Forecasting Returns: New European Evidence — **Source link** ⬈

Steven J. Jordan, Andrew Vivian, Mark E. Wohar

Institutions: University of Memphis, Loughborough University, University of Nebraska Omaha

Published on: 06 Sep 2012 - Social Science Research Network

Topics: Predictability and Market liquidity

Related papers:

- Forecasting returns: new European evidence

- Think National, Forecast Local: A Case Study of 71 German Urban Housing Markets

- Out-of-sample stock return predictability in emerging markets

- The determinants of commercial property market performance

- Revenue Forecasting Practices: Differences Across Countries and Consequences for Forecasting Performance

Forecasting Returns: New European Evidence *

Steven J. Jordan
Econometric Solutions
econometric.solutions@yahoo.com

Andrew Vivian
Loughborough University
a.j.vivian@lboro.ac.uk

Mark E. Wohar
University of Nebraska-Omaha
mwohar@mail.unomaha.edu

September 2013

1

Forecasting Returns: New European Evidence

Abstract

This paper builds on the recent debate on the in-sample and out-of-sample predictability of US aggregate returns using a wide range of predictors by providing new evidence for smaller and less market-oriented European countries.

We find evidence that macro and technical predictors can (statistically) improve forecast accuracy and (economically) generate gains to investors; in contrast to the US results, predictability in our sample of European countries exists in recent data. We also find that simple forecast combinations consistently yield substantial benefits both in forecast accuracy and economic gain. For example, the magnitude of the forecasting gains for our European countries is often larger than those found for the US and other G7 countries. We provide initial evidence on the link between country characteristics and out-of-sample forecast performance. Our empirical results indicate market development is related to the forecast performance of macro variables. There is also some evidence that forecast performance is related to market size and liquidity.

Deleted: suggest

Deleted: not strongly

Deleted: However market liquidity is related to the forecast performance of fundamental variables and market development is related to the forecast performance of macro variables.

1. INTRODUCTION

Return predictability is a hotly debated topic. The literature has concluded that in-sample (INS) return predictability is primarily found in the UK and US (Rapach and Wohar, 2009; Engsted and Pedersen, 2010), in developed countries (Hjalmarsson, 2010), or in the largest countries (Rangvid, Schmeling, and Schrimpf, 2011). However, only Rangvid, Schmeling and Schrimpf (2011) formally test the link between country characteristics and INS predictability.

We investigate whether returns can be forecast out-of-sample (OOS) for 14 European countries for which there is virtually no prior evidence. This sample includes both developed and emerging markets as well as markets of varying size, liquidity, and development. We provide initial empirical evidence on the link between country characteristics and OOS forecast performance. Compared to most prior international studies, we examine a wider range of predictor variables. In addition, two new technical variables, i.e., the ratio of rising stocks to fallers and the change in trading volume, are included in the set of predictor variables.

Many variables have been used to predict stock returns including fundamental-price ratios and macro variables; the economic foundation of these variables is well known and well established.[1] Using US data, Goyal and Welch (2008, hereafter GW), test the robustness of OOS predictability compared to a simple historical average benchmark and conclude such models would not have been useful to investors to time the market. There are few papers that examine return forecasting in an international context. Of the international studies, the majority include several macro variables[2] and

---

[1] See Goyal and Welch (2008) who provide a list of the early papers that provide this economic foundation and motivation for these variables.
[2] The exception is Jordan and Vivian (2011). Specifically, Bossaerts and Hillion (1999) use fundamental variables including long-term bond excess returns, T-bill returns, the stock market's price level, the market's dividend yield, and the market's price-earnings ratio. Rapach, Wohar, and Rangvid (2005) examine predictability of macro variables including: money market rate, T-bill rate, bond spread, long-term bond yield, inflation rate, industrial production growth, narrow money growth, broad money growth, and change in the

fundamental ratios based on dividends and earnings.[3] Further, a similar set of large economies consisting of the major developed countries is utilized.[4] Overall, the international literature confirms that OOS predictability either does not exist or if it does exist, it is not clearly utility enhancing for investors.

We extend the prior literature in several dimensions. First, we subject our predictor variables to extensive (OOS) forecasting tests on a broad subset of countries not previously studied. This includes developed and emerging markets, Euro and non-Euro currency countries, as well as small and medium-sized economies. We provide portfolio allocation evidence for more than 10 European countries; this evidence is particularly valuable because it investigates if return forecasts can help improve the asset allocation decision and provides an estimate of the magnitude of possible real-time gains. We provide initial empirical evidence on the link between market characteristics and OOS forecast performance. Second, we investigate the economic value of predictability under mean-variance optimization and under the manipulation proof methodology of Goetzmann, Ingersoll, Speigel, and Welch (2007, hereafter GISW). Third, we apply the methodology of Campbell and Thompson (2008) to determine if economically motivated parameter restrictions work in our sample of countries. Finally, we subject the predictability of technical indicators to new international data.

Our paper provides evidence on four questions. First, can any variable beat the historical average? We compare the performance of fundamental, macro, and technical variables in terms of both statistical and economic significance. We find consistent predictability of market returns. Macro

---

unemployment rate. Rapach and Wohar (2009) use the dividend-price ratio. Giot and Petitjean (2011) use fundamental variables including long-term bond spread, T-bill returns, the market's dividend yield, and the market's price-earnings ratio. Jordan and Vivian (2011) use fundamental ratios (dividend-price, consumption-price, and output-price) and their growth adjusted counterparts.

[3] The exception is Rapach, Wohar, and Rangvid (2005).

[4] Bossaerts and Hillion (1999) use data starting after 1969 to 1995 for Australia, Belgium, Canada, France, Germany, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, the UK, and the US. Rapach, Wohar, and Rangvid (2005) use data from the mid 1970s to the late 1990s for Belgium, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, the UK, and the US. Rapach and Wohar (2009) study the G7 countries. Giot and Petitjean (2011) use data starting from early 1950s to late 1960s ending in 2005 for Australia, Canada, France, Germany, Japan, Netherlands, South Africa, Sweden, the UK, and the US. Jordan and Vivian (2011) use a long data set from 1927 to 2009 for Australia, Germany, France, Italy, Sweden, the UK, and the US.

4

variables and to some extent technical variables consistently beat the historic average benchmark. However, when the traditional fundamental ratios based on dividends and earnings are used, little-to-no OOS predictability is found.

Second, can parameter restrictions improve forecast accuracy and generate economic benefits? We compare the performance of fundamental, macro, and technical variables under both the free-parameter and the restricted-parameter problem. In agreement with the findings in Jordan and Vivian (2011), we find that outside the US, parameter restrictions do not improve predictability. However, even without parameter restrictions, utility gains can be made from a subset of macroeconomic and technical variables.[5]

Thirdly, can combining forecasts using a simple average improve forecast performance?[6] We find that the most consistent model to beat the historic benchmark is a simple average over all individual models. In contrast to the moderate gains reported in Jordan and Vivian (2011), we find large economic gains in our sample of small and medium-sized economies. Jordan and Vivian (2011) propose that the gains to a simple average across models may be due to reduction in noise. Our results are consistent with the idea that single model predictions are subject to noise. There are two reasons why our data may contain more noise than that in Jordan and Vivian (2011): (1) we utilize smaller and less developed markets and (2) we use higher frequency data (i.e., monthly) than the data used in Jordan and Vivian (i.e., annual). The more noise that exists in the data the larger are the potential gains from averaging across models.

---

[5] Neely, Rapach, Tu and Zhou (forthcoming) find that technical indicators provide statistically and economically significant out-of-sample forecasting power and generate substantial utility gains. Further, technical indicators tend to detect the typical decline in the equity risk premium near cyclical peaks, while macroeconomic variables more readily pick up the typical rise near cyclical troughs. In line with this cyclical behavior, utilizing information from both technical indicators and macroeconomic variables substantially increases out-of-sample forecasting performance relative to either alone.
[6] Rapach, Strauss and Zhou (2010) and Jordan and Vivian (2011) both find forecast improvements in large economies when combining methods are used.

Finally, do US results translate well to countries with different characteristics or different sample periods? Can returns be forecast in smaller, newer, less market-oriented European economies than previously studied? We find both INS and OOS support for the ability of a simple average forecast model to predict returns for our sample of medium to small European markets, which includes both developed and emerging markets. We find evidence that the economic value of macro variable forecasts is related to economic development in both sets of tests we conduct. There is also some evidence that i) predictability performance of fundamentals is related to liquidity and market development and ii) technical variables provide larger economic gains in both larger and more developed markets. Overall, our results suggest the value of OOS forecasts can differ depending upon a countries size, liquidity and development.

From a practical perspective, our results suggest implementing these averaging strategies in medium and smaller markets could help investors to time-vary their portfolio allocations between debt and equity. Our results also could help corporate managers to time equity or debt issues. Market timing strategies based on market predictability could well be operational given that exchange traded funds (ETF's) have recently become available for many of these markets.

2. DATA DESCRIPTION

Our sample covers monthly data for 14 European and Mediterranean countries over the period January 1995 to December 2011. The sample comprises eight western European countries: Austria (OE), Finland (FN), Switzerland (SW), Luxembourg (LX), Greece (GR), Ireland (IR), Portugal (PT), and Spain (ES); three emerging European markets: Czech Republic (CZ), Hungary (HN), and Poland (PO); and three Mediterranean countries: Cyprus (CP), Israel (IS), and Turkey (TK), over a balanced

6

Deleted: . For example, we find evidence that macro variable forecasts provide larger economic gains in more developed markets, while

Deleted: have value in

Deleted: of differing

sample period from 1995-2011. The selection criteria include European and Mediterranean countries for which: there is a Datastream Total Market Index, data is available from at least 1995, and no prior OOS forecasting evidence exists.[7] We choose the January 1995 date because it means we have 60 monthly observations before beginning the OOS forecasting period in 2000, which allows for a reasonably sized OOS test period. The countries excluded by the 1995 requirement are Bulgaria [1998], Egypt [1996], Malta [2000], Romania [1996], Russia [1998], and Slovenia [1998], where the start date of the data is in brackets.

The data is monthly and primarily from Thomson Datastream. Returns have been adjusted for dividends and stock splits. We include six of the variables used by GW. The data appendix (Appendix A) provides further detail on each variable and how these variables are constructed. We include four fundamental variables from GW:

- Dividend–price ratio (log), (DP): Difference between the log of dividends paid on the market index and the log of stock prices, where dividends are measured using a one-year moving sum.
- Dividend yield (log), (DY): Difference between the log of dividends and the log of one month lagged stock prices.
- Earnings–price ratio (log), (EP): Difference between the log of earnings on the market index and the log of stock prices, where earnings are measured using a one-year moving sum.
- Dividend–payout ratio (log), (DE): Difference between the log of dividends and the log of earnings.

---

[7] Turkey and Israel both have substantial trading links with Europe. Hjalmarsson (2010), Rangvid et al. (2011) and Cheng, Jahan-Parvar and Rothman (2010) are amongst the few studies that examine the Turkish and Israeli markets.

We include two macroeconomic variables from GW:[8]

• Risk-free rate, (RF): Interest rate on a low risk short-term security.

• Aggregate stock variance, (SVAR): Sum of squared weekly returns on the market index over last 52 weeks.

We consider two new technical variables. These variables are of interest to technical traders, an important subset of investors that depend on forecasts:

• Price Pressure, (PRES): Calculated as the ratio of the number of rising stocks in the previous month divided by the number of falling stocks.

• Change in Volume, (CVm): Calculated as the monthly change in the volume of traded stocks (in the index).

Table 1 provides a summary of descriptive statistics for our sample of countries. Each line has two entries. The top line provides the mean of each variable and the second line provides the standard deviation. There are several interesting comparisons. First, the average nominal returns (r) vary substantially across countries from 0.029 (2.9% per month or 40.9% per year compounded) in Turkey, down to 0.004 (0.4% per month or 4.9% per year compounded)) in Ireland. The standard deviation of returns also varies substantially across countries from 0.137 for Turkey to 0.046 for Switzerland. This means the Sharpe ratio (return per unit of risk) varies dramatically across countries from over 0.14 in Israel, Czech Republic, and Hungary to around 0.06 in Greece, Ireland, and Cyprus. The wide variation across countries exists for most the variables we study. In particular there seems to be a wide variation in the risk-free rate (RF). From the descriptive statistics contained in Table 1, we can see that the standard deviation is often 15 times the mean value.

---

[8] Jordan (2012) demonstrates that macro variables are able to capture predictability in international market returns.

