Journal of Financial Economics* 17, 357390.

Kendall, M., 1954, Note on bias in the estimation of autocorrelation, *Biometrika* 41, 403–404.

Kho, B.C., 1996, Time-varying risk premia, volatility, and technical trading rule profits: Evidence from foreign currency futures markets, *Journal of Financial Economics* 41, 249–290.

Kilian, L., 1999, Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions?, *Journal of Applied Econometrics* 14, 491–510.

Kirby, C., 1997, Measuring the predictable variation in stock and bond returns, *Review of Financial Studies* 10, 579–630.

Kothari, S.P., and J. Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time series analysis, *Journal of Financial Economics* 44, 169–203.

Lettau, M., and S. Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

Lettau, M., and S. Van Nieuwerburgh, 2005, Reconciling the return predictability evidence, Manuscript, Stern Business School, NYU.

Lewellen, J., 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74, 209–235.

Mankiw, N.G., and M.D. Shapiro, 1986, Do we reject too often? Small sample properties of tests of rational expectations models, *Economic Letters* 20, 139–145.

Mark, N.C., 1995, Exchange rates and fundamentals: Evidence on long-horizon predictability, *American Economic Review* 85, 201–218.

McCracken, M.W., 2004, Asymptotics for out-of-sample tests of Granger causality, Manuscript, University of Missouri-Columbia.

Moon, R., A. Rubia, and R. Valkanov, 2006, Long-horizon regressions when the predictor is slowly varying, *Journal of Business and Economic Statistics,* Under revision.

Neely, C.J., 2003, Risk-adjusted, ex ante, optimal, technical trading rules in equity markets, *International Review of Economics and Finance* 12, 69–87.

Nelson, C.R., and M.J. Kim, 1993, Predictable stock returns: The role of small sample bias, *Journal of Finance* 48, 641–661.

Newey, W., and K.D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Polk, C., S. Thompson, and T. Vuolteenaho, 2006, Cross sectional forecasts of the equity premium, *Journal of Financial Economics,* Forthcoming.

Rangvid, J., D. Rapach, and M.E. Wohar, 2005, Macro variables and international stock return predictability, *International Journal of Forecasting* 21, 137–166.

Rapach, D., 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, 231–247.

Richardson, M., and J.H. Stock, 1989, Drawing inferences from statistics based on multiyear asset returns, *Journal of Financial Economics* 25, 323–348.

Shaman, P., and R.A. Stine, 1988, The bias of autoregressive coefficient estimators, *Journal of the American Statistical Association* 83, 842–848.

Sharpe, W.F., 1989, Mutual fund performance, *Journal of Business* 39, 119–138.

Shen, P., 2003, Market timing strategies that worked, *Journal of Portfolio Management* 29, 77–102.

Stambaugh, R., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Sutcliffe, S., 1997, *Stock Index Futures: Theories and International Evidence*. (2nd edition, Thomson Business Press, London).

Sweeney, R.J., 1988, Some new filter tests, methods and results, *Journal of Financial and Quantitative Analysis* 23, 285–300.

West, K.D., 1996, Asymptotic inference about predictive ability, *Econometrica* 64, 1067–1084.

Table 1
Summary statistics: Monthly Data.

Countries	Sample	RSR	RTR	PER	DPR	STY	LTY	TSP
A. Mean								
Australia	1969:07-2005:08	0.464	0.198	2.754	-1.577	0.014	-0.001	-0.140
Canada	1956:01-2005:09	0.440	0.197	2.854	-1.101	0.008	0.013	-0.265
France	1971:09-2005:10	0.536	0.218	2.812	-1.336	-0.072	-0.069	-0.192
Germany	1969:07-2005:10	0.357	0.171	2.725	-1.174	-0.020	-0.074	-0.298
Japan	1956:01-2005:09	0.531	0.075	3.286	-0.504	-0.074	-0.076	-1.036
Netherlands	1969:07-2005:10	0.775	0.348	2.343	-1.461	-0.044	-0.056	-0.346
South Africa	1960:02-2005:09	0.608	0.080	2.374	-1.265	0.040	0.029	-0.207
Sweden	1969:07-2005:09	0.740	0.191	2.603	-1.030	-0.064	-0.049	-0.216
United Kingdom	1950:01-2005:08	0.558	0.107	2.451	-1.505	0.041	0.014	-0.207
United States	1950:01-2005:07	0.625	0.092	2.718	-1.179	0.016	0.022	-0.352
B. Standard Deviation								
Australia	1969:07-2005:08	5.692	0.984	0.398	0.313	1.488	0.827	0.169
Canada	1956:01-2005:09	4.575	0.406	0.517	0.356	1.360	0.695	0.283
France	1971:09-2005:10	6.012	0.327	0.573	0.405	1.419	0.811	0.245
Germany	1969:07-2005:10	5.318	0.336	0.436	0.311	1.007	0.628	0.406
Japan	1956:01-2005:09	5.391	0.722	0.649	0.806	0.569	0.743	1.667
Netherlands	1969:07-2005:10	5.197	0.346	0.516	0.398	1.303	0.671	0.351
South Africa	1960:02-2005:09	6.414	0.572	0.316	0.324	1.678	0.913	0.212
Sweden	1969:07-2005:09	6.244	0.607	0.559	0.436	1.602	0.841	0.306
United Kingdom	1950:01-2005:08	5.213	0.611	0.463	0.283	1.319	0.737	0.414
United States	1950:01-2005:07	4.190	0.309	0.402	0.419	1.010	0.599	0.362
C. AC(1)								
Australia	1969:07-2005:08	0.048	0.259	0.979	0.984	0.890	0.902	0.943
Canada	1956:01-2005:09	0.088	0.219	0.978	0.988	0.919	0.886	0.956
France	1971:09-2005:10	0.099	0.600	0.946	0.985	0.933	0.927	0.970
Germany	1969:07-2005:10	0.064	0.270	0.979	0.979	0.926	0.907	0.981
Japan	1956:01-2005:09	0.041	0.134	0.992	0.995	0.931	0.903	0.974
Netherlands	1969:07-2005:10	0.071	0.420	0.986	0.990	0.910	0.908	0.947
South Africa	1960:02-2005:09	0.100	0.485	0.972	0.976	0.946	0.913	0.974
Sweden	1969:07-2005:09	0.140	0.230	0.951	0.984	0.894	0.919	0.951
United Kingdom	1950:01-2005:08	0.118	0.300	0.989	0.984	0.905	0.913	0.973
United States	1950:01-2005:07	0.038	0.430	0.988	0.991	0.881	0.884	0.968

RSR = log real total monthly stock returns; RTR = log real total monthly t-bills returns; PER = log price-earnings ratio; DPR = negative of the log dividend-price ratio; STY = relative treasury bill yield; LTY = relative long government bond yield; TSP = negative of the term spread. AC(1) = autocorrelation coefficient of the first order.

Table 2
Stock indices.

Country	Stock index
Australia	ASX-All Ordinaries
Canada	Toronto SE-300
France	SBF-250
Germany	CDAX
Japan	Nikko Securities Composite
Netherlands	Netherlands All-Share Index
South Africa	Johannesburg SE Return Index
Sweden	Affarsvrlden Return Index
United Kingdom	FTA All-Share
United States	S&P 500

This table lists the stock-index in each country to which the returns and dividends- and earnings-price ratios correspond.

