

# Macro Variables and International Stock Return Predictability

*(International Journal of Forecasting, forthcoming)*

David E. Rapach  
Department of Economics  
Saint Louis University  
3674 Lindell Boulevard  
Saint Louis, MO 63108-3397  
Phone: 314-977-3601  
Fax: 314-977-1478  
E-mail: [rapachde@slu.edu](mailto:rapachde@slu.edu)

Mark E. Wohar\*  
Department of Economics  
University of Nebraska-Omaha  
RH-512K  
Omaha, NE 68182-0286  
Phone: 402-554-3712  
Fax: 402-554-2853  
E-mail: [mwohar@mail.unomaha.edu](mailto:mwohar@mail.unomaha.edu)

Jesper Rangvid  
Department of Finance  
Copenhagen Business School  
Solbjerg Plads 3, 5. floor,  
DK-2000 Frederiksberg, Denmark  
Phone: +45 3815 3784  
Fax: +45 3815 3600  
E-mail: [jr.fi@cbs.dk](mailto:jr.fi@cbs.dk)

\*Corresponding author.

# Macro Variables and International Stock Return Predictability

May 7, 2004

## Abstract

In this paper, we examine the predictability of stock returns using macroeconomic variables in twelve industrialized countries. We consider both in-sample and out-of-sample tests of predictive ability, with the out-of-sample forecast period covering the 1990s for each country. We employ recently developed out-of-sample tests that have increased power, namely, the McCracken (2004) variant of the Diebold and Mariano (1995) and West (1996) test for equal predictive ability and the Clark and McCracken (2001) variant of the Harvey, Leybourne, and Newbold (1998) test for forecast encompassing. In addition to analyzing the predictive ability of each macro variable in turn, we use a procedure that combines general-to-specific model selection with out-of-sample tests of forecasting ability in an effort to identify and test the “best” forecasting model of stock returns in each country. Among the macro variables we consider, interest rates are the most consistent and reliable predictors of stock returns across countries.

*Keywords:* Return predictability; Macroeconomic variables; Out-of-sample forecasts; Data mining; General-to-specific model selection

## Biographies

**David E. Rapach** is an Assistant Professor of Economics at Saint Louis University. His research interests include time series econometrics, macroeconomics, international finance, and financial economics. He has published in numerous journals, including *Economic Inquiry*, *Journal of Applied Econometrics*, *Journal of International Economics*, *Journal of International Money and Finance*, *Journal of Macroeconomics*, and the *Journal of Money, Credit, and Banking*.

**Mark E. Wohar** is the Distinguished Enron Professor of Economics and Finance at the University of Nebraska-Omaha. His research interests include time series econometrics, macroeconomics, international finance, financial institutions, and financial economics. He has published papers in more than 65 economic and financial journals, including the *American Economic Review*, *Journal of Applied Econometrics*, *Journal of Finance*, *Journal of International Economics*, *Journal of International Money and Finance*, *Journal of Money, Credit, and Banking*, and the *Review of Economics and Statistics*.

**Jesper Rangvid** is an Associate Professor of Finance at the Copenhagen Business School. His research interests include time series econometrics, international finance, and financial economics. His papers have appeared in a number of journals, including the *European Economic Review*, *Journal of Emerging Market Finance*, *Emerging Markets Review*, and the *Journal of Economic Surveys*.

## 1. Introduction

Numerous empirical studies have investigated the predictability of stock returns using macroeconomic variables. This is not surprising, as macro variables likely exert important influences on firms' expected cash flows, as well as the rate at which these cash flows are discounted. More formally, insofar as macro variables affect future investment opportunities and consumption, they are key state variables in intertemporal asset-pricing models (Merton, 1973; Breeden, 1979; Campbell and Cochrane, 1999) and can represent priced factors in Arbitrage Pricing Theory (Ross, 1976).

In regard to U.S. equity returns, the macro variables investigated in the extant empirical literature include the inflation rate (Bodie, 1976; Jaffe and Mandelker, 1976; Nelson, 1976; Fama and Schwert, 1977; Fama, 1981; Siklos and Kwok, 1999); money stocks (Geske and Roll, 1983; Pearce and Roley, 1983, 1985; Kaul, 1987); aggregate output (Cutler et al., 1989; Balvers et al., 1990; Marathe and Shawky, 1994); unemployment rate (Boyd et al., 2001); interest rates (Campbell, 1987, 1990; Hodrick, 1992; Ang and Bekaert, 2001); term and default spreads on bonds (Campbell, 1987; Fama and French, 1989). Beginning with Chen et al. (1986), a number of studies consider a large set of macro variables (Chen, 1991; Pesaran and Timmermann, 1995; Chan, et al., 1998; Lamont, 2000; Flannery and Protopapadakis, 2002). In addition, some studies investigate the relationship between stock returns and various macro variables across countries (Solnick, 1984; Aspren, 1989; Wasserfallen, 1989; Ferson and Harvey, 1993; Conover et al., 1999, Durham, 2001).<sup>1</sup>

Many of the studies cited above present evidence that stock returns are predictable using macro variables. However, the empirical evidence is far from univocal in providing support for stock return predictability using macro variables. Some studies find that the predictive ability of certain macro variables with respect to equity returns is quite uneven over time; see, for example, Durham (2001) with respect to some measures of monetary policy. Others simply fail to find much support at all for the ability of a number of macro variables to predict returns (Chan et al., 1998). Some studies find strong evidence of predictive ability for a given variable, while others find no evidence for the same variable; see, for example, Balvers et al. (1990) and Flannery and Protopapadakis (2002) with respect to industrial

production. Furthermore, given that numerous studies have investigated the predictive ability of numerous macro variables in the extant literature, concerns about data mining naturally arise. Overall, the mixed results in the extant literature make it difficult to determine which particular macro variables (if any) are reliable indicators of stock returns.

In this paper, we re-examine the predictability of stock returns using macro variables with the aim of acquiring a better understanding of the actual nature of return predictability in international data. Our re-examination has three key features. Firstly, we examine the predictability of stock returns using a large set of macro variables in twelve industrialized countries. The macro variables we consider constitute a set of “standard” macro variables, including the inflation rate, money stocks, interest rates, term spread, industrial production, and unemployment rate. By considering a set of standard macro variables and a large number of industrialized countries, we can examine whether common patterns of return predictability using macro variables emerge across countries, or whether the predictive ability of certain macro variables is particular to only one or a few countries.

In addition, we consider both in-sample and out-of-sample tests of return predictability in the present paper. Our in-sample analysis employs a predictive regression framework, with samples typically beginning in the early 1970s and ending in the late 1990s. For our out-of-sample analysis, we reserve a period covering the bull market of the 1990s over which we analyze out-of-sample forecasts of returns. This is a period during which predictive regression models based on financial ratios, such as the price-dividend or price-earnings ratios, generally perform poorly (Lettau and Ludvigson, 2001, 2003), and we examine whether the macro variables we consider contain predictive content during this recent and “unusual” period. We analyze the out-of-sample forecasts using a pair of recently developed tests due to McCracken (2004) and Clark and McCracken (2001). The McCracken (2004) test statistic is a variant of the Diebold and Mariano (1995) and West (1996) statistic designed to test for equal predictive ability, while the Clark and McCracken (2001) test statistic is a variant of the Harvey, Leybourne, and Newbold (1998) statistic designed to test for forecast encompassing. Importantly, Clark and McCracken (2001, 2004) find the variants to be considerably more powerful than the original statistics in extensive Monte

Carlo simulations. Using these recently developed—and potentially more powerful—out-of-sample tests in the analysis of return predictability using macro variables is an important econometric contribution of the present paper.

Another contribution of the present paper is that we squarely address issues of data mining. Given our interest in testing the predictive ability of a large number of macro variables in turn, concerns about data mining naturally arise. Conventional wisdom holds that our use of out-of-sample tests helps to mitigate problems relating to data mining, and, indeed, we are very interested in whether our in-sample test results are robust to out-of-sample procedures. Nevertheless, Inoue and Kilian (2003) recently argue that out-of-sample tests of predictability may also be susceptible to data mining. The problem is that both in-sample and out-of-sample tests can consider a large number of potential predictor variables, but conclusions about predictability tend to focus on the variables that produce the “best” results. Following the recommendation of Inoue and Kilian (2003), we compute appropriate critical values for all of our in-sample and out-of-sample statistics using a data-mining-robust bootstrap procedure. We also use a procedure recently analyzed by Clark (2002) in an effort to arrive at the “best” forecasting model in each country. We first use general-to-specific model selection in order to choose the best forecasting model, where we start with a model that includes all the macro variables. Importantly, the forecasting model is selected using data from the in-sample period only. The model selected over the in-sample period is then used to compute forecasts over the out-of-sample period, and we again analyze the out-of-sample forecasts using the McCracken (2004) and Clark and McCracken (2001) test statistics. In Monte Carlo simulations, Clark (2002) finds this procedure to be an effective safeguard against model overfitting. To further guard against overfitting, we base our inferences on a data-mining-robust bootstrap procedure.<sup>2</sup>

Summarizing our results, interest rates appear to be the most common and reliable of the macro variables we consider in relation to stock return predictability. When we test the predictive ability of each macro variable in turn, one or more of the interest rates we consider has significant in-sample and out-of-sample predictive ability in nearly every country. The predictive ability of the inflation rate appears to be limited to a small number of countries (especially the Netherlands and the U.S.), but in these countries it

has significant predictive ability according to both in-sample and out-of-sample tests. Money stocks and the term spread appear to have some predictive ability in some countries, while, overall, industrial production and the unemployment rate exhibit rather limited evidence of predictive ability. We also find that the predictive ability of interest rates remains even after we control for data mining. The procedure that combines general-to-specific model selection with out-of-sample forecasting tests also points to the importance of interest rates, as one or more of the interest rate variables is often included in the forecasting model selected over the in-sample period, and the variables included in the selected model significantly improve out-of-sample forecasts in almost all countries.

The rest of the paper is organized as follows. Section 2 presents our econometric methodology; Section 3 reports the in-sample and out-of-sample predictability test results, as well as the results for the procedure that combines general-to-specific model selection with tests of out-of-sample forecasting ability procedure; Section 4 concludes.

## 2. Econometric Methodology

Following much of the extant literature, we analyze stock return predictability using a predictive regression framework. The predictive regression model takes the form,

$$y_{t+1}^k = \alpha + \beta \cdot z_t + \gamma \cdot y_t + u_{t+1}^k, \quad (1)$$

where  $y_t$  is the real return to holding stocks from period  $t-1$  to period  $t$ ,  $y_{t+1}^k = y_{t+1} + \dots + y_{t+k}$  is the real return to holding stocks from period  $t$  to  $t+k$ ,  $z_t$  is a macro variable believed to potentially predict future real returns, and  $u_{t+1}^k$  is a disturbance term. Under the null hypothesis, the variable  $z_t$  has no predictive power for future returns ( $\beta = 0$ ); under the alternative hypothesis,  $z_t$  does have predictive power for future returns ( $\beta \neq 0$ ).<sup>3</sup> Note that we include a lagged return term in equation (1) as a control variable when testing the predictive ability of  $z_t$ , as in, for example, Lettau and Ludvigson (2001).<sup>4</sup> Suppose we have observations for  $y_t$  and  $z_t$  for  $t=1, \dots, T$ . This leaves us with  $T-k$  usable observations with which to estimate the in-

sample predictive regression model. The predictive ability of  $z_t$  is typically assessed by examining the  $t$ -statistic corresponding to  $\hat{\beta}$ , the OLS estimate of  $\beta$  in equation (1), as well as the goodness-of-fit measure  $R^2$ .<sup>5</sup>

Potential econometric problems associated with estimating a predictive regression model like equation (1) include small-sample bias (Mankiw and Shapiro, 1986; Stambaugh, 1986, 1999) and overlapping observations (and thus serial correlation in the disturbance term) when  $k > 1$  (Richardson and Stock, 1989). A common procedure for dealing with the latter problem is the use of Newey and West (1987) standard errors, as these are robust to heteroskedasticity and serial correlation in the disturbance term. (In our applications in Section 3 below, we use the Bartlett kernel and a lag truncation parameter of  $[1.5 \cdot k]$ , where  $[\cdot]$  is the nearest integer function, when calculating Newey and West standard errors.) However, even when robust standard errors are used to compute  $t$ -statistics, there is the potential for serious size distortions when basing inferences on standard asymptotic distribution theory (Nelson and Kim, 1993; Goetzmann and Jorion, 1993; Kirby, 1997). To guard against potential size distortions, we follow much of the recent predictability literature and base inferences concerning  $\beta$  in equation (1) on a bootstrap procedure similar to the procedures in Nelson and Kim (1993), Mark (1995), Kothari and Shanken (1997), and Kilian (1999). The bootstrap procedure is described in detail below.

As discussed in the introduction, we also perform out-of-sample tests of return predictability. Our out-of-sample tests are based on the following recursive scheme. We first divide the total sample of  $T$  observations into in-sample and out-of-sample portions, where the in-sample portion spans the first  $R$  observations for  $y_t$  and  $z_t$ , and the out-of-sample portion spans the last  $P$  observations for the two variables. The first out-of-sample forecast for the “unrestricted” predictive regression model, equation (1), is generated in the following manner. Estimate the unrestricted predictive regression model via OLS using data available through period  $R$ ; denote the OLS estimates of  $\alpha$ ,  $\beta$ , and  $\gamma$  in equation (1) using data available through period  $R$  as  $\hat{\alpha}_{1,R}$ ,  $\hat{\beta}_{1,R}$ , and  $\hat{\gamma}_{1,R}$ . Using the OLS parameter estimates from equation (1) and  $z_R$  and  $y_R$ ,

construct a forecast for  $y_{R+1}^k$  based on the unrestricted predictive regression model using

$$\hat{y}_{1,R+1}^k = \hat{\alpha}_{1,R} + \hat{\beta}_{1,R} \cdot z_R + \hat{\gamma}_{1,R} \cdot y_R.$$

Denote the forecast error by  $\hat{u}_{1,R+1}^k = y_{R+1}^k - \hat{y}_{1,R+1}^k$ . The initial forecast for the “restricted” predictive model is generated in a similar manner, except we set  $\beta = 0$  in equation (1).

That is, we estimate the restricted regression model, equation (1) with  $\beta = 0$ , via OLS using data available through period  $R$  in order to form the forecast  $\hat{y}_{0,R+1}^k = \hat{\alpha}_{0,R} + \hat{\gamma}_{0,R} \cdot y_R$ , where  $\hat{\alpha}_{0,R}$  and  $\hat{\gamma}_{0,R}$  are the OLS estimates of  $\alpha$  and  $\gamma$  in equation (1) with  $\beta$  restricted to zero using data available through period  $R$ .

Denote the forecast error corresponding to the restricted model as  $\hat{u}_{0,R+1}^k = y_{R+1}^k - \hat{y}_{0,R+1}^k$ .

In order to generate a second set of forecasts, we update the above procedure one period by using data available through period  $R + 1$ . That is, we estimate the unrestricted and restricted predictive regression models using data available through period  $R + 1$ , 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 forecasts for  $y_{R+2}^k$  and their respective forecast errors,  $\hat{u}_{1,R+2}^k$  and  $\hat{u}_{0,R+2}^k$ . We repeat this process through the end of the available sample, leaving us with two sets of  $T - R - k + 1$  recursive forecast errors, one each for the unrestricted and restricted regression models ( $\{\hat{u}_{1,t+1}^k\}_{t=R}^{T-k}$  and  $\{\hat{u}_{0,t+1}^k\}_{t=R}^{T-k}$ ).