[INSERT TABLE 1 AROUND HERE]

3. EMPIRICAL METHODOLOGY

3.A   Predictive regressions and individual forecasts

Equation (1) is used to measure INS predictive power. r_{t+1} is the nominal continuously compounded log stock return (not excess return)[9] from t to t+1. z_t is the predictor variable for t. We estimate the following regression:

$$r_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1} \tag{1}$$

We estimate Equation (1) for a 1-month horizon and calculate bootstrapped t-statistics similar to Mark (1995) and Nelson and Kim (1993). This simulation approach helps mitigate concerns over the impact of autocorrelation and small-sample bias (Nelson and Kim, 1993; Ang and Bekaert, 2007) as well as concerns over data mining (Rapach and Wohar, 2006).

Inoue and Kilian (2005), Giacomini and White (2006), Alquist et al. (2011), Clark and McCracken (2010, 2011a, 2011b, 2012a, 2012b) have made a distinction between population-level predictability and finite-sample predictability. In comparing OOS forecasts from nested models, we examine MSPEs of the benchmark and nesting model and conduct tests of OOS population-level

---

[9] We use raw or gross returns for our tests for several reasons. First, since we make 1-month forecasts, the risk-free rate is known at the time of the forecast. Thus, it is not necessary to adjust for the risk-free rate (see Ferreira and Santa-Clara (2011) for a discussion). Second, the theoretical basis for return predictability in Campbell and Shiller (1988) is derived using "realized log gross returns" (see the definition of h_{it} on page 663 of Campbell and Shiller). Third, prior papers have used gross returns (see Ferreira and Santa-Clara, 2011 and Lettau and Van Nieuwerburgh, 2008). Lastly, we produce our main results with excess returns as robustness tests and find similar results. Thus, we report our results using log gross returns to be consistent with the motivating theory.

9

predictability. These tests are equivalent to tests of the null hypothesis that the extra parameters in the nesting model are jointly equal to zero i.e., testing whether the population beta(s) are equal to zero. This is in contrast to testing for finite-sample predictability, which focuses on testing the null hypothesis of equal (expected) OOS MSPEs. Finite sample tests of predictive ability require a larger difference in MSEs to reject the null hypothesis (see Clark and McCracken, 2012a).[10] We use population-level predictability tests because we wish to test whether the population beta(s) are equal to zero, consistent with our INS tests. Our approach of implementing OOS population level tests is also consistent with closely related US literature (Goyal and Welch, 2008; Campbell and Thompson, 2008; Rapach et al., 2010) and thus enables comparability. A final minor issue is that with implementing OOS tests of finite-sample predictability is that they either tend only to be applicable in certain environments, for example Giacomini and White (2006) can only be applied to rolling estimation windows, or they are more complex to implement than population level tests, for example inference for the Clark and McCracken (2012b) test requires a fixed regressor bootstrap to be conducted.

Out-of-sample our procedure mimics the situation faced by real-time forecasters. Forecasts from regression models are generated using only information available at period t. Time-varying coefficients of each model are estimated using a recursive (expanding window) regression technique given by Equation (2) and then forecasts are produced using Equation (3). We implement an initial window length covering five years of monthly data (60 observations). We add one observation for each subsequent time we repeat the parameter estimation. Our initial estimate utilizes 60 observations due to concerns about parameter estimation error. We use an expanding estimation window because parameter estimation error is reduced with sample size (see for example Clark and McCracken,

---

[10] This is apparent in Clark and McCracken (2012a) where in-sample tests of population predictability use critical values from chi-square distributions while in-sample finite sample tests of predictive ability us non-central chi-square critical values; thus the critical values for the finite sample test are larger.

2009).[11] Thus the February 1995 to January 2000 period provides the first coefficient estimates and the first monthly forecast is for the February 2000 return. This regression is followed for each predictor variable.

$$r_t = \alpha_t + \beta_t z_{t-1} + \varepsilon_t \tag{2}$$

$$\hat{r}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t z_t \tag{3}$$

The historical average, which is calculated over all prior observations, simply expects that next period's return is equal to the mean of all previous returns: $E(r_{t+1}) = \overline{r_t}$. This is equivalent to restricting $\beta = 0$ and thus the historical average is equivalent to the prediction of a random-walk model with drift and nested within the regression forecasts.

3.B   Forecast evaluation and application

We follow GW and Campbell and Thompson (2008) in calculating forecast evaluation measures. Campbell and Thompson (2008) propose an out-of-sample R^2 (OOS R^2), to assess the forecasting performance of each model, which is closely related to the commonly used Theil's U.[12] The OOS R^2 measure compares the performance of a specific model relative to a benchmark. The benchmark used in the literature is the historical average return, which is the forecast from a random-walk model with drift.

---

[11] Clark and McCracken (2009) note there is a trade-off between bias and variance when choosing to implement rolling or recursive schemes. Recursive schemes should have lower variance due to reduced parameter estimation error; however, rolling schemes will have less bias when there are structural breaks in the predictive regression.

[12] Note that the OOS R^2 is equal to $1-U^2$, i.e., 1- Theil's U squared.

$$OOS\ R^2 = 1 - \frac{CSE_{z,t}}{CSE_{HA,t}} = 1 - \frac{\sum_{n=1}^{N}(r_t - \hat{r}_{z,t})^2}{\sum_{n=1}^{N}(r_t - \hat{r}_{HA,t})^2} = \frac{\sum_{n=1}^{N}(r_t - \hat{r}_{HA,t})^2 - \sum_{n=1}^{N}(r_t - \hat{r}_{z,t})^2}{\sum_{n=1}^{N}(r_t - \hat{r}_{HA,t})^2} \qquad (4)$$

Equation (4) computes the ratio of cumulative squared error (CSE) of the regression model (z_t) from period 1 to period t as a proportion of the CSE of the historical average (HA) over the same period. The summation is over all forecasts made. Thus, if we make N forecasts, then we sum from $n =$ 1 to N in Equation (4), where $n = 1$ is the first forecast made and $n = N$ is the final forecast made. The OOS R^2 is then defined as one minus the ratio of cumulative squared errors. Clearly, if the OOS R^2 is positive then this indicates the regression model on average beats the historical average benchmark over the sample period. This metric also has the useful property that its value represents the proportion by which the benchmark is outperformed or underperformed. For instance a value of -0.25 indicates the cumulative mean-squared error[13] of the regression model is 25% higher than that of the historical average prediction and translates to an underperformance of 25% over the sample period considered.

To statistically assess the performance of the models, we report results from McCracken's (2007) MSE-F test.[14] The MSE-F statistic is a one-sided test for equal forecast accuracy. More specifically it is formulated under the null that the forecast error from the regression model is equal to or larger than (inferior to) that from the historical average regression. A rejection of the null hypothesis indicates that the regression model has superior forecast performance than the benchmark.

---

[13] For OOS R^2 either cumulative squared error or cumulative mean-squared error can be used, since the number of periods t is constant in both cases and thus cancels out when the ratio is taken.

[14] We also implemented the Clark and McCracken's (2001) Encompassing Test (ENC-NEW). Results are qualitatively similar to those for MSE-F. If equal forecast accuracy is rejected then the regression model forecast is not encompassed by the historical average model. In the interests of brevity we report only MSE-F.

$$MSE - F = (T - h + 1) \times \left(1 + \frac{1}{1 - [OOS\ R^2]} \right)$$

$$= (T - h + 1) \times \left(1 - \frac{CSE_{HA,T}}{CSE_{z,T}} \right)$$

(5)

h measures the degree of overlap, where h is equal to 1 for no overlap. Clark and McCracken (2005) show MSE-F have non-standard statistical distributions. Hence, critical values for MSE-F (as well as INS t-statistics) are produced via a bootstrap procedure following Mark (1995) and implemented in a similar manner to Goyal and Welch (2008) and Rapach and Wohar (2006).

$$r_{t+1} = \alpha_{t+1} + \varepsilon_{1,t+1}$$

$$z_{t+1} = \delta_{t+1} + \theta_1 z_t + \varepsilon_{2,t+1}$$

(6)

Parameters are estimated using the full sample and error terms are saved to generate pseudo series for r and z. The pseudo series for r and z have identical length to our sample and are formed by drawing from the time-series of residuals with replacement. We create pseudo series for r and z, dropping the first 100 start up series and then save the next 1000 simulated series of r and z. Bootstrapped critical values for each test are created by running the INS and OOS procedures for each set of simulated series.

We also run the bootstrap with forecasting restrictions as suggested by Campbell and Thompson (2008), i.e., we restrict i) the slope coefficient to be consistent with theory ($\beta_t \geq 0$ for all variables except for the macro variables) and ii) the stock return forecast to be positive ($r_t \geq 0$).

Our final set of empirical tests deal with the economic value of forecasts. We analyze if portfolio allocations could have improved by following the regression model rather than the historical average. Firstly, we consider a mean-variance optimizing investor in the spirit of Campbell and Viceira (2002) and Campbell and Thompson (2008). Our analysis is for log returns whilst Campbell and Thompson (2008) expound the model for simple excess returns. We take the predictive regression:

$$r_{t+1} = \alpha + \beta_t z_t + \varepsilon_{t+1} \tag{7}$$

where r_{t+1} is the log stock return. A mean-variance optimizing investor has objective function:

$$
\begin{aligned}
O &= E(r_p) + \frac{1}{2}\sigma_{r_p}^2 - \frac{\gamma}{2}\sigma_{r_p}^2 \\
&= E(r_p) - \frac{\gamma-1}{2}\sigma_{r_p}^2
\end{aligned}
\tag{8}
$$

where O is the objective, r_p is the portfolio return and γ is the coefficient of relative risk aversion. Such an investor will choose a portfolio weight (ω) of the risky asset under the prediction from the historical average and regression model (7):[15]

$$\omega_{t,HA} = \omega_{HA} = \left(\frac{1}{\gamma}\right)\left(\frac{\alpha_t + \frac{1}{2}(\sigma_t^2)}{\sigma_t^2}\right) \tag{9}$$

$$\omega_{t,z} = \left(\frac{1}{\gamma}\right)\left(\frac{\alpha_t + \beta_t z_t + \frac{1}{2}(\sigma_t^2)}{\sigma_t^2}\right) \tag{10}$$

We use 5-years of rolling monthly data to estimate volatilities; however, alternative window lengths for estimating volatility have little impact on the *change* in utility since σ_t^2 is common to both $\omega_{t,HA}$ and $\omega_{t,z}$. The weight in the historical average is determined as in Campbell and Viceira (2002, p 29),[16] whilst the weight in the regression model takes into account the prediction of z_t.

The utility gain (ΔO) from using the regression model rather than the historical average is:

$$\Delta O = r_z - r_{HA} - \frac{\gamma}{2}\left(\sigma_{r_z}^2 - \sigma_{r_{HA}}^2\right) \tag{11}$$

Secondly, we implement GISW of abnormal performance:

---

[15] Weights are recalculated in every time period and portfolio allocations adjusted accordingly.

[16] Campbell and Viceira (2002) have the same mathematical expression, but variable definitions vary [e.g., α_t is the risk premium in this paper but it is defined as the portfolio weight in Campbell and Viceira (2002)].

$$GISW = \frac{1}{1-\Gamma}\left[\ln\left(\frac{1}{T}\sum_{t=0}^{T-1}\left(\frac{1+r_{t+1,z}}{1+r_{t+1,f}} \right)^{1-\Gamma} \right) - \ln\left(\frac{1}{T}\sum_{t=0}^{T-1}\left(\frac{1+r_{t+1,HA}}{1+r_{t+1,f}} \right)^{1-\Gamma} \right) \right]$$

$$\text{where: } \Gamma = \frac{\ln[E(1+r_m)] - \ln(1+r_f)}{\text{Var}[\ln(1+r_m)]}$$

(12)

The GISW certainty equivalent measure looks at the average performance of a portfolio relative to the risk-free rate. An advantage of the GISW measure is that it can be difficult to manipulate. The parameter Γ is set to reflect the overall reward (return) to risk (variance) ratio for each country based upon the actual out-of-sample period data. This reduces the possibility of manipulation and incorrect inference.