Table 3
In-sample and out-of-sample statistical analysis, Australia.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.300	-0.124	0.110	-0.898	-0.715	-0.476	-3.174	-2.993	-2.768
U	0.997	1.091	1.254	1.000	1.026	1.081	0.999	0.988	0.994
MSE-F	1.855^{c}	-53.32	-121.1	0.143	-16.38	-47.93	0.798	8.083	3.986
ENC-NEW	-1.067	-3.316	-5.425	0.159	1.811	-3.632	0.442	5.025	2.597
DRMSE	0.015^{c}	-0.502	-1.397	0.002	-0.240	-0.759	0.023	0.228	0.113
2. DYR									
$\hat{\beta}$	-0.127	0.097	0.220	-0.340	-0.112	0.017	-1.731	-1.484	-1.353
U	1.010	1.010	1.013	1.021	1.020	1.020	0.999	0.999	0.999
MSE-F	-6.238	-6.820	-8.208	-13.23	-12.71	-13.04	0.582	0.510	0.510
ENC-NEW	-2.896	-3.086	-3.730	-6.336	-6.074	-6.223	0.318	0.280	0.280
DRMSE	-0.052	-0.057	-0.069	-0.192	-0.184	-0.189	0.017	0.015	0.015
3. STY									
$\hat{\beta}$	-0.071	-0.093	-0.360	-0.353	-0.376	-0.646	-2.838	-2.879	-3.361
U	1.011	1.012	1.021	1.022	1.023	1.034	1.010	1.011	1.021
MSE-F	-7.404	-7.979	-13.49	-13.91	-14.72	-21.57	-6.043	-7.016	-13.04
ENC-NEW	-2.788	-2.958	-4.719	-4.500	-4.689	-6.389	5.303^{b}	5.721^{b}	5.778^{b}
DRMSE	-0.062	-0.067	-0.115	-0.202	-0.215	-0.320	-0.176	-0.205	-0.386
4. LTY									
$\hat{\beta}$	-0.448^{b}	-0.481^{b}	-0.822^{bf}	-1.547^{b}	-1.587^{b}	-2.006^{be}	-5.660^{b}	-5.738^{b}	-6.522^{be}
U	1.009	1.010	1.012	1.006	1.006	1.008	1.007	1.008	1.016
MSE-F	-5.957	-6.242	-7.717	-3.775	-3.724	-5.158	-4.321	-4.775	-10.08
ENC-NEW	-1.984	-1.917	-1.388	0.586	0.918^{c}	1.863^{c}	12.65^{b}	13.26^{b}	14.26^{b}
DRMSE	-0.050	-0.052	-0.065	-0.054	-0.053	-0.074	-0.125	-0.139	-0.296
5. TSP									
$\hat{\beta}$	0.094	0.082	0.025	0.403	0.387	0.307	1.296	1.280	1.201
U	1.003	1.003	1.005	1.005	1.005	1.007	1.047	1.047	1.047
MSE-F	-1.738	-1.858	-3.230	-3.112	-3.307	-4.289	-28.47	-28.45	-28.41
ENC-NEW	-0.819	-0.861	-1.265	-1.510	-1.599	-2.006	-12.07	-11.89	-11.70
DRMSE	-0.014	-0.015	-0.027	-0.044	-0.047	-0.061	-0.875	-0.874	-0.873

The total (out-of-sample) sample period ranges from 1969:07 (1977:11) to 2005:08. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 4
In-sample and out-of-sample statistical analysis, Canada.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.006	0.067	0.213	-0.019	0.050	0.206	0.032	0.111	0.284
U	1.001	1.218	1.410	1.007	1.117	1.412	1.038	1.077	1.163
MSE-F	-0.842	-161.7	-246.5	-6.897	-97.68	-246.3	-34.95	-67.18	-126.3
ENC-NEW	-0.068	-11.11	-12.02	-2.558	-12.78	-18.40	-10.15	-16.88	-27.24
DRMSE	-0.004	-1.026	-1.929	-0.060	-0.995	-3.522	-0.698	-1.417	-2.983
2. DYR									
$\hat{\beta}$	-0.110	0.106	0.079	-0.363	-0.132	-0.147	-2.051	-1.800	-1.803
U	1.004	1.001	1.002	1.009	1.004	1.004	1.045	1.034	1.022
MSE-F	-4.036	-1.421	-1.458	-8.787	-4.137	-3.513	-40.67	-31.72	-20.77
ENC-NEW	-1.345	-0.157	-0.155	-2.370	-1.539	-1.225	-1.815	-1.782	-2.930
DRMSE	-0.019	-0.007	-0.007	-0.077	-0.036	-0.031	-0.820	-0.630	-0.405
3. STY									
$\hat{\beta}$	-0.361^{b}	-0.376^{b}	-0.545^{bf}	-0.957^{c}	-0.980^{c}	-1.251^{c}	-4.736^{b}	-4.771^{b}	-5.182^{bf}
U	0.999	0.999	0.999	1.002	1.002	1.004	1.010	1.012	1.010
MSE-F	0.939^{b}	1.113^{b}	1.271^{c}	-2.069	-2.273	-3.505	-9.555	-11.35	-9.555
ENC-NEW	0.960^{b}	1.134^{b}	1.838^{b}	0.357	0.333	0.298	3.827^{c}	4.174^{c}	3.827^{c}
DRMSE	0.004^{b}	0.005^{b}	0.006^{c}	-0.018	-0.020	-0.031	-0.183	-0.218	-0.183
4. LTY									
$\hat{\beta}$	-0.702^{a}	-0.718^{a}	-0.998^{ad}	-1.749^{a}	-1.778^{a}	-2.277^{ad}	-4.618^{b}	-4.660^{b}	-5.382^{be}
U	0.997	0.997	0.999	0.997	0.998	1.001	1.010	1.010	1.015
MSE-F	3.289^{a}	3.109^{a}	0.998^{c}	2.544^{b}	2.254^{b}	-0.486	-9.467	-9.959	-14.37
ENC-NEW	4.789^{a}	4.992^{a}	6.822^{a}	8.052^{a}	8.253^{a}	10.32^{a}	10.17^{b}	10.31^{b}	11.54^{b}
DRMSE	0.016^{a}	0.015^{a}	0.005^{c}	0.022^{b}	0.019^{b}	-0.004	-0.181	-0.191	-0.278
5. TSP									
$\hat{\beta}$	-0.306^{c}	-0.315^{c}	-0.367^{c}	-0.907^{c}	-1.806^{c}	-1.970^{c}	-2.747	-2.787	-3.006
U	0.999	0.999	1.000	1.004	1.004	1.005	1.003	1.004	1.008
MSE-F	1.055^{c}	0.931^{c}	0.248	-3.734	-3.895	-4.624	-3.105	-3.703	-7.185
ENC-NEW	1.428^{b}	1.459^{b}	1.704^{b}	1.317	1.382	1.834^{c}	4.048	4.114	4.478
DRMSE	0.005^{c}	0.004^{c}	0.001^{c}	-0.033	-0.034	-0.040	-0.059	-0.070	-0.137

The total (out-of-sample) sample period ranges from 1956:01 (1964:05) to 2005:09. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 5
In-sample and out-of-sample statistical analysis, France.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.021	0.040	0.328	-0.167	-0.097	0.230	0.460	0.5117	0.747
U	1.006	1.019	1.031	1.018	1.026	1.030	1.031	1.028	1.026
MSE-F	-3.798	-11.51	-18.55	-10.92	-15.35	-17.69	-17.40	-15.92	-14.64
ENC-NEW	-1.330	-3.673	-4.633	-3.883	-5.300	-5.757	-6.492	-5.899	-5.372
DRMSE	-0.036	-0.112	-0.183	-0.196	-0.279	-0.323	-0.727	-0.662	-0.607
2. DYR									
$\hat{\beta}$	-0.190	0.095	0.230	-0.637	-0.326	-0.185	-3.194	-2.826	-2.696
U	1.004	1.004	1.004	1.006	1.006	1.006	1.027	1.022	1.021
MSE-F	-2.288	-2.316	-2.310	-3.646	-3.538	-3.607	-15.36	-12.92	-12.35
ENC-NEW	-0.984	-0.676	-0.673	-1.507	-1.194	-1.180	-5.682	-5.076	-5.024
DRMSE	-0.022	-0.022	-0.022	-0.064	-0.062	-0.064	-0.638	-0.533	-0.509
3. STY									
$\hat{\beta}$	-0.639^{b}	-0.662^{b}	-0.809^{be}	-2.004^{b}	-2.037^{b}	-2.243^{be}	-6.284^{c}	-6.368^{c}	-6.870^{bf}
U	0.998	0.998	0.998	0.997	0.997	0.997	1.007	1.007	1.009
MSE-F	1.196^{b}	1.277^{b}	1.456^{b}	1.658^{c}	1.853^{c}	2.102^{c}	-4.266	-4.463	-5.118
ENC-NEW	0.955^{b}	1.053^{b}	1.356^{b}	2.070^{c}	2.248^{c}	2.655^{c}	0.662	0.678	0.756
DRMSE	0.011^{b}	0.012^{b}	0.014^{b}	0.029^{c}	0.032^{c}	0.037^{c}	-0.172	-0.180	-0.207
4. LTY									
$\hat{\beta}$	-0.947^{a}	-0.981^{a}	-1.217^{ad}	-2.888^{a}	-2.930^{a}	-3.212^{ad}	-8.535^{a}	-8.653^{a}	-9.431^{ad}
U	0.999	0.999	1.000	0.993	0.993	0.993	0.991	0.991	0.991
MSE-F	0.486^{c}	0.467^{c}	0.203	4.492^{a}	4.514^{a}	4.606^{a}	5.428^{b}	5.573^{b}	5.591^{b}
ENC-NEW	1.660^{b}	1.928^{b}	2.641^{b}	6.675^{a}	6.978^{a}	7.906^{a}	9.997^{b}	10.46^{b}	11.77^{b}
DRMSE	0.005^{c}	0.004^{c}	0.002	0.078^{a}	0.078^{a}	0.080^{a}	0.214^{b}	0.219^{b}	0.220^{b}
5. TSP									
$\hat{\beta}$	-0.391^{c}	-0.393^{c}	-0.396^{c}	-1.157^{c}	-1.178^{c}	-1.215^{c}	-3.641	-3.710	-3.817
U	1.000	1.000	1.000	0.999	0.999	1.000	0.995	0.995	0.996
MSE-F	-0.087	-0.058	-0.112	0.774	0.479	-0.107	2.951	2.756	2.205
ENC-NEW	0.557	0.584	0.587	1.913^{c}	1.856^{c}	1.757^{c}	4.441	4.479	4.536
DRMSE	-0.001	-0.001	-0.001	0.014	0.008	-0.002	0.117	0.109	0.088