The next step is to compare the out-of-sample forecasts from the unrestricted and restricted predictive regression models. If the unrestricted model forecasts are superior to the restricted model forecasts, then the variable  $z_t$  improves the out-of-sample forecasts of  $y_{t+1}^k$  relative to the first-order autoregressive (AR) benchmark 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.<sup>6</sup> 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 the McCracken (2004) *MSE-F* and Clark and McCracken (2001) *ENC-NEW* statistics. The first statistic is a variant of the Diebold and Mariano (1995) and West (1996) statistic designed

to test for equal predictive ability, and the second statistic is a variant of the Harvey, Leybourne, and Newbold (1998) statistic designed to test for forecast encompassing.

The *MSE-F* statistic 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) statistic, the *MSE-F* statistic is based on the loss

differential,  $\hat{d}_{t+1}^k = (\hat{u}_{0,t+1}^k)^2 - (\hat{u}_{1,t+1}^k)^2$ . Letting  $\bar{d} = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} \hat{d}_{t+1}^k = M\hat{S}E_0 - M\hat{S}E_1$ , where

$M\hat{S}E_i = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} (\hat{u}_{i,t+1}^k)^2$ ,  $i = 0, 1$ , the McCracken (2004) *MSE-F* statistic is given by

$$MSE-F = (T - R - k + 1) \cdot \bar{d} / M\hat{S}E_1. \quad (2)$$

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.<sup>7</sup> Clark and McCracken (2004) 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 (2004) 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.

Our other out-of-sample statistic, *ENC-NEW*, relates to the concept of forecast encompassing.<sup>8</sup>

Forecast encompassing is based on optimally constructed composite forecasts. Intuitively, if the forecasts from the restricted regression model encompass the unrestricted model forecasts, the macro variable included in the unrestricted model provides no useful additional information for predicting returns relative to the restricted model which excludes the macro variable; if the restricted model forecasts do not encompass the unrestricted model forecasts, then the macro variable does contain information useful for predicting returns beyond the information already contained in a model that excludes the macro 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.<sup>9</sup> The Clark and McCracken (2001) *ENC-NEW* statistic takes the form,

$$ENC-NEW = (T - R - k + 1) \cdot \bar{c} / M\hat{S}E_1, \quad (3)$$

where  $\hat{c}_{t+1}^k = \hat{u}_{0,t+1}^k (\hat{u}_{0,t+1}^k - \hat{u}_{1,t+1}^k)$  and  $\bar{c} = (T - R - k + 1)^{-1} \sum_{t=R}^{T-k} \hat{c}_{t+1}^k$ . 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, 2004) when comparing forecasts from nested models.<sup>10</sup> Again, Clark and McCracken (2004) recommend basing inference on a bootstrap procedure, given the non-pivotal limiting distribution. As indicated in the introduction, Clark and McCracken (2001, 2004) 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.

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

$$y_t = a_0 + a_1 \cdot y_{t-1} + \varepsilon_{1,t}, \quad (4)$$

$$z_t = b_0 + b_1 \cdot z_{t-1} + \dots + b_p \cdot z_{t-p} + \varepsilon_{2,t}, \quad (5)$$

where the disturbance vector  $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$  is independently and identically distributed with covariance matrix  $\Sigma$ . We first estimate equations (4) and (5) via OLS, with the lag order ( $p$ ) in equation (5) selected using the AIC (considering a maximum lag order of twelve), and compute the OLS residuals

$\{\hat{\varepsilon}_t = (\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t})'\}_{t=1}^{T-p}$ . 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 (4) by  $\hat{a}_0$  and  $\hat{a}_1$ , and the OLS estimate of  $(b_0, b_1, \dots, b_p)$  in equation (5) by  $(\hat{b}_0, \hat{b}_1, \dots, \hat{b}_p)$ . Using  $\{\hat{\varepsilon}_t^*\}_{t=1}^{T+100}$ ,  $(\hat{a}_0, \hat{a}_1, \hat{b}_0, \hat{b}_1, \dots, \hat{b}_p)$  in equations (4) and (5), and setting the initial observations for  $y_{t-1}$  and  $(z_{t-1}, \dots, z_{t-p})$  equal to zero in equations (4) and (5), we can build up a pseudo-sample of  $T + 100$  observations for  $y_t$  and  $z_t$ ,  $\{y_t^*, z_t^*\}_{t=1}^{T+100}$ . We drop the first 100 transient start-up observations in order to randomize the initial  $y_{t-1}$  and  $(z_{t-1}, \dots, z_{t-p})$  observations, leaving us with pseudo-sample of  $T$  observations, matching the original sample. For the pseudo-sample, we calculate the  $t$ -statistic corresponding to  $\beta$  in the in-sample predictive regression model given in equation (1), and the two out-of-sample statistics given in equations (2) and (3). We repeat this process 1,000 times, giving us an empirical distribution for the in-sample  $t$ -statistic and each of the out-of-sample statistics. For each statistic, the  $p$ -value is the proportion of the bootstrapped statistics that are greater than the statistic computed using the original sample. As both of the out-of-sample tests are one-sided (upper-tail), an out-of-sample statistic is significant at, say, the 10% level if the  $p$ -value is less than or equal to 0.10. As the in-sample  $t$ -test is two-sided, the in-sample  $t$ -statistic is significant at the 10% level if the  $p$ -value is less than or equal to 0.05 or greater than or equal to 0.95.

In Section 3 below, we use the in-sample and out-of-sample statistics described above to test the ability of between six and nine macro variables, in turn, to predict real stock returns in twelve industrialized countries. When testing the predictive ability of multiple variables, data mining becomes a concern. Lo and MacKinlay (1990) and Foster, Smith, and Whaley (1997) point this out with respect to in-sample tests of security return predictability. While data mining is generally considered to be a serious problem for in-sample

tests of predictability, the conventional wisdom holds that out-of-sample tests are better able to guard against data mining. However, in an interesting recent paper, Inoue and Kilian (2003) challenge the conventional wisdom and argue that “both [in-sample and out-of-sample] tests suffer from size distortions of unknown degree when standard critical values are used” (p. 2). The problem is that, for both in-sample and out-of-sample tests, researchers can consider a number of alternative potential predictor variables but are still free to focus on the “best” results. As emphasized by Inoue and Kilian (2003), the key to controlling for data mining is the use of appropriate critical values for both in-sample and out-of-sample predictability tests.

Inoue and Kilian (2003) consider a data-mining environment that is relevant for our predictability tests in the present paper. Suppose we consider  $M$  different macro variables in turn as candidate predictors in the predictive regression model, equation (1):  $z_{j,t}, j = 1, \dots, M$ . The bootstrap procedure above implicitly assumes that we analyze each macro variable in isolation. In actuality, however, we consider a large number of potential predictors, but we can focus on the variables that give the “best” results. In order to account for data mining when testing predictability, Inoue and Kilian (2003) specify the null hypothesis as

$H_0 : \beta_j = 0 \forall j$  and the alternative hypothesis as  $H_1 : \beta_j \neq 0$  for some  $j$ , where  $\beta_j$  is the coefficient

corresponding to  $z_{j,t}$  in equation (1). For an in-sample test statistic, we use  $\max_{j \in \{1, \dots, M\}} |t_{\hat{\beta}_j}|$ , where  $t_{\hat{\beta}_j}$  is

the  $t$ -statistic corresponding to  $\beta_j$ . For the out-of-sample tests statistics, we use the maximal  $MSE-F$  and

maximal  $MSE-NEW$  statistics. Inoue and Kilian (2003) derive the asymptotic distribution for the maximal in-sample and out-of-sample statistics under the null hypothesis of no predictability, as well as local alternatives, in this data-mining environment. The limiting distributions are generally data-dependent, making inferences based on asymptotic distributions difficult. Inoue and Kilian (2003) recommend that bootstrap procedures be used in practice.

We modify the bootstrap procedure described above in a straightforward manner in order to explicitly account for data mining. Supposing we have  $M$  different macro variables serving as candidate

predictors for the predictive regression model equation (1),  $z_{j,t}$ ,  $j = 1, \dots, M$ , we augment equation (5) of the bootstrap procedure described above to account of all  $M$  of the candidate predictors:

$$\begin{aligned} z_{1,t} &= b_{1,0} + b_{1,1} \cdot z_{1,t-1} + \dots + b_{1,p_1} \cdot z_{1,t-p_1} + \varepsilon_{1,2,t}, \\ &\vdots \\ z_{M,t} &= b_{M,0} + b_{M,1} \cdot z_{M,t-1} + \dots + b_{M,p_M} \cdot z_{M,t-p_M} + \varepsilon_{M,2,t}, \end{aligned} \quad (6)$$

where the disturbance vector  $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{1,2,t}, \dots, \varepsilon_{M,2,t})'$ , is independently and identically distributed with covariance matrix  $\Sigma$ . Using the system defined by equations (4) and (6), we proceed in a manner analogous to the bootstrap procedure described above in order to generate 1,000 pseudo-samples of observations for  $y_t$  and  $z_{1,t}, \dots, z_{M,t}$  under the null hypothesis of no predictability, with each pseudo-sample matching the original sample size. For each pseudo-sample, we calculate the  $t$ -statistic corresponding to  $\beta_j$  in the in-sample predictive regression model and the two out-of-sample statistics for each of the  $z_{j,t}^*$  variables ( $j = 1, \dots, M$ ) in turn. We then compute and store the largest and smallest  $t$ -statistics, as well as the maximal *MSE-F* and *ENC-NEW* statistics. After ordering the empirical distribution for each maximal out-of-sample statistic, the 900<sup>th</sup>, 950<sup>th</sup>, and 990<sup>th</sup> values serve as the 10%, 5%, and 1% critical values for each maximal out-of-sample statistic. For the in-sample  $t$ -statistic, the 950<sup>th</sup>, 975<sup>th</sup>, and 995<sup>th</sup> values of the empirical distribution for the largest  $t$ -statistic serve as the 10%, 5%, and 1% upper-tail critical values for the  $\max_{j \in \{1, \dots, M\}} \left| t_{\hat{\beta}_j} \right|$  statistic; the 50<sup>th</sup>, 25<sup>th</sup>, and 5<sup>th</sup> values of the empirical distribution for the smallest  $t$ -statistic serve as the 10%, 5%, and 1% lower-tail critical values for the  $\max_{j \in \{1, \dots, M\}} \left| t_{\hat{\beta}_j} \right|$  statistic.

In addition to analyzing each macro variable in turn, we employ a procedure recently analyzed by Clark (2002) in an effort to identify the “best” forecasting model for each country. We begin with the following predictive regression model,

$$y_{t+1}^k = \alpha + \beta_1 \cdot z_{1,t} + \dots + \beta_M z_{M,t} + \gamma \cdot y_t + u_{t+1}^k. \quad (7)$$

This model is estimated using data only from the in-sample portion of the total sample. We examine each of the  $t$ -statistics corresponding to the  $z_{j,t}$  variables in equation (7). If the smallest  $t$ -statistic (in absolute value) is greater than or equal to 1.645, we select the model that includes all  $M$  of the  $z_{j,t}$  variables. If the smallest  $t$ -statistic is less than 1.645, we exclude the  $z_{j,t}$  variable corresponding to the smallest  $t$ -statistic in the next model we consider. We proceed in this manner until all of the  $z_{j,t}$  variables included in the model have significant  $t$ -statistics. Otherwise, we select the model that excludes all of the  $z_{j,t}$  variables. Note that we always include the intercept and lagged return terms in the model, as this serves as the benchmark model. This model selection procedure is in the spirit of the LSE approach.<sup>11</sup> If the best forecasting model selected over the in-sample period includes at least one of the  $z_{j,t}$  variables, we then compare the out-of-sample return forecasts generated by the selected model to the out-of-sample forecasts generated by the benchmark model. We again form out-of-sample forecasts recursively and compare out-of-sample forecasts from the competing models using the *MSE-F* and *ENC-NEW* statistics.

In Monte Carlo simulations, Clark (2002) finds this procedure to be effective in guarding against model overfitting, as the out-of-sample statistics are close to being correctly sized under the null hypothesis that none of the  $z_{j,t}$  variables are included in the true data-generating process for  $y_t$ . The key is to select the forecasting model using data only from the in-sample portion of the total sample before analyzing the out-of-sample forecasts. Clark (2002) finds that considerable size distortions emerge if the forecasting model is selected using data from the full sample. In order to further guard against model overfitting, we generate  $p$ -values for the out-of-sample statistics by slightly modifying the data-mining robust bootstrap procedure described above. We again generate a pseudo-sample of data for  $y_t$  and all of the  $z_{j,t}$  variables under the null hypothesis that none of the  $z_{j,t}$  variables are useful in predicting returns. Using the pseudo-sample, we employ the general-to-specific model selection procedure over the in-sample period in order to select the best forecasting model, and if the selected model includes any of the  $z_{j,t}$  variables, we calculate the out-of-

sample *MSE-F* and *ENC-NEW* statistics. We repeat this process until we have empirical distributions of 1,000 bootstrap statistics for both out-of-sample statistics. For each out-of-sample statistic, the *p*-value is the proportion of the bootstrapped statistics that are greater than the statistic computed using the original sample.

### 3. Empirical Results

#### 3.1. Data

We have monthly data for twelve industrialized countries: Belgium, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, the U.K., and the U.S. The macro variables we consider are:

- relative money market rate (difference between the money market interest rate and a 12-month backward-looking moving average; RMM)
- relative 3-month Treasury bill rate (difference between the 3-month Treasury bill rate and a 12-month backward-looking moving average; RTB)
- relative long-term government bond yield (difference between the long-term government bond yield and a 12-month backward-looking moving average; RGB)
- term spread (difference between the long-term government bond yield and the 3-month Treasury bill rate; TSP)
- inflation rate (first difference in the log-levels of the consumer price index; INF)
- industrial production growth (first difference in the log-levels of the industrial production index; IPG)
- narrow money growth (first difference in the log-levels of the narrowly defined money stock; NMG)
- broad money growth (first difference in the log-levels of the broadly defined money stock; BMG)
- change in the unemployment rate (DUN)

The data are from DATASTREAM. Most of the samples begin in the early-to-mid 1970s and end in the late 1990s. For some countries, data for certain variables were not available for a sufficiently long period of time, so these macro variables are excluded from the analysis.<sup>12</sup> Following much of the extant literature, we measure interest rates as deviations from a moving average. We also work with growth rates for money stocks and industrial production and use the change in the unemployment rate, all in an effort to have macro variables that are stationary. We measure real stock returns (RSR) as the first difference in the log-levels of real stock prices. For each country, real stock prices are the Morgan Stanley Capital International stock price index deflated by the consumer price index.<sup>13</sup> Table 1 reports the available sample for each country, as well as the mean and standard deviation for each of the available variables. More details on the data, including data sources, are given in the Data Appendix.

### *3.2. Analyzing the Predictive Ability of the Macro Variables in Turn*

Tables 2-13 report in-sample regression results for equation (1) using data from the total available sample for each country when each macro variable is examined in turn. The tables also report Theil's  $U$  and the  $MSE-F$  and  $ENC-NEW$  statistics for the out-of-sample forecasts. We consider horizons of 1, 3, 12, and 24 months. For the out-of-sample tests, the in-sample portion of the sample ends in 1990:12 for each country, while the out-of-sample portion begins in 1991:01 and runs through the end of the available sample for each country. As noted in the introduction, we are interested in the out-of-sample predictive ability of the macro variables over the 1990s, as this is an "unusual" period when financial ratios, such as the price-dividend and price-earnings ratios, generally perform poorly in predicting returns. As discussed in Lettau and Ludvigson (2003), some of the poor performance over this period is undoubtedly due to changes in the way dividends and earnings have been paid out and reported. An advantage of using macro variables is that they are not subject to these measurement problems.<sup>14</sup>

Next, we briefly describe the results for each country when the macro variables are analyzed in turn. From Table 2, we see that RGB exhibits the strongest predictive ability among the macro variables in Belgium. The in-sample  $t$ -statistic for RGB is significant at each of the horizons considered. Significant

results also characterize the out-of-sample tests, as both out-of-sample statistics for RGB are significant at horizons of 1 and 3 months, and the *ENC-NEW* statistic is significant at the 12-month horizon. TSP also exhibits in-sample and out-of-sample predictive ability at longer horizons, and there is some evidence of predictive ability for the money growth rates. Turning to the results for Canada in Table 3, there is significant in-sample and out-of-sample evidence of predictive ability at the 1-month horizon for RGB and the 12-month horizon for INF. There is also some evidence of predictive ability for IPG at shorter horizons. For Denmark (see Table 4), RMM, RGB, and INF evince significant in-sample and out-of-sample predictive ability at horizons of 1 and 3 months. There is no evidence of predictive ability at longer horizons for any macro variable for Denmark.