4. IN-SAMPLE RETURN PREDICTABILITY

In this section, we consider INS predictability at the 1-month horizon. We explore the robustness of the US INS predictability results to new European data covering a range of countries. Statistical evidence of INS predictability is found more frequently than evidence of OOS forecast accuracy in the aggregate stock return literature and in the empirical finance literature more generally. Given these empirical results, many researchers have placed greater emphasis on OOS tests than INS tests. It has also been suggested that INS tests are more susceptible to data mining or dynamic mis-specification. In an important and influential article Inoue and Kilian (2005) provide theoretical analysis that questions these conjectures about the superiority of OOS tests. Results from Inoue and Kilian (2005) indicate that in almost all the settings they consider INS tests of predictability are not

less powerful than OOS tests of forecast accuracy.[17] This result is robust to the data mining adjustment that they consider and the forms of dynamic mis-specification which they consider. However, the loss of power for OOS tests is much greater when compared to the one-sided t-test than the INS F-test where the loss of power is more modest. In our study the INS analysis is conducted using two-sided t-tests (consistent with the overwhelming majority of empirical finance studies) and almost all our INS analysis is conducted on a single predictor variable; Inoue and Kilian note this case is equivalent to the INS F-test for which they report results. The implications for our study are as follows: firstly, in general, the INS two-sided t-tests should be at least as reliable as the OOS forecast accuracy tests. Consequently, both tests should provide similar results. Secondly, it is possible for either the INS test or the OOS test to falsely reject the null hypothesis (or falsely not reject the null hypothesis). This has two consequences: a) it is useful to implement both tests to check if the general results are robust and b) when the results between INS tests and OOS tests differ, on a case-by-case basis (for each country-variable pair), it is not possible to determine which result is false. We emphasize that simply because the OOS test has somewhat less power than the INS two-sided t-test, in general, this does not mean for an individual case where the results differ that the OOS inference is anomalous and the INS inference is correct (or vice-versa). In this paper we implement both INS tests and OOS tests emphasizing results that are generally found whichever test is considered.[18]

Table 2 provides the magnitude of predictability for each country-fundamental pair. Our INS predictability tests consist of regressions of one period ahead stock returns on current predictor variables. Several observations are relevant to the predictability literature. First, traditional fundamental ratios, e.g., dividend-price (DP), dividend-yield (DY), earnings-price (EP), and the

---

[17] Inoue and Kilian (2005) conclusion is stronger, i.e., INS is typically more powerful than OOS, because they emphasize their INS one-sided t-test results. However, we implement their two-sided tests, for which INS and OOS are more comparable.
[18] Inoue and Kilian (2005, page 372) state that "If in-sample and out-of-sample tests of predictability tended to give the same answer, when applied to the same data set, it would not matter much, which one we use." This is true for our data, thus our INS and OOS tests complement each other and one test can be viewed as robustness test for the other.

dividend-payout (DE) generally perform poorly. Most of the coefficients are statistically insignificant at the 5% two-tailed significance level. In total 83.9% (47 of the 56) estimated coefficients on DP, DY, EP, and DE are statistically insignificant at the 5% significance level. Two significant coefficients on EP (Austria and Ireland) have the wrong sign. Thus, 87.5% of the estimated coefficients conflict with the prior evidence from US data. The two work horses of the fundamental predictability literature, DP and DY, are positive and significant at the 5% level for only one country, Turkey.

Second, macro variables show mixed performance across countries. The short term interest rate (RF) performs reasonably as it is significantly linked to returns in 6 of the 14 countries. Results for Stock Variance (SVAR) are statistically significant in 5 of 14 countries, however only 3 are significant at the 5% level.

Thirdly, some technical indicators, which could be utilized by practitioners, perform well in relation to fundamentals. Price pressure (PRES), the number of rising stocks divided by number of falling stocks, is also a consistent predictor across markets. The coefficient on PRES is positive and significant at the 5% level in 8 of 13 countries.

Finally, the last row of Table 2 provides results for the strategy of combining all individual predictor model forecasts (AVall) using a simple average; note the full sample parameter estimates are used in this in-sample analysis. Predictability is found in 8 of 14 countries and the R-squares of the regressions tend to be large compared to most of the individual predictor models, e.g., the largest R-squared for an individual predictor for Finland (FN) is 0.026 for SVAR while it is 0.059 for AVall.

[INSERT TABLE 2 AROUND HERE]

17

Lettau and Van Nieuwerburgh (2008) and Paye and Timmermann (2006) note the US dividend-price ratio predictive ability deteriorates during the 1990s; they suggest this could be due to model instability.[19] Our results verify that the US poor dividend-price predictability in recent years is common across a wide range of European countries consistent with Rangvid, Schmeling and Schrimpf (2011). We extend Rangvid et al. by demonstrating this is not unique to just the dividend-price ratio. In our sample, all fundamental variables have virtually no predictability. Even including the years around the financial crisis when returns mean revert and valuation ratios should do well, our evidence does not support INS predictability.

Overall our results presented in Table 2 suggest that there are cases of 1-month stock-return predictability in European markets. The predictability is not by traditional fundamental ratios (DP, DY, EP, DE). Instead technical indicators and macro variables are the variables that exhibit predictive ability. We also find that the simple average of all forecasts (based on full-sample parameter values) performs well in-sample.

5. OUT-OF-SAMPLE STOCK RETURN FORECASTS

Could investors actually utilize regression models in order to benefit from more accurate predictions of future stock returns? This issue is of importance to both practitioners and academics alike. Asset managers, economic policymakers, as well as pension providers and contributors all need accurate estimates of future market returns.

We examine a range of fundamental-price ratios as well as macro and technical variables for a range of European countries. Following Rapach, Strauss, and Zhou (2010) and Stock and Watson

---

[19] Kellard, Nankervis and Papadimitriou (2010) suggest the weaker performance of US dividend-price compared to UK dividend-price is due to greater disappearance of dividends in the US.

(2004) we consider if combining forecasts using a simple average can improve forecast accuracy over individual models. Whilst there is a large literature on forecast combinations, the simpler forecast methods provide the best results (Clemen, 1989). We use the historical average return as our benchmark.

5.A   OOS Forecast Accuracy (without Restrictions)

Table 3 reports the OOS R^2 in percentage points. Perhaps the most striking finding from Table 3 is that the average of all forecasts (AVall) outperforms the benchmark (is positive) in 11 of the 14 countries. Statistical outperformance of the benchmark is assessed using McCracken's (2007) MSE-F test under the null that the regression forecast is not better than the benchmark (see Section 3.B for a fuller explanation). The MSE-F test is statistically significant at the 1% (5%) level in Luxemburg, Hungary, Turkey (Austria, Switzerland, Greece, Ireland, Portugal, and Spain), which indicates the regression forecast mean-squared error is statistically smaller than the benchmark. With the exception of price pressure (PRES), AVall outperforms single-variate models including fundamental, macro, and technical variables.

[INSERT TABLE 3 AROUND HERE]

The performance of traditional fundamental-price ratios in univariate regressions is dismal. The only country that consistently shows predictability from fundamentals is Turkey. For DP in 12 of

19

the 14 countries and for DY in 12 of the 14 countries, there is no evidence that fundamental-price

ratios beat the historical average benchmark.[20]

OOS forecasts for macro and technical variables outperform forecasts from fundamentals. The

risk-free rate (RF) and price pressure (PRES) all beat the historical average benchmark in about half of

the countries. Not all variables exist for Cyprus, leaving 13 countries in our tests. In 10 of the 13

countries at least one of these variables beats the benchmark providing strong evidence that macro and

technical variables provide consistent predictability of returns in our sample. Overall, these OOS

forecast results largely confirm and corroborate our INS findings.

We provide two robustness tests in Appendix B. First, we exclude the years from 2007

onwards, since the onset of the financial crisis. Next, we conduct our tests with excess returns rather

than raw returns. In both instances the results are in strong agreement with the results reported in

Table 3.

5.B   OOS Forecast Accuracy (with Restrictions)

Table 4 reports results applying the restrictions of Campbell and Thompson (2008) that (i)

coefficients have the correct sign and (ii) returns are positive. Campbell and Thompson find the

restrictions generally improve forecast accuracy, especially when unrestricted forecasts substantially

underperform the benchmark. We find consistent evidence that these parameter restrictions do not

improve forecasts (using OOS R^2) in the set of European countries we examine. Interestingly,

statistical significance when restrictions are applied is very similar to the unrestricted results in Table

3.

---

[20] At the 10% level of statistical significance

[INSERT TABLE 4 AROUND HERE]

At the 5% significance level, the univariate model using price pressure (PRES) provides the most consistent evidence of forecast outperformance. PRES outperforms the benchmark in 10 of 13 countries. Compared to AVall, PRES generally shows larger gains, e.g., when PRES is significant at the 5% level it has larger gains in 9 countries compared to only 3 countries with larger gains using AVall. The average of all forecasts also performs well. AVall, outperforms the historical average for all countries (i.e., all values are positive) except for Portugal and Israel. There is statistical outperformance of the benchmark in the 9 out of 14 markets. Interestingly, the magnitude of predictability for PRES and AVall is often reduced by the parameter restrictions.

The parameter restrictions reduce the performance of both the macro and technical variables. In all variables, predictability is found in less than half of the countries studied. The evidence presented in Table 4 indicates that the restrictions proposed by Campbell and Thompson (2008) do not appear to improve forecast accuracy for our sample of European countries during the mixed market conditions experienced over 2000-2011. This could reflect either those restrictions do not work outside the US and / or restrictions do not work in periods that are not primarily bull markets. There is little evidence of substantial improvements in forecast evidence from applying restrictions. This suggests, in practice, applying such restrictions may be of limited assistance to asset managers.[21]

---

[21] We further explore why the Campbell and Thompson (2008) restrictions do not work in our sample of European data. Appendix C reports the percent of times the restrictions apply for each country-predictor combination.

There are some countries for which the restrictions bind often, yet there is little effect. For OE, the only difference is that after restrictions SVAR is also significant. Restrictions applied in only 13.87% of the forecasts. Restrictions did not matter for DP, DY, EP, DE, or CVm. For these variables, restrictions were applied to 66.69%, 66.42%, 62.77%, 82.48%, and 27.74% of the forecasts.

There are other countries for which restrictions don't bind often, but when the restrictions strongly apply, opposite to Campbell and Thompson's predictions, predictability disappears. For GR, restrictions bind only 1.46%, 0.00%, and 10.95% of the time for DP, DY, and SVAR so understandably there is no difference before and after restrictions. However, restrictions bind 97.81% of the time for RF and predictability disappears once restrictions are imposed.

Overall, we find that OOS predictability is not by traditional fundamental ratios (DP, DY, EP, DE). Instead technical indicators and macro variables are the variables that exhibit predictive ability. We also find that the simple average of all forecasts (based on full-sample parameter values) performs well out-of-sample. These results are broadly consistent with our INS predictability findings. Inoue and Kilian (2005) conclude that if INS and OOS tests give similar conclusions, then it is both tests provide relevance. Since our INS and OOS results are consistent, both results can be viewed as corroborating evidence that predictability exists.

6. ECONOMIC SIGNIFICANCE OF STOCK RETURN FORECASTS

This section examines whether regression forecasts could enhance the risk-return trade-off. First, we follow Campbell and Thompson (2008) where portfolios comprise a mix of equity and the risk-free asset with an equity weight between 0 and 1.5 of the total portfolio. We conduct tests with and without slope or sign restrictions. We apply a utility gain measure and the manipulation proof measure of performance (GISW). The reported significance levels in Table 5 to Table 8 are bootstrapped to account for finite sample and data mining biases. The utility gain measure has a stronger foundation in economic theory, while GISW is a better measure of market timing ability. Both utility gains and GISW provide very consistent results in our application.

Finally, there are a few instances where restrictions rarely bind yet they affect predictability, while when the restrictions strongly bind there is little effect. For example in TK, restrictions bind only 4.38% of the time for DE yet predictability is now found. On the other hand, restrictions bind 100%, 73.73%, 87,59%, and 100% for RF, SVAR, PRES, and CVm with no impact on predictability.

We conclude that imposing the restrictions of Campbell and Thompson (2008) do not materially affect significance in our out-of-sample European data. More importantly, the fact that the Campbell and Thompson restrictions do not affect predictability is not related to how often the restrictions bind. One possible explanation is the US results were just discovered by chance and that the Campbell and Thompson restrictions are not effective in other markets. It is also possible that although the Campbell and Thompson (2008) restrictions appear to be effective in the US they do not present statistical tests for OOS metrics; therefore we don't know whether in the US the restrictions lead to a statistically significant improvement in OOS performance."

6.A   Utility Gains (with and without Restrictions)

Applying the portfolio allocation methodology of Campbell and Thompson, we find that the results are supportive of our prior results. There are consistent utility gains for several models and a few instances of large utility gain. Table 5 reports the results of the economic significance of regression forecasts when sign and slope restrictions are not applied.

[INSERT TABLE 5 AROUND HERE]

The most notable feature of Table 5 is the combination of all forecasts (AVall) offers portfolio gains over the historical average in 11 of 14 countries. Only stock variance (SVAR) and price pressure (PRES) have comparable performance in consistently beating the benchmark. The AVall gains are substantial, e.g., large annualized gains of 13.92 in the Luxemburg and of 10.99 in Finland over the historical average are achieved.

For the traditional valuation ratios (DP, DY, EP, and DE), there is utility loss about as often as there is utility gain. Macro variables (RF and SVAR) do well; in most country-macro variable pairs (21 of 27) there is utility gain. In many of these instances the gains are quite substantial. For example, there is an 18.27 gain for Luxemburg-RF and a 10.84 gain for Ireland-SVAR. Stock variance (SVAR) has positive utility gains in 13 of the 14 European markets. SVAR does extremely well and is almost as consistent and performs about as well overall as the average of all forecasts (AVall). This could be because SVAR reduces the risky portfolio weight during periods of high volatility. If volatility is persistent then this could reduce the utility penalty from holding equity in such periods, i.e., SVAR could do well because it reduces portfolio variance rather than increasing return. The technical variable

(PRES) also performs well. Price pressure (PRES) has utility gains in 10 of the 13 markets. Again, the gains can be large. For example, PRES has a gain of 10.72 for Greece.