The total (out-of-sample) sample period ranges from 1971:09 (1980:02) to 2005:10. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 6
In-sample and out-of-sample statistical analysis, Germany.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.366	-0.236	-0.072^{c}	-0.998	-0.877	-0.727	-2.487	-2.453	-2.412
U	1.002	1.047	1.210	1.003	1.014	1.054	1.013	1.018	1.023
MSE-F	-0.975	-29.40	-106.0	-2.189	-9.195	-33.41	-8.201	-11.16	-14.42
ENC-NEW	-0.107	0.238	2.602	-0.401	-0.895	-0.532	-0.892	-2.208	-3.486
DRMSE	-0.008	-0.262	-1.169	-0.034	-0.145	-0.558	-0.292	-0.400	-0.521
2. DYR									
$\hat{\beta}$	-0.202	-0.004	0.169	-0.653	-0.447	-0.267	-2.700	-2.481	-2.348
U	1.004	1.001	1.001	1.014	1.001	1.002	1.055	1.044	1.043
MSE-F	-2.411	-0.445	-0.486	-9.301	-0.763	-1.196	-33.13	-26.94	-26.04
ENC-NEW	-0.383	0.226	0.218	-2.025	0.016	-0.207	-8.374	-6.595	-6.370
DRMSE	-0.020	-0.004	-0.004	-0.147	-0.012	-0.019	-1.254	-1.004	-0.968
3. STY									
$\hat{\beta}$	-0.532^{b}	-0.535^{b}	-0.560^{be}	-1.606^{b}	-1.616^{b}	-1.686^{be}	-5.367^{a}	-5.413^{a}	-5.747^{ae}
U	0.997	0.997	0.997	0.995	0.995	0.995	0.998	0.998	0.997
MSE-F	1.963^{b}	2.019^{b}	2.157^{b}	3.072^{b}	3.232^{b}	3.633^{b}	1.381	1.544	1.914
ENC-NEW	1.357^{b}	1.407^{b}	1.557^{b}	2.503^{b}	2.628^{b}	2.992^{b}	2.290	2.437	2.893
DRMSE	0.016^{b}	0.017^{b}	0.018^{b}	0.047^{b}	0.050^{b}	0.056^{b}	0.048	0.054	0.067
4. LTY									
$\hat{\beta}$	-0.209	-0.199	-0.100	-0.685	-0.675	-0.578	-1.770	-1.779	-1.867
U	1.000	1.000	1.000	0.999	0.999	0.999	1.000	1.000	1.000
MSE-F	0.347	0.321	0.114	0.797^{c}	0.783^{c}	0.652^{c}	0.227	0.201	0.169
ENC-NEW	0.180	0.167	0.060	0.414	0.407	0.337	0.153	0.138	0.114
DRMSE	0.003	0.003	0.001	0.012^{c}	0.012^{c}	0.010^{c}	0.008	0.007	0.006
5. TSP									
$\hat{\beta}$	-0.1256	-0.143	-0.163	-0.252	-0.285	-0.321	0.128	0.052	-0.023
U	1.000	1.001	1.001	1.000	1.000	1.000	1.005	1.006	1.006
MSE-F	-0.244	-0.499	-0.512	0.323	0.166	0.353	-3.480	-3.641	-3.546
ENC-NEW	-0.105	-0.192	-0.202	0.231	0.244	0.339	-0.118	0.078	0.154
DRMSE	-0.002	-0.004	-0.004	0.005	0.003	0.005	-0.123	-0.128	-0.125

The total (out-of-sample) sample period ranges from 1969:07 (1977:11) to 2005:10. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 7
In-sample and out-of-sample statistical analysis, Japan.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.418^{c}	-0.219^{c}	-0.189^{bf}	-1.374^{c}	-1.117^{c}	-1.077^{c}	-6.421^{b}	-6.075^{b}	-6.031^{bf}
U	0.997	1.054	1.066	0.990	1.027	1.024	0.960	0.979	0.967
MSE-F	2.733^{c}	-49.85	-59.72	10.18^{b}	-25.56	-22.79	41.69^{b}	21.17^{c}	34.03^{b}
ENC-NEW	2.413	-7.001	-5.359	8.021	-5.011	-0.530	30.91^{c}	21.46	28.91
DRMSE	0.014^{c}	-0.284	-0.346	0.098^{b}	-0.259	-0.230	0.900^{b}	0.471^{c}	0.742^{b}
2. DYR									
$\hat{\beta}$	-0.387^{c}	-0.120	-0.211^{c}	-1.274^{c}	-0.914	-1.029^{c}	-5.519	-5.053	-5.210
U	0.998	0.998	0.997	0.998	0.997	0.996	0.986	0.985	0.985
MSE-F	2.064^{c}	2.003^{b}	2.665^{b}	1.854	3.483^{c}	4.066^{c}	13.87	14.498^{c}	14.480^{c}
ENC-NEW	1.877	2.093^{c}	2.450^{c}	3.304	4.558	4.891	14.20	14.694	14.730
DRMSE	0.011^{c}	0.011^{b}	0.014^{b}	0.018	0.034^{c}	0.039^{c}	0.312^{c}	0.326^{c}	0.325^{c}
3. STY									
$\hat{\beta}$	-0.426^{b}	-0.435^{b}	-0.529^{be}	-1.366^{c}	-1.379^{c}	-1.511^{c}	-4.320^{c}	-4.368^{c}	-4.832^{bf}
U	1.005	1.005	1.006	1.013	1.014	1.016	1.020	1.021	1.028
MSE-F	-4.571	-4.778	-6.061	-12.82	-13.11	-15.07	-18.99	-20.05	-26.19
ENC-NEW	0.370	0.396	0.525	1.431	1.467	1.540	9.282^{c}	9.364^{c}	9.640^{c}
DRMSE	-0.024	-0.025	-0.032	-0.128	-0.131	-0.150	-0.449	-0.475	-0.627
4. LTY									
$\hat{\beta}$	-0.555^{a}	-0.559^{a}	-0.616^{ae}	-1.502^{b}	-1.514^{b}	-1.697^{be}	-2.626	-2.668	-3.253
U	0.999	0.999	0.999	0.995	0.995	0.996	0.996	0.996	0.995
MSE-F	1.473^{b}	1.461^{b}	1.011^{c}	4.800^{b}	4.815^{b}	4.070^{b}	3.995^{c}	4.227^{c}	4.769^{c}
ENC-NEW	1.640^{b}	1.661^{b}	1.643^{b}	4.376^{b}	4.437^{b}	4.468^{b}	2.762	2.931	3.598
DRMSE	0.008^{b}	0.008^{b}	0.005^{c}	0.047^{b}	0.047^{b}	0.040^{b}	0.091^{c}	0.096^{c}	0.109^{c}
5. TSP									
$\hat{\beta}$	-0.010	-0.002	-0.008	0.271	0.257	0.266	2.392	2.394	2.394
U	1.022	1.022	1.022	1.053	1.054	1.054	1.105	1.106	1.107
MSE-F	-20.682	-20.97	-20.99	-48.77	-49.03	-49.11	-87.68	-88.64	-88.96
ENC-NEW	-2.950	-3.025	-2.972	-7.904	-7.978	-7.999	-13.10	-13.38	-13.36
DRMSE	-0.112	-0.114	-0.114	-0.514	-0.517	-0.518	-2.335	-2.365	-2.375

The total (out-of-sample) sample period ranges from 1956:01 (1964:05) to 2005:09. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript 'd' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 8
In-sample and out-of-sample statistical analysis, Netherlands