The results for France in Table 5 indicate that RGB has significant in-sample and out-of-sample predictive ability at horizons of 1, 3, and 12 months, while TSP has significant in-sample and out-of-sample predictive ability at the 24-month horizon. There is only in-sample evidence of predictive ability for RTB at horizons of 1 and 3 months. Similar to the results for France, from Table 6 we see that RGB has significant in-sample and out-of-sample predictive ability at horizons of 1, 3, and 12 months for Germany. In addition, there is significant in-sample and out-of-sample predictive ability for RTB at the 12-month horizon, IPG at the 3-month horizon, and BMG at horizons of 12 and 24 months, while RMM demonstrates in-sample predictive ability at horizons of 12 and 24 months and RTB at horizons of 3 and 24 months for Germany. For Italy (see Table 7), RTB and RGB exhibit in-sample and out-of-sample predictive ability at the 1-month horizon. Interestingly, RGB exhibits significant out-of-sample, but not in-sample, predictive ability at horizons of 3, 12, and 24 months.<sup>15</sup> There is also evidence of predictive ability for NMG and BMG at various horizons.

From Table 8, we see that there is very little evidence of out-of-sample predictive ability for the macro variables in Japan (only BMG at horizons of 1 and 3 months.) There is in-sample evidence of predictive ability for RMM (horizons of 1 and 3 months) and RGB (horizons of 1, 3, and 12 months). For the Netherlands (see Table 9), RGB and INF evince in-sample and out-of-sample predictive ability at almost all horizons, and there is some evidence of predictive ability at a few points for the other variables.

For Norway (see Table 10), there is in-sample and/or out-of-sample predictive ability at all horizons for RGB and at longer horizons for INF and BMG.

Looking at Table 11 for Sweden, RMM and RTB demonstrate in-sample and out-of-sample evidence of predictive ability at the 1-month horizon, and TSP does likewise at the 12-month horizon. There is also some out-of-sample predictive ability evident for IPG and BMG at certain horizons. For the U.K. (see Table 12), RGB is the only variable that demonstrates in-sample or out-of-sample predictive ability, and this occurs at shorter horizons. From Table 13, we see that the interest rate variables all exhibit significant predictive ability for the U.S. However, with the exception of RGB, the evidence is mostly in-sample. RGB and INF both show significant in-sample and out-of-sample predictive ability at most horizons. There is also some in-sample and/or out-of-sample evidence of predictive ability for TSP, NMG, BMG, and DUN at certain horizons for the U.S.

Taking the results in Tables 2-13 together, the interest rate variables—RMM, RTB, and, especially, RGB—appear to be the most consistent and reliable in-sample and out-of-sample predictors of stock returns across countries.<sup>16</sup> The strong performance of interest rates across a number of countries is in agreement with the recent study of Ang and Bekaert (2001). They test for in-sample stock return predictive ability from the mid-1970s to the late 1990s for France, Germany, Japan, the U.K., and the U.S. In addition to interest rate variables, they consider price-dividend, price-earnings, and payout ratios as potential predictors. They find that the interest rate is the only robust predictor of stock returns. We find that interest rate variables also have out-of-sample predictive ability over the 1990s, and that interest rates are generally more robust than a host of other macro variables in terms of predictive power across many industrialized countries. Overall, our results and those of Ang and Bekaert (2001) point to the reliability of interest rates as predictors of returns in industrialized countries, especially compared to financial ratios and other macro variables.<sup>17</sup>

INF has significant in-sample and out-of-sample predictive ability at multiple horizons in the Netherlands, Norway, and the U.S. We find it interesting that there is strong evidence of predictive ability for the inflation rate in three countries, but not in the other industrialized countries we consider. It may be

that certain types of supply-side shocks are especially important in a few particular countries. With regard to the other macro variables we consider, few stand out as strong in-sample and out-of-sample predictors of stock returns. Broad money growth and the term spread appear to have in-sample and out-of-sample predictive power in a number of countries at certain horizons. Industrial production growth and the change in unemployment are mostly notable for their overall lack of in-sample and out-of-sample predictive ability across countries.<sup>18,19</sup>

It is interesting to observe that our in-sample and out-of-sample test results are often in agreement in Tables 2-13. This is in contrast with some of the extant empirical literature on stock returns (Bossaerts and Hillion, 1999; Goyal and Welch, 2003). The general agreement between our in-sample and out-of-sample test results relative to other studies is most likely due to the increased power of the recently developed McCracken (2004) and Clark and McCracken (2001) tests that we employ.<sup>20</sup>

Table 14 reports critical values for the  $\max_{j \in \{1, \dots, M\}} |t_{\hat{\beta}_j}|$ , maximal *MSE-F*, and maximal *ENC-NEW* statistics reported in Table 2-13. The critical values were generated using the data-mining-robust bootstrap procedure described in Section 2 above. We can use these critical values to check if the significance of the “best” statistics reported in Tables 2-13 is due primarily to data mining. For Italy, we cannot rule out the likelihood that the best evidence of in-sample and out-of-sample predictive ability in Table 7 is due to data mining. For Japan, the significant evidence of out-of-sample predictive ability in Table 8 also appears due to data mining. For the other countries, the critical values in Table 14 suggest that the best evidence of in-sample and out-of-sample predictability at most horizons is not due to data mining.

While we do find evidence of out-of-sample forecasting ability in Tables 2-13, it is important to note that the forecasting gains appear to be limited according to a relative RMSE criterion as embodied in the *U* values reported in Tables 2-13. In the situations where  $U < 1$ , so that the out-of-sample forecasts corresponding to a model that includes a given macro variable have a lower RMSE than the benchmark model, the reduction in RMSE is almost never greater than 10%.<sup>21</sup> Together with the relatively low in-

sample  $R^2$  values in Tables 2-13, the small reductions in RMSE underscore the notion from the extant empirical literature that the predictive component in stock returns is small. Nevertheless, the significant *MSE-F* statistics in Table 2-13 indicate that the reduction in MSE is significant in a number of cases. Furthermore, the *ENC-NEW* statistic is based on forecast encompassing and does not rely on a relative MSE metric. In fact, there are cases in Tables 2-13 where  $U > 1$  but the *ENC-NEW* statistic is still significant. This indicates that a given macro variable contains information that is useful in forecasting stock returns beyond that contained in the benchmark model, despite the fact that  $U > 1$ .

### *3.3. General-to-Specific Model Selection and Out-of-Sample Forecasting Ability*

In Table 15, we report results for the procedure analyzed by Clark (2002) that combines in-sample general-to-specific model selection with tests of out-of-sample forecasting ability. The in-sample period ends in 1990:12 and the out-of-sample period begins in 1991:01 for each country. Again pointing to the importance of the interest rate variables, one or more of the interest rate variables is almost always included among the explanatory variables in the model selected over the in-sample period for each country. Indeed, it is often the case that multiple nominal interest rate variables appear in the selected model. Other macro variables appear more sporadically in the selected models. For ten of the twelve countries, the model selected over the in-sample period generates significant out-of-sample statistics at one or more horizons. Most of the evidence of out-of-sample predictability is concentrated at shorter horizons. Consider the case of the U.S., which is fairly representative of a number of the other countries. Both out-of-sample statistics for the forecasting model selected over the in-sample period are significant at horizons of 1 and 3 months, and neither is significant at horizons of 12 and 24 months. The forecasting model selected over the in-sample period at the 1-month horizon includes two interest rate variables (RMM and RGB), as well as INF and NMG; at the 3-month horizon, the selected model includes three interest rate variables (RMM, RTB, RGB) and NMG. These macro variables appear to significantly improve out-of-sample forecasts of U.S. stock returns over much of the 1990s. We should again note that

the forecasting gains are typically small according to a relative RMSE criterion, as the  $U$  values at horizons of 1 and 3 months are each 0.99 in Table 15 for the U.S. This reiterates the idea that the predictable component in stock returns is small. However, forecast encompassing tests, which are not based on a relative MSE metric, indicate that the selected model contains information that is useful for forecasting beyond that contained in the benchmark model.

#### 4. Conclusion

In this paper, we examine the predictability of stock return using macro variables in twelve industrialized countries. The macro variables we consider include various interest rates, the terms spread, inflation rate, industrial production, money stocks, and unemployment rate. We consider both in-sample and out-of-sample tests of predictability, and we also check our test results using a data-mining-robust bootstrap procedure. For our out-of-sample analysis, we use the recently developed McCracken (2004) and Clark and McCracken (2001) tests, as these appear to be more powerful than other tests in the literature. Our out-of-sample period covers much of the 1990s, a period for which it is notoriously difficult to predict equity returns.

When we analyze each of the macro variables in turn for each country, interest rates stand out in terms of predictive ability—both in-sample and out-of-sample—across countries. The inflation rate also appears to have important in-sample and out-of-sample predictive ability in certain countries at multiple horizons. In general, the evidence of predictive ability for the other macro variables is limited in most countries, especially with regard to industrial production and the unemployment rate. The evidence of significant in-sample and out-of-sample predictability is generally robust to a bootstrap procedure that controls for data mining. When we combine general-to-specific model selection with out-of-sample tests of forecasting ability in an effort to identify the “best” forecasting model, we again find that interest rates are helpful in predicting stock returns in a number of countries. The forecasting models selected over the in-sample period typically contain one or more interest rate variables, and the selected forecasting models often generate out-of-sample forecasts that have predictive content according to the out-of-sample tests in many

countries. The predictive ability of the interest rate variables is primarily concentrated at shorter horizons. Our results are reminiscent of—and add considerable value—to those in Ang and Bekaert (2001). Ang and Bekaert (2001) find that interest rates are robust in-sample predictors of stock returns in five countries, while financial ratios are not. We find that interest rates are generally more consistent and reliable predictors of stock returns than a number of other macro variables, and that this is true for a large number of industrialized countries. In addition, we find that interest rates have significant out-of-sample predictive ability over the 1990s in many countries. However, it is important to keep in mind that the forecasting gains associated with interest rates are typically small according to a relative RMSE metric, which underscores the notion that stock returns have only a small predictable component and are inherently difficult to predict.

Finally, we note some areas for future research. In the present paper, we are primarily concerned with testing for return predictability using macro variables in population, and not necessarily whether an investor in real time could have used one or more macro variables to earn extra-normal profits. While our use of out-of-sample tests somewhat mimics the situation of an investor in real time, we would also have to take account of data revisions by using real-time data when forming a portfolio.<sup>22</sup> It would be interesting in future research to test for predictability using macro variables in real time. However, it is worth noting that interest rates—the macro variable that displays the strongest evidence of return predictability in our analysis—are typically not subject to revision and are available immediately, so that our results pertaining to interest rates are likely to be relevant in real time. It would also be interesting to examine whether some macro variables perform significantly better if one allows for time-varying effects of macro variables on stock returns (McQueen and Roley, 1993; Boyd et al., 2001). This would allow, for example, the effects of macro variables on returns to vary across the stage of the business cycle.<sup>23</sup>

## **Data Appendix**

The data used in this paper are monthly time series and come from three different sources, all available and extracted from DATASTREAM: (i) the International Monetary Fund's *International Financial Statistics* for interest rates, consumer price indices, and industrial production indices; (ii)

OECD for unemployment rates and the money stock series; (iii) Morgan Stanley Capital International (MSCI) for stock return data. The following twelve countries are analysed: Belgium (BG), Canada (CN), Denmark (DK), France (CN), Germany (BD), Italy (IT), Japan (JP), Netherlands (NL), Norway (NW), Sweden (SD), the UK (UK), and the US (US).

- *Money market rate.* The series codes in DATASTREAM are: BGI60B, CNI60B, DKI60B, FRI60B, BDI60B, ITI60B, JPI60B, NLI60B, NWI60B, SDI60B, UKI60B, USI60B.
- *Treasury bill rate.* The series codes in DATASTREAM are: BGI60C, CNI60C, DKI60C, FRI60C, BDI60C, ITI60C, ITI60C, JPI60BS (interbank money rate), NLI60C, NWI60A (average cost for central bank funding), SDI60C, UKI60C, USI60C.
- *Long-term government bond yield.* The series codes in DATASTREAM are: BGI61, CNI61, DKI61, FRI61, BDI61, ITI61, JPI61, NLI61, NWI61, SDI61, UKI61, USI61.
- *Consumer price index.* The series codes in DATASTREAM are: BGI64...F, CAI64...F, DKI64...F, FRI64...F, BDI64...F, ITI64...F, JPI64...F, NLI64...F, NWI64...F, SDI64...F, UKI64...F, USI64...F.
- *Industrial production.* The series codes for the series in DATASTREAM are: BGI66..BH, CNI66..IG, DKI66..IG, FRI66..IG, BDI66..IG, ITI66..IG, JPI66..IG, NLI66..IG, NWI66..IG, SDI66..IG, UKI66..IG, USI66..IG.
- *Narrow money stock.* The narrow money stocks are M1 indices. The series codes in DATASTREAM are: BGOMA027A, CNOCM1MNE, DKOMA033G, FROMA027B, BDOMA027B, ITOMA027B, JPOMA027B, NLOMA027A, NWOMA033G, SDOMA013G, UKOMA033G, USOMA033G.
- *Broad money stock.* The broad money stocks are the M2 or M3 indices. The series codes in DATASTREAM are: BGOMA013A, CNOMA004B, DKXMA001G, FROMA013B, BDOMA002B, ITOMA002B, JPOMA007B, NLOMA014A, NWOMA002A, SDOMA001G, UKOMA002B, USOMA001G.

- *Unemployment rate.* The unemployment rate is registered unemployment as a percent of the total labor force. The series codes for the series in DATASTREAM are: BGXUN013Q, CNOUN014Q, DKOUN013Q, FROCSUN%E, BDOCSUN%E, ITOCSUN%E, JPOCSUN%E, NLOCSUN%E, NWOCSUN%E, SDOCSUN%E, UKOCSUN%E, USOCSUN%E.
- *Stock price index.* The series codes in DATASTREAM are: MSBELGL(PI), MSCNDAL(PI), MSDNMKL(PI), MSFRNCL(PI), MSGERML(PI), MSITALL(PI), MSJPANL(PI), MSNETHL(PI), MSNWAYL(PI), MSSWDNL(PI), MSUTDKL(PI), MSUSAML(PI).

### **Acknowledgments**

We thank Mike McCracken and two anonymous referees for very helpful comments. The usual disclaimer applies. The results reported in this paper were generated using GAUSS 5.0. The GAUSS programs are available at <http://pages.slu.edu/faculty/rapachde/Research.htm>.