Table 6 reports utility gain results applying the Campbell and Thompson (2008) method with slope and sign of forecast restrictions, that is, (i) coefficients have the correct sign and (ii) returns are positive. The results are very similar to those reported in Table 5. The traditional fundamental ratios still perform poorly, having negative incremental utility in over half of the country-fundamental pairs. There is some evidence that restrictions are applied throughout the period and thus some gains of 0 are reported, e.g., RF for Spain and SVAR for Poland. Nevertheless, SVAR and AVall still perform extremely well.

[INSERT TABLE 6 AROUND HERE]

6.B   Manipulation Proof Utility Gains (with and without Restrictions)

We also apply the GISW manipulation proof measure of portfolio performance. It should be noted that the manipulation proof adjustment is made to provide a measure that purely captures an investors' ability to time the market. However, the manipulation proof measure also implies there are differences across countries in the investor risk parameters. This implication does not appear to be too concerning given that there are cross-country differences in risk preferences (Weber and Hsee, 1998).[22]

---

[22] Risk preference is also connected to the uncertainty avoidance cultural characteristic and the openness to experience personality trait, both of which vary substantially across countries.

If the GISW analysis is done with constant risk aversion parameters across countries, then the results should be similar to the utility gains results reported in the previous section.[23]

Table 7 contains the results. The average of all forecasts (AVall) leads to outperformance of the benchmark in 13 of 14 countries. Again annualized gains can be substantial, e.g., there is an 12.01% gain for Luxemburg. Stock variance (SVAR) still performs well with gains in 11 of the 14 countries. Perhaps the biggest difference between the GISW results and the utility results is that dividend-price and dividend-yield ratios have more consistent utility gains under the manipulation proof measure.

[INSERT TABLE 7 AROUND HERE]

Table 8 applies the regression restrictions of Campbell and Thompson (2008) to forecasts before they are used for portfolio choice. The regression restrictions make little difference to the portfolio allocation results compared to the unrestricted results. In fact, very few of the individual portfolio results are substantially affected by applying the regression restrictions at all; not even in cases where there was large underperformance of the benchmark do the restrictions make a material difference. One possible reason for this is that constraining the equity portfolio weight to be between 0 and 1.5, i.e., restricting extreme positions, may be quite effective in reducing the impact of counter-intuitive regression signs or negative return forecasts. Overall, the restricted results presented in Table 8 are consistent with the unrestricted results in Table 7.

---

[23] It does not make sense to use a constant risk aversion across all countries as this violates the manipulation proof concept. The whole idea is that using an average slope coefficient allows spurious results of profits to occur as a local investor can then take advantage of the difference between the average slope and their country specific slope. More generally, if Γ is not set as in equation 12, then the GISW measure becomes a function of both i) timing ability of the investor and ii) the average weight they place in equity. Therefore, the GISW measure in no longer "manipulation proof" since it is no longer purely dependent on the timing ability of the investor. However, the constant risk parameter results are available by request.

25

[INSERT TABLE 8 AROUND HERE]

Our analysis here indicates once more that forecasts from macro and technical variables are generally stronger than those from fundamental ratios consistent with our INS results and OOS forecast accuracy results. Building on the previous section's finding that macro and technical variables enable forecast errors to be reduced relative to the historical average, the results in this section suggest that large portfolio gains can be made and economic value derived by using macro and technical variables. More specifically, this section demonstrates there are cases where regression forecasts would enable an investor to tilt their portfolio so that utility could be increased and the risk-return trade-off enhanced. Simply, when the regression model predicts high (low) returns the portfolio is tilted towards equities (T-bills). Our empirical results suggest there is some economic value from forecasting returns with macro and technical variables that could potentially be exploited by practitioners. Our results also indicate that noise may be an issue and that results are best and most consistently obtained with an average over all individual forecast models.

7. COUNTRY CHARACTERISTICS AND FORECAST PERFORMANCE

In this section we provide initial evidence on whether country characteristics are linked to OOS forecast performance. Prior literature has reported that INS predictability evidence primarily exists in countries with characteristics such as large financial markets or developed markets. However, the link between characteristics and INS predictability is only tested in Rangvid, Schmeling and Schrimpf (2011) for the dividend-price ratio. We provide some new evidence on the link between various OOS forecast performance measures and country characteristics for a range of different predictor variables.

7.A   Correlation Tests

The approach adopted in this section is in the spirit of Cenesizoglu and Timmermann (2012) who examine the correlation between forecast accuracy and economic value. We examine the correlation between country characteristics and forecast performance (i.e., forecast accuracy and economic value).

The country characteristics examined proxy for market development, market size and liquidity. These characteristics could be linked to the degree of market efficiency since one would anticipate equity prices to adjust slowly in less developed, small, or illiquid markets. Prior empirical literature suggests that INS predictability is connected to market development (Hjalmarsson, 2010) and market size (Rangvid et al., 2011). We use two proxies for market development: GDP per capita and Stock Market Capitalization to Total GDP ratio. We use two proxies for market size: Stock Market Capitalization and Total Volume traded. Finally we use the Turnover ratio (Total Volume to Stock Market Capitalization) to proxy for liquidity. These proxies are calculated as averages over the OOS period of 2000-2011.

Table 9 Panel A (B) reports the results for the correlation between country characteristics and forecast accuracy for the unrestricted (restricted) models. Overall, there is limited evidence supporting a relationship between country characteristics and OOS R^2. In Panel A, the ability of traditional fundamental variables, DP and DY, to forecast market returns are positively and significantly related to a market's liquidity. The more liquid a market, the better performance one can expect from the

traditional fundamental variables.[24] This may be an important factor in explaining the different results we find in our less liquid sample of countries compared to results reported for highly liquid countries like the US and UK.[25] Interestingly, AVall is not significantly related to any market characteristic. This result is consistent with the noise theory proposed by Jordan and Vivian (2011). If averaging is an effective control for noise in individual model forecasts, and noise is related to market characteristics, then one would expect weak correlation between AVall and market characteristics. In Panel B after model restrictions, DP and DY predictability remain positive and significantly related to liquidity. In addition, predictability from volatility is positively and significantly related to the level of market development. Returns in more developed markets are better predicted by volatility. This is an intuitive result, as less developed markets should experience higher levels of noise, reducing the relationship between volatility and expected return.

[INSERT TABLE 9 AROUND HERE]

Table 9 Panel C (D) reports the results for the correlation between country characteristics and utility gains for the unrestricted (restricted) models. There is stronger support that utility gains are significantly correlated with market development, market size, and liquidity proxies; this supports the INS results of Rangvid et al. (2011) for DP. These results for utility gains provide stronger support for the thesis that fundamentals are more useful in predicting larger, more developed, and more liquid markets. Again, this may provide important economic intuition for why our subset of smaller and less

---

[24] This results reflects the fact that pricing effect of fundamentals is related to standard finance theory, which depends heavily on market efficiency. Efficient markets assume trade is possible for any quantity, instantaneously, and only impacts price if the trade contains new information. Thus, liquidity is an important characteristic for efficient markets. Higher liquidity should be related to more efficiency. Thus, more liquid markets should exhibit stronger relationships with fundamentals as predicted by theory.

[25] For example, Kellard, Nankervis and Papadimitriou (2010) provide evidence that DP and DY have OOS predictive power in the UK and US over a fairly similar OOS period to this study.

liquid economies exhibit little fundamental predictability compared to the stronger results reported with US data. Panels C and D additionally show that macro variables (RF and SVAR) are positively and significantly related to market development. Again, given the higher noise levels in less developed markets, these results are intuitive. These results do provide some support concerning the thesis in Jordan and Vivian (2011) that AVall should realize higher gains in less developed markets.

Table 9 Panel E (F) reports the correlations between country characteristics and GISW manipulation proof performance measure for the unrestricted (restricted) models. The results for the GISW measure are generally similar to those for utility gains, albeit slightly weaker. The most notable difference from the utility gains is that the correlation between GISW gains and market size proxies are now statistically insignificant for DP, DY and EP.

To summarize, there are two strong and robust links between univariate forecasts and market characteristics. First, there is a strong and consistent relationship between traditional fundamental variables and market liquidity. More liquid markets are more predictable using traditional fundamentals. There is weaker evidence that traditional fundamentals work better in more developed and larger markets. Second, there is strong and consistent evidence that predictability from macro variables is related to market development. More developed markets are more easily predicted by macro variables. In Sections 5 and 6 we find strong OOS predictability for two technical variables and AVall. However, this predictability appears not to be well captured by our correlation tests.

7.B   Risk Factor Principal Component Analysis

In this section, we explore the relationship of the first principal component for various risk classes to the out-of-sample R-squared measure of forecast accuracy, the utility gain, and GISW

measure of economic value. The country characteristics examined again are: market development, market size and market liquidity. For each market characteristic we calculate the principal components of the space spanned by our set of variables. For example, we utilize two market development variables (GDP and market value to GDP). We then use the first principal component as our explanatory variable. The results are reported in Table 10.

Our dependent variables are: OOS R-squared (Panel A), utility gain (Panel B), and the GISW manipulation proof utility gain (Panel C). The dependent variables are pooled across all 14 variables and for up to all nine variables. Since we use principal components, the results in Table 10 are the change in the dependent variable given one standard deviation of the explanatory variable. For example in Panel C for ALL, if the development level is one standard deviation above average, then on average one expects a 0.735 percentage point increase in economic value as measured by GISW gains.

[INSERT TABLE 10 AROUND HERE]

Panel A explores whether the OOS R-squared is related to market characteristics. There is little consistency between the raw results and the results that impose reasonable restrictions on the coefficients in line with Campbell and Thompson (2008). [26] For example, with no parameter restrictions, predictability is found for valuation variables, but this predictability is not found once reasonable restrictions are applied. The opposite is true for macro variable predictability, i.e., there is no predictability without restrictions and predictability is found when restrictions are applied. Tech predictability does not exist (at the 5% significance level) regardless of whether restrictions are applied or not. Overall, this is agreement with what we learned in Table 9 with the correlation results.

---

[26] In Panel A the negative sign on DEV for the valuation variables (VALUATION row) becomes statistically insignificant once reasonable restrictions are applied.

Panels B and C explore whether economic gains (utility gain and GISW manipulation proof measures, respectively) are related to market characteristics. If we consider the results that are consistent across unrestricted and restricted regressions, then the results are nearly identical for both panels. We find that the macro economic variables performance is related to the development level of the market. In both Panels B and C, the coefficient is positive and significant for the MACRO row and DEV column. This is true in both the unrestricted and restricted test results. The technical variables are able to explain both the market's development and size. Again, this is true in both Panels and both for unrestricted and restricted coefficients.

To summarize, there are two strong and robust links between principal component forecasts and market characteristics. First, there is a relationship between macroeconomic variables and market development. More developed markets are more predictable using macroeconomic variables; this result is consistent with the correlation results (see Section 7.A, Table 9). Plausibly, this result could be due to the expected return information contained by macro variables being more efficiently incorporated into equity index prices in more developed countries. Our principal component results help us understand the predictability contained in the technical variables documented in Sections 5 and 6. The technical variables are related to both market development and to market size. Thus, the evidence suggests that technical variables are more useful in prediction models for larger and more developed markets.

8. CONCLUSION

There are few open questions that are fundamental and core to every field of finance. One such question is whether expected future stock returns can be accurately predicted and forecasted. This

study examines in-sample (INS) predictability but focuses on out-of-sample (OOS) forecasting of stock returns using a sample of 14 European countries with differing characteristics. These countries have none or very little prior OOS forecasting evidence. We consider if three types of predictors (fundamental ratios, macro variables, and technical variables) can forecast stock returns. Additionally we examine if methods proposed to overcome poor return predictability can improve forecast accuracy or increase economic value.

The key contributions of the paper provide evidence on the questions: (i) Can regression forecasts of market returns beat the historical average in real time? (ii) Can parameter restrictions improve forecast accuracy and generate economic benefits? We compare the performance under both the free-parameter problem, the restricted-parameter problem, and under the manipulation proof measure of Goetzmann et al. (2007). (iii) Can combining forecasts using a simple average improve statistical and economic value of forecasts? (iv) Do US results translate well to countries with different characteristics? Can returns be forecast in smaller, newer, less market-oriented European countries than previously studied? Is the value of OOS forecasts related to country characteristics?