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.188	-0.025	0.051	-0.743	-0.538	-0.448	-3.825	-3.610	-3.520
U	0.999	0.999	1.008	1.003	0.994	0.991	1.018	1.013	1.003
MSE-F	0.750	0.438	-5.003	-2.064	3.969	6.036^{c}	-11.508	-8.123	-2.196
ENC-NEW	0.589	2.527	4.636	-0.276	3.731	7.071	0.131	1.841	5.024
DRMSE	0.006	0.003	-0.040	-0.029	0.056	0.085^{c}	-0.368	-0.257	-0.069
2. DYR									
$\hat{\beta}$	-0.249	0.024	-0.004	-0.879	-0.565	-0.598	-3.643	-3.338	-3.372
U	1.002	1.007	1.007	1.006	1.014	1.015	1.011	1.014	1.022
MSE-F	-1.118	-4.682	-4.534	-4.164	-9.277	-9.506	-7.094	-9.158	-13.562
ENC-NEW	0.392	-1.008	-0.952	0.545	-2.446	-2.890	4.126	1.470	-1.287
DRMSE	-0.009	-0.037	-0.036	-0.060	-0.135	-0.138	-0.224	-0.291	-0.435
3. STY									
$\hat{\beta}$	-0.566^{b}	-0.577^{b}	-0.677^{be}	-1.480^{b}	-1.504^{b}	-1.736^{be}	-4.472^{c}	-4.516^{c}	-4.919^{c}
U	0.995	0.995	0.994	0.990	0.989	0.988	0.993	0.993	0.991
MSE-F	3.354^{a}	3.503^{a}	3.972^{a}	6.719^{b}	7.149^{b}	8.518^{b}	4.402	4.728	6.032^{b}
ENC-NEW	2.274^{b}	2.394^{b}	2.856^{b}	4.524^{b}	4.813^{b}	5.898^{b}	4.122	4.333	5.286^{b}
DRMSE	0.026^{b}	0.027^{a}	0.031^{a}	0.094^{b}	0.100^{b}	0.119^{b}	0.136	0.145	0.185^{b}
4. LTY									
$\hat{\beta}$	-0.991^{a}	-1.000^{a}	-1.079^{ad}	-2.753^{a}	-2.782^{a}	-3.0508^{ad}	-6.208^{a}	-6.272^{a}	-6.881^{ad}
U	0.991	0.991	0.991	0.961	0.960	0.958	0.967	0.965	0.959
MSE-F	6.422^{a}	6.429^{a}	6.103^{a}	27.69^{a}	28.39^{a}	29.81^{a}	22.50^{b}	23.65^{b}	28.03^{a}
ENC-NEW	7.580^{a}	7.802^{a}	8.555^{a}	23.27^{a}	24.19^{a}	27.48^{a}	19.94^{b}	20.81^{b}	24.73^{b}
DRMSE	0.050^{a}	0.050^{a}	0.047^{a}	0.370^{a}	0.379^{a}	0.397^{a}	0.665^{b}	0.697^{b}	0.819^{a}
5. TSP									
$\hat{\beta}$	-0.008	-0.019	-0.075	0.114	0.097	0.012	1.217	1.189	1.055
U	1.000	1.000	0.999	0.996	0.996	0.996	0.983	0.983	0.983
MSE-F	0.195	0.305	0.584	2.416	2.573^{c}	2.904^{c}	11.29^{c}	11.40^{c}	11.51^{c}
ENC-NEW	0.151	0.216	0.402	1.600	1.702	1.702	10.30	10.40	10.60
DRMSE	0.002	0.002	0.005	0.034	0.036^{c}	0.036^{c}	0.342^{c}	0.345^{c}	0.349^{c}

The total (out-of-sample) sample period ranges from 1969:07 (1977:11) to 2005:10. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript 'd' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 9
In-sample and out-of-sample statistical analysis, South Africa.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.555^{b}	-0.412^{b}	-0.017^{b}	-1.784^{c}	-1.632^{c}	-1.214^{b}	-5.990^{c}	-5.851^{c}	-5.477^{c}
U	1.001	1.089	1.153	1.005	1.054	1.071	1.044	1.070	1.068
MSE-F	-0.486	-69.95	-110.5	-4.766	-44.58	-57.08	-36.07	-55.17	-53.96
ENC-NEW	0.514	-12.81	-16.51	-0.302	-7.655	-10.33	-5.751	-9.926	-8.267
DRMSE	-0.004	-0.608	-1.044	-0.068	-0.683	-0.896	-1.137	-1.803	-1.760
2. DYR									
$\hat{\beta}$	-0.408^{c}	-0.324^{c}	-0.132^{c}	-1.446^{c}	-1.3478^{c}	-1.1309^{c}	-5.777	-5.700	-5.533
U	1.002	1.002	1.002	1.011	1.009	1.009	1.071	1.072	1.070
MSE-F	-2.072	-1.522	-1.710	-9.918	-7.451	-7.719	-56.13	-56.37	-55.48
ENC-NEW	-0.529	0.057	0.184	-3.840	-2.772	-2.884	-21.03	-21.68	-21.27
DRMSE	-0.016	-0.012	-0.013	-0.143	-0.107	-0.111	-1.838	-1.847	-1.814
3. STY									
$\hat{\beta}$	-0.561^{b}	-0.576^{b}	-0.677^{be}	-1.464^{b}	-1.500^{b}	-1.736^{be}	-2.091	-2.155	-2.577
U	1.001	1.001	1.001	0.993	0.993	0.993	0.985	0.985	0.982
MSE-F	-0.805	-0.883	-0.959	6.388^{a}	6.399^{a}	6.783^{a}	13.77^{a}	13.79^{a}	15.78^{a}
ENC-NEW	1.052^{b}	1.027^{c}	1.070^{c}	7.639^{a}	7.792^{a}	8.448^{a}	14.80^{a}	15.14^{a}	16.90^{a}
DRMSE	-0.006	-0.007	-0.007	0.090^{a}	0.090^{a}	0.095^{a}	0.398^{a}	0.398^{a}	0.778^{a}
4. LTY									
$\hat{\beta}$	-0.378^{c}	-0.399^{c}	-0.6393^{c}	-1.198^{b}	-1.224^{b}	-1.518^{be}	-1.478	-1.527	-2.082
U	1.004	1.005	1.006	1.003	1.003	1.005	1.017	1.017	1.017
MSE-F	-3.703	-4.003	-5.123	-2.240	-2.468	-4.135	-14.14	-14.06	-14.299
ENC-NEW	-0.554	-0.597	-0.706	2.929^{b}	3.023^{b}	2.955^{b}	1.246	1.479	1.931
DRMSE	-0.029	-0.031	-0.040	-0.032	-0.035	-0.059	-0.428	-0.426	-0.433
5. TSP									
$\hat{\beta}$	-0.305	-0.303	-0.297	-0.708	-0.740	-0.822	-0.803	-0.870	-1.035
U	1.002	1.002	1.003	1.008	1.009	1.010	1.028	1.030	1.037
MSE-F	-2.076	-2.114	-2.184	-7.262	-7.823	-8.880	-23.30	-25.06	-30.35
ENC-NEW	-0.059	-0.050	0.026	-0.819	-0.910	-0.718	-5.680	-6.128	-6.816
DRMSE	-0.016	-0.016	-0.017	-0.104	-0.112	-0.128	-0.717	-0.774	-0.946

The total (out-of-sample) sample period ranges from 1960:02 (1968:06) to 2005:09. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 10
In-sample and out-of-sample statistical analysis, Sweden.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.059	-0.008	0.220	-0.507	-0.437	-0.127	-3.424	-3.361	-3.095
U	1.001	1.015	1.036	0.999	1.001	1.009	0.997	0.995	0.994
MSE-F	-0.312	-9.676	-22.74	0.384	-0.465	-5.693	1.765	3.109	3.890
ENC-NEW	0.101	1.341	4.357	0.677	0.467	0.206	1.678	2.291	2.579
DRMSE	-0.003	-0.097	-0.235	0.007	-0.009	-0.109	0.078	0.137	0.171
2. DYR									
$\hat{\beta}$	-0.187	0.049	0.237	-0.689	-0.427	-0.220	-4.461	-4.125	-3.891
U	1.002	1.002	1.002	1.002	1.003	1.003	1.001	1.002	1.003
MSE-F	-1.033	-1.343	-1.255	-1.014	-1.699	-1.993	-0.674	-1.225	-1.596
ENC-NEW	-0.348	-0.148	-0.117	-0.013	-0.166	-0.263	0.754	0.413	0.195
DRMSE	-0.010	-0.013	-0.012	-0.019	-0.032	-0.038	-0.030	-0.054	-0.071
3. STY									
$\hat{\beta}$	-0.599^{b}	-0.619^{b}	-0.840^{be}	-0.802	-0.848	-1.377^{c}	-0.205	-0.259	-0.857
U	0.998	0.998	0.997	0.999	0.998	0.996	1.000	1.000	0.998
MSE-F	1.558^{b}	1.660^{b}	2.217^{b}	0.977	1.123^{c}	2.931^{b}	0.093	0.245	1.537
ENC-NEW	1.088^{b}	1.177^{b}	1.718^{b}	0.615	0.722	1.839^{c}	0.056	0.133	0.793
DRMSE	0.015^{b}	0.016^{b}	0.022^{b}	0.018^{c}	0.021^{c}	0.055^{b}	0.004	0.011	0.068
4. LTY									
$\hat{\beta}$	-0.195	-0.221	-0.435	-0.386	-0.426	-0.748	-1.656	-1.698	-2.037
U	1.001	1.001	1.001	1.002	1.002	1.002	1.002	1.002	1.001
MSE-F	-0.949	-0.910	-0.559	-1.453	-1.390	-1.185	-1.561	-1.424	-0.760
ENC-NEW	-0.334	-0.273	0.251	-0.568	-0.508	-0.172	-0.640	-0.562	-0.138
DRMSE	-0.009	-0.009	-0.006	-0.028	-0.026	-0.022	-0.069	-0.063	-0.034
5. TSP									
$\hat{\beta}$	-0.116	-0.126	-0.165	0.263	0.221	0.056	4.777	4.733	4.562
U	1.000	1.000	1.000	1.000	1.000	0.999	1.006	1.006	1.006
MSE-F	-0.011	0.016	0.151	-0.148	0.025	0.623	-3.644	-3.696	-3.894
ENC-NEW	0.017	0.033	0.116	0.136	0.245	0.659	0.088	0.073	0.022
DRMSE	-0.000	0.000	0.002	-0.003	0.001	0.012	-0.163	-0.165	-0.174