## References

- Ang, A. & Bekaert, G. (2001). Stock return predictability: Is it there?" Working Paper No. 8207, National Bureau of Economic Research.
- Asprem, M. (1989). Stock prices, asset portfolios and macroeconomic variables in ten European countries, *Journal of Banking and Finance*, 13, 589-612.
- Baker, M. & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns, *Journal of Finance*, 55, 2219-2257.
- Balvers, R., Cosimano, T., & McDonald, T. (1990). Predicting stock returns in an efficient market, *Journal of Finance*, 45, 1109-1118.
- Bodie, Z. (1976). Common stocks as a hedge against inflation, *Journal of Finance*, 3, 459-470.
- Bossaerts, P. & Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies*, 12, 405-428.
- Boyd, J.H., Jaganathan, R., & Hu, J. (2001). The stock market's reaction to unemployment news: Why bad news is usually good for stock," Working Paper No. 8092, National Bureau of Economic Research.
- Breeden, D. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics*, 7, 265-296.
- Campbell, J.Y. (1987). Stock returns and the term structure, *Journal of Financial Economics*, 18, 373-399.
- Campbell, J.Y. (1990). Measuring the persistence of expected returns, *American Economic Review Papers and Proceedings*, 43-47.
- Campbell, J.Y. & Cochrane, J.H. (1999). Forces of habit: A consumption-based explanation of aggregate stockholder behavior, *Journal of Political Economy*, 107, 205-251.
- Chan, K.C., Karceski, J., & Lakonishok, J. (1998). The risk and return from factors, *Journal of Financial and Quantitative Analysis*, 33, 159-188.
- Chen, N.F. (1991). Financial investment opportunities and the macroeconomy, *Journal of Finance*, 46, 529-554.
- Chen, N.F., Roll, R., & Ross, S. (1986). Economic forces and the stock market, *Journal of Business*, 59, 383-403.
- Clark, T.E. (2002). Can out-of-sample forecast comparisons help prevent overfitting? *Journal of Forecasting*, forthcoming.
- Clark, T.E. & McCracken, M.W. (2001). Tests of equal forecast accuracy and encompassing for nested models, *Journal of Econometrics*, 105, 85-110.
- Clark, T.E. & McCracken, M.W. (2003). Forecast-based model selection in the presence of structural breaks, *Journal of Econometrics*, forthcoming.

- Clark, T.E. & McCracken, M.W. (2004). Evaluating long-horizon forecasts, Manuscript, University of Missouri-Columbia.
- Clements, M.P. & Hendry, D.F. (1998). *Forecasting Economic Time Series*, Cambridge, U.K.: Cambridge University Press.
- Conover, C.M, Jensen, G.R., & Johnson, R.R. (1999). Monetary environments and international stock returns, *Journal of Banking and Finance*, 23, 1357-1381.
- Cooper, M., Gutierrez Jr., R.C., & Marcum, W. (2001). On the predictability of stock returns in real time, *Journal of Business*, forthcoming.
- Cutler, D.M., Poterba, J.M., & Summers, L.H. (1989). What moves stock prices, *Journal of Portfolio Management*, 15, 4-12.
- Diebold, F.X. & Mariano, R.S. (1995). Comparing predictive accuracy, *Journal of Economics and Business Statistics*, 13, 253-263.
- Durham, J.B. (2001). The effects of monetary policy on monthly and quarterly stock market returns: Cross-country evidence and sensitivity analyses, Finance and Economics Discussion Series Paper 2001-42, Federal Reserve Board of Governors.
- Fama, E.F. (1981). Stock returns, real activity, inflation and money, *American Economic Review*, 71, 545-565.
- Fama, E.F. & French, K. (1989). Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics*, 25, 23-49.
- Fama, E.F. & Schwert, G.W. (1977). Asset returns and inflation, *Journal of Financial Economics*, 5, 115-146.
- Ferson, W. & Harvey, C. (1993). The risk and predictability of international equity returns, *Review of Financial Studies*, 6, 527-566.
- Flannery, M.J. & Protopapadakis, A.A. (2002). Macroeconomic factors do influence aggregate stock returns, *Review of Financial Studies*, 15, 751-782.
- Foster, F.D., Smith, T., & Whaley, R.E. (1997). Assessing goodness-of-fit of asset pricing models: The distribution of the maximal  $R^2$ , *Journal of Finance*, 53, 591-607.
- Friedman, M. (1988). Money and the stock market, *Journal of Political Economy*, 96, 221-245.
- Geske, R. & Roll, R. (1983). The fiscal and monetary linkages between stock returns and inflation, *Journal of Finance*, 38, 1-34.
- Goetzmann, W.N. & Jorion, P. (1993). Testing the predictive power of dividend yields, *Journal of Finance*, 48, 663-679.
- Goyal, A. & Welch, I. (2003). Predicting the equity premium with dividend ratios, *Management Science*, 49, 639-654.

- Harvey, D.I., Leybourne, S.J., & Newbold, P. (1998). Tests for forecast encompassing, *Journal of Business and Economic Statistics*, 16, 254-259.
- Hodrick, R.J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies*, 5, 357-386.
- Hoover, K.D. & Perez, S. (1999). Data mining reconsidered: Encompassing and the general-to-specific approach to specification search, *Econometrics Journal*, 2, 1-25.
- Inoue, A. & Kilian, L. (2003). In-sample or out-of-sample tests of predictability: Which one should we use? Manuscript, University of Michigan.
- Jaffe, J. & Mandelker, G. (1976). The 'Fisher Effect' for risky assets: An empirical investigation, *Journal of Finance*, 31, 447-458.
- Kaul, G. (1987). Stock returns and inflation: The role of the monetary sector, *Journal of Financial Economics*, 18, 253-276.
- 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. & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time series analysis, *Journal of Financial Economics*, 44, 169-203.
- Lamont, O. (2001). Economic tracking portfolios, *Journal of Econometrics*, 105, 161-184.
- Lettau, M. & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns, *Journal of Finance*, 56, 815-849.
- Lettau, M. & Ludvigson, S. (2003). Measuring and modeling variation in the risk-return tradeoff. In: Y. Ait-Shalia & L.P. Hansen (Eds.), *Handbook of Financial Econometrics*, forthcoming.
- Lo, A.W. & MacKinlay, A.C. (1990). Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies*, 3, 431-467.
- Mankiw, N.G. & Shapiro, M.D. (1986). Do we reject too often? *Economic Letters*, 20, 139-145.
- Marathe, A. & Shawky, H. (1994). Predictability of stock returns and real output, *Quarterly Review of Economics and Finance*, 34, 317-331.
- 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.

- McQueen, G. & Roley, V.V. (1993). Stock prices, news, and business conditions, *Review of Financial Studies*, 6, 683-707.
- Merton, R. (1973). An intertemporal asset pricing model, *Econometrica*, 41, 867-887.
- Nelson, C.R. (1976). Inflation and the rates of return on common stock, *Journal of Finance*, 31, 471-483.
- Nelson, C.R. & Kim, M.J. (1993). Predictable stock returns: The role of small sample bias, *Journal of Finance*, 48, 641-661.
- Newey, W. & West, K.J. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-708.
- Paye, B. & Timmermann, A. (2003). Instability of return prediction models, Manuscript, University of California-San Diego.
- Pearce, D.K. & Roley, V.V. (1983). The reaction of stock prices to unanticipated changes in money: A note, *Journal of Finance*, 38, 1323-1333.
- Pearce, D.K. & Roley, V.V. (1985). Stock prices and economic news, *Journal of Business*, 58, 49-67.
- Pesaran, M.H. & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance, *Journal of Finance*, 50, 1201-1228.
- Rapach, D.E. & Wohar, M.E. (2003). Structural change and the predictability of stock returns, Manuscript, University of Nebraska-Omaha.
- Rapach, D.E. & Wohar, M.E. (2004). In-sample and out-of-sample tests of stock return predictability in the context of data mining, Manuscript, University of Nebraska-Omaha.
- Richardson, M. & Stock, J.H. (1989). Drawing inferences from statistics based on multiyear asset returns, *Journal of Financial Economics*, 25, 323-348.
- Ross, S. (1976). The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, 13, 341-360.
- Siklos, P. & Kwok, B. (1999). Stock returns and inflation: A new test of competing hypotheses, *Applied Financial Economics*, 9, 567-581.
- Solnick, B. (1984). The relationship between stock prices and inflationary expectations: The international evidence, *Journal of Finance*, 38, 35-48.
- Stambaugh, R.F. (1986). Biases in regressions with lagged stochastic regressors, Working Paper No. 156, Graduate School of Business, University of Chicago.
- Stambaugh, R.F. (1999). Predictive regressions, *Journal of Financial Economics*, 54, 375-421.
- Wasserfallen, W. (1989). Macroeconomic news and the stock market: Evidence for Europe, *Journal of Banking and Finance*, 13, 613-626.
- West, K.D. (1996). Asymptotic inference about predictive ability, *Econometrica*, 64, 1067-1084.

Table 1: Summary statistics, monthly data

Country	Sample	RMM	RTB	RGB	TSP	INF	IPG	NMG	BMG	DUN	RSR
A. Mean											
Belgium	1980:01-1998:12	-0.14	-0.15	-0.13	0.86	3.43	1.41	4.32	6.34	0.02	8.72
Canada	1975:12-2001:12	-0.07	-0.08	-0.06	0.95	6.81	2.30	7.46	8.37	0.003	1.55
Denmark	1974:02-2000:12	-0.06	-	-0.10	1.13	5.36	2.68	-	6.96	0.01	4.91
France	1978:02-1998:12	-	-0.12	-0.12	0.64	4.95	1.28	5.58	6.53	0.03	8.76
Germany	1978:02-1998:12	-0.01	-0.02	-0.03	1.38	2.67	1.07	7.41	7.04	0.02	6.47
Italy	1978:02-1998:12	-0.18	-0.21	-0.22	0.17	7.86	1.78	9.36	8.02	-	8.47
Japan	1970:02-2002:07	-0.11	-	-0.08	0.62	3.54	2.29	9.16	8.35	0.01	2.18
Netherlands	1970:02-1998:12	-0.08	-	-0.05	1.60	4.06	1.98	8.30	9.85	0.01	3.94
Norway	1972:07-2001:04	0.05	-	-0.00	-0.46	5.82	3.85	-	9.30	-	2.31
Sweden	1975:02-1995:04	-0.02	-0.01	0.06	0.71	7.32	1.15	-	7.20	0.03	7.58
U.K.	1975:02-2002:06	-0.03	-0.12	-0.18	0.70	6.41	1.01	-	-	0.01	6.63
U.S.	1970:02-2002:06	-0.10	-0.08	-0.03	1.56	4.81	2.65	5.41	7.97	0.01	2.22
B. Standard deviation											
Belgium	1980:01-1998:12	1.18	0.83	0.67	1.26	4.31	46.42	33.42	13.08	0.23	64.40
Canada	1975:12-2001:12	1.64	1.44	0.77	1.76	20.56	15.03	17.50	6.07	0.23	63.64
Denmark	1974:02-2000:12	2.48	-	1.26	2.96	7.27	36.39	-	13.47	0.19	61.48
France	1978:02-1998:12	-	1.23	0.83	1.31	4.70	15.40	15.92	7.66	0.09	75.28
Germany	1978:02-1998:12	0.93	0.87	0.62	0.96	3.88	21.93	14.01	10.53	0.16	65.43
Italy	1978:02-1998:12	1.29	1.23	1.18	1.18	6.25	28.25	17.07	11.14	-	87.39
Japan	1970:02-2002:07	1.10	-	0.57	1.33	8.41	17.15	13.35	6.41	0.08	66.81
Netherlands	1970:02-1998:12	1.93	-	0.60	2.15	5.45	31.30	31.23	40.06	0.17	60.46
Norway	1972:07-2001:04	1.89	-	0.67	1.96	6.13	52.89	-	25.04	-	92.44
Sweden	1975:02-1995:04	4.66	1.75	0.86	2.04	7.57	38.05	-	9.01	0.20	78.69
U.K.	1975:02-2002:06	1.72	1.39	0.73	2.23	8.16	15.68	-	-	0.13	68.65
U.S.	1970:02-2002:06	1.43	1.12	0.74	1.31	4.03	9.45	7.44	4.85	0.18	53.30

Notes: RMM = relative money market rate; RTB = relative Treasury bill rate; RGB = relative government bond yield; TSP = term spread; INF = inflation rate; IPG = industrial production growth; NMG = narrow money growth; BMG = broad money growth; DUN = change in unemployment rate; RSR = real stock return

Table 2: In-sample and out-of-sample predictability test results, 1991:01-1998:12 out-of-sample period, Belgium

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	2.58	-4.28	-50.57	-63.61	-3.33	-6.83	-83.67	-122.00	-17.39	-48.45	-140.05	-138.93
$t$ -statistic	0.61	-0.47	-1.70	-1.26	-0.78	-0.63	-1.93	-2.15	<b>-4.08</b>	<b>-5.39</b>	<b>-5.94</b>	<b>-3.48</b>
	[0.30]	[0.70]	[0.93]	[0.82]	[0.79]	[0.71]	[0.93]	[0.93]	[1.00]	[1.00]	[1.00]	[0.98]
$R^2$	0.03	0.00	0.05	0.05	0.03	0.00	0.14	0.17	0.09	0.15	0.36	0.20
$U$	1.00	1.01	1.05	0.99	1.01	1.04	1.40	1.24	0.95	0.92	1.01	1.07
$MSE-F$	-0.28	-2.00	-8.49	1.14	-1.19	-6.98	-41.67	-25.17	<b>9.43</b>	<b>17.37</b>	-1.24	-9.20
	[0.37]	[0.72]	[0.90]	[0.25]	[0.76]	[0.94]	[1.00]	[0.97]	[0.00]	[0.00]	[0.36]	[0.77]
$ENC-NEW$	-0.09	-0.85	-1.72	1.85	-0.21	-2.20	-6.26	1.16	<b>7.37</b>	<b>17.77</b>	<b>18.25</b>	2.50
	[0.47]	[0.81]	[0.81]	[0.22]	[0.58]	[0.94]	[0.97]	[0.29]	[0.00]	[0.00]	[0.02]	[0.27]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	-0.88	1.67	71.48	120.53	1.29	-12.03	-32.24	-16.70	4.93	-0.82	5.27	4.05
$t$ -statistic	-0.21	0.20	<b>2.41</b>	<b>3.63</b>	0.30	-1.36	-1.26	-0.53	1.16	-0.21	0.79	0.82
	[0.60]	[0.42]	[0.05]	[0.02]	[0.37]	[0.90]	[0.85]	[0.67]	[0.12]	[0.61]	[0.19]	[0.21]
$R^2$	0.03	0.00	0.11	0.15	0.03	0.01	0.02	0.00	0.03	0.00	0.00	0.00
$U$	1.01	1.02	1.26	0.84	1.00	1.00	1.00	1.01	1.01	1.00	1.00	1.00
$MSE-F$	-2.14	-3.26	-31.35	<b>31.62</b>	-0.46	0.34	0.40	-1.36	-1.04	-0.26	-0.09	-0.02
	[0.89]	[0.73]	[0.97]	[0.03]	[0.43]	[0.17]	[0.22]	[0.60]	[0.75]	[0.79]	[0.65]	[0.46]
$ENC-NEW$	0.08	0.14	<b>12.01</b>	<b>58.47</b>	-0.21	0.87	1.20	-0.61	-0.10	-0.12	0.03	0.03
	[0.32]	[0.35]	[0.06]	[0.01]	[0.56]	[0.13]	[0.15]	[0.67]	[0.44]	[0.86]	[0.28]	[0.28]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	-2.56	5.01	-2.13	-7.07	-2.08	14.82	-1.47	-10.56	-4.17	-9.23	-17.81	-4.55
$t$ -statistic	-0.60	0.71	-0.30	-0.93	-0.49	<b>1.78</b>	-0.10	-0.72	-0.98	-1.17	-1.00	-0.22
	[0.74]	[0.24]	[0.68]	[0.86]	[0.69]	[0.03]	[0.56]	[0.69]	[0.84]	[0.85]	[0.78]	[0.56]
$R^2$	0.03	0.00	0.00	0.00	0.03	0.02	0.00	0.00	0.03	0.01	0.01	0.00
$U$	1.01	1.00	1.00	1.00	1.01	0.98	1.00	1.00	1.01	1.00	0.99	1.01
$MSE-F$	-1.82	0.08	<b>0.38</b>	<b>0.39</b>	-2.13	<b>3.39</b>	0.18	0.13	-1.12	-0.47	<b>1.37</b>	-1.44
	[0.87]	[0.20]	[0.02]	[0.03]	[0.91]	[0.01]	[0.26]	[0.31]	[0.76]	[0.58]	[0.08]	[0.71]
$ENC-NEW$	0.03	0.06	<b>0.20</b>	<b>0.21</b>	-0.50	<b>2.62</b>	0.11	0.08	-0.34	0.04	0.78	-0.68
	[0.33]	[0.31]	[0.05]	[0.04]	[0.84]	[0.02]	[0.35]	[0.43]	[0.79]	[0.38]	[0.15]	[0.81]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 3: In-sample and out-of-sample predictability test results, 1991:01-2001:12 out-of-sample period, Canada

