Combination Forecasts of Bond and Stock Returns: An Asset Allocation Perspective

Ekaterini Panopoulou[*]   Sotiria Plastira

University of Kent   University of Piraeus

May 2014

Abstract

We investigate the out-of-sample forecasting ability of the HML, SMB, momentum, short-term and long-term reversal factors along with their size and value decompositions on U.S. bond and stock returns for a variety of horizons ranging from the short run (1 month) to the long run (2 years). Our findings suggest that these factors contain significantly more information for future bond and stock market returns than the typically employed financial variables. Combination of forecasts of the empirical factors turns out to be particularly successful, especially from an an asset allocation perspective. Similar findings pertain to the European and Japanese markets.

JEL Classification: C53; G11; G12.

Keywords: Combination forecasts; Fama French factors; Stock return predictability; Bond return predictability; Asset allocation;

[*]*Correspondence to:* Ekaterini Panopoulou, Kent Business School, University of Kent, Canterbury CT2 7PE, United Kingdom, Tel.: 0044 1227 824469, Email: A.Panopoulou@kent.ac.uk.

1 Introduction

The asset allocation decision, i.e. how much to allocate wealth in asset classes such as cash, stocks and bonds, is a key determinant of investors' portfolio performance. The importance of this decision has further been highlighted by empirical findings suggesting that stock and bond returns contain a sizeable predictable component that needs to be addressed. The degree to which bond and stock returns are predictable is a subject of ongoing debates and intensive empirical research.

The seminal contribution of Goyal and Welch (2008), who show that their long list of predictors, consisting of both macroeconomic and financial variables, can not deliver consistently superior out-of-sample performance for US stock returns, renewed the interest on stock return predictability. Contributions to this field include Campbell and Thompson (2008) who show that when imposing simple restrictions, suggested by economic theory, on predictive regressions' coefficients, the out-of-sample performance improves. The authors show that market timing strategies can deliver profits to investors (see also Ferreira and Santa-Clara (2011)). Ludvigson and Ng (2007) and Neely et al. (2013) adopt a diffusion index approach, which can conveniently track the key movements in a large set of predictors, and find evidence of improved equity premium forecasting ability.[1]

In a similar manner, various financial and macroeconomic variables are also employed to predict US government bond returns. For example, Keim and Stambaugh (1986), Fama and French (1989) and Campbell and Shiller (1991) show that yield spreads have predictive power. Cochrane and Piazzesi (2005) employed a linear combination of five forward rates and find a high degree of predictability, while Ludvigson and Ng (2009) show that the impressive predictive power, found by Cochrane and Piazzesi (2005), can be improved with five macroeconomic factors estimated from a set of 132 macroeconomic variables that measure a wide range of economic activities. More recently, Goh et al. (2013) take another route and study the predictive ability of technical indicators vis-a-vis economic variables for bond returns and find that technical indicators have both in- and out-of-sample forecasting power.

In our analysis, we also take an alternative route and investigate the forecasting ability of value, size and momentum empirical factors vis-a-vis typically employed financial variables for US bond and stock market returns. Specifically, we employ the value premium (High minus Low; HML), the size premium (Small minus Big; SMB), the momentum (Winners minus Losers over the past year; MOM), the long term

---

[1]Rapach and Zhou (2012) offer a detailed review on the issue of equity return predictability.

reversal (Winners minus Losers over the past one to five years; LT) and the short term reversal (Winners minus Losers over the past one to one month; ST). Following Fama and French (2012), we decompose the aforementioned factors into their size and value counterparts. In this way, we can disentangle the value effect on the size premium and the size effect on the remaining factors. Our paper also relates to the broad literature of forecast combinations by considering whether combinations of individual model forecasts based on the empirical factors can further improve the predictability of bond and stock returns. Rapach, Strauss and Zhou (2010) show that combination of individual financial variables forecasts improve equity premium forecasts. The authors argue that single variable forecasts cannot generate reliable forecasts over time due to parameter instability and complexity of the real economy. To this end, they show that the success of combination forecasts is attributed mainly to their link with the real economy and their ability to stabilize forecasts. In a similar manner, we also employ a variety of combination methods applied to individual empirical factors forecasts. The performance of the proposed models is assessed not only statistically, but also economically from an asset allocation perspective.

To anticipate our key results, we find that the proposed empirical factors, aggregate and decomposed, display superior forecasting ability for bond and stock market returns compared to the financial variables, not only in the U.S. market, but also in other markets, such as Europe and Japan. From an economic perspective, the empirical factors lead to significant performance fees that an investor would be willing to pay in order to have access to the information offered by the proposed factors.

The remainder of the paper is organized as follows. Section 2 describes in detail the construction of the forecasts and the corresponding statistical significance of our results. Section 3 presents the data and the empirical results concerning the forecasting ability of the empirical factors and financial variables, when employed individually or through combining methods. The asset allocation framework along with empirical results are discussed in Section 4. Section 5 reports the results of the robustness checks and Section 6 summarizes and concludes.

2   Forecast Methodology

2.1   AutoRegressive Distributed Lag (ARDL) models

Following Rapach and Weber (2004), the predictive ability of the empirical factors and financial variables is evaluated by means of the following predictive AutoRe-

gressive Distributed Lag (ARDL) model:

$$z_{t+h} = a + \sum_{i=0}^{q_1-1} \beta_i r_{t-i} + \sum_{i=0}^{q_2-1} \gamma_i x_{t-i} + \epsilon_{t+h} \tag{1}$$

where $z_{t+h} = \sum_{i=1}^{h} r_{t+i}$ is the return to be predicted from period t to $t+h$ with h the forecast horizon, r_t is the one-period return at time t, x_t the candidate predictor variable, ϵ_{t+h} the disturbance term, a the intercept, q_1 and q_2 the data-determined lag orders for r_t and x_t.[2] A heteroscedasticity and autocorrelation-consistent (HAC) covariance matrix should be employed when multi-step forecasts are concerned, i.e. $h > 1$, since cumulative returns z_{t+h} overlap and this induces serial correlation to the disturbance term (see e.g. Newey and West, 1987).

In order to study the out-of-sample forecasting ability, the total sample T is divided into the first R in-sample observations and the last P out-of-sample observations. In order to create the first out-of-sample forecast, we make use of the in-sample portion of the sample and get the estimated parameters $\widehat{a}_{1,R}$, $\widehat{\beta}_{1,R,i}$ and $\widehat{\gamma}_{1,R,i}$ of the ARDL equation via ordinary least squares (OLS) for the unrestricted form of the model. Then, the estimated equation: $\widehat{z}_{1,R+h} = \widehat{a}_{1,R} + \sum_{i=0}^{q_1-1} \widehat{\beta}_{1,R,i} r_{R-i} + \sum_{i=0}^{q_2-1} \widehat{\gamma}_{1,R,i} x_{R-i}$ creates the first out-of-sample forecast for the unrestricted form of the model, as well as, the forecast error: $\widehat{u}_{1,R+h} = z_{R+h} - \widehat{z}_{1,R+h}$.

Following the same procedure, we estimate the equation for the restricted form of the model: $\widehat{z}_{0,R+h} = \widehat{a}_{0,R} + \sum_{i=0}^{q_1-1} \widehat{\beta}_{0,R,i} r_{R-i}$, where $\widehat{a}_{0,R}$ and $\widehat{\beta}_{0,R,i}$ are the OLS parameter estimates and compute the forecast error: $\widehat{u}_{0,R+h} = z_{R+h} - \widehat{z}_{0,R+h}$. This restricted model forms the benchmark model in the forecast evaluation and we refer to it as the benchmark AR model. In order to create the next forecasts, we expand recursively the in-sample portion of the sample and repeat the whole procedure through the end of the available sample, generating $P = T - R - h + 1$ out-of-sample forecast errors for the unrestricted and the restricted form of the predictive model, $\{\widehat{u}_{1,t+h}\}_{t=R}^{T-h}$ and $\{\widehat{u}_{0,t+h}\}_{t=R}^{T-h}$, respectively.

2.2 Combination Forecasts

Combination forecasts, denoted by $\widehat{z}_{CB,t+h/t}$, are linear combinations of the n individual ARDL model forecasts, $\widehat{z}_{i,t+h}$, which are constructed by employing one factor at a time at the predictive ARDL model (Equation 1). Specifically, combination

---

[2]The maximum lag value is 8 and is selected by means of the SIC criterion.

forecasts are formed as follows:

$$\widehat{z}_{CB,t+h/t} = \sum_{i=1}^{n} w_{i,t}\widehat{z}_{i,t+h/t} \qquad (2)$$

where $\sum_{i=1}^{n} w_{i,t} = 1$. The weights, $w_{i,t}$, allocated to each of the individual forecasts are estimated by both simple and more complicated methods.

We employ three simple combination methods, namely the mean, the median and the trimmed mean one. The mean combination forecast imposes equal weights on all individual predictive models i.e., $w_{i,t} = 1/n$ $(i = 1, ..., n)$. The median combination forecast is just the sample median of $\{\widehat{z}_{i,t+h/t}\}_{t=1}^{n}$, while the trimmed mean combination forecast sets $w_{i,t} = 1/(n-2)$ for all the individual forecasts, excluding the smallest and the largest one at time t.

We also employ the discount Mean Square Forecast Error (DMSE) combining method of Stock and Watson (2004), which assigns weights based on the historical performance of the individual ARDL models, as follows:

$$w_{i,t} = m_{i,t}^{-1}/\sum_{j=1}^{n} m_{j,t}^{-1}, m_{i,t} = \sum_{s=R}^{t-h} \psi^{t-h-s}(z_{s+h} - \widehat{z}_{i,s+h/s})^2 \qquad (3)$$

where ψ is a discount factor that makes the recent forecasting accuracy of the individual ARDL models more important in the cases where $\psi < 1$. In particular, forecasts based on individual factors with lower MSFEs are given greater weights, and as such more accurate models are more important for the formation of this combination forecast. DMSE forecasts require a holdout out-of-sample period in order to calculate the weights attributed to each individual forecast. We employ the last P_0 observations of the in-sample period as the initial holdout window. The values of ψ we consider are 1.0 and 0.9.