The total (out-of-sample) sample period ranges from 1969:07 (1977:11) to 2005:09. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 11
In-sample and out-of-sample statistical analysis, United Kingdom.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.404^{b}	-0.213^{c}	-0.068^{ae}	-1.285^{c}	-1.079	-0.925^{c}	-5.423^{b}	-5.227^{c}	-5.088^{be}
U	0.997	1.035	1.045	0.998	1.004	0.994	0.982	0.982	0.981
MSE-F	3.206^{b}	-38.05	-47.78	1.802	-4.136	6.680^{b}	20.20^{c}	20.82^{b}	21.68^{b}
ENC-NEW	2.581^{c}	-4.970	3.181	2.935	0.681	9.987^{b}	15.32^{c}	16.33^{c}	16.99^{c}
DRMSE	0.015^{b}	-0.191	-0.243	0.016	-0.037	0.058^{b}	0.376^{c}	0.387^{b}	0.403^{b}
2. DYR									
$\hat{\beta}$	-0.589^{a}	-0.465^{a}	-0.3130^{ad}	-1.788^{b}	-1.664^{b}	-1.510^{be}	-7.616^{a}	-7.502^{a}	-7.358^{ad}
U	0.998	1.000	0.995	0.996	1.000	0.998	0.977	0.979	0.985
MSE-F	2.857^{a}	0.498	5.278^{b}	4.970^{b}	0.521	1.975	25.97^{b}	23.76^{b}	16.87^{b}
ENC-NEW	2.692^{b}	2.206	4.981^{b}	4.807^{c}	2.422	3.423	26.67^{b}	23.04^{b}	18.63^{c}
DRMSE	0.014^{a}	0.002	0.025^{b}	0.044^{b}	0.005	0.017	0.480^{b}	0.441^{b}	0.316^{b}
3. STY									
$\hat{\beta}$	-0.261^{c}	-0.278^{c}	-0.563^{c}	-0.872^{c}	-0.893^{c}	-1.229^{c}	-3.661	-3.694	-4.204
U	1.003	1.003	1.006	1.006	1.006	1.008	1.038	1.039	1.045
MSE-F	-3.659	-3.873	-6.238	-6.238	-6.303	-8.559	-39.92	-40.47	-46.34
ENC-NEW	-0.158	0.005	0.350	1.698^{c}	2.009^{b}	2.560^{c}	0.515	0.698	0.984
DRMSE	-0.018	-0.019	-0.030	-0.055	-0.056	-0.076	-0.808	-0.819	-0.946
4. LTY									
$\hat{\beta}$	-0.430^{b}	-0.451^{b}	-0.759^{be}	-1.224	-1.254	-1.681	-3.670	-3.720	-4.431^{c}
U	1.007	1.007	1.010	1.010	1.011	1.015	1.019	1.021	1.025
MSE-F	-7.738	-8.267	-11.30	-11.42	-12.25	-16.11	-21.00	-22.28	-26.73
ENC-NEW	1.231^{b}	1.368^{b}	1.795^{b}	7.245^{a}	7.487^{a}	8.144^{a}	14.40^{a}	14.88^{a}	15.95^{a}
DRMSE	-0.037	-0.040	-0.055	-0.102	-0.110	-0.014	-0.414	-0.439	-0.531
5. TSP									
$\hat{\beta}$	-0.033	-0.057	-0.139	-0.058	-0.093	-0.211	0.361	0.321	0.186
U	1.000	1.000	1.000	1.001	1.001	1.001	1.013	1.013	1.013
MSE-F	-0.409	-0.244	-0.369	-1.012	-0.830	-1.053	-14.25	-14.57	-14.455
ENC-NEW	-0.150	0.001	-0.051	-0.171	0.087	-0.041	-4.011	-4.063	-4.007
DRMSE	-0.002	-0.001	-0.002	-0.009	-0.007	-0.009	-0.278	-0.284	-0.282

The total (out-of-sample) sample period ranges from 1950:01 (1958:05) to 2005:07. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript 'd' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 12
In-sample and out-of-sample statistical analysis, United States.

	k=1			k=3			k=12		
	OLS	STAM	LEW	OLS	STAM	LEW	OLS	STAM	LEW
1. PER									
$\hat{\beta}$	-0.286	-0.134	-0.008^{be}	-0.862	-0.713	-0.586^{c}	-3.721	-3.560	-3.428
U	1.000	1.149	1.180	1.002	1.077	1.085	0.997	1.031	1.033
MSE-F	0.028	-137.4	-159.3	-1.981	-77.70	-85.00	3.029	-33.25	-35.03
ENC-NEW	1.218	-15.06	-18.34	2.722	-18.18	-16.20	22.86	-0.343	2.999
DRMSE	0.000	-0.643	-0.775	-0.014	-0.591	-0.654	0.045	-0.523	-0.552
2. DYR									
$\hat{\beta}$	-0.334^{c}	-0.103	-0.097^{ad}	-1.023^{c}	-0.793	-0.792^{be}	-4.375	-4.126	-4.130
U	0.997	0.999	0.999	0.991	0.996	0.998	0.959	0.964	0.959
MSE-F	3.732^{b}	1.183	1.177	10.61^{b}	4.102	2.719	49.02^{b}	41.98^{b}	48.80^{b}
ENC-NEW	3.702^{b}	4.372^{b}	4.325^{b}	10.50^{b}	11.371^{b}	9.333^{b}	48.14^{b}	50.41^{b}	48.76^{b}
DRMSE	0.014^{b}	0.004	0.004	0.071^{b}	0.028	0.018	0.690^{b}	0.596^{b}	0.687^{b}
3. STY									
$\hat{\beta}$	-0.624^{a}	-0.627^{a}	-0.687^{ad}	-1.491^{a}	-1.501^{a}	-1.720^{ad}	-4.256^{b}	-4.282^{b}	-4.816^{be}
U	0.997	0.997	0.998	0.997	0.997	1.000	1.034	1.036	1.046
MSE-F	3.511^{a}	3.412^{a}	1.922^{b}	3.442^{b}	3.184^{b}	-0.251	-35.99	-37.55	-48.18
ENC-NEW	9.091^{a}	9.104^{a}	9.626^{a}	18.76^{a}	18.88^{a}	20.20^{a}	40.45^{a}	40.68^{a}	41.69^{a}
DRMSE	0.013^{a}	0.013^{a}	0.007^{b}	0.023^{b}	0.022^{b}	-0.002	-0.568	-0.594	-0.774
4. LTY									
$\hat{\beta}$	-0.689^{a}	-0.699^{a}	-0.903^{ad}	-1.698^{a}	-1.717^{a}	-2.108^{ad}	-3.619^{b}	-3.654^{b}	-4.353^{ae}
U	1.000	1.000	1.005	1.000	1.000	1.006	1.019	1.020	1.031
MSE-F	0.315	-0.021	-5.280	0.403	-0.105	-6.799	-20.94	-21.83	-32.57
ENC-NEW	10.95^{a}	11.06^{a}	13.08^{a}	24.33^{a}	24.57^{a}	28.12^{a}	31.87^{a}	32.27^{a}	35.11^{a}
DRMSE	0.001	-0.000	-0.020	0.003	-0.001	-0.047	-0.324	-0.337	-0.511
5. TSP									
$\hat{\beta}$	-0.402^{a}	-0.401^{a}	-0.393^{ae}	-1.109^{b}	-1.117^{b}	-1.150^{bf}	-3.662^{c}	-3.692^{c}	-3.802^{c}
U	1.011	1.011	1.011	1.033	1.033	1.036	1.070	1.072	1.083
MSE-F	-12.31	-12.26	-11.74	-35.29	-35.91	-38.64	-69.65	-72.24	-82.24
ENC-NEW	1.702^{c}	1.749^{c}	1.850^{c}	3.318	3.402	3.279	16.86^{c}	16.60^{c}	14.97^{c}
DRMSE	-0.048	-0.047	-0.045	-0.252	-0.257	-0.277	-1.155	-1.203	-1.391

The total (out-of-sample) sample period ranges from 1950:01 (1958:05) to 2005:07. The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model (Equation 1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$; U is Theil's U as defined in Section 2.2.1; MSE-F, DRMSE, and ENC-NEW are given in Equations (7, 8, and 9. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. No significance test is carried out on U. An additional $d/e/f$ subscript on $\hat{\beta_{LEW}}$ indicates *joint* significance (given by the modified Bonferroni's upper bound) at the 1/5/10% level respectively (see Section 2.1.2). For example, subscript '*d*' indicates joint significance at the 1% level, meaning that the null hypothesis ($\beta = 0$) is rejected by using either Stambaugh's or Lewellen's approach at the 1% level.