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	0.52	7.67	-25.54	-20.59	-4.14	-2.46	-39.80	-11.42	-13.65	-22.63	-45.52	-58.58
$t$ -statistic	0.14	0.61	-0.67	-0.69	-1.15	-0.21	-0.93	-0.34	<b>-3.85</b>	<b>-2.33</b>	-1.00	-1.44
	[0.42]	[0.31]	[0.72]	[0.72]	[0.85]	[0.56]	[0.76]	[0.58]	[1.00]	[0.98]	[0.78]	[0.83]
$R^2$	0.00	0.01	0.01	0.02	0.00	0.00	0.03	0.01	0.05	0.04	0.03	0.05
$U$	1.00	1.00	1.00	1.02	1.00	1.00	1.00	1.02	0.99	1.01	1.04	1.00
$MSE-F$	-0.21	-0.80	-0.29	-4.80	0.52	-0.34	-1.15	-3.28	<b>1.56</b>	-3.17	-8.09	-0.84
	[0.32]	[0.41]	[0.29]	[0.73]	[0.14]	[0.26]	[0.38]	[0.58]	[0.05]	[0.78]	[0.78]	[0.40]
$ENC-NEW$	-0.10	-0.23	0.07	-1.92	0.47	-0.14	0.32	-1.32	<b>2.77</b>	0.21	-2.36	0.76
	[0.45]	[0.44]	[0.38]	[0.79]	[0.19]	[0.40]	[0.37]	[0.66]	[0.01]	[0.31]	[0.76]	[0.35]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	0.70	1.80	40.41	53.88	2.86	-7.92	-24.30	-8.28	5.49	9.68	4.99	12.90
$t$ -statistic	0.19	0.17	0.88	1.44	0.75	-1.06	<b>-1.84</b>	-0.60	<b>1.52</b>	<b>1.30</b>	0.40	0.98
	[0.41]	[0.45]	[0.22]	[0.13]	[0.22]	[0.86]	[0.95]	[0.71]	[0.07]	[0.13]	[0.35]	[0.22]
$R^2$	0.00	0.00	0.03	0.04	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.01
$U$	1.00	1.00	1.02	1.04	1.01	1.00	0.99	1.00	1.00	1.00	1.00	1.00
$MSE-F$	-0.31	-0.97	-3.62	-8.59	-3.03	-0.31	<b>1.63</b>	-0.18	0.21	<b>1.18</b>	-0.01	0.15
	[0.32]	[0.39]	[0.47]	[0.61]	[0.97]	[0.37]	[0.08]	[0.39]	[0.18]	[0.07]	[0.27]	[0.26]
$ENC-NEW$	-0.14	-0.45	-0.51	-1.31	-1.26	-0.06	0.96	-0.07	0.27	0.68	0.00	0.10
	[0.46]	[0.53]	[0.45]	[0.53]	[0.98]	[0.41]	[0.13]	[0.49]	[0.23]	[0.12]	[0.39]	[0.35]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	0.60	-1.42	-3.12	8.48	-2.47	0.63	-18.02	-59.33	-4.57	-7.30	4.43	-4.90
$t$ -statistic	0.17	-0.25	-0.24	0.67	-0.68	0.07	-0.48	-1.20	-1.26	-0.99	0.24	-0.41
	[0.43]	[0.55]	[0.58]	[0.27]	[0.75]	[0.46]	[0.62]	[0.80]	[0.90]	[0.82]	[0.42]	[0.62]
$R^2$	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.05	0.01	0.01	0.00	0.01
$U$	1.01	1.00	1.00	1.00	1.00	1.01	1.02	1.00	1.01	1.00	1.00	1.00
$MSE-F$	-1.51	-0.48	-0.74	-0.99	-0.17	-3.20	-3.99	0.86	-2.15	-0.01	-1.16	-0.53
	[0.84]	[0.57]	[0.74]	[0.79]	[0.27]	[0.86]	[0.61]	[0.25]	[0.93]	[0.26]	[0.56]	[0.42]
$ENC-NEW$	-0.53	-0.11	-0.33	-0.30	-0.03	-1.28	-1.70	0.93	-0.43	0.14	-0.35	-0.21
	[0.86]	[0.51]	[0.81]	[0.79]	[0.36]	[0.90]	[0.70]	[0.32]	[0.81]	[0.29]	[0.59]	[0.52]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 4: In-sample and out-of-sample predictability test results, 1991:01-2000:12 out-of-sample period, Denmark

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-10.52	-26.51	-3.74	7.38					-13.58	-30.41	-27.48	30.10
$t$ -statistic	<b>-3.01</b>	<b>-3.50</b>	-0.13	0.19					<b>-3.89</b>	<b>-3.08</b>	-1.10	0.66
	[1.00]	[1.00]	[0.56]	[0.45]					[1.00]	[1.00]	[0.84]	[0.31]
$R^2$	0.03	0.07	0.01	0.00					0.05	0.08	0.02	0.01
$U$	0.99	0.99	1.00	1.00					0.98	0.97	0.99	1.02
$MSE-F$	<b>2.91</b>	<b>2.80</b>	-0.88	-0.49					<b>5.81</b>	<b>8.18</b>	1.27	-4.65
	[0.01]	[0.06]	[0.43]	[0.45]					[0.00]	[0.01]	[0.19]	[0.79]
$ENC-NEW$	<b>2.13</b>	<b>3.18</b>	-0.41	-0.22					<b>4.17</b>	<b>6.44</b>	0.82	-1.82
	[0.02]	[0.04]	[0.54]	[0.54]					[0.00]	[0.01]	[0.28]	[0.84]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	2.73	7.84	-0.12	27.42	-10.69	-11.87	-17.58	-21.21	2.39	-5.30	-11.30	-15.66
$t$ -statistic	0.79	1.17	-0.00	0.63	<b>-3.11</b>	<b>-1.66</b>	-0.87	-0.51	0.69	-1.03	-1.22	<b>-2.01</b>
	[0.20]	[0.14]	[0.51]	[0.31]	[1.00]	[0.95]	[0.79]	[0.64]	[0.25]	[0.86]	[0.89]	[0.97]
$R^2$	0.01	0.02	0.01	0.01	0.04	0.03	0.01	0.00	0.01	0.02	0.01	0.00
$U$	1.00	1.01	1.01	1.00	0.98	0.99	1.00	1.01	1.00	1.00	1.00	1.00
$MSE-F$	-0.54	-2.52	-2.39	0.34	<b>3.96</b>	<b>1.43</b>	0.37	-1.36	0.00	<b>0.35</b>	-0.18	-0.16
	[0.55]	[0.79]	[0.54]	[0.29]	[0.00]	[0.07]	[0.20]	[0.56]	[0.23]	[0.09]	[0.64]	[0.66]
$ENC-NEW$	-0.14	-0.68	-1.07	0.58	<b>2.69</b>	<b>0.96</b>	0.21	-0.65	0.03	0.30	0.06	0.03
	[0.52]	[0.73]	[0.62]	[0.36]	[0.01]	[0.09]	[0.29]	[0.65]	[0.32]	[0.12]	[0.27]	[0.33]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$					2.42	4.54	-34.67	-58.05	3.46	10.39	44.29	7.04
$t$ -statistic					0.71	0.52	-2.06	-1.70	1.01	1.33	1.68	0.17
					[0.24]	[0.31]	[0.95]	[0.89]	[0.18]	[0.12]	[0.09]	[0.47]
$R^2$					0.01	0.02	0.02	0.03	0.01	0.03	0.04	0.00
$U$					1.00	1.00	1.01	1.00	1.00	1.01	1.06	1.08
$MSE-F$					-0.53	-0.43	-2.74	0.51	0.68	-1.40	-11.70	-14.47
					[0.56]	[0.39]	[0.70]	[0.26]	[0.13]	[0.73]	[0.96]	[0.96]
$ENC-NEW$					-0.17	0.04	-0.46	1.68	0.44	0.02	-3.22	-6.16
					[0.57]	[0.34]	[0.59]	[0.20]	[0.18]	[0.36]	[0.95]	[0.98]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 5: In-sample and out-of-sample predictability test results, 1991:01-1998:12 out-of-sample period, France

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$					-8.17	-15.21	-58.26	-102.91	-21.11	-46.64	-142.91	-152.51
$t$ -statistic					<b>-1.70</b>	<b>-2.11</b>	-1.92	-2.05	<b>-4.40</b>	<b>-3.98</b>	<b>-6.25</b>	<b>-2.84</b>
					[0.96]	[0.96]	[0.94]	[0.92]	[1.00]	[1.00]	[1.00]	[0.96]
$R^2$					0.01	0.01	0.05	0.09	0.07	0.12	0.28	0.18
$U$					1.00	1.00	1.11	1.14	0.98	0.96	1.01	1.08
$MSE-F$					-0.80	0.16	-16.44	-17.28	<b>4.24</b>	<b>6.98</b>	-1.17	-10.49
					[0.67]	[0.22]	[0.95]	[0.95]	[0.01]	[0.02]	[0.38]	[0.79]
$ENC-NEW$					0.21	0.89	-3.74	-2.82	<b>4.63</b>	<b>8.24</b>	<b>13.20</b>	1.38
					[0.27]	[0.18]	[0.90]	[0.83]	[0.00]	[0.01]	[0.04]	[0.32]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	-1.17	-0.03	43.36	103.97	-1.32	-7.01	-55.90	-70.48	-1.07	5.09	4.86	3.47
$t$ -statistic	-0.25	-0.00	1.22	2.35	-0.28	-0.52	-1.44	-0.94	-0.22	0.90	0.52	0.39
	[0.57]	[0.52]	[0.16]	[0.07]	[0.63]	[0.67]	[0.87]	[0.73]	[0.58]	[0.17]	[0.35]	[0.39]
$R^2$	0.00	0.00	0.03	0.09	0.00	0.00	0.05	0.04	0.00	0.00	0.00	0.00
$U$	1.01	1.01	1.06	0.94	1.00	1.01	1.02	1.04	1.01	1.00	1.00	1.00
$MSE-F$	-1.10	-2.65	-9.54	<b>10.17</b>	-0.58	-1.91	-3.13	-5.10	-1.21	0.16	-0.22	-0.21
	[0.75]	[0.72]	[0.77]	[0.11]	[0.49]	[0.67]	[0.59]	[0.65]	[0.80]	[0.18]	[0.59]	[0.58]
$ENC-NEW$	-0.35	-0.94	0.42	<b>20.83</b>	-0.27	-0.60	2.42	0.74	-0.51	0.12	-0.09	-0.07
	[0.73]	[0.73]	[0.36]	[0.04]	[0.64]	[0.64]	[0.17]	[0.36]	[0.86]	[0.24]	[0.67]	[0.63]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	-1.05	7.28	4.13	-0.43	-2.24	-3.15	-24.37	-30.71	-2.90	-5.77	-14.99	-12.94
$t$ -statistic	-0.22	1.06	0.35	-0.02	-0.46	-0.32	-0.80	-0.56	-0.61	-0.55	-0.65	-0.41
	[0.58]	[0.12]	[0.40]	[0.50]	[0.68]	[0.64]	[0.74]	[0.70]	[0.73]	[0.70]	[0.67]	[0.57]
$R^2$	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01
$U$	1.00	1.01	1.00	1.00	1.00	1.00	1.05	1.07	1.00	1.01	1.00	1.00
$MSE-F$	-0.54	-1.22	0.21	-0.14	-0.47	-0.86	-7.73	-8.65	-0.48	-1.27	-0.46	0.38
	[0.56]	[0.94]	[0.23]	[0.43]	[0.47]	[0.58]	[0.89]	[0.82]	[0.55]	[0.64]	[0.42]	[0.28]
$ENC-NEW$	-0.24	-0.17	0.14	-0.03	-0.03	-0.36	-2.49	-3.59	-0.19	-0.53	-0.10	0.20
	[0.65]	[0.77]	[0.30]	[0.47]	[0.38]	[0.65]	[0.88]	[0.88]	[0.64]	[0.72]	[0.47]	[0.39]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 6: In-sample and out-of-sample predictability test results, 1991:01-1998:12 out-of-sample period, Germany

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-4.06	-14.29	-56.07	-98.16	-6.15	-17.63	-61.29	-83.5	-12.00	-30.09	-102.65	-101.59
$t$ -statistic	-0.98	-1.63	<b>-2.29</b>	<b>-3.85</b>	-1.48	<b>-1.84</b>	<b>-2.36</b>	<b>-2.71</b>	<b>-2.90</b>	<b>-3.03</b>	<b>-4.63</b>	<b>-2.98</b>
	[0.84]	[0.93]	[0.95]	[0.98]	[0.94]	[0.95]	[0.95]	[0.95]	[0.99]	[1.00]	[1.00]	[0.96]
$R^2$	0.02	0.02	0.06	0.08	0.02	0.02	0.07	0.06	0.04	0.06	0.17	0.08
$U$	1.00	0.99	0.97	0.93	1.00	0.99	0.96	0.95	0.98	0.97	0.94	0.96
$MSE-F$	0.33	0.99	5.52	10.95	0.66	1.85	<b>6.53</b>	8.01	<b>3.23</b>	<b>6.57</b>	<b>10.67</b>	6.40
	[0.15]	[0.18]	[0.14]	[0.13]	[0.11]	[0.11]	[0.10]	[0.12]	[0.01]	[0.03]	[0.06]	[0.12]
$ENC-NEW$	0.26	0.92	4.34	7.60	0.52	1.45	4.72	5.02	<b>2.85</b>	<b>5.91</b>	<b>13.87</b>	6.31
	[0.21]	[0.22]	[0.18]	[0.16]	[0.15]	[0.14]	[0.16]	[0.18]	[0.01]	[0.02]	[0.03]	[0.12]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	0.36	5.79	30.69	71.01	-3.42	-7.88	-12.33	-11.26	2.26	6.85	0.65	2.75
$t$ -statistic	0.09	0.47	0.84	1.28	-0.83	-1.04	-0.63	-0.30	0.55	<b>1.42</b>	0.08	0.36
	[0.44]	[0.30]	[0.29]	[0.20]	[0.81]	[0.82]	[0.70]	[0.59]	[0.29]	[0.08]	[0.47]	[0.37]
$R^2$	0.01	0.01	0.02	0.04	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00
$U$	1.00	1.00	1.00	0.97	1.02	1.00	1.00	1.01	1.00	1.00	1.00	1.00
$MSE-F$	-0.42	-0.71	0.47	5.26	-3.83	0.19	-0.04	-1.17	-0.24	<b>0.61</b>	-0.09	0.06
	[0.44]	[0.36]	[0.28]	[0.18]	[0.98]	[0.21]	[0.31]	[0.55]	[0.36]	[0.06]	[0.46]	[0.28]
$ENC-NEW$	-0.20	-0.27	0.49	3.47	-0.78	0.57	0.08	-0.46	-0.07	0.37	-0.04	0.04
	[0.59]	[0.47]	[0.39]	[0.27]	[0.94]	[0.19]	[0.37]	[0.61]	[0.46]	[0.11]	[0.58]	[0.36]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	4.20	-6.43	-11.89	-24.93	-0.69	-11.85	-49.72	-101.27	4.79	9.10	16.75	66.16
$t$ -statistic	1.01	-0.64	-0.91	-1.49	-0.16	-1.38	<b>-3.81</b>	<b>-3.96</b>	1.16	1.37	0.66	1.69
	[0.16]	[0.72]	[0.75]	[0.85]	[0.59]	[0.90]	[1.00]	[0.99]	[0.10]	[0.16]	[0.34]	[0.15]
$R^2$	0.02	0.01	0.01	0.01	0.01	0.01	0.05	0.08	0.02	0.01	0.01	0.04
$U$	1.01	1.01	1.00	1.00	1.00	1.00	0.96	0.92	1.00	0.99	1.07	1.28
$MSE-F$	-1.64	-1.28	-0.59	-0.41	-0.32	-0.87	<b>6.33</b>	<b>12.76</b>	-0.03	1.19	-10.96	-28.56
	[0.87]	[0.79]	[0.58]	[0.52]	[0.43]	[0.61]	[0.02]	[0.01]	[0.28]	[0.07]	[0.94]	[0.97]
$ENC-NEW$	-0.54	-0.35	-0.23	0.10	-0.16	0.26	<b>5.37</b>	<b>11.67</b>	0.06	0.66	-1.14	<b>4.02</b>
	[0.89]	[0.77]	[0.66]	[0.38]	[0.56]	[0.26]	[0.03]	[0.01]	[0.33]	[0.15]	[0.86]	[0.06]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 7: In-sample and out-of-sample predictability test results, 1991:01-1998:12 out-of-sample period, Italy