Finally, we employ the cluster combining method, introduced by Aiolfi and Timmermann (2006). In order to create the cluster combining forecasts, we form K clusters of equal size based on the past MSFE performance with the first one being that with the lowest MSFE values. Then, the first combination forecast is the average of the ARDL model forecasts in the first cluster. This procedure begins over the initial holdout period and goes through the end of the available out-of-sample period using a rolling window. In our analysis, we consider $K = 2, 3$, leading to the CL(2) and CL(3) combination schemes.

2.3 Statistical Forecast Evaluation

The accuracy of forecasts is evaluated by the Campbell and Thompson (2008) out-of-sample R^2 (R^2_{os}) and the Clark and West (2007) CW-t statistic. The R^2_{os} statistic measures the proportional reduction in mean squared forecast error (MSFE) for the unrestricted model forecast relative to the benchmark AR specification and is defined as follows:

$$R^2_{os} = 1 - (MSFE_1/MSFE_0) \tag{4}$$

where $MSFE_1/MSFE_0$ is the ratio of the MSFE of either the individual unrestricted models or any of the combination schemes over the MSFE of the benchmark AR model. When $R^2_{os} > 0$, the forecast of the unrestricted model is more accurate than the AR model's forecast, suggesting that the candidate variable/combination scheme can improve forecasts.

In order to statistically test the ability of a candidate variable or combination scheme to improve forecasts over the benchmark model, we use the Clark and West (2007) statistic, CW-t, for equal forecasting ability. The CW-t is a modified Diebold and Mariano (1995) and West (1996) statistic and tests the null hypothesis that both the unrestricted model and the restricted one have equal MSFEs ($H_0 : R^2_{os} = 0$, i.e. $MSFE_1 = MSFE_0$) against the one-sided (upper-tail) alternative hypothesis that the MSFE of the unrestricted model is smaller than the restricted one ($H_A : R^2_{os} > 0$, i.e. $MSFE_i < MSFE_0$). The statistic can be easily calculated by first defining the following quantity:

$$\widehat{f_{t+h}} = (z_{t+h} - \widehat{z}_{0,t+h})^2 - [(z_{t+h} - \widehat{z}_{1,t+h})^2 - (\widehat{z}_{0,t+h} - \widehat{z}_{1,t+h})^2] \tag{5}$$

The first two terms in (5) are the sample MSFEs of the unrestricted and restricted models respectively, while the last term is an adjustment term that normalizes the bias produced in the MSFE by the nonzero parameters of the unrestricted model. The CW-t statistic is the t-statistic for a zero coefficient calculated by regressing $\widehat{f_{t+h}}$ on a constant and has an asymptotic distribution well approximated by the standard normal. In this respect, if the t-statistic is greater than 1.282, we reject the null hypothesis that the models have equal MSFEs at 10% level of significance (for a one-sided test). For forecast horizons greater than 1, an autocorrelation consistent standard error should be employed. (Newey and West,1987). In extensive Monte Carlo simulations, Clark and West (2007) demonstrate that the CW-t statistic performs reasonably well in terms of size and power when comparing forecasts from linear nested models.

3 Empirical Results

3.1 Data

The data used in our analysis are monthly observations for the period from July 1963 to December 2010 (570 observations). The series of interest are US long-term bond returns and stock market returns. Long-term bond returns are sourced from *Ibbotson's Stocks, Bonds, Bills and Inflation Yearbook* and stock market returns are returns on the S&P 500 index sourced from the Center for Research in Security Press (CRSP).[3]

The empirical factors employed are taken from Professor Kenneth French's website.[4] The SMB and HML factors are constructed from 6 value-weighted portfolios formed on size and book/market. Specifically, the intersections of the big/small and the value/neutral/growth portfolios form the 6 value-weighted portfolios, namely the small value (SV), small neutral (SN), small growth (SG), big value (BV), big neutral (BN) and big growth (BG) portfolio.[5] The average return of the three small portfolios minus that of the three big portfolios forms the SMB portfolio, whereas the average return of the two value portfolios minus the average return of the two growth portfolios forms the HML portfolio. The ST, LT and MOM factors are formed from 6 value-weighted portfolios formed on size and prior returns (small low, small medium, small high, big low, big medium, and big high). These prior-return portfolios are constructed on prior (1-1), (13-60), and (2-12) returns, respectively.[6] The average return on the two low prior-return portfolios (big and small) minus the average return on the two high prior-return portfolios (big and small) forms the ST and LT factors, while the MOM factor is the average of the returns on the two high prior-return portfolios (big and small) minus the average return on the two low prior-return portfolios (big and small).

Following Fama and French (2012), we decompose all the factors (except for SMB) into their small and big counterparts. For example, the difference between the small (big) value portfolio and the small (big) growth one forms the HML_s

---

[3]Both series are available at Prof. Goyal's website at: http://www.hec.unil.ch/agoyal/.

[4]Tha data are downloadable at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[5]The breakpoint for year t for size is the median NYSE market equity at the end of June of year t, while for the book/market is the 30th and 70th NYSE percentile. The book/market ratio for June of year t is the book equity for the last fiscal year end in t-1 divided by market equity for December of t-1. The portfolios for July of year t to June of t+1 include all NYSE, AMEX, and NASDAQ stocks for which we have market equity data for December of t-1 and June of t, and (positive) book equity data for t-1.

[6]The breakpoint for the equity is the median NYSE market equity, while for the prior returns is the 30th and 70th NYSE percentile.

(HML_b) portfolio, as follows:

$$HML\_s = SV - SG$$
$$HML\_b = BV - BG \tag{6}$$

Decompositions of the LT, ST and MOM factors are formed according to the following formulas:

$$LT\_s = SL - SH, \quad LT\_b = BL - BH$$
$$ST\_s = SL - SH, \quad ST\_b = BL - BH \tag{7}$$
$$MOM\_s = SH - SL, \quad MOM\_b = BH - BL$$

For the SMB factor, we construct a value decomposition. Specifically, we decompose the size premium into its value, neutral and growth components, denoted by SMB_v, SMB_n and SMB_g, respectively, and calculated as follows:

$$SMB\_v = SV - BV$$
$$SMB\_n = SN - BN$$
$$SMB\_g = SG - BG \tag{8}$$

In addition, we employ fourteen financial variables, which have been shown in the literature to exhibit predictive ability on returns. The data for the financial variables, which are used by Rapach and Zhou (2012), are described in detail by Goyal and Welch (2008)[7]. These are the dividend/price ratio (DP), dividend yield (DY), earnings/price ratio (EP), dividend/earnings ratio (DE), stock variance (SVAR), book/market ratio (BM), net equity expansion (NTIS), treasury bill rate (TBL), long-term yield (LTY), long-term government bond returns (LTR),term spread (TMS), the default yield spread (DFY), default return spread (DFR), stock market return (SP500),and the inflation rate (INF).[8]

The total sample of the 570 monthly observations is divided into the estimation period consisting of the first $R = 380$ in-sample observations (July 1963 to February 1995) and the evaluation period with the last $P = 190$ (corresponding to the 1/3 of our sample) out-of-sample observations (March 1995 to December 2010). The holdout period for the combining methods that require one is set to 7 years (84 months) prior to the start of the out-of-sample evaluation period.

---

[7]This set of data can be downloaded from htts://www.hec.unil.ch/AGoyal.

[8]Please refer to Goyal and Welch (2008) for details on the construction and the sources of the series.

3.2 Forecasting U.S. Bond Returns

We begin our analysis by evaluating the forecasting ability of the candidate predictors for US long-term government bond returns for horizons ranging from 1 to 24 months. Table 1 reports the R^2_{os} associated with individual ARDL models for both the empirical factors (Panel A) and the financial/macroeconomic variables (Panel B). Bold indicates a statistically superior forecast relative to the benchmark AR(1) model on the basis of the CW-t statistic. As is evident, the momentum and short term reversal factors display significant predictive ability for a variety of horizons, while the forecasting ability of the value premium, the size premium and the long-term reversal is rather muted. Specifically, momentum displays significant predictive ability at horizons of 1-3, 6, and 12 months, while short-term reversal for horizons less than 3 months. Examining closely the performance of the size decompositions of the momentum factor, we note that the whole information is attributed to its small component, which appears to be a significant predictor for horizons ranging from 1 month to 1 year. On the other hand, the momentum of big companies improves bond forecasts in 1- and 6-month horizons ahead. Similarly, the small component of short-term reversal emerges as a significant predictor for all the horizons considered with the exception of the 6-month one. With respect to the big component of the short term reversal, its predictability appears at horizons of 2 and 3 months ahead.

[TABLE 1 AROUND HERE]

Turning to the predictive ability of the financial variables employed (Panel B), we observe that only the stock market return improves bond return forecasts only the stock market return improves bond return forecasts, performance which is evident only in the short run, at horizons of 1-3 months. Quite interestingly, the remaining financial variables exhibit hardly any significant predictive ability.

We next examine whether combining individual forecasts can result in superior predictive ability. We consider forecast combinations of (i) the five aggregate empirical factors (HML, SMB, MOM, LT and ST), reported in Panel A of Table 2, (ii) the eleven decomposed factors, reported in Panel B of Table 2 and (iii) the 14 financial variables, reported in Panel C of Table 2. As already discussed in Section 2, we employ the mean, median, trimmed mean, DMSE and cluster combining methods. For the DMSE, we employ two discount factors of $\psi = 0.90$ (DMSE(0.9)) and $\psi = 1.00$ (DMSE(1)), while for the cluster combining method, we employ 2 clusters (CL(2)) and 3 clusters (CL(3)). We observe that when the five aggregate

(HML, SMB, MOM, LT, and ST) factors are combined, out-of-sample predictive ability appears only short term. Specifically, at horizons of 1-3 months, the median, trimmed mean and CL(2) combining method forecasts display significant forecasting ability, while the mean, DMSE(1), DMSE(0.9) and CL(3) combining methods exhibit predictive ability at horizons of 1 and 3 months.