Table 13
Out-of-sample economic analysis, 1-month excess returns ($k = 1$).

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST		0.5095	-0.4121	-3.3873	-0.1158	-2.2642	2.9704	-4.8351	-0.9801	-1.5092
1. PER										
EXT					-3.2858				-0.6876	
OLS					-2.9449				-2.6735	
STAM										
LEW										
2. DYR										
EXT					-3.1223				0.2363	-1.4964
OLS					-3.2664				-2.8072	-1.1978
STAM										
LEW					-4.7273				-3.5970	-5.1925
3. STY										
EXT		0.2102	-0.2840	0.1156		0.4093	0.7829	0.3484		0.3606
OLS		0.6285	-0.5378	-0.3981		-2.3137	4.1552	-3.1140		0.4696
STAM		0.9525	-0.1155	-0.4180		-2.3505	4.1416	-2.4580		0.4696
LEW		1.1024	-0.1349	-0.3663		-3.1927	4.3700	-2.3523		0.3893
4. LTY										
EXT		2.7656	1.1868		1.2616	0.6308			1.1671	0.3459
OLS		1.3232	1.5704		0.8531	-0.9612			-0.7237	0.4105
STAM		1.0202	1.8222		0.8990	-0.9612			-1.0132	0.3378
LEW		1.3516	2.6414		0.7694	-1.3408			-0.5984	0.8042
5. TSP										
EXT		0.2630								1.7433
OLS		-0.2113								0.1805
STAM		-0.0293								0.2793
LEW		0.1765								0.4371

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 14
Out-of-sample economic analysis, 1-month horizon ($k = 1$), Sharpe Ratio.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
BH		0.1543	0.2457	0.1778	0.2586	0.3976	0.2139	0.4421	0.2632	0.3283
REST		0.3526^{a}	0.3830^{a}	0.0032^{c}	0.2638^{a}	0.3357^{a}	0.5753^{c}	0.3799^{a}	0.3228^{a}	0.2929^{a}
1. PER										
EXT					0.0968^{a}				0.2712^{a}	
OLS					0.1233^{a}				0.2468^{b}	
STAM										
LEW										
2. DYR										
EXT					0.1177^{a}				0.3244^{a}	0.2864^{a}
OLS					0.1057^{a}				0.2226^{a}	0.4144^{a}
STAM										
LEW					-0.0050				0.1913^{a}	0.2192
3. STY										
EXT		0.1867^{a}	0.2336^{a}	0.1856^{a}		0.4311^{a}	0.2825	0.4650^{a}		0.3930^{a}
OLS		0.3412^{a}	0.3399^{a}	0.2025^{b}		0.3103^{a}	0.6314^{b}	0.4622^{a}		0.4539^{a}
STAM		0.3721^{a}	0.3671^{a}	0.2020^{b}		0.3080^{a}	0.6272^{b}	0.5001^{a}		0.4539^{a}
LEW		0.3831^{a}	0.3628^{a}	0.2055^{a}		0.2574^{a}	0.6431^{b}	0.5055^{a}		0.4474^{a}
4. LTY										
EXT		0.3961^{a}	0.3229^{a}		0.3529^{b}	0.4544^{a}			0.4168^{a}	0.4219^{a}
OLS		0.3642^{a}	0.4538^{a}		0.3377^{b}	0.4189^{a}			0.3582^{a}	0.4707^{a}
STAM		0.3322^{a}	0.4642^{a}		0.3406^{b}	0.4189^{a}			0.3357^{a}	0.4645^{a}
LEW		0.3551^{a}	0.5165^{a}		0.3330^{b}	0.3946^{a}			0.3665^{a}	0.5143^{a}
5. TSP										
EXT		0.1936^{a}								0.4942^{a}
OLS		0.2680^{a}								0.4888^{a}
STAM		0.2883^{a}								0.4974^{a}
LEW		0.3093^{a}								0.5110^{a}

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 15
Out-of-sample economic analysis, 1-month horizon ($k = 1$), X^* metric.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST		0.2437^{c}	0.2538	-0.0424	0.0015	0.0466	0.5909^{c}	0.1085	0.1934	0.0233
1. PER										
EXT					-0.1049				0.1157	
OLS					-0.0381				0.1566	
STAM										
LEW										
2. DYR										
EXT					-0.0762				0.1678^{c}	0.0452
OLS					-0.0779				0.1280	0.1633^{c}
STAM										
LEW					-0.1651				0.1114	-0.0285
3. STY										
EXT		0.0475	-0.0077	0.0255		0.0558	0.1438	0.0766		0.1150^{c}
OLS		0.2181	0.2194	0.1435		0.0204	0.6453^{b}	0.2061		0.1770
STAM		0.2411^{c}	0.2433	0.1422		0.0146	0.6413^{b}	0.2554		0.1770
LEW		0.2431^{c}	0.2404	0.1434		-0.0361	0.6582^{b}	0.2651		0.1711
4. LTY										
EXT		0.2769^{a}	0.1273		0.1268	0.1063			0.1632^{c}	0.1337^{b}
OLS		0.2419	0.3286		0.1252	0.1215			0.1845	0.2015^{c}
STAM		0.2161	0.3396		0.1306	0.1215			0.1595	0.1962^{c}
LEW		0.2408	0.4022		0.1231	0.0984			0.1854	0.2387^{b}
5. TSP										
EXT		0.0653								0.2262^{b}
OLS		0.1452								0.2175^{c}
STAM		0.1591								0.2250^{c}
LEW		0.1710								0.2351^{c}

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 16

Out-of-sample economic analysis, 1-month horizon ($k = 1$), X_{eff} metric.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
BH		-0.9409	-1.8273	-1.6088	-1.1738	-1.0976	-2.0401	-1.9604	-1.2178	0.1513
REST		-0.1108	-0.2782	-0.9091	-1.0634	-0.6652	-0.3120	-0.4568	-0.4088	0.0938
1. PER										
EXT					-0.7846				-0.6837	
OLS					-0.7679				-0.2137	
STAM										
LEW										
2. DYR										
EXT					-0.6671				-0.6360	0.0946
OLS					-0.7659				-0.2210	0.1308
STAM										
LEW					-0.8025				-0.1336	-0.0270
3. STY										
EXT		-0.6978	-1.8534	-1.5603		-1.0350	-1.4185	-1.8957		0.1809
OLS		-0.1812	-0.5439	-0.8004		-1.0357	-0.2780	-0.5417		0.1963
STAM		-0.1662	-0.5069	-0.7719		-1.0387	-0.2898	-0.5461		0.1963
LEW		-0.1712	-0.5294	-0.7673		-1.0866	-0.2634	-0.5286		0.1922
4. LTY										
EXT		-0.4010	-1.4850		-0.9012	-0.9890			-0.5190	0.1867
OLS		-0.2007	-0.4474		-0.8767	-0.7725			-0.3365	0.1993
STAM		-0.2405	-0.4641		-0.8733	-0.7725			-0.3468	0.1956
LEW		-0.2432	-0.4283		-0.8814	-0.7985			-0.3240	0.2214
5. TSP										
EXT		-0.6663								0.2536
OLS		-0.1885								0.1919
STAM		-0.1786								0.1968
LEW		-0.1703								0.2052

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 17

Out-of-sample economic analysis, 1-month horizon ($k = 1$), Jensen α.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST		2.3144^{b}	2.8297^{c}	-1.7885	0.2202	0.1930	6.2680^{c}	1.5387	1.8754	0.3939
1. PER										
EXT					-1.5664				0.8367	
OLS					-1.1068				1.1166	
STAM										
LEW										
2. DYR										
EXT					-0.9499				1.6362	0.2447
OLS					-1.1440				0.8714	1.911^{b}
STAM										
LEW					-2.4382				0.1114	-0.3633
3. STY										
EXT		0.6922	-0.2158	0.1831		0.7526	1.9877	0.5224		1.2219
OLS		2.3231^{b}	2.3568	0.9667		-0.5548	7.2748^{c}	2.5281		2.2098^{b}
STAM		2.6262^{b}	2.7184	0.9639		-0.5896	7.2407^{c}	3.0760		2.2098^{b}
LEW		2.7533^{b}	2.6651	1.0153		-1.3123	7.4695^{c}	3.1544		2.1356^{c}
4. LTY										
EXT		3.5630^{a}	1.7140^{b}		1.8090^{c}	1.2949^{c}			3.1195^{b}	1.7609^{b}
OLS		2.7630^{b}	3.9918^{b}		1.6217	1.3688			2.2955	2.4353^{c}
STAM		2.4416^{c}	4.1629^{b}		1.6684	1.3688			2.0305	2.3675^{c}
LEW		2.7197^{b}	4.9375^{b}		1.5501	1.0221			2.3962	2.9193^{b}
5. TSP										
EXT		0.8123								2.5473^{b}
OLS		1.5925^{c}								2.6420^{b}
STAM		1.7696^{c}								2.7310^{b}
LEW		1.9612^{c}								2.8713^{b}

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 18

Out-of-sample economic analysis, 3-month excess returns ($k = 3$).