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-1.60	-0.13	18.16	-15.38	-9.03	-13.37	-34.55	-82.14	-8.46	-23.14	-90.03	-162.30
$t$ -statistic	-0.29	-0.01	0.27	-0.24	-1.61	-0.89	-0.57	-0.99	-1.52	-1.44	-1.75	-2.00
	[0.64]	[0.53]	[0.43]	[0.63]	[0.94]	[0.79]	[0.67]	[0.76]	[0.93]	[0.89]	[0.92]	[0.90]
$R^2$	0.01	0.01	0.02	0.00	0.02	0.02	0.02	0.03	0.02	0.03	0.07	0.11
$U$	1.01	1.01	1.04	1.04	0.99	1.00	0.99	1.09	0.99	0.98	0.93	0.98
$MSE-F$	-1.25	-1.88	-6.15	-4.89	<b>1.73</b>	0.72	1.36	-11.56	<b>1.38</b>	<b>4.16</b>	<b>14.10</b>	2.95
	[0.81]	[0.60]	[0.74]	[0.68]	[0.04]	[0.17]	[0.22]	[0.79]	[0.07]	[0.04]	[0.04]	[0.23]
$ENC-NEW$	-0.56	-0.68	-1.28	-2.08	<b>1.14</b>	0.47	0.83	-1.75	<b>0.92</b>	<b>2.79</b>	<b>10.35</b>	<b>10.78</b>
	[0.89]	[0.65]	[0.66]	[0.75]	[0.07]	[0.26]	[0.32]	[0.68]	[0.11]	[0.07]	[0.06]	[0.08]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	-1.05	-13.83	-26.55	60.47	2.88	5.88	7.32	35.32	-1.81	-0.53	-6.41	-0.11
$t$ -statistic	-0.19	-1.09	-0.46	0.63	0.52	0.40	0.14	0.52	-0.33	-0.08	-0.62	-0.01
	[0.55]	[0.82]	[0.65]	[0.31]	[0.29]	[0.37]	[0.44]	[0.36]	[0.63]	[0.55]	[0.74]	[0.50]
$R^2$	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
$U$	1.00	0.99	1.00	1.09	1.00	1.01	1.05	1.07	1.00	1.00	1.00	1.00
$MSE-F$	-0.38	1.24	0.11	-11.41	-0.50	-1.06	-8.42	-9.81	-0.69	-0.12	0.01	-0.35
	[0.41]	[0.13]	[0.29]	[0.80]	[0.45]	[0.54]	[0.89]	[0.82]	[0.65]	[0.41]	[0.29]	[0.83]
$ENC-NEW$	-0.18	1.00	0.39	-3.69	-0.12	-0.37	-3.43	-3.56	-0.31	-0.05	0.03	-0.15
	[0.52]	[0.18]	[0.37]	[0.81]	[0.45]	[0.60]	[0.93]	[0.87]	[0.74]	[0.50]	[0.32]	[0.87]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	5.64	13.27	14.57	37.97	8.42	23.24	34.53	77.97				
$t$ -statistic	1.01	<b>1.97</b>	1.00	1.55	1.49	<b>2.73</b>	1.46	1.73				
	[0.17]	[0.02]	[0.19]	[0.10]	[0.07]	[0.01]	[0.12]	[0.10]				
$R^2$	0.01	0.02	0.02	0.01	0.02	0.03	0.02	0.02				
$U$	1.00	0.99	1.00	1.03	0.99	0.98	1.00	1.10				
$MSE-F$	-0.31	<b>1.33</b>	-0.78	-4.45	<b>1.17</b>	<b>4.55</b>	-0.42	-12.28				
	[0.36]	[0.03]	[0.74]	[0.98]	[0.07]	[0.00]	[0.45]	[0.98]				
$ENC-NEW$	-0.07	<b>0.95</b>	-0.17	-1.10	0.79	<b>3.61</b>	0.66	-1.38				
	[0.42]	[0.05]	[0.65]	[0.97]	[0.11]	[0.01]	[0.21]	[0.83]				

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 8: In-sample and out-of-sample predictability test results, 1991:01-2002:07 out-of-sample period, Japan

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-9.50	-28.36	-76.79	-97.03					-13.43	-31.37	-81.71	-108.00
$t$ -statistic	<b>-2.80</b>	<b>-2.85</b>	-1.76	-1.68					<b>-3.98</b>	<b>-3.00</b>	<b>-2.53</b>	-2.09
	[1.00]	[0.99]	[0.92]	[0.90]					[1.00]	[1.00]	[0.97]	[0.94]
$R^2$	0.03	0.06	0.09	0.06					0.05	0.07	0.10	0.08
$U$	1.00	1.02	1.00	1.01					1.04	1.07	1.09	1.11
$MSE-F$	-1.07	-5.14	-0.97	-3.17					-11.68	-17.52	-21.16	-21.89
	[0.79]	[0.92]	[0.39]	[0.52]					[1.00]	[1.00]	[0.96]	[0.95]
$ENC-NEW$	-0.17	-1.41	1.35	0.35					-2.18	-2.75	-3.13	-2.30
	[0.58]	[0.86]	[0.31]	[0.41]					[1.00]	[0.97]	[0.82]	[0.74]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	3.53	13.42	12.80	-13.24	-3.71	-8.16	-38.84	-44.15	0.81	5.87	-1.62	-30.53
$t$ -statistic	1.04	1.36	0.39	-0.29	-1.08	-1.18	-1.50	-1.19	0.24	0.93	-0.12	-1.62
	[0.15]	[0.11]	[0.40]	[0.55]	[0.85]	[0.88]	[0.91]	[0.83]	[0.40]	[0.18]	[0.55]	[0.89]
$R^2$	0.01	0.02	0.01	0.00	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.01
$U$	1.00	1.01	1.02	1.02	1.01	1.02	1.04	1.07	1.00	1.00	1.01	1.01
$MSE-F$	-0.37	-3.12	-3.91	-3.56	-2.16	-4.18	-10.33	-14.73	-0.74	0.22	-1.70	-2.14
	[0.45]	[0.77]	[0.55]	[0.48]	[0.91]	[0.97]	[0.99]	[0.96]	[0.69]	[0.18]	[0.79]	[0.81]
$ENC-NEW$	-0.04	-0.83	-1.55	-1.69	-0.83	-1.64	-3.59	-5.50	-0.33	0.15	-0.73	-0.26
	[0.41]	[0.73]	[0.62]	[0.59]	[0.94]	[0.98]	[0.98]	[0.97]	[0.77]	[0.25]	[0.85]	[0.62]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	2.62	3.44	-3.20	-20.97	7.06	13.21	37.46	25.51	-2.05	0.04	15.58	32.96
$t$ -statistic	0.77	0.44	-0.12	-0.42	<b>2.09</b>	1.16	0.93	0.45	-0.60	0.01	1.52	<b>2.01</b>
	[0.22]	[0.34]	[0.53]	[0.61]	[0.02]	[0.15]	[0.24]	[0.36]	[0.73]	[0.50]	[0.09]	[0.05]
$R^2$	0.01	0.01	0.01	0.00	0.02	0.02	0.03	0.00	0.01	0.01	0.01	0.01
$U$	1.00	1.00	1.00	1.01	1.00	0.99	1.00	1.12	1.00	1.00	1.00	1.00
$MSE-F$	0.29	0.02	-0.66	-2.73	<b>1.28</b>	<b>2.19</b>	0.93	-22.74	-0.64	-0.65	-0.53	-0.62
	[0.17]	[0.26]	[0.45]	[0.71]	[0.07]	[0.10]	[0.24]	[0.88]	[0.63]	[0.70]	[0.63]	[0.67]
$ENC-NEW$	0.20	0.04	-0.30	-1.01	<b>1.66</b>	<b>1.96</b>	1.11	-8.34	-0.12	-0.29	0.04	0.26
	[0.24]	[0.36]	[0.56]	[0.74]	[0.05]	[0.10]	[0.29]	[0.90]	[0.47]	[0.77]	[0.36]	[0.25]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 9: In-sample and out-of-sample predictability test results, 1991:01-1998:12 out-of-sample period, Netherlands

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-6.28	-16.82	-62.23	-59.21					-16.85	-44.04	-110.19	-101.66
$t$ -statistic	<b>-1.93</b>	<b>-2.13</b>	-1.70	-1.66					<b>-5.23</b>	<b>-5.01</b>	<b>-4.06</b>	<b>-3.26</b>
	[0.97]	[0.96]	[0.92]	[0.92]					[1.00]	[1.00]	[1.00]	[0.98]
$R^2$	0.02	0.03	0.09	0.03					0.08	0.16	0.23	0.09
$U$	1.00	1.00	0.98	0.98					0.97	0.95	0.90	0.96
$MSE-F$	0.55	0.31	2.93	2.65					<b>6.47</b>	<b>9.57</b>	<b>21.10</b>	5.89
	[0.12]	[0.26]	[0.13]	[0.15]					[0.00]	[0.00]	[0.01]	[0.12]
$ENC-NEW$	0.34	0.30	1.81	1.39					<b>6.05</b>	<b>10.33</b>	<b>16.96</b>	4.39
	[0.17]	[0.31]	[0.17]	[0.21]					[0.00]	[0.00]	[0.01]	[0.13]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	3.72	10.11	29.94	-7.16	-6.99	-19.63	-51.91	-85.12	-0.95	-4.02	-11.83	-19.36
$t$ -statistic	1.14	<b>1.45</b>	0.93	-0.12	<b>-2.15</b>	<b>-3.42</b>	<b>-2.78</b>	<b>-2.73</b>	-0.29	-0.84	-1.33	-1.57
	[0.11]	[0.12]	[0.23]	[0.52]	[0.99]	[1.00]	[0.99]	[0.98]	[0.60]	[0.83]	[0.90]	[0.94]
$R^2$	0.01	0.02	0.03	0.00	0.02	0.04	0.06	0.06	0.01	0.01	0.01	0.01
$U$	1.00	1.00	1.00	1.04	0.99	1.00	0.96	0.95	1.00	1.00	1.00	1.00
$MSE-F$	-0.09	-0.53	-0.56	-5.09	<b>1.61</b>	0.15	<b>6.92</b>	<b>8.00</b>	-0.49	0.20	-0.27	-0.30
	[0.33]	[0.47]	[0.41]	[0.73]	[0.04]	[0.23]	[0.01]	[0.02]	[0.64]	[0.15]	[0.84]	[0.88]
$ENC-NEW$	0.05	-0.05	-0.06	-2.28	<b>1.35</b>	<b>1.65</b>	<b>4.80</b>	<b>4.88</b>	-0.20	0.15	0.06	0.05
	[0.32]	[0.43]	[0.45]	[0.79]	[0.03]	[0.04]	[0.01]	[0.03]	[0.67]	[0.18]	[0.23]	[0.22]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	2.72	17.92	8.81	1.58	1.58	6.53	8.29	8.31	6.62	12.09	11.62	21.40
$t$ -statistic	0.84	<b>3.07</b>	0.98	0.14	0.49	2.28	0.63	0.40	<b>2.05</b>	1.93	0.61	0.77
	[0.21]	[0.00]	[0.12]	[0.44]	[0.34]	[0.08]	[0.33]	[0.40]	[0.02]	[0.09]	[0.29]	[0.28]
$R^2$	0.01	0.04	0.01	0.00	0.01	0.01	0.01	0.00	0.02	0.02	0.01	0.01
$U$	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.01	1.02	1.01	1.02	1.02
$MSE-F$	0.22	<b>1.99</b>	0.09	0.03	0.07	-0.14	-0.62	-0.77	-3.20	-1.97	-2.62	-2.87
	[0.23]	[0.03]	[0.23]	[0.34]	[0.25]	[0.49]	[0.74]	[0.79]	[0.95]	[0.85]	[0.84]	[0.84]
$ENC-NEW$	0.17	<b>2.14</b>	0.09	0.01	0.04	-0.05	-0.30	-0.38	-1.13	-0.59	-1.15	-1.30
	[0.28]	[0.02]	[0.20]	[0.42]	[0.36]	[0.59]	[0.81]	[0.87]	[0.97]	[0.86]	[0.90]	[0.89]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 10: In-sample and out-of-sample predictability test results, 1991:01-2001:04 out-of-sample period, Norway

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-6.06	-12.87	-74.27	-53.54					-15.69	-36.12	-92.38	-66.28
$t$ -statistic	-1.21	-1.09	-1.66	-1.15					<b>-3.15</b>	<b>-3.75</b>	<b>-3.37</b>	-1.78
	[0.88]	[0.85]	[0.93]	[0.82]					[1.00]	[1.00]	[1.00]	[0.94]
$R^2$	0.02	0.02	0.05	0.01					0.04	0.06	0.07	0.02
$U$	1.00	1.00	1.14	1.11					0.98	0.98	1.06	0.96
$MSE-F$	-0.45	-0.62	-26.25	-18.92					<b>4.89</b>	<b>4.74</b>	-12.22	<b>7.65</b>
	[0.54]	[0.44]	[1.00]	[0.98]					[0.00]	[0.03]	[0.90]	[0.07]
$ENC-NEW$	-0.02	-0.01	-0.21	-1.47					<b>5.76</b>	<b>11.56</b>	<b>10.55</b>	<b>6.98</b>
	[0.39]	[0.38]	[0.47]	[0.78]					[0.00]	[0.00]	[0.04]	[0.06]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	2.81	4.77	26.06	-27.33	1.68	-18.14	-63.37	-88.92	-0.89	-6.55	1.09	5.13
$t$ -statistic	0.56	0.42	0.53	-0.57	0.33	-1.71	<b>-2.29</b>	-1.93	-0.18	-1.19	0.17	0.65
	[0.30]	[0.35]	[0.33]	[0.64]	[0.41]	[0.96]	[0.97]	[0.93]	[0.58]	[0.90]	[0.44]	[0.30]
$R^2$	0.01	0.01	0.01	0.00	0.01	0.02	0.04	0.03	0.01	0.01	0.01	0.00
$U$	1.01	1.03	1.22	1.04	1.00	1.00	1.01	0.95	1.00	1.00	1.00	1.00
$MSE-F$	-3.39	-7.41	-36.71	-6.97	-0.77	-0.64	-1.82	<b>11.67</b>	-0.14	-0.04	-0.03	-0.03
	[0.98]	[0.97]	[1.00]	[0.86]	[0.68]	[0.55]	[0.70]	[0.02]	[0.33]	[0.32]	[0.44]	[0.48]
$ENC-NEW$	-1.25	-2.49	-11.16	-3.23	-0.33	0.69	1.77	<b>8.90</b>	-0.07	0.09	-0.01	0.01
	[0.99]	[0.97]	[1.00]	[0.91]	[0.75]	[0.14]	[0.11]	[0.01]	[0.44]	[0.25]	[0.51]	[0.43]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$					-5.18	-7.22	-17.02	-17.26				
$t$ -statistic					-1.04	-1.10	<b>-1.67</b>	-1.31				
					[0.86]	[0.87]	[0.96]	[0.89]				
$R^2$					0.01	0.02	0.01	0.00				
$U$					1.00	1.00	1.00	1.00				
$MSE-F$					0.10	0.40	<b>0.76</b>	<b>0.72</b>				
					[0.20]	[0.11]	[0.02]	[0.05]				
$ENC-NEW$					0.31	0.37	<b>0.55</b>	<b>0.46</b>				
					[0.18]	[0.13]	[0.03]	[0.06]				