Firstly, the statistical and economic value of predictability varies across countries and across models. This finding is robust to a thorough analysis of methods that have been proposed to improve predictability such as parameter restrictions and forecast combinations. Specifically, we find that traditional fundamental-price ratios perform poorly. However, some macro variables and technical variables exhibit consistent predictability. A new technical indicator, i.e., the ratio of rising stocks to fallers, performs particularly well. The economic significance of predictability can be substantial. We provide initial evidence that OOS forecast performance is linked to some country characteristics such as liquidity and development. These results provide further evidence that supports one possible

explanation for prior findings that INS return predictability primarily exists in the US and large economies (Rapach and Wohar, 2009; Hjalmarsson, 2010; Rangvid et al., 2011).

Secondly, parameter restrictions have been proposed to solve the poor predictability in the US (Campbell and Thompson, 2008). Our European market results generally provide evidence that such parameter restrictions do not (consistently) improve forecast accuracy. Also, the economic importance of predictability, i.e., the gains from portfolio allocation, is only marginally affected when parameter restrictions are imposed. The poor results of these economic restrictions may be due to the fact that the empirical implementation ignores risk of loss and is not related to the theoretical justification used to motivate positive only returns. The restriction which ignores loss is probably more severe in our 200 to 2011 OOS sample than in Campbell and Thompson's 1925 to 2005 sample as their sample contains many long bull markets, while during our sample down markets were well represented.

Thirdly, a simple forecast procedure using a combination of univariate models leads to an improvement in forecast accuracy and portfolio performance in the majority of countries relative to the historical average benchmark. This finding supports prior US evidence that forecast combination methods often perform well when a wide range of different predictors are used (Rapach et al., 2010). Our results indicate that forecast combination methods are effective in small and medium-sized European markets with different characteristics to the US. This suggests that forecast combinations could generate gains universally. There is some evidence that economic gains from forecast combinations are as large or even larger in the less liquid and less developed markets we consider than those reported in the US studies.

Finally, we provide some new evidence on the relationship between OOS forecast performance and market characteristics. We find evidence that macro variable forecasts provide larger economic gains in more developed markets. Both the correlation and principal component tests support this

finding. We find evidence from principal component tests that technical variables are helpful for forecasting economic gains in both larger and more developed markets. Finally, we also find from correlation tests that traditional fundamentals provide larger forecast gains in more liquid markets and to a lesser extent in more developed and larger markets. This result provide possible insight into why, in contrast to the US evidence, we do not generally find significant forecast ability for traditional fundamental variables in our sample of European countries.

In summary, the statistical and economic value of forecasts from individual predictors tends to vary across European countries. Consistent gains are realized when forecast combinations are implemented. We find that during the 2000-2011 OOS period traditional valuation ratios (such as dividend-price) do not perform well. There are few instances where fundamental ratios would have generated substantial benefits to either investors or corporate managers. On the other hand, some macro and technical variables do have value for predicting future market returns. We find that forecast combination methods are particularly effective for forecasting fluctuating equity index returns. The gains to European investors in terms of improved forecast accuracy and economic value are potentially as large as or even larger than in the US. Substantial benefits to European investors could be generated by time-varying portfolio allocations between debt and equity and substantial benefits to European corporate managers could be generated by timing equity or debt issues.

BIBLIOGRAPHY

Ang, A. and Bekaert, G. (2007), "Stock return predictability: Is it there?" *Review of Financial Studies,* Vol. 20, No. 3, pp. 651-707.

Alquist, R., Kilian, L., and Vigfusson, R. J. (forthcoming), "Forecasting the price of oil." In: Graham Elliott, Clive W. J. Granger, and Allan Timmermann (eds.), Handbook of Economic Forecasting. North Holland, Amsterdam..

Bossaerts, P. and Hillion, P. (1999), "Implementing statistical criteria to select return forecasting models: What do we learn?" *Review of Financial Studies*, Vol. 12, No. 2, pp. 405-428.

Campbell, J. Y. and Shiller, R. J. (1998), "Valuation ratios and the long-run stock market outlook." *Journal of Portfolio Management*, Vol. 24, No. 2, pp. 11–26.

Campbell J. Y. and Thompson S. (2008), "Predicting excess stock returns out of sample: Can anything beat the historical average?" *Review of Financial Studies*, Vol. 21, No.4, pp.1509-1531.

Campbell, J.Y. and Viceira, L. M. (2002), "Strategic asset allocation." Clarendon Lectures in Economics, Oxford University Press.

Cenesizoglu, T. and Timmermann, A. (2012). "Do return prediction models add economic value?" *Journal of Banking and Finance*, forthcoming, Vol. 36, No. 11, pp.2974-2984

Cheng, A., Jahan-Parvar, M. R., and Rothman, P. (2010), "An empirical investigation of stock market behavior in the Middle East and North Africa." *Journal of Empirical Finance*, Vol. 17, No. 3, pp. 413-427.

Clark T. E. and McCracken, M. W. (2001), "Tests of equal forecast accuracy and encompassing for nested models." *Journal of Econometrics*, Vol. 105, pp. 85-110.

Clark T. E. and McCracken, M. W. (2005), "Evaluating direct multistep forecasts." *Econometric Reviews*, Vol. 24, No. 4, pp. 369-404.

Clark T. E. and McCracken, M. W. (2009), "Improving forecast accuracy by combining recursive and rolling forecasts." *International Economic Review*, Vol. 50, No. 2, pp. 363-395.

Clark, T.E. and McCracken, M.W. (2011a), "Testing for unconditional predictive ability." in Oxford Handbook on Economic Forecasting, edited by Michael P. Clements and David F. Hendry, pp. 415-440. Oxford: Oxford University Press.

Clark, T.E., and McCracken, M. W. (2011b), "Advances in forecast evaluation." Working Paper 2011-025, Federal Reserve Bank of St. Louis.

Clark, T.E., and McCracken, M. W. (2012a), "In-sample tests of predictive ability: A new approach." *Journal of Econometrics,* Vol. 170, No. 1, pp. 1-14

Clark, T. E., McCracken, M. W. (2012b), "Nested forecast model comparisons: a new approach to testing equal accuracy." Working Paper 2009-050, Federal Reserve Bank of St. Louis, revised

Clemen R. T. (1989), "Combining forecasts - A review and annotated-bibliography." *International Journal of Forecasting*, Vol. 5, No. 4, pp. 559-583.

Engsted, T. and Pedersen, T. Q. (2010), "The dividend-price ratio does predict dividend growth: International evidence." *Journal of Empirical Finance*, Vol. 17, No. 4, pp. 585-605.

Ferreira, M. A. and Santa-Clara, P. (2011), "Forecasting stock market returns: The sum of the parts is more than the whole." *Journal of Financial Economics*, Vol. 100, No. 3, pp. 514–537.

Giacomini, R. and White, H. (2006), "Tests of conditional predictive ability." *Econometrica*, Vol. 74, No. 6, pp. 1545-1578.

Giot, P. and Petitjean, M. (2011), "On the statistical and economic performance of stock return predictive regression models: an international perspective." *Quantitative Finance*, Vol 11, No. 2, pp. 175-193.

Goetzmann, W., Ingersoll, J, Speigel, M., and Welch, I. (2007), "Portfolio performance manipulation and manipulation-proof performance measures." *Review of Financial Studies*, Vol. 20, No. 5, pp. 1503-1546.

Giacomini, R., White, H., (2006), "Tests of conditional predictive ability." *Econometrica* Vol. 74, No. 6, pp. 1545-1578.

Goyal A. and Welch I. (2008), "A comprehensive look at the empirical performance of the equity premium prediction.", *Review of Financial Studies*, Vol. 21, No.4, pp.1455-1508.

Hjalmarsson E. (2010), "Predicting global stock returns." *Journal of Financial and Quantitative Analysis* Vol. 45, No. 1, pp. 49-80.

Inoue A. and Kilian L. (2005), "In-sample or out-of-sample tests of predictability: Which one should we use?" *Econometric Reviews*, Vol. 23, No. 4, pp. 371-402.

Jordan, S. J. (2012), "Time-varying risk and long-term reversals: A re-examination of the international evidence." *Journal of International Business Studies*, Vol. 43, No. 2, pp. 123-142.

Jordan, S. J. and Vivian, A. J. (2011), "Forecasting stock returns internationally: Can fundamental-price models beat the historical average?" IFABS 2011 Conference (Rome) Paper

Kellard, N., Nankervis, J. and Papadimitriou, F. (2010), "Predicting the equity premium with dividend ratios: Reconciling the evidence." *Journal of Empirical Finance*, Vol. 17, No. 4, pp. 539-551.

Lettau M. and Van Nieuwerburgh S. (2008), "Reconciling the return predictability evidence." *Review of Financial Studies*, Vol. 21, No.4, pp.1607-1652.

Mark N. C. (1995), "Exchange-rates and fundamentals – Evidence on long-horizon predictability." *American Economic Review*, Vol. 85, No. 1, pp. 201-218.

McCracken M. (2007) "Tests of equal forecast accuracy and encompassing for nested models." *Journal of Econometrics*, Vol. 140, No.2 (October), pp. 719-752.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (forthcoming), "Forecasting the equity risk premium: The role of technical indicators." *Management Science*.

Nelson C. and Kim, M. (1993), "Predictable stock returns: The role of small sample bias." *Journal of Finance*, Vol. 48, No. 2, pp. 641-661.

Paye, B. and Timmermann, A. (2006). "Instability of return prediction models." *Journal of Empirical Finance*, Vol. 13,No. 3, pp. 274–315.

Rangvid J., Schmeling, M. and Schrimpf, A. (2011), "Dividend predictability around the world." Working Paper SSRN id 1542592.

Rapach, D. E., Strauss, J. K., and Zhou, G. F. (2010), " Out-of-sample equity premium prediction: Combination forecasts and links to the real economy." *Review of Financial Studies*, Vol. 23, No. 2. pp. 821-862.

Rapach, D. E. and Wohar, M. E. (2006), "In-sample vs. out-of-sample tests of stock return predictability in the context of data mining." *Journal of Empirical Finance*, Vol. 13, No. 2, pp. 231-247.

Rapach, D. E., Wohar, M. E., and Rangvid, J. (2005), "Macro variables and international stock return predictability." *International Journal of Forecasting*, Vol. 21, No. 1, pp. 137-166.

Rapach, D. E. and Wohar, M. E. (2009), "Multi-period portfolio choice and the intertemporal hedging demands for stocks and bonds: International evidence." *Journal of International Money and Finance*, Vol. 28, No. 3, pp. 427-453.

Stock, J. and Watson, M. (2004), "Combination forecasts of output growth in a seven-country data set." *Journal of Forecasting*, Vol. 23, No. 6, pp. 405-430.

Weber, E. U. and Hsee, C. (1998), "Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk." *Management Science*, Vol. 44, No. 9, pp. 1205-1217.

Table 1: Descriptive Statistics – Mean and Standard Deviations of Variables from January 1995 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
R	0.006	0.009	0.007	0.006	0.005	0.004	0.006	0.009	0.010	0.013	0.009	0.007	0.010	0.029
	0.056	0.086	0.057	0.046	0.083	0.061	0.055	0.057	0.067	0.089	0.085	0.116	0.059	0.137
DP	-3.941	-3.616	-3.823	-4.066	-3.608	-3.804	-3.556	-3.582	-3.374	-3.705	-4.058	-3.052	-3.540	-3.819
	0.326	0.490	0.276	0.281	0.369	0.338	0.319	0.379	0.512	0.377	0.568	1.113	0.445	0.513
DY	-3.937	-3.609	-3.818	-4.061	-3.605	-3.801	-3.553	-3.576	-3.367	-3.694	-4.050	-3.053	-3.533	-3.791
	0.316	0.482	0.274	0.274	0.360	0.329	0.309	0.377	0.515	0.377	0.567	1.093	0.439	0.525
EP	-2.729	-2.771	-2.738	-2.886	-2.800	-2.630	-2.843	-2.732	-2.686	-2.665	-2.692	-2.157	-2.194	-2.596
	0.261	0.442	0.508	0.196	0.366	0.353	0.383	0.307	0.344	0.256	0.355	1.183	0.538	0.314
DE	-1.212	-0.845	-1.086	-1.180	-0.807	-1.174	-0.713	-0.850	-0.688	-1.040	-1.365	-0.894	-1.346	-1.223
	0.306	0.342	0.493	0.145	0.204	0.148	0.241	0.166	0.588	0.314	0.595	0.897	0.514	0.419
RF	0.002	0.003	0.002	0.001	0.005	0.003	0.006	0.003	0.005	0.010	0.009		0.006	0.035
	0.001	0.001	0.001	0.001	0.004	0.001	0.002	0.002	0.004	0.006	0.006		0.004	0.033
SVAR	0.022	0.040	0.022	0.023	0.035	0.028	0.023	0.027	0.030	0.037	0.038	0.040	0.028	0.059
	0.011	0.017	0.009	0.008	0.013	0.013	0.010	0.008	0.008	0.012	0.013	0.019	0.007	0.021
PRES	1.082	1.051	1.073	1.041	0.965	1.055	0.978	0.992	1.203	1.003	1.019		1.038	1.043
	0.222	0.233	0.253	0.202	0.278	0.262	0.236	0.203	0.335	0.236	0.265		0.284	0.333
CVm	0.058	0.059							0.155	0.081	0.073		0.190	0.094
	0.309	0.279							0.865	0.386	0.361		1.675	0.406

This table reports the mean (upper value) and standard deviation (lower value). For example, for nominal return R, Greece has a monthly return of 0.005 and a standard deviation of 0.083. OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

R is the nominal return, DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency.