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST	-4.5736	-0.3708	-3.0145	-2.5757	-0.3137	-2.0776	-2.3099	-4.5996	-1.6163	-0.7921
1. PER										
EXT					-2.7029				-0.1839	
OLS					-2.9301					
STAM										
LEW									0.3203	
2. DYR										
EXT					-2.9462				0.3897	-1.0277
OLS									-1.5320	-1.6154
STAM										
LEW					-4.3035					-4.3147
3. STY										
EXT			-0.1303	0.0525		-0.1461	0.1639	-0.7841	-0.5388	0.0916
OLS			0.5879	-0.7166		-2.0810	0.8228		-1.0104	1.2156
STAM			0.4818	-0.6739		-2.2006	0.8228		-1.0695	1.1814
LEW			0.3779	-0.6933		-2.2428	0.7942	-5.8759	-0.7899	1.1761
4. LTY										
EXT	-0.0141	1.8610	0.7223		1.0654	-0.2200	1.4920			0.5808
OLS		1.5828	2.2527		1.0759	0.5317	-0.3491			0.5019
STAM	-1.0548	1.4987	2.4632		1.0759	0.5317	-0.0836			0.5019
LEW	-0.8989	1.4404	2.8413		1.0769	0.5213	0.3006			0.4915
5. TSP										
EXT		0.3961	-0.0565							
OLS			-3.9861							
STAM			-4.0065							
LEW		0.9798	-3.6486							

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 19
Out-of-sample economic analysis, 3-month horizon ($k=3$), Sharpe Ratio.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
BH	0.2617	0.1442	0.2419	0.1706	0.2331	0.3931	0.1985	0.3875	0.2448	0.3154
REST	0.0341^{c}	0.2296^{a}	0.1551^{a}	0.0603^{b}	0.2213^{a}	0.3051^{a}	0.1605	0.3117^{a}	0.2681^{a}	0.3137^{a}
1. PER										
EXT					0.1257^{a}				0.2894^{a}	
OLS					0.1035^{a}					
STAM										
LEW									0.3412^{a}	
2. DYR										
EXT					0.1095^{a}				0.3158^{a}	0.3237^{a}
OLS									0.3325^{a}	0.3108^{a}
STAM										
LEW					0.0109^{a}					0.1059^{a}
3. STY										
EXT			0.2360^{a}	0.1742^{a}		0.3854^{a}	0.2267	0.3622^{a}	0.2369^{a}	0.3496^{a}
OLS			0.3801^{a}	0.1608^{b}		0.3038^{a}	0.3191		0.3075^{b}	0.4844^{a}
STAM			0.3653^{a}	0.1630^{b}		0.2973^{a}	0.3191		0.3034^{b}	0.4817^{a}
LEW			0.3551^{a}	0.1606^{b}		0.2949^{a}	0.3177	0.2218^{a}	0.3217^{b}	0.4813^{a}
4. LTY										
EXT	0.2800^{a}	0.2954^{a}	0.2829^{a}		0.3093^{a}	0.3881^{a}	0.2800^{a}			0.4295^{a}
OLS		0.3592^{a}	0.4724^{a}		0.3188^{b}	0.5055^{a}	0.2455^{b}			0.4589^{a}
STAM	0.2749^{b}	0.3478^{a}	0.4832^{a}		0.3188^{b}	0.5055^{a}	0.2628^{b}			0.4589^{a}
LEW	0.2855^{a}	0.3398^{a}	0.5031^{a}		0.3211^{b}	0.5048^{a}	0.2821^{b}			0.4653^{a}
5. TSP										
EXT		0.1868^{a}	0.2436^{a}							
OLS			0.0781^{a}							
STAM			0.0762^{a}							
LEW		0.4149^{a}	0.0985^{a}							

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 20
Out-of-sample economic analysis, 3-month horizon ($k = 3$), X^* metric.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST	-0.2234	0.3476	0.0142	-0.0732	-0.0624	-0.0120	0.3021	0.1625	0.1751	0.0915
1. PER										
EXT					-0.1978				0.4457	
OLS					-0.1552					
STAM										
LEW									0.5099	
2. DYR										
EXT					-0.2158				0.5125	0.2200
OLS									0.5943	0.2533
STAM										
LEW					-0.4326					0.0128
3. STY										
EXT			0.0092	0.0514		0.0243	0.2513	-0.0761	0.0154	0.2401
OLS			0.8252^c	0.2519		0.0601	0.8846		0.2966	0.6655^c
STAM			0.7875^c	0.2533		0.0343	0.8846		0.2888	0.6591^c
LEW			0.7462^c	0.2392		0.0278	0.8774	-0.2045	0.3503	0.6607^c
4. LTY										
EXT	0.1336	0.5704^b	0.2565		0.3145	0.1025	0.6247^b			0.4166^b
OLS		0.6976^c	1.0341^b		0.3810	0.7160^b	0.5874			0.5560^c
STAM	0.3611	0.6784^c	1.0825^b		0.3810	0.7160^b	0.6555			0.5560^c
LEW	0.3971	0.6614^c	1.1686^b		0.3966	0.7244^b	0.7394			0.5670^c
5. TSP										
EXT		0.2095	0.1259							
OLS			-0.2354							
STAM			-0.2474							
LEW		0.6448^b	-0.1821							

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 21
Out-of-sample economic analysis, 3-month horizon ($k = 3$), X_{eff} metric.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
BH	-6.5390	-5.6090	-10.355	-9.4915	-7.7828	-7.1826	-12.068	-14.206	-7.4931	0.5377
REST	-0.9213	-0.9933	-4.3605	-4.6580	-7.9263	-6.4529	-5.2002	-4.4906	-2.4984	0.4457
1. PER										
EXT					-4.2439				-4.6037	
OLS					-4.3151					
STAM										
LEW									-3.3703	
2. DYR										
EXT					-4.1970				-4.7009	0.4204
OLS									-1.3564	0.3511
STAM										
LEW					-4.3167					0.0123
3. STY										
EXT			-10.361	-9.2785		-7.1802	-9.6392	-13.844	-5.4991	0.5689
OLS			-5.0791	-6.2839		-6.8457	-5.7911		-2.5275	0.7671
STAM			-5.1685	-6.2525		-6.9059	-5.7911		-2.5465	0.7620
LEW			-5.2782	-6.5724		-6.9211	-5.7874	-5.5974	-2.5629	0.7611
4. LTY										
EXT	-5.1447	-3.3990	-9.4194		-6.8606	-7.2194	-10.407			0.6634
OLS		-1.5667	-3.8927		-6.6043	-5.3013	-6.8847			0.6681
STAM	-2.7569	-1.5982	-3.9071		-6.6043	-5.3013	-6.7901			0.6681
LEW	-2.7263	-1.6433	-3.9532		-6.5829	-5.3117	-6.7352			0.6695
5. TSP										
EXT		-3.8930	0.1259							
OLS			-0.2354							
STAM			-0.2474							
LEW		-0.8222	-0.1821							

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 22
Out-of-sample economic analysis, 3-month horizon ($k = 3$), Jensen α.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST	-0.6112	1.3754	-0.0102	-0.9951	-0.2815	-0.8948	0.5287	1.0003	1.4438	0.5010
1. PER										
EXT					-0.6908				1.4424	
OLS					-1.2059					
STAM										
LEW									2.3157	
2. DYR										
EXT					-0.8777				1.8189	0.8598
OLS									2.1044	0.9448
STAM										
LEW					-2.3144					0.0629
3. STY										
EXT			-0.0942	0.1202		-0.0745	1.1090	-0.5330	0.2029	0.6571
OLS			3.1739	0.3140		-0.9435	3.1817		1.9187	2.6309^c
STAM			2.9579	0.3492		-1.0424	3.1817		1.8678	2.6000^c
LEW			2.8010	0.2922		-1.0830	3.1557	-0.7764	2.0949	2.5942^c
4. LTY										
EXT	0.6394	2.4753^c	0.9993		1.4753	0.1000	2.2565^c			1.9631^c
OLS		2.9344^b	4.6069^b		1.7087	2.7913	1.8363			2.3921^c
STAM	1.1278	2.8274^b	4.7823^b		1.7087	2.7913	2.1658			2.3921^c
LEW	1.2703	2.7519^b	5.1041^b		1.7626	2.7807	2.5178			2.4716^c
5. TSP										
EXT		0.8751	0.1528							
OLS			-1.4395							
STAM			-1.4991							
LEW		2.8258^b	-1.1797							

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 23
Out-of-sample economic analysis, 1-year excess returns ($k = 12$).