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 11: In-sample and out-of-sample predictability test results, 1991:01-1995:04 out-of-sample period, Sweden

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-10.12	3.45	19.72	57.85	-10.07	-6.05	62.31	107.80	-5.24	-5.48	-4.12	65.62
$t$ -statistic	<b>-1.97</b>	0.24	1.01	1.66	<b>-1.96</b>	-0.44	1.42	2.21	-1.03	-0.43	-0.09	0.79
	[0.98]	[0.45]	[0.26]	[0.13]	[0.97]	[0.68]	[0.16]	[0.07]	[0.88]	[0.67]	[0.55]	[0.28]
$R^2$	0.04	0.02	0.02	0.02	0.04	0.02	0.04	0.05	0.03	0.02	0.01	0.01
$U$	1.08	1.22	1.05	2.42	0.98	1.01	1.02	1.00	1.00	1.01	1.08	1.04
$MSE-F$	-7.10	-16.41	-3.99	-24.07	<b>1.93</b>	-0.87	-1.89	-0.23	0.01	-1.12	-5.70	-1.97
	[0.97]	[1.00]	[0.95]	[1.00]	[0.02]	[0.64]	[0.67]	[0.48]	[0.31]	[0.64]	[0.88]	[0.75]
$ENC-NEW$	<b>2.58</b>	-4.83	<b>1.01</b>	<b>2.63</b>	<b>1.18</b>	-0.41	0.62	1.96	0.23	-0.45	-2.46	1.23
	[0.01]	[1.00]	[0.06]	[0.01]	[0.04]	[0.71]	[0.29]	[0.14]	[0.21]	[0.67]	[0.91]	[0.19]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	-2.22	-23.52	-120.30	-133.01	-0.38	1.10	-11.88	-27.21	-4.31	-8.34	-1.96	-4.54
$t$ -statistic	-0.44	-2.13	<b>-3.81</b>	-2.09	-0.08	0.10	-0.50	-0.56	-0.86	-1.15	-0.20	-0.42
	[0.64]	[0.96]	[1.00]	[0.91]	[0.51]	[0.48]	[0.68]	[0.65]	[0.80]	[0.86]	[0.57]	[0.65]
$R^2$	0.03	0.04	0.14	0.08	0.03	0.02	0.01	0.00	0.03	0.02	0.01	0.00
$U$	1.00	1.00	1.01	0.80	1.00	1.01	1.00	0.99	1.00	1.00	1.00	1.00
$MSE-F$	-0.51	0.50	-0.87	<b>16.47</b>	-0.42	-0.75	-0.14	0.59	0.32	<b>0.33</b>	-0.11	-0.05
	[0.69]	[0.22]	[0.52]	[0.02]	[0.67]	[0.76]	[0.47]	[0.21]	[0.19]	[0.09]	[0.71]	[0.66]
$ENC-NEW$	-0.21	1.13	<b>4.96</b>	<b>16.45</b>	-0.19	-0.35	-0.04	0.30	0.19	<b>0.20</b>	-0.05	-0.02
	[0.73]	[0.13]	[0.08]	[0.02]	[0.72]	[0.82]	[0.49]	[0.27]	[0.24]	[0.11]	[0.76]	[0.70]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$					-2.37	-5.70	-26.36	-60.39	-0.07	9.11	34.58	48.59
$t$ -statistic					-0.47	-0.60	-0.93	-1.32	-0.01	1.02	1.41	1.45
					[0.67]	[0.70]	[0.77]	[0.83]	[0.49]	[0.16]	[0.12]	[0.17]
$R^2$					0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.01
$U$					1.03	1.01	0.99	0.94	1.00	1.01	1.01	0.97
$MSE-F$					-3.00	-1.26	0.64	<b>4.13</b>	-0.39	-0.73	-0.67	1.66
					[1.00]	[0.81]	[0.22]	[0.03]	[0.64]	[0.82]	[0.61]	[0.11]
$ENC-NEW$					-1.16	-0.53	0.52	<b>2.94</b>	-0.18	-0.27	-0.29	0.90
					[1.00]	[0.84]	[0.24]	[0.04]	[0.71]	[0.83]	[0.67]	[0.15]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 12: In-sample and out-of-sample predictability test results, 1991:01-2002:06 out-of-sample period, U.K.

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-0.41	2.63	-0.22	-23.51	-3.57	-0.64	-0.08	-10.46	-9.29	-10.63	-8.31	-12.75
$t$ -statistic	-0.12	0.41	-0.02	-1.25	-1.02	-0.09	-0.00	-0.43	<b>-2.62</b>	-1.29	-0.46	-0.60
	[0.56]	[0.36]	[0.52]	[0.86]	[0.85]	[0.54]	[0.52]	[0.67]	[1.00]	[0.88]	[0.65]	[0.68]
$R^2$	0.00	0.00	0.01	0.02	0.01	0.00	0.01	0.01	0.03	0.01	0.01	0.01
$U$	1.00	1.00	1.00	1.02	1.00	1.00	1.00	1.01	0.98	0.99	1.00	1.01
$MSE-F$	-0.12	-0.17	-0.68	-4.05	0.51	-0.43	-0.76	-1.23	<b>5.30</b>	<b>3.46</b>	-0.17	-1.18
	[0.28]	[0.25]	[0.35]	[0.70]	[0.14]	[0.30]	[0.33]	[0.44]	[0.00]	[0.07]	[0.29]	[0.45]
$ENC-NEW$	-0.06	-0.07	-0.34	-1.30	0.41	-0.21	-0.37	-0.45	<b>3.41</b>	1.95	-0.03	-0.42
	[0.41]	[0.38]	[0.46]	[0.69]	[0.20]	[0.43]	[0.45]	[0.52]	[0.00]	[0.13]	[0.41]	[0.51]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	0.69	1.38	11.37	17.18	0.45	-3.61	-9.42	-27.69	-2.01	1.11	4.23	-0.50
$t$ -statistic	0.20	0.17	0.59	0.47	0.13	-0.60	-0.68	-1.80	-0.56	0.27	0.59	-0.09
	[0.40]	[0.42]	[0.31]	[0.37]	[0.44]	[0.74]	[0.75]	[0.93]	[0.70]	[0.39]	[0.32]	[0.55]
$R^2$	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.03	0.01	0.00	0.01	0.01
$U$	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1.00
$MSE-F$	-0.21	-0.72	0.20	-0.00	-0.34	-0.95	-1.99	-1.00	-0.04	-0.21	-0.12	-0.11
	[0.32]	[0.35]	[0.25]	[0.33]	[0.40]	[0.59]	[0.66]	[0.45]	[0.25]	[0.41]	[0.36]	[0.42]
$ENC-NEW$	-0.11	-0.35	0.12	0.28	-0.15	-0.23	-0.54	0.68	0.02	-0.10	-0.04	-0.05
	[0.45]	[0.49]	[0.37]	[0.42]	[0.54]	[0.52]	[0.63]	[0.27]	[0.34]	[0.53]	[0.45]	[0.54]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$									2.09	2.39	5.83	20.64
$t$ -statistic									0.60	0.34	0.32	1.03
									[0.27]	[0.40]	[0.38]	[0.23]
$R^2$									0.01	0.00	0.01	0.02
$U$									1.00	1.00	1.00	1.02
$MSE-F$									0.09	-0.39	-1.25	-3.90
									[0.23]	[0.35]	[0.40]	[0.57]
$ENC-NEW$									0.14	-0.11	-0.55	-1.33
									[0.30]	[0.44]	[0.52]	[0.63]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 13: In-sample and out-of-sample predictability test results, 1991:01-2002:06 out-of-sample period, U.S.

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Relative money market rate				B. Relative Treasury bill rate				C. Relative government bond yield			
$\hat{\beta}$	-11.40	-26.94	-72.74	-64.64	-11.09	-23.37	-67.99	-61.98	-14.13	-30.82	-52.73	-48.95
$t$ -statistic	<b>-4.23</b>	<b>-3.87</b>	<b>-2.96</b>	-1.97	<b>-4.12</b>	<b>-3.70</b>	<b>-2.84</b>	-1.88	<b>-5.24</b>	<b>-5.16</b>	<b>-2.77</b>	<b>-2.13</b>
	[1.00]	[1.00]	[0.99]	[0.92]	[1.00]	[1.00]	[0.99]	[0.93]	[1.00]	[1.00]	[0.99]	[0.96]
$R^2$	0.05	0.08	0.13	0.05	0.04	0.06	0.11	0.04	0.07	0.10	0.07	0.03
$U$	1.01	1.04	1.05	1.03	1.01	1.03	1.05	1.02	1.00	1.00	1.04	0.99
$MSE-F$	-2.58	-9.12	-10.98	-5.79	-2.43	-7.15	-11.49	-4.55	<b>0.95</b>	0.05	-10.04	2.05
	[0.95]	[0.98]	[0.92]	[0.76]	[0.96]	[0.95]	[0.90]	[0.67]	[0.08]	[0.26]	[0.86]	[0.20]
$ENC-NEW$	0.43	-1.27	-2.86	-2.26	<b>0.95</b>	-0.17	-2.47	-1.56	<b>4.59</b>	<b>6.88</b>	-1.42	1.49
	[0.18]	[0.85]	[0.87]	[0.82]	[0.08]	[0.46]	[0.83]	[0.69]	[0.00]	[0.01]	[0.67]	[0.25]
	D. Term spread				E. Inflation rate				F. Industrial production growth			
$\hat{\beta}$	6.38	17.84	61.37	56.91	-10.97	-24.77	-64.76	-86.75	-1.60	-6.61	-22.60	-46.33
$t$ -statistic	<b>2.36</b>	<b>2.40</b>	<b>2.17</b>	1.76	<b>-4.08</b>	<b>-4.19</b>	<b>-3.29</b>	<b>-3.50</b>	-0.59	-1.05	-1.83	-2.00
	[0.01]	[0.02]	[0.06]	[0.11]	[1.00]	[1.00]	[1.00]	[0.99]	[0.74]	[0.83]	[0.94]	[0.95]
$R^2$	0.02	0.04	0.09	0.04	0.04	0.07	0.10	0.08	0.00	0.01	0.02	0.02
$U$	1.01	1.03	1.02	1.00	1.00	0.98	0.95	0.94	1.00	1.01	1.03	1.03
$MSE-F$	-3.38	-8.02	-4.37	-0.82	<b>1.20</b>	<b>5.46</b>	<b>15.14</b>	<b>15.77</b>	-0.12	-3.02	-8.07	-6.13
	[0.98]	[0.96]	[0.59]	[0.36]	[0.07]	[0.02]	[0.01]	[0.05]	[0.30]	[0.91]	[0.96]	[0.89]
$ENC-NEW$	-0.20	-0.45	3.19	1.07	<b>4.34</b>	<b>9.25</b>	<b>12.98</b>	<b>9.50</b>	-0.01	-1.09	-3.32	-2.49
	[0.60]	[0.55]	[0.21]	[0.35]	[0.00]	[0.00]	[0.01]	[0.07]	[0.37]	[0.92]	[0.97]	[0.93]
	G. Narrow money growth				H. Broad money growth				I. Change in unemployment rate			
$\hat{\beta}$	-2.36	-8.46	-34.10	-78.29	0.11	-6.10	-42.04	-90.03	2.72	7.34	32.64	28.53
$t$ -statistic	-0.87	-1.60	-1.21	-1.47	0.04	-1.08	<b>-2.24</b>	-2.24	1.00	1.20	<b>3.35</b>	1.36
	[0.80]	[0.93]	[0.81]	[0.85]	[0.49]	[0.81]	[0.95]	[0.94]	[0.15]	[0.14]	[0.00]	[0.14]
$R^2$	0.00	0.01	0.03	0.05	0.00	0.00	0.04	0.09	0.01	0.01	0.03	0.01
$U$	1.00	0.99	1.00	1.00	1.00	1.01	0.98	0.97	1.00	1.01	1.03	1.02
$MSE-F$	-0.09	<b>2.63</b>	-0.20	-0.32	-0.74	-1.54	4.34	8.46	0.01	-2.27	-6.75	-4.06
	[0.31]	[0.04]	[0.33]	[0.37]	[0.65]	[0.70]	[0.11]	[0.11]	[0.24]	[0.89]	[0.95]	[0.81]
$ENC-NEW$	0.08	<b>2.24</b>	1.30	0.94	-0.25	0.70	<b>7.01</b>	<b>11.69</b>	0.16	-0.56	-2.04	-1.60
	[0.33]	[0.04]	[0.19]	[0.29]	[0.65]	[0.21]	[0.07]	[0.06]	[0.25]	[0.81]	[0.93]	[0.84]

Notes:  $\hat{\beta}$  and  $t$ -statistic are the OLS estimate of  $\beta$  in equation (1) and its corresponding  $t$ -statistic;  $R^2$  is the goodness-of-fit measure in equation (1);  $U$  the ratio of the RMSE for the out-of-sample forecasts for the unrestricted model to the RMSE for the out-of-sample forecasts for the restricted model;  $MSE-F$  and  $ENC-NEW$  are the out-of-sample statistics given in equations (2) and (3);  $p$ -values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates  $< 0.005$ .