Table 2: In-sample Predictability of Stock Returns - 1 month ahead over February 1995 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	-0.085 *	-0.013	0.060	-0.088 *	0.037	-0.037	-0.073	0.065	0.101	0.149 *	0.039	-0.042	-0.002	0.241 ***
	{0.007}	{0.000}	{0.004}	{0.008}	{0.001}	{0.001}	{0.005}	{0.004}	{0.010}	{0.023}	{0.002}	{0.002}	{0.000}	{0.058}
DY	-0.038	0.029	0.102	-0.057	0.075	0.004	-0.038	0.085	0.111	0.164	0.030	-0.019	0.011	0.222 ***
	{0.001}	{0.001}	{0.010}	{0.003}	{0.006}	{0.000}	{0.001}	{0.007}	{0.013}	{0.028}	{0.001}	{0.000}	{0.001}	{0.050}
EP	-0.143 **	0.014	-0.077	-0.101 *	-0.045	-0.130 **	-0.014	-0.024	0.081	0.162 **	0.038	-0.083	0.014	0.226 ***
	{0.021}	{0.000}	{0.006}	{0.010}	{0.002}	{0.017}	{0.000}	{0.001}	{0.007}	{0.027}	{0.002}	{0.007}	{0.000}	{0.051}
DE	0.032	-0.036	0.113	-0.034	0.147 **	0.227 ***	-0.072	0.192 ***	0.041	0.047	0.014	0.058	-0.017	0.125 *
	{0.001}	{0.001}	{0.013}	{0.001}	{0.022}	{0.051}	{0.005}	{0.037}	{0.002}	{0.002}	{0.000}	{0.003}	{0.000}	{0.016}
RF	-0.234 ***	-0.150 **	-0.224 ***	-0.099	0.157 **	-0.017	-0.172 **	0.045	-0.081	0.181 **	0.024		0.057	0.111
	{0.054}	{0.022}	{0.050}	{0.010}	{0.025}	{0.000}	{0.030}	{0.002}	{0.007}	{0.034}	{0.001}		{0.003}	{0.012}
SVAR	-0.028	-0.163 **	-0.241 ***	-0.055	-0.082	-0.160 **	-0.122 *	-0.064	-0.055	0.094	0.096	-0.153 *	0.071	0.079
	{0.001}	{0.026}	{0.058}	{0.003}	{0.007}	{0.025}	{0.015}	{0.004}	{0.003}	{0.009}	{0.009}	{0.023}	{0.005}	{0.006}
PRES	0.189 ***	-0.007	0.193 ***	0.192 ***	0.222 ***	0.158 **	0.272 ***	0.153 **	0.093	0.127 *	-0.004		0.182 ***	-0.020
	{0.036}	{0.000}	{0.037}	{0.037}	{0.049}	{0.025}	{0.074}	{0.023}	{0.009}	{0.017}	{0.000}		{0.033}	{0.000}
CVm	-0.010	0.092							0.004	0.091	-0.041		0.034	-0.093
	{0.000}	{0.009}							{0.000}	{0.009}	{0.002}		{0.001}	{0.009}
AVall	4.748 ***	7.949 **	3.229 ***	3.844	4.886 ***	3.591 **	5.272 ***	4.352 **	4.207	2.845	5.275	3.347	6.463	3.007 **
	{0.072}	{0.059}	{0.082}	{0.039}	{0.078}	{0.062}	{0.098}	{0.048}	{0.026}	{0.053}	{0.010}	{0.024}	{0.035}	{0.076}

Notes: This table reports the in-sample predictability of stock returns at the 1-month horizon. Our in-sample predictability tests consist of regressions of one period ahead stock returns on current predictor variables. Each country-fundamental pairing the top value is the slope coefficient and the adjusted R-squared is given in curly brackets. ***, ** and * denotes statistical significance of slope coefficient at the 1%, 5%, and 10% level respectively for a two-sided test; critical values are bootstrapped from the empirical distribution.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean; The AVall results are based on regressing the stock return for each country on a constant and the average of all individual model forecasts combined using the mean (in this in-sample exercise individual model forecasts for each period are based on the full-sample parameter estimates).

Table 3: Out-of-sample Forecasts of Stock Returns - 1 month ahead from February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	-3.15	-2.34	0.00	-1.42	-0.49	-3.68	-0.59	-1.84	-1.51	0.37 *	-1.96	-3.14	-3.28	8.35 ***
DY	-3.29	-2.00	0.04	-2.19	0.15	-3.98	-0.98	-1.51	-1.27	2.00 **	-1.54	-3.12	-3.38	6.38 ***
EP	-0.77	-3.78	-3.98	-0.75	-0.72	-0.63	-1.34	-2.27	-0.61	-2.50	-2.29	-0.90	-1.49	5.18 ***
DE	-1.18	-5.68	-2.64	-2.66	1.71 **	3.72 ***	-5.11	2.41 **	-1.76	-0.05	-1.47	-1.73	-1.32	-1.10
RF	5.57 ***	0.20 *	3.70 ***	0.12	1.86 **	-1.87	3.80 ***	-0.73	-0.83	4.51 ***	-1.59		-2.37	-60.55
SVAR	-6.86	-0.56	3.48 ***	-1.72	0.08	-0.57	-1.13	-1.19	-0.63	0.25	-1.74	-0.63	-1.05	-0.79
PRES	3.66 ***	-0.68	2.11 **	4.53 ***	6.20 ***	2.15 **	7.36 ***	1.91 **	-0.24	2.76 ***	-2.84		2.68 ***	-0.80
CVm	-1.70	-1.06							-0.29	0.32	-3.34		0.12	-0.02
AVall	2.73 **	0.17	4.42 ***	0.84 **	2.77 **	1.70 **	2.09 **	2.02 **	0.20	3.15 ***	-0.68	-1.33	-0.80	5.73 ***

Notes: This table reports the out-of-sample R-squared (OOS R^2) in percentage points. OOS R^2 gives the percentage by which the regression model beats the historical average benchmark. Note that the link between OOS R^2 and Theil's U is OOS $R^2 = 1 - U^2$. Statistical inference is based on McCracken's (2007) MSE-F test, which assesses if the forecast error from the regression model is smaller than the forecast error from the historical average regression. Critical values are based on a bootstrap procedure under the null hypothesis of equal forecast accuracy. ***, ** and * denotes statistical significance at the 1%, 5%, and 10% level respectively for a one-sided test.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

Table 4: Out-of-sample Forecasts of Stock Returns with Restrictions applied - 1 month ahead over February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	-0.60	-1.13	-0.44	-0.77	-0.38	-2.49	0.00	-1.84	-1.22	0.44	-1.86	-0.10	-2.41	8.00***
DY	-0.49	-1.33	0.13	-1.36	0.15	-2.90	-0.13	-1.53	-1.09	2.04**	-1.17	-0.40	-2.49	8.40***
EP	-0.31	-0.15	-0.22	-0.91	-0.59	-0.97	-0.48	-1.36	-0.56	-2.87	-2.08	0.00	-1.17	4.68***
DE	-0.38	-5.66	3.31***	-1.40	1.75**	4.76***	-0.18	4.32***	-0.92	-0.05	-1.15	-0.62	-0.61	1.10**
RF	4.04***	-0.63	1.01*	0.04	-0.17	-0.36	3.69***	0.00	-0.41	0.00	-0.11		-0.90	0.00
SVAR	1.77**	2.87***	3.94***	0.30	0.71	3.35**	1.78**	-0.26	0.87*	-0.61	0.00	3.83***	-0.03	-0.85
PRES	2.61**	-0.10	1.64**	3.58***	4.44***	1.90**	6.30***	2.33**	1.70**	2.79***	0.00		2.25**	-0.48
CVm	-1.52	-1.20							-0.24	0.32	-0.14		0.12	0.00
AVall	1.98**	0.64	3.54***	0.69	1.32**	1.16*	2.46***	1.62**	0.64*	1.70**	-0.08	0.83	-0.39	4.81***

Notes: This table reports the out-of-sample R-squared (OOS R^2) in percentage points. OOS R^2 gives the percentage by which the regression model beats the historical average benchmark. Note that the link between OOS R^2 and Theil's U is OOS $R^2 = 1 - U^2$. The Campbell and Thompson (2008) restrictions are applied. If the slope coefficient has a counter-intuitive sign, it is set to 0 and the minimum forecasted return is zero. Statistical inference is based on McCracken's (2007) MSE-F test, which assesses if the forecast error from the regression model is smaller than the forecast error from the historical average regression. Critical values are based on a bootstrap procedure under the null hypothesis of equal forecast accuracy. ***, ** and * denotes statistical significance of MSE-F at the 1%, 5%, and 10% level respectively for a one-sided test.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

Table 5: Portfolio Allocation Gains to a Mean-Variance Optimizing Investor - 1 month ahead over February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	1.58	-0.31	-1.32	1.71	-2.85	-2.51	0.19 **	5.02 **	0.40	-0.65	-4.13	-3.23	-3.14	0.78
DY	1.88 *	0.22	1.37	0.93	-1.60	-2.17	-0.95	5.73 **	0.35	-1.38	-3.51	-4.16	-2.99	2.67 **
EP	4.02 **	-2.53	4.26 **	2.10	-1.15	0.14	-2.03	-1.16	-2.33	-0.20	-4.95	1.09	-1.71	1.96 **
DE	-0.16	0.59	10.87 ***	-1.25	8.15 ***	13.72 ***	-2.65	9.36 ***	-1.41	-2.70	-1.61	-1.95	-1.58	-0.52
RF	13.53 ***	13.37 ***	18.27 ***	3.39 *	6.45 ***	-3.21	3.27 ***	-2.22	-0.04	9.85 ***	-0.82		-0.41	3.21 ***
SVAR	4.08 **	10.99 ***	13.92 ***	1.32	3.56 **	10.84 ***	4.16 ***	3.78 **	2.85 **	1.49 **	1.37 **	5.97 **	-2.35	2.06 *
PRES	5.99 ***	-2.02	8.38 ***	8.33 ***	10.72 ***	7.24 ***	7.40 ***	8.70 ***	3.00 **	6.63 ***	-3.64		8.21 ***	-1.86
CVm	-2.95	0.34							-0.94	0.25	-1.67		0.18	1.12 *
AVall	4.75 **	5.49 ***	14.06 ***	2.94 *	6.27 ***	6.86 **	3.49 ***	3.67 **	0.89	-1.52	-2.29	0.03	-0.48	0.30

Notes: This table reports the economic significance of regression forecasts. The table reports bootstrap significance levels that have been adjusted for data mining *** indicates significance at the 1% level, ** at the 5% significance level, and * at the 10% significance level. Utility gains are reported as percentage points and are annualized by multiplying monthly values by 12. Utility gains are calculated for an investor with mean-variance preferences under assumptions of a relative risk aversion coefficient of 3 and limiting the weight in the risky asset to be no less than 0 and no more than 1.5 (as in Campbell and Thompson, 2008). We forecast the return and use the current market risk-free rate to calculate the excess return. AVall combines forecasts using the average of all individual forecasts. AVall takes the average (unsmoothed) portfolio weight of each portfolio and then applies the weight limits on the risky asset.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

Table 6: Portfolio Allocation Gains to a Mean-Variance Optimizing Investor with Restrictions Applied - 1 month ahead over February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	-1.90	-0.11	-1.32	-0.32	-2.59	-1.53	0.00*	5.02**	0.40	-0.65	-3.94	-0.13	-2.27	0.78
DY	-1.63	0.30	1.37*	0.10	-1.60	-1.64	-0.60	5.73**	0.35	-1.38	-3.05	-0.50	-2.17	2.67**
EP	-1.21	-1.97	-0.78	-0.52	-1.55	-0.85	-0.88	-1.18	-1.99	-0.20	-4.95	0.00	-1.36	1.96**
DE	-0.12	0.34	10.87***	0.01	9.25***	13.72***	-0.70	9.36***	-1.51	-2.70	-1.25	-0.03	-0.71	1.35**
RF	13.80***	13.37***	18.27***	3.39**	-0.18	-0.63	3.36***	0.00	-0.02	0.00	-0.34		-0.86	0.00
SVAR	4.08**	11.02***	13.92***	1.32	4.15**	10.84***	4.16***	3.78**	2.85**	-0.45	0.00	8.73***	-0.04	-0.35
PRES	6.05***	0.05	8.38***	8.33***	10.72***	7.24***	7.40***	8.70***	3.00**	6.63***	0.00		8.21***	-0.07
CVm	-2.96	0.34							-0.94	0.25*	-0.36		0.18	0.00
AVall	4.64***	4.66**	10.17***	1.62*	4.01***	2.12**	4.74***	2.92**	0.86	-3.52	-1.57	2.12**	-0.12	-0.39