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST	-4.3580	-0.7687	-2.4460		0.0653	-2.7065			-2.0147	-1.521
1. PER										
EXT					-1.0032				0.8204	
OLS					-1.7087				-3.3196	
STAM					-2.0736				-2.9820	
LEW					-1.3939				-3.0719	
2. DYR										
EXT									1.8275	
OLS									-1.2806	
STAM									-1.9545	
LEW									-2.6619	
3. STY										
EXT		0.4914			0.7139	-0.2353				0.2540
OLS		1.2339			0.6101					0.3757
STAM		1.3375			0.6101					0.3757
LEW		1.3156			0.6456	-1.5351				0.4445
4. LTY										
EXT	0.6993	0.7850	0.3175			-0.2656			0.5719	0.3286
OLS	-1.4172	0.7205	-0.1053			-1.1442				-0.5389
STAM	-1.5159	0.7205	-0.0765			-1.0733				-0.6239
LEW	-1.4887	0.6674	0.1553			-1.1295			-0.8476	-0.6052
5. TSP										
EXT										1.0167
OLS										-0.0982
STAM										-0.0924
LEW										0.0047

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 24

Out-of-sample economic analysis, 1-year horizon ($k = 12$), Sharpe Ratio.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
BH	0.2554	0.1208	0.2155		0.2048	0.3754			0.2177	0.2855
REST	0.0547^{c}	0.1426^{a}	0.1651^{a}		0.2148^{a}	0.2574^{a}			0.1889^{b}	0.2224
1. PER										
EXT					0.2104^{a}				0.3232^{a}	
OLS					0.1847^{a}				0.1700^{b}	
STAM					0.1330^{a}				0.1854^{a}	
LEW					0.1765^{a}				0.1937^{a}	
2. DYR										
EXT									0.3814^{a}	
OLS									0.3080^{a}	
STAM									0.2622^{a}	
LEW									0.2041^{a}	
3. STY										
EXT		0.1736^{a}			0.2616^{b}	0.3582^{a}				0.3518^{a}
OLS		0.3144^{a}			0.2603^{c}					0.4064^{a}
STAM		0.3179^{a}			0.2603^{c}					0.4064^{a}
LEW		0.3173^{a}			0.2624^{c}	0.3144^{a}				0.4146^{a}
4. LTY										
EXT	0.3178^{a}	0.1922^{a}	0.2320^{a}			0.3596^{a}			0.3447^{a}	0.3716^{a}
OLS	0.3057^{a}	0.2613^{a}	0.2749^{a}			0.3269^{a}				0.3288^{a}
STAM	0.2910^{a}	0.2613^{a}	0.2765^{a}			0.3316^{a}				0.3233^{a}
LEW	0.2941^{a}	0.2547^{a}	0.2872^{a}			0.3288^{a}			0.2938^{a}	0.3351^{a}
5. TSP										
EXT										0.4235^{a}
OLS										0.4104^{a}
STAM										0.4110^{a}
LEW										0.4122^{a}

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 25
Out-of-sample economic analysis, 1-year horizon ($k = 12$), X^* metric.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST	-0.5563	0.6335	0.6808		0.0948	-0.8827			-0.0387	-0.4770
1. PER										
EXT					0.9312				2.7838^{c}	
OLS					0.8596				0.9605	
STAM					-0.1509				1.0822	
LEW					0.6081				1.0598	
2. DYR										
EXT									3.4415^{a}	
OLS									2.2942^{b}	
STAM									1.8861^{c}	
LEW									1.4159^{c}	
3. STY										
EXT		0.7173			1.2151	-0.0301				1.0397^{c}
OLS		2.3122^{b}			1.2519					1.6997^{c}
STAM		2.4114^{b}			1.2519					1.6997^{c}
LEW		2.3929^{b}			1.3067	0.7292				1.8012^{c}
4. LTY										
EXT	1.1918	1.1114^{c}	0.5971			0.3498			1.1980^{b}	0.3716^{a}
OLS	0.9239	1.7838^{c}	2.0686^{c}			1.0560				0.3288^{a}
STAM	0.8125	1.7838^{c}	2.0804^{c}			1.1284^{c}				0.3233^{a}
LEW	0.8538	1.7154^{c}	2.2615^{c}			1.0945			1.2047	0.3351^{a}
5. TSP										
EXT										1.8447^{c}
OLS										1.8526^{c}
STAM										1.8511^{c}
LEW										1.8795^{c}

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 26
Out-of-sample economic analysis, 1-year horizon ($k = 12$), X_{eff} metric.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
BH	-46.264	-44.843	-85.727		-74.605	-66.644			-70.729	2.3526
REST	-1.3800	-9.6962	-44.503		-73.679	-56.744			-26.241	1.4467
1. PER										
EXT					-40.026				-38.658	
OLS					-31.145				-4.880	
STAM					-47.944				-6.6716	
LEW					-44.656				-6.0122	
2. DYR										
EXT									-35.939	
OLS									-14.507	
STAM									-12.066	
LEW									-10.318	
3. STY										
EXT		-27.997			-63.774	-66.969				2.5811
OLS		-10.817			-61.897					2.7060
STAM		-11.586			-61.897					2.7060
LEW		-11.335			-61.849	-57.305				2.7514
4. LTY										
EXT	-33.446	-29.396	-82.519			-66.662			-27.2493	2.6397
OLS	-10.301	-12.128	-45.090			-61.494				2.1317
STAM	-10.578	-12.128	-45.109			-61.237				2.0808
LEW	-10.471	-12.292	-46.383			-61.779			-19.8503	2.1028
5. TSP										
EXT										3.0766
OLS										2.4480
STAM										2.4517
LEW										2.5035

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.

Table 27

Out-of-sample economic analysis, 1-year horizon ($k = 12$), Jensen α.

Models	AU	CN	FR	GE	JP	NL	SA	SW	UK	US
REST	0.0317	0.7408	0.3672		0.0864	-1.1040			0.6478	-0.4822
1. PER										
EXT					0.7951				2.3150^{c}	
OLS					0.5808				0.6991	
STAM					-0.7680				0.8373	
LEW					0.0951				0.8854	
2. DYR										
EXT									3.1976^{a}	
OLS									2.0744^{c}	
STAM									1.5738	
LEW									1.0080	
3. STY										
EXT		1.0564			1.2246	-0.2100				1.1448
OLS		2.5344^{b}			1.2457					2.0231
STAM		2.6037^{b}			1.2457					2.0231
LEW		2.5896^{b}			1.2863	-0.3621				2.1185
4. LTY										
EXT	1.4400	1.3258	0.4087			-0.1177			2.6784^{c}	1.4866
OLS	1.7482	2.0073^{c}	1.9612			-0.3578				1.1550
STAM	1.6002	2.0073^{c}	1.9890			-0.2580				1.1018
LEW	1.6324	1.9451^{c}	2.1632			-0.3025			2.0009	1.2911
5. TSP										
EXT										2.0853^{b}
OLS										2.1635^{b}
STAM										2.1695^{b}
LEW										2.1715^{b}

The predictive variables are: PER = Price-Earnings Ratio; DYR = Dividend Yield Ratio; STY = Short-Term Yield; LTY = Long-Term Yield; TSP = Term Spread. OLS (STAM/LEW) refers to the predictive regression model of Equation (1) with no explicit (Stambaugh's/Lewellen's) correction for the small-sample bias in $\hat{\beta}$. REST is the restricted model, where $\beta = 0$ in Equation (1). EXT compares the current level of the predictive variable to the 90th percentile of its unconditional distribution, as defined in Section 2.3.1. The $a/b/c$ subscript indicates significance at the 1/5/10% level respectively, using the bootstrapping procedure described in Section 2.2.2. AU = Australia; CN = Canada; FR = France; GE = Germany; JP = Japan; NL = Netherlands; SA = South Africa; SW = Sweden; UK = United Kingdom; US = United States.