[TABLE 2 AROUND HERE]

Our findings for the combination forecasts constructed with the decomposed factors (reported in Panel B) are quite interesting since predictability reaches the 18-month horizon. More in detail, the median and trimmed mean combining methods improve forecasts at horizons of 3 to 18 months and for 1-3, 6, 12, and 18 months, respectively. CL(2) exhibits significant forecasting ability on bond returns at horizons of 1, 2, and 9 months, while CL(3) improves forecasts for 1-3 months and 9 months ahead. Moreover, the mean and DMSE(0.9) combining schemes are associated with superior predictive ability only at the 1-month horizon, while the DMSE(1) one for horizons of 1 and 12 months. More importantly, there are no benefits associated with combination forecasts of financial variables, as suggested by Panel C of Table 2.

Overall, our findings so far suggest that combining empirical factors can lead to improved predictability for bond returns and that size and value decompositions of the empirical factors can further enhance it. This latter finding suggests that the disaggregated factors contain significant information for the evolution of future bond returns which is rather hidden when considering aggregate factors.[9]

3.3 Forecasting US stock returns

We now examine whether the forecasting ability of the candidate predictors is maintained for US stock returns (S&P500 index returns). Panel A (Table 3) reports the out-of-sample performance of the empirical factors and their components, while Panel B reports the related findings for the financial variables.

[TABLE 3 AROUND HERE]

Among the 30 candidate predictors, the momentum factor emerges as the most powerful one, as it improves forecasts over the AR benchmark at horizons of 6,

---

[9]Unreported results suggest that combinations of both the empirical factors and the financial variables fail to improve the accuracy of forecasts relative to the performance of the AR model. This set of results are available from the authors upon request.

9, 12, and 18 months. This performance is consistent with the one for bond returns and is mainly attributed to the momentum of big companies. Moreover, at a 3-month horizon, the long-term reversal factor along with both its components displays significant predictive ability. Turning to the financial variables, we have to note that their ability is rather weak and limited to horizons of 6-12 months and 18-24 months for the book to market ratio and the term spread, respectively.

Given the rather limited individual variable predictability, we do not expect combination methods to work impressively well, since they aggregate over weak predictors. Our findings, reported in Table 4 (Panels A to C), support this conjecture. Specifically, when considering combination forecasts of the five empirical factors, we find improved forecasting ability only at the 3-month horizon and on the basis of the mean, DMSE and CL(3) combination schemes. Similar findings pertain when the decomposed factors are considered (Panel B), since significant forecasting ability is evident for the trimmed mean combining method at the horizon of 3 months, as well. Finally, as expected, combination forecasts of financial variables do not improve stock returns forecasts over the AR benchmark.

[TABLE 4 AROUND HERE]

To sum up, the evidence in this section suggests that the proposed empirical factors exhibit strong forecasting ability for US bond returns and are weaker when it comes to stock returns. Their size and value decompositions further enhance their ability especially when combination of forecasts are considered. Given that statistical significance does not always imply economic significance, we next assess whether this forecasting ability can be useful from an asset allocation perspective.

4   Asset allocation benefits of combination forecasts

A utility-based evaluation of forecasts was first proposed by West et al. (1993) in assessing exchange rate volatility forecasts (see also Abhyankar et al. (2005) and Della Corte et al. (2009), Rime et al. (2010)). Following Fleming et al. (2001) and Della Corte et al. (2008, 2009), Thorton and Valente (2012) quantify how much a risk-averse investor is willing to pay to switch from a dynamic portfolio strategy based on a model with no predictable bond excess returns to a model that uses either forward spreads or the term structure of forward rates. Campbell and Thomson (2008), Rapach et al. (2010), Ferreira and Santa-Clara (2011), Dangl and

Halling (2012) and Neely et al. (2013) provide evidence that investors who rely on equity premium forecasts based on economic variables can gain profit relative to those who just rely on the historical average forecast.

In our analysis, we investigate whether the forecasting ability of the proposed empirical factors/combination schemes can lead to significant economic gains for a mean-variance investor, who incorporates them to asset allocation decisions.

4.1 The Framework

We consider a mean-variance investor with relative risk aversion (RRA), γ, who re-balances her portfolio every month. Her portfolio maximization problem, described in detail in Campbell and Viceira (2002), is the following:

$$\max_{w_t} \mathbf{w}_t'(E_t\mathbf{R}_{t+h} - R_{f,t\to t+h}\boldsymbol{\iota}) - \frac{\gamma}{2}\mathbf{w}_t'\boldsymbol{\Sigma}_{t+h}^{-1}\mathbf{w}_t \tag{9}$$

where $E_t\mathbf{R}_{t+h} - R_{f,t\to t+h}$ is the vector of expected excess returns on the risky assets over the risk-free interest rate $(R_{f,t\to t+h})$ prevailing from time t to $t+h$, $\boldsymbol{\iota}$ is a vector of ones, $\mathbf{w}_t$ is the vector of portfolio weights on risky assets, and $\mathbf{w}_t'\boldsymbol{\Sigma}_{t+h}^{-1}\mathbf{w}_t$ is the expected variance of the portfolio return. The solution to this maximization problem is:

$$w_{i,t} = \frac{1}{\gamma}\boldsymbol{\Sigma}_{t+h}^{-1}(E_t\mathbf{R}_{t+1} - R_{f,t}\boldsymbol{\iota}), i = b, s \tag{10}$$

where b, s stand for bond and stock returns, respectively.

The conditional expectation $E_t\mathbf{R}_{t+h}$ is given by the bond and stock return combination forecasts for each horizon and combining scheme we employed in the previous section. The expected variance/covariance matrix for bond and stock market returns, $\boldsymbol{\Sigma}_{t+h}$, is computed using a rolling window of 40 past observations.[10] The optimal weights allocated to government bonds and the stock market are winsorized to $0 < w_{i,t} < 1.5$, thus preventing short selling and extreme allocation to any of the risky assets. The investor's taste of risk, controlled by the RRA coefficient, is set equal to 3 and 5. Having estimated the optimal weights, the resulting portfolio return is equal to:

$$R_{p,t} = (1 - w_{1,t} - w_{2,t}) * R_{f,t} + w_{1,t} * R_{b,t} + w_{2,t} * R_{s,t} \tag{11}$$

where $R_{b,t}$ and $R_{s,t}$ are the realized bond and stock returns at each point of time, t, over the out-of-sample evaluation period (P observations). Over the forecast

---

[10]Campbell and Thomson (2008) and Goh et al. (2013) consider a 5-year rolling window of past returns.

evaluation period the investor with initial wealth of $W_o = 1$ realizes an average utility of

$$\overline{U} = \frac{1}{P} \sum_{t=1}^{P} \left[(1 + R_{p,t}) - \frac{\gamma}{2} * \left(R_{p,t} - \overline{R}_p \right)^2 \right] \tag{12}$$

where $\overline{R}_p$ denotes the average portfolio return over the evaluation period. In a similar way, we calculate the utility associated with the benchmark AR specification, given by the following equation:

$$\overline{U}^{AR} = \frac{1}{P} \sum_{t=1}^{P} \left[\left(1 + R_{p,t}^{AR}\right) - \frac{\gamma}{2} * \left(R_{p,t}^{AR} - \overline{R}_p^{AR} \right)^2 \right]$$

where $R_{p,t}^{AR}$ refers to the portfolio returns constructed based on the benchmark model forecasts and $\overline{R}_p^{AR}$ is the respective average portfolio return over the evaluation period. The difference (ΔU) between the average utility realized from the proposed specification and the one of the benchmark specification is calculated as follows:

$$\Delta U = \overline{U} - \overline{U}^{AR} \tag{13}$$

It can be interpreted as the annual percentage portfolio management fee that an investor would be willing to pay to have access to our proposed forecasting methodology relative to the AR benchmark.

We also employ an alternative economic evaluation measure, which is the manipulation-proof performance measure (MPPM), proposed by Goetzmann et al. (2007). This measure takes into account the effect of non-normality, the underestimation of the performance of dynamic strategies and the choice of the utility function. It can be interpreted as the portfolio's premium return after adjusting for risk and is defined as follows:

$$MPPM = \frac{1}{1 - \gamma} \ln \left[\frac{1}{P} \sum_{t=1}^{P} \left(\frac{1 + R_{p,t}}{1 + R_{f,t}} \right)^{1-\gamma} \right] \tag{14}$$

The proposed specification performs better than the benchmark one when the difference between the $MPPM$ of the proposed model and that of the benchmark one, Θ, defined as follows:

$$\Theta = \frac{1}{1 - \gamma} \ln \left[\frac{1}{P} \sum_{t=1}^{P} \left(\frac{1 + R_{p,t}}{1 + R_{f,t}} \right)^{1-\gamma} \right] - \frac{1}{1 - \gamma} \ln \left[\frac{1}{P} \sum_{t=1}^{P} \left(\frac{1 + R_{p,t}^{AR}}{1 + R_{f,t}} \right)^{1-\gamma} \right] \tag{15}$$

is positive.

4.2 Asset Allocation: Empirical Results

We consider a mean-variance investor who allocates her wealth among bonds, stocks and the risk-free interest rate, and rebalances her portfolio monthly over the 1995:03 - 2010:12 out-of-sample evaluation period.[11] As already mentioned, we assume two values for the investor's RRA, $\gamma = 3$ and $\gamma = 5$, and calculate the variance covariance matrix between stocks and bond returns by employing a rolling 40-month window of past observations. Consistent with the statistical evaluation, we assess the economic value for horizons of 1, 3, 6 and 12 months.