Notes: This table reports the economic significance of regression forecasts with slope restrictions and sign of forecast restrictions applied. The table reports bootstrap significance levels that have been adjusted for data mining *** indicates significance at the 1% level, ** at the 5% significance level, and * at the 10% significance level. Utility gains are reported as percentage points and are annualized by multiplying monthly values by 12. Utility gains are calculated for an investor with mean-variance preferences under assumptions of a relative risk aversion coefficient of 3 and limiting the weight in the risky asset to be no less than 0 and no more than 1.5 (as in Campbell and Thompson, 2008). We forecast the return and use the current market risk-free rate to calculate the excess return. AVall combines forecasts using the average of all individual forecasts. AVall takes the average (unsmoothed) portfolio weight of each portfolio and then applies the weight limits on the risky asset.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

Table 7: Portfolio Allocation Gains using GISW Manipulation Proof Measure - 1 month ahead over February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	0.62	0.33	1.93	-0.31	-0.73	-1.06	0.05	6.01**	2.80	4.62**	4.15**	1.40	-1.53	3.52*
DY	0.84	3.95*	5.92**	-0.16	1.35	-1.21	-0.66	6.61**	2.99*	4.47**	3.45*	-0.05	-1.34	2.80*
EP	2.81*	-1.34	3.10*	-0.57	-1.50	2.16*	-1.24	-2.36	-3.09	5.52**	-3.51	5.74**	-0.24	10.41***
DE	0.02	-0.90	9.03***	-1.93	10.21***	4.79**	-1.90	8.60***	-0.40	-0.44	2.96	-0.31	-1.27	-4.73
RF	14.24***	15.39***	17.55***	4.19*	-0.54	-6.23	4.20***	-5.76	2.08	0.54	-1.38		-2.92	-3.84
SVAR	2.51**	9.57**	11.26***	-0.07	4.01*	7.36**	6.90***	0.30	2.58	1.18	0.26	8.73**	-1.34	-1.83
PRES	4.92**	-1.84	7.17***	5.74**	11.51***	3.77**	10.34***	5.67**	2.41*	4.85**	-2.47		5.96***	-0.02
CVm	-1.65	4.16*							-0.76	2.13*	-1.47		0.11	4.42**
AVall	3.73**	3.43**	12.01***	1.25	3.81***	0.91	3.80***	3.16**	1.44	0.64	1.14	3.40*	-0.23	2.78***

Notes: This table reports the economic significance of regression forecasts. The table reports bootstrap significance levels that have been adjusted for data mining *** indicates significance at the 1% level, ** at the 5% significance level, and * at the 10% significance level. We limit the weight in the risky asset to be no less than 0 and no more than 1.5 (as in Campbell and Thompson, 2008). This test utilizes the certainty equivalent measure of Goetzmann, Ingersoll, Spiegel and Welch (2007). The values reported are in percentage points and are annualized. Gains are measured using a certainty equivalent measure of abnormal performance, which is more robust to manipulation motives of agents.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

Table 8: Portfolio Allocation Gains using GISW Manipulation Proof Measure with Restrictions Applied - 1 month ahead over February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	-1.84	1.28*	1.93*	-0.10	-0.41	-0.51	0.00*	6.01***	2.80*	4.62**	4.34***	-0.11	-0.35	3.52**
DY	-1.62	4.07***	5.92***	0.54	1.35*	-0.72	-0.30	6.61***	2.99**	4.47**	3.94***	-0.40	-0.23	2.80*
EP	-1.28	-0.54	-0.35	-0.36	-1.41	0.96	-0.73	-0.74	-2.77	5.52**	-3.51	0.00	-0.02	10.41***
DE	-0.02	-0.40	9.03***	0.26	10.61***	4.79**	-0.28	8.60***	-0.49	-0.44	3.34**	0.34	-0.28	-3.90
RF	14.49***	15.39***	17.55***	4.19**	-0.07	-0.04	4.35***	0.00	2.11*	0.00	-0.09		0.11	0.00
SVAR	2.51***	9.59***	11.26***	-0.07	4.20***	7.36***	6.90***	0.30	2.58**	-0.15	0.00	9.18***	-0.03	-0.26
PRES	4.98**	-0.29	7.17***	5.74**	11.51***	3.77**	10.34***	5.67***	2.41**	4.85**	0.00		5.96***	-0.17
CVm	-1.65	4.16**							-0.76	2.13**	-0.37		0.11	0.00
AVall	3.95***	4.04***	9.36***	1.67**	4.15***	-0.09	5.09***	3.66***	1.52*	1.04**	2.47***	1.97**	0.87**	3.24***

Notes: This table reports the economic significance of regression forecasts when restrictions to slope coefficients and sign of forecast have been applied. We limit the weight in the risky asset to be no less than 0 and no more than 1.5 (as in Campbell and Thompson, 2008). The values reported are in percentage points and are annualized. Gains are measured by the Goetzmann, Ingersoll, Spiegel and Welch (2007) certainty equivalent measure of abnormal performance), which is more robust to manipulation motives of agents. The table reports bootstrap significance levels that have been adjusted for data mining *** indicates significance at the 1% level, ** at the 5% significance level, and * at the 10% significance level.

OE is Austria, FN is Finland, LX is Luxembourg, SW is Switzerland, GR is Greece, IR is Ireland, PT is Portugal, ES is Spain, CZ is Czech Republic, HN is Hungary, PO is Poland, CP is Cyprus, IS is Israel, and TK is Turkey.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

45

Table 9: Forecast performance and Country characteristics

Panel A: Correlation between unrestricted forecast accuracy and country characteristics

| | DEVELOPMENT | | SIZE | | LIQUIDITY |
	GDP	MV/GDP	MV	Vol($)	Turnover
DP	-0.31	-0.17	-0.05	-0.01	0.51 ***
DY	-0.32	-0.23	-0.13	-0.08	0.44 **
EP	-0.43	-0.23	-0.02	-0.03	0.41 *
DE	-0.10	-0.34	0.03	0.07	-0.01
RF	0.34	0.18	0.05	0.00	-0.56
SVAR	0.28	0.12	-0.15	-0.12	-0.13
PRES	0.20	0.20	0.16	0.08	-0.23
VCm	-0.13	0.02	-0.18	0.04	0.42
AVall	0.18	-0.17	-0.07	-0.03	0.23

Panel B: Correlation between restricted forecast accuracy and country characteristics

| | DEVELOPMENT | | SIZE | | LIQUIDITY |
	GDP	MV/GDP	MV	Vol($)	Turnover
DP	-0.29	-0.17	-0.12	-0.09	0.47 **
DY	-0.33	-0.26	-0.19	-0.14	0.43 **
EP	-0.07	-0.06	-0.09	-0.05	0.41 *
DE	0.23	-0.28	0.00	0.04	-0.18
RF	0.12	-0.16	-0.15	-0.18	-0.26
SVAR	0.62 ***	0.16	-0.29	-0.31	-0.45
PRES	0.04	0.11	0.15	0.07	-0.28
CVm	-0.87	-0.46	-0.59	-0.45	-0.03
AVall	0.15	-0.17	-0.12	-0.06	0.25

Panel C: Correlation between unrestricted utility gains and country characteristics

| | DEVELOPMENT | | SIZE | | LIQUIDITY |
	GDP	MV/GDP	MV	Vol($)	Turnover
DP	0.08	0.22	0.61 ***	0.70 ***	0.59 ***
DY	0.22	0.15	0.49 ***	0.60 ***	0.53 ***
EP	0.61 ***	0.25	0.11	0.06	-0.14
DE	0.55 ***	0.01	0.07	0.12	-0.13
RF	0.51 ***	0.18	-0.21	-0.23	-0.21
SVAR	0.69 ***	0.16	-0.18	-0.13	-0.15
PRES	0.39	0.18	0.23	0.18	-0.36
CVm	-0.34	0.31	0.23	0.45	0.73 ***
AVall	0.85 ***	0.28	0.03	0.03	-0.27

Panel D: Correlation between restricted utility gains and country characteristics

| | DEVELOPMENT | | SIZE | | LIQUIDITY |
	GDP	MV/GDP	MV	Vol($)	Turnover
DP	-0.06	0.11	0.45 **	0.62 ***	0.63 ***
DY	0.13	0.14	0.44 **	0.60 ***	0.56 ***
EP	0.07	0.09	0.07	0.09	0.38 *
DE	0.53 ***	0.02	0.07	0.12	-0.12
RF	0.72 ***	0.29	-0.07	-0.08	-0.25
SVAR	0.71 ***	0.21	-0.17	-0.14	-0.27
PRES	0.39	0.19	0.24	0.19	-0.39
CVm	-0.42	0.36	0.19	0.34	0.46
AVall	0.76 ***	0.26	0.01	0.01	-0.32

Panel E: Correlation between unrestricted GISW gains and country characteristics

| | DEVELOPMENT | | SIZE | | LIQUIDITY |
	GDP	MV/GDP	MV	Vol($)	Turnover
DP	-0.35	-0.33	0.04	0.18	0.30
DY	0.08	-0.10	0.06	0.20	0.22
EP	-0.01	-0.18	-0.28	-0.26	0.17
DE	0.39 [*]	-0.08	0.02	0.06	-0.37
RF	0.60 [***]	0.32	-0.10	-0.14	-0.28
SVAR	0.55 [***]	0.13	-0.33	-0.32	-0.35
PRES	0.23	0.06	0.06	0.01	-0.38
CVm	-0.05	0.50	0.54 [*]	0.76 [***]	0.96 [***]
AVall	0.67 [***]	0.11	-0.15	-0.13	-0.36

Panel F: Correlation between restricted GISW gains and country characteristics

| | DEVELOPMENT | | SIZE | | LIQUIDITY |
	GDP	MV/GDP	MV	Vol($)	Turnover
DP	-0.34	-0.24	0.10	0.24	0.39 [*]
DY	0.06	-0.01	0.12	0.25	0.25
EP	-0.24	-0.15	-0.10	-0.04	0.54 [***]
DE	0.42 [*]	0.01	0.10	0.13	-0.36
RF	0.69 [***]	0.31	-0.07	-0.08	-0.22
SVAR	0.55 [***]	0.13	-0.34	-0.33	-0.33
PRES	0.23	0.07	0.05	-0.01	-0.43
CVm	0.22	0.80 [***]	0.57 [*]	0.70 [***]	0.58 [**]
AVall	0.52 [***]	0.08	-0.10	-0.07	-0.24

Notes: This table reports the correlations between measures of forecast performance and country characteristics. Panel A (B) reports results for the correlation between country characteristics and out-of-sample R-squared (OOS R^2) for the unrestricted (restricted) model. Panel C (D) reports results for the correlation between country characteristics and utility gain for the unrestricted (restricted) model. Panel E (F) reports results for the correlation between country characteristics and GISW measure for the unrestricted (restricted) model. The country characteristics are GDP per capita (GDP), stock market development (MV / GDP), total equity market capitalisation (MV), value of total market volume (Vol($)) and turnover (calculated as Vol($)/ MV($)). Country characteristics are measured over the 2000-2011 OOS period. ***, **, * indicate statistical significance at the 1%, 5% and 10% level respectively.

DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean.

Table 10: Univariate regressions – Using principal components for Development proxies and Size proxies.