Table 14: 10% data mining critical values for the  $\max |t_{\hat{\beta}_j}|$ , maximal  $MSE-F$ , and maximal  $ENC-NEW$  statistics

Horizon:	1	3	12	24	Horizon:	1	3	12	24
A. Belgium (Table 2)					G. Japan (Table 8)				
$t$ -statistic	<b>-2.62, 2.59</b>	<b>-2.85, 2.90</b>	<b>-3.60, 3.46</b>	-4.25, 4.13	$t$ -statistic	<b>-2.55, 2.51</b>	<b>-2.83, 2.67</b>	-3.19, 3.22	-3.69, 3.52
$MSE-F$	<b>3.16</b>	<b>5.95</b>	15.47	<b>21.15</b>	$MSE-F$	2.95	5.36	17.89	26.23
$ENC-NEW$	<b>2.87</b>	<b>5.43</b>	<b>14.14</b>	<b>18.63</b>	$ENC-NEW$	2.40	4.85	15.66	25.01
B. Canada (Table 3)					H. Netherlands (Table 9)				
$t$ -statistic	<b>-2.59, 2.49</b>	-2.84, 2.92	-3.43, 3.30	-3.81, 3.76	$t$ -statistic	<b>-2.68, 2.50</b>	<b>-3.17, 3.25</b>	<b>-3.50, 3.50</b>	-4.10, 3.99
$MSE-F$	3.18	6.17	17.69	23.50	$MSE-F$	<b>2.49</b>	<b>4.82</b>	<b>9.41</b>	14.42
$ENC-NEW$	<b>2.70</b>	5.53	15.98	22.19	$ENC-NEW$	<b>2.12</b>	<b>3.76</b>	<b>7.72</b>	11.12
C. Denmark (Table 4)					I. Norway (Table 10)				
$t$ -statistic	<b>-2.40, 2.56</b>	<b>-2.84, 2.75</b>	-3.04, 3.09	-3.31, 3.43	$t$ -statistic	<b>-2.35, 2.41</b>	<b>-2.75, 2.71</b>	<b>-2.80, 2.81</b>	-3.10, 3.20
$MSE-F$	<b>2.82</b>	<b>4.77</b>	11.10	12.18	$MSE-F$	<b>2.30</b>	<b>3.89</b>	8.57	<b>10.52</b>
$ENC-NEW$	<b>2.40</b>	<b>4.07</b>	9.01	10.20	$ENC-NEW$	<b>1.86</b>	<b>3.47</b>	<b>7.11</b>	<b>8.51</b>
D. France (Table 5)					J. Sweden (Table 11)				
$t$ -statistic	<b>-2.53, 2.51</b>	<b>-2.99, 2.93</b>	<b>-3.57, 3.48</b>	-4.17, 4.20	$t$ -statistic	-2.60, 2.55	-3.22, 3.15	<b>-3.60, 3.67</b>	-4.32, 4.24
$MSE-F$	<b>2.83</b>	<b>6.59</b>	16.51	22.50	$MSE-F$	2.36	4.02	9.70	<b>9.59</b>
$ENC-NEW$	<b>2.53</b>	<b>5.55</b>	14.09	<b>20.78</b>	$ENC-NEW$	<b>1.77</b>	3.08	7.59	<b>7.65</b>
E. Germany (Table 6)					K. U.K. (Table 12)				
$t$ -statistic	<b>-2.68, 2.52</b>	-3.04, 3.17	<b>-3.79, 3.66</b>	-4.58, 4.38	$t$ -statistic	<b>-2.46, 2.41</b>	-2.68, 2.68	-3.26, 3.27	-3.61, 3.67
$MSE-F$	<b>2.83</b>	<b>6.03</b>	20.65	29.73	$MSE-F$	<b>2.67</b>	6.10	16.50	24.79
$ENC-NEW$	<b>2.66</b>	<b>5.70</b>	18.65	26.23	$ENC-NEW$	<b>2.39</b>	5.57	14.85	21.85
F. Italy (Table 7)					L. U.S. (Table 13)				
$t$ -statistic	-2.47, 2.42	-2.71, 2.92	-3.59, 3.76	-3.81, 4.26	$t$ -statistic	<b>-2.51, 2.49</b>	<b>-2.93, 2.87</b>	<b>-3.44, 3.31</b>	-3.68, 3.55
$MSE-F$	2.75	5.90	17.33	24.93	$MSE-F$	2.87	6.62	17.72	25.08
$ENC-NEW$	2.40	5.00	15.58	22.47	$ENC-NEW$	<b>2.35</b>	<b>5.59</b>	14.52	19.97

Notes: Bold entries indicate that the  $\max |t_{\hat{\beta}_j}|$ , maximal  $MSE-F$ , or maximal  $ENC-NEW$  statistic in the table indicated in parentheses is significant according to the critical values in the table.

Table 15: General-to-specific model selection results

Horizon:	1	3	12	24	1	3	12	24	1	3	12	24
	A. Belgium				B. Canada				C. Denmark			
Included Variables	RMM,RGB	RTB,RGB, TSP	RMM,RTB, RGB,TSP	RMM,RTB, RGB,NMG, BMG	RMM,RGB, INF,DUN	RMM,RGB, BMG	-	RMM,RTB, RGB,TSP, DUN	RGB,TSP, INF	RMM,RGB, DUN	RGB,BMG, DUN	TSP,BMG, DUN
<i>U</i>	0.96	1.04	1.31	1.15	1.02	1.04	-	1.05	0.96	0.98	1.06	1.07
<i>MSE-F</i>	<b>7.13</b>	-6.62	-35.20	-17.82	-4.21	-10.45	-	-10.50	<b>10.13</b>	<b>5.98</b>	-12.80	-12.20
	[0.00]	[0.58]	[0.92]	[0.66]	[0.57]	[0.75]		[0.45]	[0.00]	[0.03]	[0.71]	[0.67]
<i>ENC-NEW</i>	<b>10.23</b>	<b>28.28</b>	<b>40.38</b>	0.60	1.30	-1.36	-	1.71	<b>6.96</b>	<b>8.07</b>	-2.54	-3.38
	[0.00]	[0.00]	[0.01]	[0.46]	[0.20]	[0.64]		[0.37]	[0.00]	[0.02]	[0.69]	[0.76]
	D. France				E. Germany				F. Italy			
Included Variables	RTB,RGB	RTB,RGB	RTB,RGB, TSP,NMG, BMG	RGB	RMM,RGB, NMG	RGB,NMG	RGB	RMM,RTB, RGB,TSP, BMG,DUN	RMM,RTB	-	-	RMM,RGB, BMG
<i>U</i>	0.98	1.01	0.95	1.08	1.00	0.97	0.94	1.21	0.99	-	-	1.04
<i>MSE-F</i>	<b>3.21</b>	-1.11	<b>8.55</b>	-10.49	-0.59	<b>5.08</b>	<b>10.67</b>	-23.03	<b>2.60</b>	-	-	-5.68
	[0.02]	[0.24]	[0.07]	[0.50]	[0.16]	[0.04]	[0.07]	[0.64]	[0.04]			[0.37]
<i>ENC-NEW</i>	<b>6.19</b>	<b>9.64</b>	<b>45.93</b>	1.38	<b>2.55</b>	<b>5.55</b>	13.87	<b>22.34</b>	1.91	-	-	16.32
	[0.00]	[0.03]	[0.01]	[0.43]	[0.09]	[0.09]	[0.11]	[0.10]	[0.15]			[0.11]
	G. Japan				H. Netherlands				I. Norway			
Included Variables	RBG,BMG	RBG	RGB,INF, DUN	INF,DUN	RGB,INF, DUN	RGB,INF, NMG,DUN	RGB,INF, IPG	RGB,INF, IPG	RMM,RGB	RGB	RMM,RGB, INF	RMM,INF
<i>U</i>	1.03	1.07	1.12	1.08	0.98	0.96	0.87	0.92	0.99	0.98	1.21	1.08
<i>MSE-F</i>	-7.15	-17.52	-26.70	-16.60	<b>4.74</b>	<b>8.80</b>	<b>28.37</b>	<b>13.35</b>	<b>2.06</b>	<b>4.74</b>	-35.67	-13.89
	[0.93]	[0.97]	[0.82]	[0.57]	[0.01]	[0.01]	[0.01]	[0.05]	[0.05]	[0.04]	[0.99]	[0.80]
<i>ENC-NEW</i>	-0.01	-2.75	-4.86	-5.32	<b>5.55</b>	<b>11.75</b>	<b>22.35</b>	<b>9.11</b>	<b>4.51</b>	<b>11.56</b>	4.90	5.67
	[0.43]	[0.86]	[0.75]	[0.74]	[0.01]	[0.01]	[0.01]	[0.10]	[0.01]	[0.01]	[0.17]	[0.15]
	J. Sweden				K. U.K.				L. U.S.			
Included Variables	RMM,BMG	RMM,TSP	TSP	RGB,TSP	RGB	-	-	RMM	RMM,RGB, INF,NMG	RMM,RTB, RGB,NMG	RMM,BMG, DUN	TSP,IPG, BMG
<i>U</i>	1.15	1.37	1.01	0.68	0.98	-	-	1.02	0.99	0.99	1.05	1.04
<i>MSE-F</i>	-12.75	-23.38	-0.87	<b>34.28</b>	<b>5.30</b>	-	-	-4.05	<b>1.89</b>	<b>4.03</b>	-11.71	-9.14
	[0.99]	[1.00]	[0.40]	[0.01]	[0.01]			[0.33]	[0.05]	[0.05]	[0.52]	[0.43]
<i>ENC-NEW</i>	1.71	-5.14	4.96	<b>26.30</b>	<b>3.41</b>	-	-	-1.30	<b>7.71</b>	<b>13.55</b>	1.94	7.40
	[0.11]	[1.00]	[0.15]	[0.01]	[0.06]			[0.55]	[0.00]	[0.01]	[0.34]	[0.25]

Notes: *U* is the ratio of the RMSE for the out-of-sample forecasts for the selected model to the RMSE for the out-of-sample forecasts for the restricted model; *MSE-F* and *ENC-NEW* are the out-of-sample statistics given in equations (2) and (3); *p*-values given in brackets; bold entries indicate significance at the 10% level; 0.00 indicates < 0.005; - indicates that the selected model does not include any of the macro variables.

---

<sup>1</sup> The studies cited are not intended to constitute an exhaustive list, but are simply examples of relevant studies.

<sup>2</sup> Ang and Bekaert (2001) test for stock return predictability in five countries in an effort to mitigate problems relating to data mining. By examining returns predictability in twelve countries, the present paper can also be viewed as attempting to control for data mining in this manner.

<sup>3</sup> Inoue and Kilian (2003) recommend using a one-sided alternative hypothesis if theory makes strong predictions about the sign of  $\beta$  in equation (1), as this increases the power of in-sample tests. For the macro variables we consider, theory does not always make strong predictions as to the sign of  $\beta$ , so we use a two-sided alternative hypothesis.

<sup>4</sup> 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.

<sup>5</sup> Following Baker and Wurgler (2000), we first divide  $z_t$  by its standard deviation over the full sample when we analyze the macro variables in turn. This normalization has no effect on our in-sample or out-of-sample statistical inferences, but it makes it easier to compare the estimated  $\beta$  coefficient in equation (1) across macro variables, as the coefficient can be interpreted as the change in expected returns given a one-standard-deviation change in the macro variable.

<sup>6</sup> Strictly speaking, Theil's  $U$  compares the predictions from a given model to a random walk model. Even though we include a lagged return term in our benchmark model, we still use the term Theil's  $U$ .

<sup>7</sup> While West (1996) shows that the Diebold and Mariano (1995) and West (1996) statistic has a standard normal limiting distribution when comparing forecasts from non-nested models, McCracken (2004) shows that it has a non-standard limiting distribution when comparing forecasts from nested models.

<sup>8</sup> See Clements and Hendry (1998) for a textbook discussion of forecast encompassing.

---

<sup>9</sup> The composite forecast takes the form of a convex combination of the restricted and unrestricted model forecasts.

<sup>10</sup> As pointed out by Clark and McCracken (2004), the Harvey, Leybourne, and Newbold (1998) statistic has a standard normal limiting distribution when comparing forecasts from non-nested models according to the theory in West (1996). However, for nested models, Clark and McCracken (2001, 2004) show that it has a non-standard limiting distribution.

<sup>11</sup> See Hoover and Perez (1999) for a recent exposition of the LSE general-to-specific modeling approach.

<sup>12</sup> When Treasury bill rate data are not available for a country, we measure the term spread as the difference between the long-term government bond yield and the money market rate.

<sup>13</sup> Our results do not depend critically on our use of capital gains to approximate returns, as total returns are dominated by capital gains. In addition, we obtain very similar results when we use excess returns, defined as gross stock returns minus a short-term interest rate, as real stock returns and excess returns are highly correlated for all of the countries in our sample. The real stock returns-excess returns correlations are as follows: Belgium, 0.998; Canada, 0.945; Denmark, 0.993; France, 0.999; Germany, 0.998; Italy, 0.998, Japan, 0.993; Netherlands, 0.995; Norway, 0.998; Sweden, 0.995; U.K., 0.994; U.S., 0.998.

<sup>14</sup> Of course, some macro variables are subject to data revisions. We discuss the relevance of this in Section 4 below.

<sup>15</sup> In Monte Carlo simulations, although Clark and McCracken (2001, 2004) find that in-sample tests are typically more powerful than out-of-sample tests, they do find that the *MSE-F* and *ENC-NEW* statistics can have power comparable to the in-sample tests in some instances; also see Inoue and Kilian (2003).

<sup>16</sup> As indicated above, we are following most of the extant literature in using the relative interest rate, that is, the nominal interest rate minus a 12-month backward-looking moving average. We may suspect that real interest rates are important determinants of real stock returns. 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). We thank an anonymous referee for pointing this out.

<sup>17</sup> Like Ang and Bekaert (2001), our evidence of return predictability using interest rates is concentrated at shorter horizons. While Ang and Bekaert (2001) fail to find evidence of long-horizon return predictability using interest rates, we find some evidence of return predictability at longer horizons when we use the relative government bond yield. Ang and Bekaert (2001) only consider the 3-month Treasury bill rate, and this may explain why they fail to find evidence of predictability at longer horizons.

<sup>18</sup> We applied unit root tests to all of the variables. The unit root tests indicate that RSR, RMM, RTB, RGB, TSP, IPG, NMG, and DUN are stationary for every country (where relevant). However, there is evidence that INF and BMG are nonstationary for some countries. In light of this, we conducted the in-sample and out-of-sample predictability tests for every country using the change in INF and the change in BMG. With respect to the change in INF, the results are qualitatively similar to those for INF in Tables 2-13. The most notable exception is that, in contrast to the results for INF for Belgium reported in Table 2, the change in INF evinces in-sample and out-of-sample predictive ability at the 1-month horizon and out-of-sample predictive ability at the 3-month horizon for Belgium. Overall, there is somewhat less evidence of in-sample and out-of-sample predictive ability when we use the change in BMG in place of BMG. The complete unit root test results and in-sample and out-of-sample predictability test results for the change in INF and the change in BMG are available at [Hhttp://pages.slu.edu/faculty/rapachde/Research.htm](http://pages.slu.edu/faculty/rapachde/Research.htm)H. Note that we also experimented with conducting the in-sample and out-of-sample predictability tests using relative INF and relative BMG, so inflation and broad money growth are measured as deviations from a 12-month backward-looking moving average. The results are similar to those that obtain when we use the change in INF and the change in BMG.

<sup>19</sup> Paye and Timmermann (2003) and Rapach and Wohar (2003) find evidence of structural breaks in a number of predictive regression models of stock returns based on financial variables. However, apart from the interest rate, these studies do not consider the macro variables used in the present paper. Our use

---

of out-of-sample tests should help to signal breakdowns in predictive ability due to structural instability over the 1990s. See Clark and McCracken (2003) for an analysis of the behavior of in-sample and out-of-sample tests of predictive ability in the presence of structural breaks.

<sup>20</sup> Lettau and Ludvigson (2001) and Rapach and Wohar (2004) also employ the McCracken (2004) and Clark and McCracken (2001) out-of-sample statistics and find considerable agreement between in-sample and out-of-sample test results.

<sup>21</sup> The only case where the reduction in RMSE is greater than 10% is the term spread for Sweden, where  $U = 0.80$ .

<sup>22</sup> See Lettau and Ludvigson (2003) and Inoue and Kilian (2003) for interesting discussions on testing for predictability in population versus real-time predictability. For recent empirical work on stock return predictability in real time, see Cooper et al. (2001).

<sup>23</sup> It may also be interesting in future research to estimate the predictive regression model using a panel approach for all twelve countries and/or to examine the ability of U.S. macro variables to predict returns in non-U.S. countries; on the latter, see Ang and Bekaert (2001). In addition, given that interest rates generally converged during the 1990s among European countries, we may be able to identify a common factor that accounts for return predictability based on interest rates across the European continent. Identifying such a factor is another interesting area for future research.