Table 5 (Panels A to D) reports the performance fees (ΔU) that a mean-variance investor would be willing to pay to have access to our models along with the risk-adjusted measure Θ for an investor with a risk aversion coefficient of 3. Our findings for an investment horizon of 1 month are given in Panel A. Overall, combination forecasts of both the aggregate factors and the disaggregated ones always generate positive utility gains. Utility gains range from 0.765% per year (Trimmed mean combination forecast of the disaggregate factors) to 3.029% per year (CL(3) combination method of the disaggregate factors). The best performance is achieved by the CL(3) combination method, closely followed by CL(2). However, the simplest combining method, i.e. the mean one, attains a satisfactory performance of 1.348% and 1.417% for the aggregate and disaggregated factors, respectively. Similar findings pertain when forecasts are evaluated on the basis of the risk-adjusted measure Θ. More importantly, combination forecasts of financial variables fail to generate profits to the investor in excess of the ones already contained in the benchmark AR model, with the exception of the CL(2) method.

[TABLE 5 AROUND HERE]

Turning to the forecast horizon of 3 months (Panel B), our findings suggest that combination forecasts of either the five or the eleven factors can generate positive utility gains that reach 6.414% for the CL(3) method, with the exception of the median combining scheme of the disaggregate factors. When combining the aggregate factors, the cluster combining methods rank first followed by the mean and the DMSE ones. However, on the basis of the disaggregated factors, the mean and DMSE methods rank first followed by the trimmed mean and the cluster ones. Our findings with respect to Θ, are quite similar. Moreover, similar to the 1-month forecast horizon, all the combining methods (with the exception of the median one) point to negative gains and thus greater average utility for the AR benchmark

---

compared to the combination methods.

Longer investment horizons of 6 months (Panel C) and 12 months (Panel D) do not consistently generate profits to the investor. Specifically, for the 6-month horizon an investor would be willing to have access to the forecasts generated by the cluster combinations of the five factors or the median and trimmed mean combinations of the eleven factors. The difference in MPPMs, Θ, points to benefits when a pool of the financial variables is employed. Specifically, Θ is positive at the horizon of 6 months for the median and the CL combination methods generating premium returns of up to 3.265% per year. Turning to the 12-month horizon, we note that the ability of the proposed models to generate utility gains to an investor is rather limited to the case of the CL(2) combination of the disaggregated factors.

When we allow for a more conservative investor, our findings are qualitatively similar. More in detail, Table 6 reports the respective findings for an investor with RRA of 5. For a short-term horizon of 1 and 3 months (Panels A and B), the investor would be willing to pay a performance fee to utilize forecasts from our combining methods on the basis of the empirical factors (both aggregate and disaggregate ones). As expected, these fees are lower compared to the ones for the less risk averse investor (Table 5). On the other hand, when turning to the medium investment horizon of 6 months (Panel C), the investor can still benefit from our combination forecasts of the empirical factors and in some cases of the financial variables, as well. The combination methods of aggregate factors are all successful and generate fees up to 1.217% (CL(2) method), whereas when disaggregate factors are considered, all but the cluster combining methods accrue benefits of up to 1.144% to the investor. More importantly, median and cluster combinations of the financial variables can generate positive utility gains of 2.842%. Employing Θ leads to similar findings for the financial variables pool but not for the factor ones. In some cases positive utility gains are associated with negative Θs. The opposite is true for the longer horizon of 12 months and the case of the pool of financial variables. Specifically, while positive utility gains and Θs are associated with the median and CL(2) methods, positive Θs prevail for all the combination methods at hand.

[TABLE 6 AROUND HERE]

5 International Evidence

So far we have provided evidence for significant forecasting ability of combination forecasts of empirical factors for US bond and stock returns both in statistical and

economic evaluation terms. In this section, we test whether these factors exhibit similar forecasting ability on European and Japanese stock and bond returns.

We use monthly observations of the empirical factors for the period November 1990 to April 2012.[12] European and Japanese market returns along with the aggregate factor returns and their decompositions are taken from Professor Kenneth French's website.[13] Long-term bond returns are downloaded from DataStream.[14] The total sample consists of 258 observations, 86 are reserved for the out-of-sample evaluation period.[15] The horizons examined are 1-24 months, but, for brevity, we present the results for horizons of 1-3, 6, 9, 12, 18 and 24 months. As previously, we assess the out-of-sample forecasting ability of the proposed models from an asset allocation perspective as well. In particular, we consider a mean-variance investor who allocates her wealth among bonds, stocks and the risk-free interest rate and rebalances her portfolio monthly based on information through period t over the 2005:03 - 2012:04 out-of-sample evaluation period. The investor's relative risk aversion (RRA) is set equal to $\gamma = 3$.

5.1   Forecasting European Bond and Stock Returns

Table 7 (Panels A to C) reports the forecasting ability of empirical aggregate and decomposed factors for European bond returns along with combinations of them. The only factor that appears valuable in forecasting bond returns is the HML factor who is significant both in the short run and in the long run. Specifically, the aggregate value premium is a significant predictor at horizons of 1-3, 18 and 24 months, while its big component is successful at horizons of 2, 3, 6 and 24 months. On the other hand, the predictive ability of the small component (HML_s) is restricted only at the horizon of 1 month. The remaining factors exhibit hardly any significant forecasting ability on bond returns.

[TABLE 7 AROUND HERE]

Turning to combination forecasts of the aggregate factors (Panel B), we have to note that our findings vary with the combination method employed. Specifically, the

---

[12]All returns are given in U.S. dollars.

[13]The European factors and portfolios include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom.

[14]The series concerning the long-term government bond returns for Europe is the series: BOFA ML PAN EUROPE GVT 10+Y ($) - TOT RETURN IND, while for Japan is the series:BOFA ML JAPAN GVT 10+Y ($) - TOT RETURN IND.

[15]The holdout period is 3 years (36 months).

median and trimmed mean combining methods display significant forecasting ability at horizons of 3 and 9 months, while the DMSE combining methods outperform the AR model at horizons of 18 and 24 months (DMSE(1)) and at the horizon of 2 years (DMSE (0.9)). Both cluster combining methods are associated with short run predictability of 1 month. Quite interestingly and in sharp contrast with the US market, combinations of the decomposed factors completely fail to outperform the autoregressive benchmark.

We continue by examining the level of predictability for the European stock market, which is reported in Table 10 (Panels A to C). Our findings suggest that the SMB factor is the dominant predictor with significant forecasting ability at horizons of 2, 6, 9 and 18 months. This performance is attributed partly to the neutral and growth component of the factor. The growth component of the size premium appears to contain useful information for 3-, 6-, 9- and 18-month future returns, while the neutral component for 1, 2, 6 and 18 months ahead. In addition, momentum along with its small and big decompositions contain useful information for the European market at the horizon of 3 months.

[TABLE 8 AROUND HERE]

Similar to European bond returns, combination forecasts do not appear very successful. When we combine the individual forecasts of the aggregate factors, both the median and trimmed mean combing methods exhibit significant forecasting ability at the horizon of 3 months, while CL(3) improves forecasts at horizons of 3 and 18 months. Considering the forecasts of combinations of decomposed factors, both cluster combining methods appear significant at horizons of 3 and 18 months, while the median combining method exhibits forecasting ability at the horizon of 2 months.

Our asset allocation exercise paints a starkly different picture. Despite the anaemic statistical significance of combination forecasts of both stock and bond returns, the gains for a european investor can be sizable. Table 9 reports the average utility gains of a mean-variance investor who allocates her wealth between stock, bonds and the risk free interest rate along with the manipulation-proof measure of the competing models for horizons up to 1 year. For a short term horizon of 1-month, mean and DMSE combinations of aggregate factors can lead to utility gains of up to 4.622%. Increasing the horizon to 3 months can lead to gains of 8.865% for the CL(3) method. This horizon is also associated with significant profitability of up to 10.707% on the basis of the CL(3) combinations of disaggregated factors. However, longer horizons of 6 and 12 months do not consistently generate profits to the investor.

17

[TABLE 9 AROUND HERE]

5.2 Forecasting the Japanese Bond and Stock Market

In this section, we investigate the robustness of our results employing data for the Japanese bond and stock market. Tables 10 and 11 report the forecasting performance of empirical factors for bond and stock returns, respectively. With respect to bond returns our findings suggest that single factor models prove successful at short horizons. Specifically, at the 1-month horizon both the size and the value premium improve bond return forecasts, mainly stemming from the small value component and the growth and neutral size component, respectively. Additionaly, the small value component improves forecasts for the 2-month horizon as well, while the neutral value component for the 3-month and 6-month horizons. Our findings with respect to combination forecasts are more reassuring. Specifically, with the exception of the 24-month horizon, the remaining horizons are characterized with a high degree of predictability. The 1-month and 3-month bond returns can be predicted with almost all the methods at hand and on the basis of both the aggregate and decomposed factors. Overall, combinations of the aggregate factors perform better than the decomposed ones.

Similar findings pertain with respect to stock returns where the level of predictability is higher. The value premium is successful in improving forecasts for all the horizons up to the 9-month one. This forecasting ability is equally split between its big and small component which contains useful information for the long-run as well. Quite interestingly, the value and growth decompositions of the size premium emerge as powerful predictors for horizons greater than 18 and 24 months. As expected, this individual forecasting ability is recorded in the success of forecast combinations. On the basis of forecast combinations of aggregate factors, the mean, DMSE and cluster combining methods improve forecasts for the majority of horizons considered. Quite interestingly, the 24-month horizon is associated with a high degree of predictability of combination methods of both aggregate and decomposed factors.