Panel A: Out-of-sample R-squared

OOSU	DEV	SIZE	LIQ		OOSR	DEV	SIZE	LIQ
ALL	0.285	0.008	-2.190		**ALL**	-0.080	-0.120	0.479
	{0.003}	{0.000}	{0.012}			{0.002}	{0.005}	{0.005}
VALUATION	-0.588 **	-0.064	2.711 *		**VALUATION**	-0.351	-0.147	2.059
	{0.080}	{0.001}	{0.108}			{0.032}	{0.007}	{0.070}
MACRO	2.099	0.047	-15.323		**MACRO**	0.279	-0.275 ***	-1.866 **
	{0.049}	{0.000}	{0.162}			{0.046}	{0.055}	{0.127}
TECH	0.677 *	0.489	-1.090		**TECH**	0.268	0.324	-1.375
	{0.079}	{0.042}	{0.015}			{0.025}	{0.037}	{0.049}

Panel B: Utility gain

UGU	DEV	SIZE	LIQ		UGR	DEV	SIZE	LIQ
ALL	1.301 ***	0.319	-0.929		**ALL**	1.155 ***	0.290	-0.937
	{0.109}	{0.008}	{0.004}			{0.097}	{0.007}	{0.004}
VALUATION	0.802 **	0.690 **	1.391		**VALUATION**	0.442	0.528	1.941
	{0.070}	{0.065}	{0.013}			{0.024}	{0.043}	{0.030}
MACRO	1.862 *	-0.765	-3.267		**MACRO**	2.414 **	-0.453	-4.642
	{0.174}	{0.036}	{0.033}			{0.290}	{0.013}	{0.066}
TECH	1.475 ***	1.154 ***	-2.950		**TECH**	1.357 ***	1.095 ***	-3.060
	{0.134}	{0.084}	{0.040}			{0.138}	{0.092}	{0.052}

Panel C: GISW gain

GISWU	DEV	SIZE	LIQ		GISWR	DEV	SIZE	LIQ
ALL	0.735 **	-0.203	-1.751		**ALL**	0.677 **	-0.013	-0.891
	{0.043}	{0.004}	{0.016}			{0.042}	{0.000}	{0.005}
VALUATION	-0.103	-0.057	0.213		**VALUATION**	-0.084	0.203	1.515
	{0.001}	{0.001}	{0.000}			{0.001}	{0.007}	{0.020}
MACRO	2.251 **	-0.838	-5.931		**MACRO**	2.074 *	-0.682	-4.505
	{0.204}	{0.035}	{0.088}			{0.229}	{0.030}	{0.067}
TECH	0.889 *	0.557 *	-0.925		**TECH**	1.010 ***	0.542 *	-2.455
	{0.068}	{0.028}	{0.006}			{0.099}	{0.029}	{0.044}

Notes: This table reports pooled regressions across our sample of 14 countries and across a subset of variables. ALL consistes of all the variables in our study, including: DP is log dividend-price ratio, DY is log dividend-yield, EP is log earnings-price ratio, DE is log payout ratio, RF is the risk-free rate, SVAR is stock variance, PRES is price pressure measured by number of rising stocks divided by number of falling stocks, and CVm is change in volume at monthly frequency, and AVall is an average of all forecasts combined using the mean. VALUATION is the subset of valuation variables, which includes DP, DY, EP, and the payout ratio. MACRO is the subset of macro variables, which includes GDP and Market value to GDP. TECH is the subset of technical predictors, which includes the ratio of price rises to fallers and the change in trading volume. The top row of each entry is the coefficient estimate, while the R-squared is provided in brackets.

The explanatory variables are defined as follows: DEV (development) is the first principle component to our development variable space, which spans GDP and Market Value to GDP. LIQ is the same as in Table II which is turnover. SIZE is the first principle component of market value and trading volume. The left side of the table (i.e., the columns 3 to 4) is unrestricted GISW gains, while right side of the table (i.e., the last three columns) uses the restricted GISW gain. Non-negativity restrictions are imposed on the coefficients according to the method of Campbell and Thompson (1998).

Panel A uses the OOS R-square as the dependent variable. Panel B uses utility gain as the dependent variable. Panel C uses the GISW gain as the dependent variable. ***, **, * indicate statistical significance at the 1%, 5% and 10% level respectively.

.

Appendix A: Data Definitions

Data in this paper are monthly time series and come from two main sources, which are extracted from DATASTREAM: (i) Datastream compiled stock indices for equity market data and (ii) the International Monetary Fund's International Financial Statistics for interest rates. The following 14 countries are analysed: Austria (OE), Finland (FN), Switzerland (SW), Luxembourg (LX), Greece (GR), Ireland (IR), Portugal (PT), Spain (ES), Czech Republic (CZ), Hungary (HN), Poland (PO), Cyprus (CP), Israel (IS), and Turkey (TK).

- Stock returns are calculated as the log change in the stock return index (RI). Thus $R_t = \ln[RI_t/RI_{t-1}]$). The stock return index codes are: TOTMKOE(RI), TOTMKFN(RI), TOTMKSW(RI), TOTMKLX(RI), TOTMKGR(RI), TOTMKIR(RI), TOTMKPT(RI), TOTMKES(RI), TOTMKCZ(RI), TOTMKHN(RI), TOTMKPO(RI), TOTMKCP(RI), TOTMKIS(RI), and TOTMKTK(RI).

- Dividend-price ratio. In the empirical analysis we take the log of this ratio. This is the sum dividends paid over the last 12 months on firms in the stock index divided by the current price of the stock index. Datastream reports this value in percent and so we convert it to decimal by dividing by 100. Then the natural logarithm is taken. The codes of the raw dividend-price ratio (DY) are:

  TOTMKOE(DY), TOTMKFN(DY), TOTMKSW(DY), TOTMKLX(DY), TOTMKGR(DY), TOTMKIR(DY), TOTMKPT(DY), TOTMKES(DY), TOTMKCZ(DY), TOTMKHN(DY), TOTMKPO(DY), TOTMKCP(DY), TOTMKIS(DY), and TOTMKTK(DY).

- Dividend-yield ratio. This is the sum dividends paid over the last 12 months on firms in the stock index divided by the previous months stock index price. This is calculated by taking Datastream dividend-price ratio (DY) and multiplying it by its current price index (PI) and dividing by the previous months price index. The natural logarithm of this ratio is then taken.

- Earnings-price ratio. This is derived in a similar manner to the dividend-price ratio. Datastream provides the price-earnings ratio (PE). The current stock index price divided by the sum of earnings generated by index firms over the last months months. PE is expressed as a percentage by Datastream so we convert it to a decimal by dividing by 100. We then take the log of the reciprocal to generate the log earnings-price ratio: i.e. ln (100/PE). The codes of the raw price-earnings ratio (PE) are:

  TOTMKOE(PE), TOTMKFN(PE), TOTMKSW(PE), TOTMKLX(PE), TOTMKGR(PE), TOTMKIR(PE), TOTMKPT(PE), TOTMKES(PE), TOTMKCZ(PE), TOTMKHN(PE), TOTMKPO(PE), TOTMKCP(PE), TOTMKIS(PE), and TOTMKTK(PE).

- Payout ratio (DE). This is the log dividend-price ratio minus the log earnings-price ratio.

- The risk-free rate proxy. This is taken from IMF international financial statistics. Once a country joins the Euro the German money market rate (BDI60B..) is used. The German money market rate is very highly correlated with the individual country money market rate for the period before each country joins the Euro. The raw money market rates are:

  OEI60B.., FNI60B.., SWI60B.., LXI60B.., GRI60B.., IRI60B.., PTI60B.., ESI60B.., CZI60B.., HNI60B.., POI60B.., CPI60B.., ISI60B.., TKI60B..

- Stock variance. This is standard deviation of the weekly Wednesday to Wednesday returns on the return index (RI) for each country over the past 52 weeks. This is taken for the 52 weeks up to the last Wednesday of each month.

- Ratio of risers to fallers (PRES). This is the ratio of the number of risers (RS) to the number of fallers (FS). These two series are extracted from Datastream at a monthly frequency.

  The risers codes are: TOTMKOE(RS), TOTMKFN(RS), TOTMKSW(RS), TOTMKLX(RS), TOTMKGR(RS), TOTMKIR(RS), TOTMKPT(RS), TOTMKES(RS), TOTMKCZ(RS), TOTMKHN(RS), TOTMKPO(RS), TOTMKCP(RS), and TOTMKIS(RS), TOTMKTK(RS).

  The fallers codes are: TOTMKOE(FS), TOTMKFN(FS), TOTMKSW(FS), TOTMKLX(FS), TOTMKGR(FS), TOTMKIR(FS), TOTMKPT(FS), TOTMKES(FS), TOTMKCZ(FS), TOTMKHN(FS), TOTMKPO(FS), TOTMKCP(FS), TOTMKIS(FS), and TOTMKTK(FS).

- Change in Trading Volume (CVm). This is calculated as the monthly change in trading volume (turnover) i.e. ($[VO_t / VO_{t-1}]$ -1). The total trading volume (VO) in each month is extracted.

  TOTMKOE(VO), TOTMKFN(VO), TOTMKSW(VO), TOTMKLX(VO), TOTMKGR(VO), TOTMKIR(VO), TOTMKPT(VO), TOTMKES(VO), TOTMKCZ(VO), TOTMKHN(VO), TOTMKPO(VO), TOTMKCP(VO), TOTMKIS(VO), and TOTMKTK(VO).

Table A: Out-of-sample Forecasts of Stock Returns - 1 month ahead from February 2000 to December 2006

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	0.95	-2.19	-0.45	-1.95	0.22	2.99**	-1.11	8.97***	-0.05	10.64***	4.43***	-5.72	-5.04	11.34***
DY	0.35	-0.77	1.06*	-0.89	0.78	3.19**	-0.57	9.56***	0.03	10.74***	3.36**	-5.60	-5.26	8.34***
EP	-1.14	-4.33	-5.44	-2.41	-0.88	0.51	-1.65	-0.08	-0.80	2.65**	-2.39	-1.19	-2.22	18.35***
DE	3.06**	-7.76	-4.20	-1.52	-1.08	-1.61	-12.12	6.85***	-1.99	1.79**	-0.02	-3.23	-1.44	-1.69
RF	9.97***	-0.59	5.36***	0.59	0.26	-2.55	3.12**	-0.61	0.19	4.31**	-2.59		-3.40	-79.74
SVAR	6.50***	1.41**	-1.75	-1.91	0.11	-2.42	-4.38	-1.34	0.83*	-8.59	-3.90	-0.41	-1.05	-2.79
PRES	4.18***	-0.19	2.87**	4.56***	3.64***	-0.37	6.60***	-1.20	-0.09	1.93**	-3.82		-0.22	-0.73
CVm	-0.12	-0.61							-0.19	3.10**	-6.50		0.09	
AVall	2.61**	0.28	4.58***	0.59	2.75**	2.57**	1.25*	6.56***	0.97	7.43***	0.70	-2.78	-1.48	8.07***

Table B: Out-of-sample Forecasts of Excess Stock Returns - 1 month ahead from February 2000 to June 2011

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	-3.29	-2.38	-0.51	-1.55	-0.47	-3.49	-0.64	-1.41	-0.30	0.00	-0.74	0.18	-3.28	4.08***
DY	-3.38	-1.96	0.35*	-2.24	0.21	-3.82	-1.04	-1.03	0.03	1.33**	-0.66	-0.30	-3.38	2.49**
EP	-0.63	-3.91	-3.97	-0.77	-0.70	-0.63	-1.23	-1.99	-1.13	-1.76	-2.08	1.38*	-1.49	7.26***
DE	-1.10	-5.96	-2.48	-2.74	1.89**	3.15***	-4.98	2.77***	-1.07	-0.39	-0.44	-2.37	-1.32	-3.35
RF	6.64***	0.58*	4.57***	0.68*	0.00	-1.51	5.29***	-0.82	0.69*	1.36**	-1.06		-2.37	-57.59
SVAR	-3.66	-3.43	-2.48	-3.31	-3.28	-2.41	-3.66	-2.43	-1.47	-5.67	-3.81	-11.35	-2.27	-3.37
PRES	3.80***	-0.68	2.21**	4.61***	6.20***	2.00**	7.55***	1.99**	-0.53	2.00**	-2.91		2.68**	-0.84
CVm	-1.69	-1.08							-0.27	0.03	-3.49		0.12	-0.02
AVall	2.62**	0.30	4.84***	0.95*	2.24***	1.48**	2.36**	1.88**	0.86*	2.24***	-0.68	1.30*	-0.73	4.78***

For table description, please see Table 3

Appendix C: Restrictions –percentage of forecast applied to.

	OE	FN	LX	SW	GR	IR	PT	ES	CZ	HN	PO	CP	IS	TK
DP	65.69	27.01	5.84	48.91	1.46	28.47	100.00	0.00	0.73	4.38	2.92	92.70	49.64	21.90
DY	66.42	4.38	3.65	33.58	0.00	22.63	59.12	0.73	0.73	4.38	6.57	85.40	44.53	24.09
EP	62.77	8.03	91.24	84.67	24.09	26.28	78.10	23.36	18.25	8.76	21.17	100.00	70.80	10.22
DE	82.48	33.58	23.36	42.34	18.25	44.53	88.32	27.74	26.28	0.73	41.61	12.41	35.04	4.38
RF	18.98	7.30	24.09	0.73	97.81	81.02	9.49	100.00	13.87	100.00	94.89		93.43	100.00
SVAR	13.87	28.47	37.23	8.76	10.95	29.20	18.25	10.22	14.60	94.16	100.00	25.55	95.62	73.72
PRES	32.12	92.70	32.12	17.52	20.44	14.60	30.66	5.84	13.14	1.46	99.27		3.65	87.59
CVm	27.74	12.41							0.73	0.00	47.45		0.00	100.00

Notes: This table reports the percentage of observations that the Campbell and Thompson (2008) restrictions are applied. For variable definitions please see Table 4.