[TABLES 10 & 11 AROUND HERE]

Finally, the most striking result appears in Table 12 that reports the forecast combination benefits from an asset allocation perspective. Specifically, a mean-variance investor who employs our forecast combination methodology can always

enjoy significant gains for all the horizons up to 6 months. The benefits accrued by including the information of the aggregate factors reach 3.067% at the 1-month horizon and increase to 5.990% at the horizon of 3 months, while they can even exceed 20% for the 6-month horizon. The 1-year horizon is associated with benefits reaching 10.479% for combinations of the aggregate factors and exceed 20% for combination forecasts of the decomposed ones. Our findings with respect to the MPPMs of the respective portfolios are fully consistent with the ones of utility gains.

[TABLE 12 AROUND HERE]

6 Conclusions

In this paper, we study the forecasting ability of empirical factors; namely, the value premium (HML), the size premium (SMB) and the momentum factors (MOM, LT and ST) along with widely employed financial variables on U.S. bond and stock returns. One of our contributions to the literature consists of the decomposition of these empirical factors to their size and value components, investigating thus the size effect on the value and momentum premium and the value effect on the size premium. Our findings suggest that these empirical factors contain significantly more information for future bond and stock market returns than the typically employed financial variables, but the extent to which this forecasting ability appears differs.

To address the instability and time-variability of individual forecasts, we go one step further and combine them by employing a variety of combination methods. Specifically, we construct forecasts on the basis of three simple combining methods; namely, the mean, median and trimmed mean and two more advanced ones; the Discount Mean Square forecast Error (DMSE) combining method, which is based on the historical performance of the individual models, and the Cluster Combining method (CL), which is based on equal-sized clusters related to past forecasting performance. The forecasting ability of combination forecasts is assessed not only statistically, by means of the R^2_{os} statistic, which measures the improvement of the MSFE of the proposed model over the MSFE of the benchmark AR model forecast, but also economically by computing the performance fee that investors would be willing to pay to have access to our methodology. In addition, we calculate the risk-adjusted portfolio's premium return (manipulation-proof performance measure) in order to assess the most valuable model among the competing ones.

Our results provide evidence that combination forecasts based on decomposed factors display superior forecasting ability relative to the forecasts based on typically employed financial variables at horizons ranging from the short run to the long run. This performance is also evident from an asset allocation perspective. In particular, investors can accrue positive utility gains by employing trading strategies based on forecasts produced by the empirical factors, irrespective of the degree of relative risk aversion and borrowing constraints. Finally, the robustness of our results is assessed by conducting the same tests for markets outside the US. By employing data for the European and Japanese bond and stock market, we find that the forecasting ability of combination forecasts formed on the basis of the empirical factors is rather pervasive in these markets, as well.

References

[1] Abhyankar A., Sarno L., and G. Valente (2005), 'Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability', Journal of International Economics, 66(2), 325-348

[2] Aiolfi M., and A. Timmermann (2006), 'Persistence in Forecasting Performance and Conditional Combination Strategies', Journal of Econometrics, 135(1-2), 31-53

[3] Campbell, J. Y., and R. J. Shiller (1991), 'Yield Spreads and Interest Rates: A Bird's Eye View', Review of Economic Studies, 58(3), 495–514

[4] Campbell J. Y., and S. B. Thomson (2008), 'Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?', Review of Financial Studies, 21(4), 1509-1531

[5] Campbell J.Y., and L. M. Viceira (2002), 'Strategic Asset Allocation', Oxford: Oxford University Press

[6] Clark E. T., and D. K. West (2007), 'Approximately Normal Tests for Equal Predictive Accuracy in Nested Models', Journal of Econometrics, 138(1), 291-311

[7] Cochrane, J. H., and M. Piazzesi (2005), 'Bond Risk Premia', American Economic Review, 95(1), 138–160

[8] Dangl T., and M. Halling (2012), 'Predictive Regressions with Time-Varying Coefficients', Journal of Financial Economics, 106(1), 157-181

[9] Della Corte P., Sarno L., and D. L. Thorton (2008), 'The Expectations Hypothesis of the Term Structure of Very Short-Term Rates: Statistical Tests and Economic Value', Journal of Financial Economics, 89(1), 158-174

[10] Della Corte P., Sarno L., and I. Tsiakas (2009), 'An Economic Evaluation of Empirical Exchange Rate Models', Review of Financial Studies, 22(9), 3491-3530

[11] Diebond X. F., and S. R. Mariano (1995), 'Comparing Predictive Accuracy', Journal of Business & Economic Statistics, 13(3), 253-263

[12] Fama, E. F., and K. R. French (1989), 'Business Conditions and Expected Returns on Stocks and Bonds', Journal of Financial Economics, 25(1), 23–49

[13] Fama F. E., and R. K. French (2012), 'Size, Value and Momentum in International Stock Returns', Journal of Financial Economics, 105(3), 457-472

[14] Ferreira M. I., and P. Santa-Clara (2011), 'Forecasting Stock Market Returns: The Sum of the Parts os More than the Whole', Journal of Financial Economics, 100(3), 514-537

[15] Fleming J., Kirby C., and B. Ostdiek (2001), 'The Economic Value of Volatility Timing Using Realized Volatility', Journal of Financial Economics, 67(3), 473-509

[16] Goetzmann W., Ingersoll J., Spiegel M., and I. Welch (2007), 'Portfolio Performance Manipulation and Manipulation-proof Performance Measures', Review of Financial Studies, 20(5), 1503-1546

[17] Goh J., Jiang F., Tu J., and G. Zhou (2013), 'Forecasting Bond Risk Premia Using Technical Indicators', available http://ssrn.com/abstract=1914227

[18] Goyal A., and I. Welch (2008), 'A Comprehensive Look at The Empirical Performance of Equity Premium Prediction', Review of Financial Studies, 21(4), 1455-1508

[19] Keim, D. B., and R. F. Stambaugh (1986), 'Predicting Returns in the Stock and Bond Markets' Journal of Financial Economics, 17(2), 357–390

[20] Ludvigson S. C., and S. Ng (2007), 'The Empirical Risk-Return Relation: A factor Analysis Approach', Journal of Financial Econometrics, 83(1), 171-222

[21] Ludvigson S. C., and S. Ng (2009), 'Macro Factors in Bond Risk Premia', Review of Financial Studies, 22(12), 5027-5067

[22] Neely C., Rapach E. D., Tu J., and G. Zhou (2013), 'Forecasting the Equity Risk Premium: The Role of Technical Indicators', Management Science, forthcoming.

[23] Newey K. W., and D. K. West (1987), 'A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix', Econometrica, 55(3), 703-708

[24] Rapach E. D., Strauss K. J., and G. Zhou (2010), 'Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy', Review of Financial Studies, 23(2), 821-862

[25] Rapach E. D., and E. C. Weber (2004), 'Financial Variables and the Simulated Out-of-Sample Forecastability of U.S. Output Growth Since 1985: An Encompassing Approach', Economic Inquiry, 42(4), 717-738

[26] Rapach E. D., and G. Zhou (2012), 'Forecasting Stock Returns', Forthcoming in the Handbook of Economic Forecasting, Volume 2

[27] Rime D., Sarno L., and E. Sojli (2010), 'Exchange Rate Forecasting, Order Flow and Macroeconomic Information', Journal of International Economics, 80(1), 72-88

[28] Stock H. J., and W. M. Watson (2004), 'Combination Forecasts of Output Growth in a Seven-Country Data Set', Journal of Forecasting, 23(6), 405-430

[29] Thorton D. L., and G. Valente (2012), 'Out of Sample Predictions of Bond Excess Returns and Forward Rates: An Asset Allocation Perspective', Review of Financial Studies, 25(10), 3141-3168

[30] West D. K. (1996), 'Asymptotic Inference About Predictive Ability', Econometrica, 64(5), 1067-1084

[31] West D. K., Edison H. J., and D. Cho (1993), 'A Utility-Based Comparison of Some Models of Exchange Rate Volatility', Journal of International Economics, 35(1-2), 23-45

Table 1. Out-of-sample performance of individual ARDL models -US bond returns

Panel A. Empirical factors								
Predictor/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
HML	0.103	-0.710	-0.705	-1.625	-4.572	-7.880	-6.604	-2.450
SMB	-0.013	-3.076	-1.663	-4.906	-5.405	-7.139	-10.113	-19.116
MOM	**1.328**	**1.169**	**1.238**	**2.084**	0.865	**1.114**	-2.744	-3.021
LT	-0.295	-0.804	-1.460	-1.815	-0.106	-0.734	-0.430	-0.664
ST	**1.446**	**2.374**	**3.255**	-0.822	-0.818	-0.278	0.423	0.362
HML_b	-0.340	-0.577	-0.517	-0.258	-0.248	-1.775	-0.410	-0.936
HML_s	-0.377	-0.583	-0.669	-0.919	-0.613	-15.025	-10.867	-16.900
SMB_g	-0.179	-5.168	-4.604	-4.303	-5.374	**0.031**	-2.470	-9.054
SMB_n	0.091	-3.820	-4.534	-1.452	-2.533	-1.593	0.137	-0.378
SMB_v	-0.323	-0.167	-4.660	-2.279	-5.076	-2.469	-1.776	-4.679
MOM_b	**0.533**	0.144	0.444	**1.608**	-0.132	-13.215	-48.097	-8.835
MOM_ s	**1.930**	**2.351**	**1.968**	**1.835**	**1.184**	**1.796**	-0.211	0.261
LT_b	-0.131	-0.157	-0.273	-0.420	-0.494	0.178	-0.181	0.189
LT_s	-1.354	-3.479	-4.959	-4.416	-0.031	-2.565	-0.614	-3.516
ST_b	0.592	**1.433**	**1.718**	-1.259	-1.410	-1.302	-0.205	-0.233
ST_s	**1.995**	**2.694**	**4.042**	-1.516	**0.858**	**2.044**	**1.527**	**1.577**

Panel B. Financial variables								
Predictor/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
DP	-2.899	-1.851	-6.431	-19.903	<-20	<-20	<-20	<-20
DY	-2.749	-5.354	-7.519	-16.785	<-20	<-20	<-20	<-20
EP	-0.845	-5.853	-7.692	-3.765	-7.797	-12.603	<-20	<-20
DE	-18.583	-43.518	-2.228	-2.225	-6.869	-4.849	-4.171	-7.814
SVAR	-0.031	0.702	-1.837	-2.117	-3.302	-5.249	-8.999	<-20
BM	-0.311	-1.780	-2.310	-1.443	-1.569	-0.246	-0.110	-16.653
NTIS	-2.164	-4.310	-7.041	-10.494	-18.703	<-20	<-20	-16.078
TBL	-1.025	-2.062	-3.458	-7.604	-14.095	<-20	<-20	<-20
LTY	-0.953	-2.229	-5.124	-12.607	<-20	<-20	<-20	<-20
SP500	**2.662**	**3.186**	**2.010**	-0.413	-0.430	0.084	-1.470	-0.951
TMS	-0.153	-0.003	-0.734	-0.844	-4.128	-7.831	-14.524	<-20
DFY	-3.552	-4.332	-7.326	-16.942	<-20	<-20	<-20	<-20
DFR	-1.881	-2.088	-4.054	-10.076	-19.944	<-20	<-20	<-20
INFL	0.444	-0.526	-0.635	-0.931	-2.984	-3.227	-0.827	-0.136

Notes: The table reports the out-of-sample R^2_{os} of the individual ARDL models relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$ statistic, which tests the null hypothesis: $R^2_{os}=0$ against the alternative: $R^2_{os}>0$.

Table 2. Out-of-sample performance of combination methods -U.S. bond returns

Panel A. Five empirical factors

Method/ Horizon	1	2	3	6	9	12	18	24
Mean	**0.824**	0.543	**1.188**	-0.341	-0.768	-0.839	-2.030	-2.539
Median	**0.934**	**0.667**	**1.102**	-0.227	-0.273	-0.131	-1.650	-2.199
Trimmed mean	**0.844**	**0.540**	**1.184**	-0.008	-0.432	-0.577	-1.911	-1.606
DMSE(1)	**0.822**	0.548	**1.190**	-0.323	-0.758	-0.840	-2.022	-2.514
DMSE(0.9)	**0.827**	0.567	**1.179**	-0.366	-0.791	-0.917	-1.973	-2.152
CL(2)	**0.974**	**1.217**	**1.729**	0.269	-0.320	-1.207	-1.548	-0.896
CL(3)	**1.053**	0.381	**1.235**	0.097	-1.931	-3.067	-2.296	-0.871

Panel B. Eleven decomposed factors

Method/ Horizon	1	2	3	6	9	12	18	24
Mean	**0.554**	0.357	0.529	0.338	0.414	1.464	0.203	-0.562
Median	0.006	-0.008	**0.542**	**0.522**	**0.630**	**1.169**	**0.680**	0.076
Trimmed mean	**0.353**	**0.568**	**0.759**	**0.656**	0.462	**1.338**	**1.025**	-0.336
DMSE(1)	**0.555**	0.377	0.560	0.358	0.453	**1.407**	0.300	-0.389
DMSE(0.9)	**0.560**	0.268	0.458	0.255	0.451	0.868	0.124	-0.766
CL(2)	**0.736**	**0.985**	0.358	-0.087	**1.140**	-0.057	-1.121	-1.588
CL(3)	**0.980**	**1.120**	**1.099**	-0.981	**1.017**	-2.234	-2.262	-1.227

Panel C. Financial variables

Method/ Horizon	1	2	3	6	9	12	18	24
Mean	-0.654	-1.358	-0.529	-2.707	-4.788	-6.910	-10.098	<-20
Median	-0.419	-0.871	-1.458	-2.629	-3.572	-4.940	-9.041	-12.924
Trimmed mean	-0.122	-0.311	-0.903	-3.071	-5.556	-7.098	-9.534	-18.198
DMSE(1)	-0.693	-1.393	-0.524	-2.524	-4.383	-6.021	-7.232	-13.543
DMSE(0.9)	-1.074	-1.780	-1.327	-4.220	-8.241	-11.206	-12.412	-18.159
CL(2)	-2.540	-1.316	-2.371	-3.714	-4.339	-4.286	-2.943	-8.036
CL(3)	-4.711	-1.473	-3.471	-6.037	-8.342	-9.982	-6.133	-12.111

Notes: The table reports the out-of-sample R^2_{os} of the combination schemes relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW - t$ statistic, which tests the null hypothesis: $R^2_{os} = 0$ against the alternative: $R^2_{os} > 0$.

Table 3. Out-of-sample performance of individual ARDL models -US stock returns

Panel A. Empirical factors								
Predictor/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
HML	-0.168	0.098	-0.302	-0.103	-0.405	-0.716	-0.550	-0.510
SMB	-0.476	-0.161	0.373	-2.351	0.199	-0.118	-0.084	0.190
MOM	-0.777	-0821	0.247	**0.553**	**0.744**	**0.782**	**0.545**	-0.019
LT	0.025	0.906	**1.836**	1.079	-0.182	0.103	0.016	-0.039
ST	-0.768	-0.583	-0.438	-0.436	-0.084	-0.312	-0.277	-0.259
HML_b	0.016	0.078	-0.383	-1.174	-0.365	-0.311	-0.034	-0.392
HML_s	-0.720	-0.517	-0.290	-0.444	-0.294	-0.219	-0.364	-0.559
SMB_g	-1.307	-1.065	-0.762	-1.388	-0.718	-1.155	-0.301	-0.250
SMB_n	-0.518	-0.207	-0.210	-0.428	-0.260	-0.517	-0.129	-0.210
SMB_v	0.093	0.033	-0.269	-0.623	-0.667	-0.848	-0.269	-0.534
MOM_b	-0.532	-0.231	0.212	**0.816**	**0.859**	**0.899**	**0.595**	-0.056
MOM_ s	-0.976	-1.061	-0.006	-0.086	0.209	0.143	0.072	-0.138
LT_b	-0.248	0.444	**0.916**	1.032	-0.047	0.060	-0.057	-0.387
LT_s	0.300	0.657	**1.181**	0.261	-1.154	0.211	0.068	-0.080
ST_b	0.003	-0.351	-0.341	-0.479	-0.224	-0.410	-0.551	-0.602
ST_s	-0.960	-0.802	-0.465	-0.370	-0.089	-0.534	-0.471	-0.610

Panel B. Financial variables								
Predictor/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
DP	-2.182	-3.740	-5.791	-12.380	-19.530	-25.158	-33.246	-38.658
DY	-2.090	-3.430	-4.872	-9.531	-17.119	-21.031	-31.108	-39.367
EP	-0.935	-4.014	-9.161	-9.452	-16.894	-29.060	-11.259	-4.831
DE	-2.897	-12.300	-19.963	-26.787	-27.196	-47.349	-17.752	-15.982
SVAR	0.226	-2.549	-8.813	-4.349	-2.327	-2.477	-6.047	-4.312
BM	-0.275	-0.200	0.139	**0.603**	**1.114**	**0.837**	0.432	-0.271
NTIS	-3.299	-6.284	-10.276	-21.521	-28.734	-30.707	-33.192	-26.437
TBL	-1.178	-0.895	-1.583	-3.296	-2.747	-5.189	-5.748	-5.234
LTY	-0.973	-1.005	-0.710	-1.364	-0.835	-1.174	-0.560	**1.655**
SP500	-0.391	-1.374	-0.841	-2.344	-0.636	-0.948	-0.852	-0.720
TMS	-1.256	-1.964	-2.971	-3.861	-3.854	-1.598	**3.186**	**2.245**
DFY	-3.800	-6.382	-8.097	-8.216	-6.336	-4.923	-4.869	-12.719
DFR	0.032	-0.373	-0.444	-0.345	-0.767	-1.070	-4.318	-7.006
INFL	-2.226	-3.162	-4.528	-0.590	-0.089	0.337	-1.188	-1.814

Notes: The table reports the out-of-sample R^2_{os} of the individual ARDL models relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$ statistic, which tests the null hypothesis: $R^2_{os}=0$ against the alternative: $R^2_{os}>0$.

Table 4. Out-of-sample performance of combination methods -U.S. stock returns

Panel A. Five empirical factors								
Method/ Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	-0.133	0.244	**0.612**	0.169	0.319	0.180	0.104	0.006
Median	-0.340	-0.324	0.112	-0.228	0.159	-0.315	-0.314	0.052
Trimmed mean	-0.202	-0.107	0.449	-0.201	0.000	-0.173	-0.031	0.097
DMSE(1)	-0.138	0.246	**0.613**	0.170	0.317	0.179	0.106	0.011
DMSE(0.9)	-0.162	0.236	**0.606**	0.197	0.338	0.183	0.091	-0.008
CL(2)	-0.831	-0.069	0.635	0.624	0.847	0.487	0.208	0.134
CL(3)	-1.017	-0.140	**1.260**	0.323	0.261	0.352	-0.074	-0.048

Panel B. Eleven decomposed factors								
Method/ Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	-0.140	0.043	0.232	0.130	-0.018	-0.023	0.082	-0.158
Median	0.075	-0.021	0.058	0.068	0.201	-0.056	0.110	-0.130
Trimmed mean	-0.088	0.056	**0.285**	0.007	-0.043	-0.100	0.048	-0.156
DMSE(1)	-0.141	0.046	0.235	0.133	-0.018	-0.019	0.088	-0.158
DMSE(0.9)	-0.141	0.046	0.229	0.150	-0.002	-0.013	0.099	-0.164
CL(2)	-0.176	-0.270	0.191	-0.040	-0.037	0.053	0.069	-0.269
CL(3)	-0.079	-0.394	0.073	0.034	-0.121	0.098	0.170	-0.178

Panel C. Financial variables								
Method/ Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	-0.736	-2.057	-3.462	-4.770	-6.033	-6.778	-7.053	-8.032
Median	-0.182	-0.344	-0.619	-0.487	-0.430	-0.841	-2.403	-4.436
Trimmed mean	-0.739	-1.475	-2.481	-2.959	-3.543	-3.637	-4.722	-6.102
DMSE(1)	-0.727	-2.021	-3.317	-4.363	-4.866	-5.396	-5.053	-5.355
DMSE(0.9)	-0.624	-1.880	-2.958	-3.624	-4.718	-6.017	-6.448	-6.946
CL(2)	-0.316	-1.570	-2.229	-2.110	-1.340	-4.130	-3.459	-3.817
CL(3)	-0.279	-2.222	-2.518	-1.737	-1.258	-2.862	-2.660	0.660

Notes: The table reports the out-of-sample R^2_{os} of the combination schemes relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$ statistic, which tests the null hypothesis: $R^2_{os} = 0$ against the alternative: $R^2_{os} > 0$.

Table 5. Asset allocation benefits for a US Investor ($\gamma = 3$)

	Panel A. Horizon 1						**Panel B. Horizon 3**					
	Five Factors		Eleven Factors		Financial Variables		Five Factors		Eleven Factors		Financial Variables	
Method	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**
Mean	**1.348**	**1.630**	**1.417**	**1.652**	-1.591	-1.673	**2.326**	**3.309**	**2.155**	**2.432**	-4.883	-4.860
Median	**0.351**	**0.570**	**0.567**	**0.604**	-0.650	-0.748	**1.998**	**2.737**	-0028	**0.038**	**1.427**	**1.803**
Tr.Mean	**0.845**	**1.051**	**0.765**	**0.888**	-2.291	-2.576	**1.908**	**2.684**	**1.871**	**2.071**	-4.011	-3.733
DMSE(1)	**1.340**	**1.622**	**1.425**	**1.662**	-1.502	-1.580	**2.323**	**3.310**	**2.110**	**2.401**	-4.297	-4.267
DMSE(0.9)	**1.235**	**1.514**	**1.395**	**1.625**	-0.994	-1.005	**2.200**	**3.136**	**2.248**	**2.513**	-3.231	-3.047
CL(2)	**1.524**	**1.867**	**2.178**	**2.507**	**0.111**	**0.295**	**4.807**	**6.281**	**1.009**	**1.596**	-1.859	-0.767
CL(3)	**2.299**	**2.728**	**3.029**	**3.396**	-0.211	-0.049	**6.414**	**8.113**	**1.535**	**2.498**	-2.032	-1.041

	Panel C. Horizon 6						**Panel D. Horizon 12**					
	Five Factors		Eleven Factors		Financial Variables		Five Factors		Eleven Factors		Financial Variables	
Method	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**	**ΔU**	**Θ**
Mean	-1.440	-1.197	-0.489	-0.343	-14.087	-13.411	-2.529	-2.393	-1.313	-0.948	-16.267	-14.031
Median	-2.347	-2.064	**0.643**	**0.636**	-0.043	**2.418**	-1.209	-1.112	-0.877	-0.754	-1.632	-1.360
Tr. Mean	-1.877	-1.681	**0.292**	**0.402**	-11.103	-9.465	-2.108	-1.998	-0.984	-0.726	-13.491	-11.423
DMSE(1)	-1.375	-1.117	-0.504	-0.360	-11.525	-10.997	-2.510	-2.373	-1.310	-0.997	-12.796	-10.784
DMSE(0.9)	-1.303	-1.084	-0.496	-0.351	-8.546	-8.254	-2.306	-2.171	-1.480	-1.051	-12.909	-11.003
CL(2)	**1.099**	**1.509**	-0.638	-0.474	-0.853	**3.258**	-2.569	-2.482	**0.979**	**0.809**	-5.939	-5.114
CL(3)	**3.055**	**4.277**	-0.567	-0.386	-1.787	**3.265**	-3.491	-3.654	-0.261	**0.045**	-1.680	-1.614

Notes: The table reports the average utility gain (ΔU) and the difference between the manipulation-proof performance measure (Θ) of the proposed specification relative to the benchmark AR model. Figures are reported in annualized percentage points. Portfolio weights are constrained to lie between 0 and 1.5.

Table 6. Asset allocation benefits for a US Investor ($\gamma = 5$)

	Panel A. Horizon 1						**Panel B. Horizon 3**					
	Five Factors		Eleven Factors		Financial Variables		Five Factors		Eleven Factors		Financial Variables	
Method	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ
Mean	**1.160**	**1.374**	**0.803**	**0.979**	-2.017	-2.335	**1.927**	**2.086**	**1.299**	**1.118**	-4.669	-5.906
Median	**0.274**	**0.448**	**0.139**	**0.164**	-1.150	-1.534	**1.152**	**1.141**	**0.141**	**0.132**	**1.378**	**1.123**
Tr.Mean	**0.613**	**0.779**	**0.307**	**0.406**	-2.418	-2.951	**1.615**	**1.745**	**1.128**	**0.989**	-3.509	-4.361
DMSE(1)	**1.156**	**1.370**	**0.806**	**0.983**	-1.973	-2.286	**1.921**	**2.077**	**1.260**	**1.086**	-4.264	-5.553
DMSE(0.9)	**1.104**	**1.315**	**0.795**	**0.967**	-1.703	-1.949	**1.836**	**1.976**	**1.361**	**1.177**	-2.785	-3.942
CL(2)	**1.367**	**1.626**	**1.252**	**1.510**	-1.028	-1.062	**3.274**	**2.876**	**0.316**	**0.097**	-0.688	-0.939
CL(3)	**1.964**	**2.269**	**1.757**	**2.053**	-1.209	-1.312	**3.893**	**3.284**	**0.134**	-0.202	-1.209	-2.241

	Panel C. Horizon 6						**Panel D. Horizon 12**					
	Five Factors		Eleven Factors		Financial Variables		Five Factors		Eleven Factors		Financial Variables	
Method	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ
Mean	**0.206**	-0.646	**0.127**	-0.408	-9.437	-10.347	-2.657	<-20.0	-3.041	<-20.0	-11.074	**6.670**
Median	**0.334**	**2.267**	**1.144**	**1.319**	**2.842**	**5.186**	-2.848	<-20.0	-3.152	<-20.0	**5.063**	>20.0
Tr. Mean	**0.179**	**1.324**	**0.347**	-0.125	-4.439	-2.333	-2.771	<-20.0	-2.900	<-20.0	-2.722	>20.0
DMSE(1)	**0.190**	-0.719	**0.121**	-0.425	-7.744	-8.478	-2.662	<-20.0	-3.061	<-20.0	-8.161	**12.412**
DMSE(0.9)	**0.415**	-0.148	**0.237**	-0.163	-3.407	-1.870	-2.501	<-20.0	-2.878	<-20.0	-10.095	**7.504**
CL(2)	**1.217**	-1.182	-0.298	-2.008	**2.141**	**6.996**	-1.839	-11.166	-2.007	<-20.0	-0.626	>20.0
CL(3)	**1.134**	-5.571	-0.398	-4.087	**1.359**	**5.934**	-2.238	<-20.0	-2.670	-15.469	**4.206**	>20.0

Notes: The table reports the average utility gain (ΔU) and the difference between the manipulation-proof performance measure (Θ) of the proposed specification relative to the benchmark AR model. Figures are reported in annualized percentage points. Portfolio weights are constrained to lie between 0 and 1.5.

Table 7. Out-of-sample performance - European bond returns

Panel A. Individual ARDL models

Predictor/Horizon	1	2	3	6	9	12	18	24
HML	**3.170**	**1.745**	**3.150**	1.359	-0.287	-0.744	**3.227**	**2.793**
SMB	-0.658	-0.463	0.322	-0.293	0.140	-0.302	-1.309	-1.400
MOM	-1.756	-2.424	-5.312	-5.983	-0.371	-1.469	-3.036	-1.391
HML_b	1.001	**0.840**	**2.809**	**1.713**	-0.193	-1.354	3.568	**1.344**
HML_s	**1.998**	0.914	1.263	-0.456	-0.327	0.307	-0.688	2.104
SMB_g	-0.868	-1.321	-2.736	-1.144	-0.245	-1.494	-7.749	0.096
SMB_n	-1.791	-1.816	0.347	-0.216	-0.217	-0.366	-1.077	-4.230
SMB_v	-1.591	-0.799	0.678	-0.588	-0.040	-0.358	-1.336	-4.667
MOM_b	-0.283	-1.380	-3.196	-5.550	-1.194	-2.407	-4.693	-2.364
MOM_s	-2.764	-3.373	-7.089	-4.978	0.410	-0.333	-0.894	-0.195

Panel B. Combination forecasts -Empirical factors

Method/Horizon	1	2	3	6	9	12	18	24
Mean	0.414	-0.072	0.484	-1.106	0.202	0.167	1.975	2.424
Median	-0.095	-0.050	**1.460**	-0.370	**0.427**	0.413	0.170	0.189
Trimmed mean	-0.095	-0.050	**1.460**	-0.370	**0.427**	0.413	0.170	0.189
DMSE(1)	0.423	-0.054	0.538	-1.087	0.204	0.163	**2.167**	**2.562**
DMSE(0.9)	0.441	-0.068	0.542	-1.063	0.165	0.034	1.960	**2.518**
CL(2)	**1.439**	-0.185	0.542	-0.016	-1.484	-2.208	0.088	2.197
CL(3)	**1.987**	0.753	-1.469	0.615	-2.795	-4.077	-5.596	-2.365

Panel C. Combination forecasts- Decomposed factors

Method/Horizon	1	2	3	6	9	12	18	24
Mean	-0.409	-0.576	-0.330	-1.248	0.044	-0.186	-0.279	0.491
Median	-0.261	-0.502	0.426	-1.001	0.056	-0.028	-0.431	0.141
Trimmed mean	-0.402	-0.697	-0.297	-1.270	-0.122	-0.158	-0.280	0.037
DMSE(1)	-0.407	-0.726	-0.311	-1.242	0.043	-0.192	-0.227	0.582
DMSE(0.9)	-0.410	-0.877	-0.304	-1.224	0.023	-0.274	-0.430	0.554
CL(2)	-0.393	-2.088	-1.255	-0.012	-0.483	-0.861	-1.889	1.101
CL(3)	0.040	-0.810	-2.104	-0.227	-0.077	-2.951	-4.800	2.074

Notes: The table reports the out-of-sample R^2_{os} of the individual ARDL models and combination schemes relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$ statistic, which tests the null hypothesis: $R^2_{os}=0$ against the alternative: $R^2_{os}>0$.

Table 8. Out-of-sample performance - European stock returns

Panel A. Individual ARDL models

Predictor/Horizon	1	2	3	6	9	12	18	24
HML	-3.547	-1.420	-2.117	-3.340	-1.007	-1.321	-1.630	-0.811
SMB	0.320	**1.141**	1.060	**1.706**	**0.714**	0.822	**1.316**	-0.088
MOM	1.727	-0.120	**2.813**	-0.382	1.042	0.128	0.188	-0.397
HML_b	-3.926	-2.609	-2.294	-1.348	-1.652	-2.493	-2.094	-1.639
HML_s	-3.189	0.227	1.552	-1.396	0.741	0.641	-0.118	-0.110
SMB_g	-0.448	1.114	**4.132**	**1.291**	**1.123**	1.217	**1.541**	-0.381
SMB_n	**1.436**	**0.861**	0.394	**1.483**	0.529	0.781	**1.579**	-0.648
SMB_v	-2.719	-0.895	-0.125	-0.793	0.242	-0.198	-0.146	0.080
MOM_b	1.641	-0.265	**2.547**	-0.720	1.010	0.027	0.164	-1.014
MOM_s	0.836	0.070	**2.812**	-0.384	0.796	0.052	0.063	0.265

Panel B. Combination forecasts -Empirical factors

Method/Horizon	1	2	3	6	9	12	18	24
Mean	0.212	0.025	0.893	-0.371	0.461	-0.017	0.045	-0.398
Median	0.553	0.392	**1.792**	0.031	-0.034	0.031	0.112	-0.694
Trimmed mean	0.553	0.392	**1.792**	0.031	-0.034	0.031	0.112	-0.694
DMSE(1)	0.194	0.019	0.888	-0.363	0.451	-0.009	0.094	-0.404
DMSE(0.9)	0.192	0.028	0.909	-0.322	0.463	-0.023	0.060	-0.400
CL(2)	1.256	0.027	1.925	0.281	0.892	-0.684	0.591	-0.455
CL(3)	0.983	0.292	**1.861**	-0.004	1.658	-0.062	**1.126**	-0.340

Panel C. Combination forecasts- Decomposed factors

Method/Horizon	1	2	3	6	9	12	18	24
Mean	-0.302	-0.056	1.605	-0.078	0.533	0.110	0.250	-0.409
Median	-0.195	**0.460**	2.375	-0.051	0.468	0.003	0.156	-0.100
Trimmed mean	-0.352	0.329	2.044	-0.010	0.534	0.106	0.268	-0.314
DMSE(1)	-0.309	-0.061	1.611	-0.073	0.535	0.123	0.273	-0.413
DMSE(0.9)	-0.310	-0.054	1.614	-0.059	0.541	0.124	0.269	-0.409
CL(2)	0.292	0.298	**2.463**	0.325	0.893	0.406	**0.648**	-0.221
CL(3)	0.150	-0.391	**2.800**	0.277	0.790	0.351	**1.456**	-0.446

Notes: The table reports the out-of-sample R^2_{os} of the individual ARDL models and combination schemes relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$ statistic, which tests the null hypothesis: $R^2_{os}=0$ against the alternative: $R^2_{os}>0$.

Table 9. Asset allocation benefits - European Investor

| | Panel A. Horizon 1 | | | | Panel B. Horizon 3 | | | |
| | Three Factors | | Seven Factors | | Three Factors | | Seven Factors | |
Method	Δ**U**	Θ	Δ**U**	Θ	Δ**U**	Θ	Δ**U**	Θ
Mean	**4.622**	**4.094**	**0.730**	**0.008**	-2.355	-1.849	**2.601**	**0.266**
Median	**1.897**	**0.594**	**0.743**	-0.969	**5.070**	**1.465**	**9.144**	**2.898**
Tr.Mean	**1.897**	**0.594**	**0.011**	-1.757	**5.070**	**1.465**	**6.855**	**1.887**
DMSE(1)	**4.541**	**4.022**	**0.701**	-0.017	-2.477	-1.907	**2.618**	**0.263**
DMSE(0.9)	**4.501**	**3.954**	**0.710**	-0.057	-2.219	-1.776	**2.698**	**0.298**
CL(2)	**0.601**	-3.386	-2.975	-8.019	**7.313**	**2.238**	**9.810**	**2.582**
CL(3)	-3.644	-14.210	-2.370	-7.804	**8.865**	**3.285**	**10.707**	**2.231**

| | Panel C. Horizon 6 | | | | Panel D. Horizon 12 | | | |
| | Three Factors | | Seven Factors | | Three Factors | | Seven Factors | |
Method	Δ**U**	Θ	Δ**U**	Θ	Δ**U**	Θ	Δ**U**	Θ
Mean	-8.515	-1.228	-8.368	**0.092**	-7.847	**0.530**	**2.386**	**0.554**
Median	-11.396	-0.962	-2.395	-0.456	**7.431**	**0.756**	-10.422	**0.847**
Tr. Mean	-11.396	-0.962	-6.103	**0.076**	**7.431**	**0.756**	<-20.0	**0.965**
DMSE(1)	-8.485	-1.214	-8.613	**0.110**	-8.617	**0.552**	<-20.0	**0.973**
DMSE(0.9)	-8.054	-0.978	-8.386	**0.196**	-8.566	**0.520**	**0.293**	**0.027**
CL(2)	**2.997**	**1.321**	-2.902	**2.358**	-9.017	-0.980	-14.996	**1.766**
CL(3)	-0.416	-2.354	**1.718**	**2.607**	-11.891	**0.451**	<-20.0	**1.895**

Notes: The table reports the average utility gain (ΔU) and the difference between the manipulation-proof performance measure (Θ) of the proposed specification relative to the benchmark AR model. Figures are reported in annualized percentage points. Portfolio weights are constrained to lie between 0 and 1.5 and RRA is set equal to 3.

Table 10. Out-of-sample performance - Japanese bond returns

Panel A. Individual ARDL models								
Predictor/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
HML	**2.031**	1.384	0.821	-0.231	0.326	-0.486	0.538	-0.720
SMB	**1.629**	-1.518	0.326	0.159	0.609	0.310	-0.492	-0.862
MOM	0.080	1.548	1.653	-0.397	0.487	-0.140	-1.112	-0.367
HML_b	0.801	-0.676	-0.021	-0.069	0.656	-0.293	0.700	-0.293
HML_s	**1.989**	**1.661**	1.088	-0.394	-0.593	-0.535	0.104	-0.271
SMB_g	**0.599**	-3.395	-0.677	-0.148	-0.110	-0.469	-0.079	-0.282
SMB_n	**1.526**	0.101	**1.769**	**0.824**	1.461	1.032	-0.391	-0.730
SMB_v	1.560	0.331	0.131	-0.325	0.559	0.584	-1.150	-1.187
MOM_b	-0.372	1.118	1.686	-0.242	0.625	-0.169	0.053	-0.180
MOM_s	0.194	0.638	0.498	-0.912	-0.374	-0.070	-2.936	-0.612
Panel B. Combination forecasts -Empirical factors								
Method/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	**1.357**	**0.846**	**1.294**	0.000	0.651	0.002	-0.268	-0.578
Median	**1.590**	-0.118	**0.762**	**0.264**	**0.875**	**0.470**	-0.106	-0.274
Trimmed mean	**1.590**	-0.118	**0.762**	**0.264**	**0.875**	**0.470**	-0.106	-0.274
DMSE(1)	**1.356**	0.823	**1.282**	-0.006	0.651	-0.004	-0.254	-0.584
DMSE(0.9)	**1.387**	0.855	**1.309**	-0.034	**0.660**	-0.014	-0.259	-0.581
CL(2)	**1.388**	0.818	0.052	-0.446	0.348	-0.483	-0.325	-1.689
CL(3)	**1.343**	-2.148	-1.716	-1.379	**0.749**	-0.051	**1.053**	-1.163
Panel C. Combination forecasts- Decomposed factors								
Method/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	**1.040**	0.316	**0.981**	0.005	**0.564**	0.196	-0.387	-0.430
Median	**1.096**	0.352	**0.890**	0.051	0.118	0.245	0.257	-0.258
Trimmed mean	**1.032**	0.259	**0.908**	0.052	0.302	0.261	-0.087	-0.427
DMSE(1)	**1.040**	0.294	**0.969**	-0.002	**0.559**	0.188	-0.382	-0.425
DMSE(0.9)	**1.056**	0.323	**0.989**	-0.017	**0.571**	0.174	-0.389	-0.415
CL(2)	**1.087**	-0.378	0.051	-0.596	0.570	-0.427	-0.851	-0.344
CL(3)	0.489	-2.042	-1.095	-1.873	0.170	-0.566	-1.302	-0.408

Notes: The table reports the out-of-sample R^2_{os} of the individual ARDL models and combination schemes relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$, which tests the null hypothesis: $R^2_{os}=0$ against the alternative: $R^2_{os}>0$.

Table 11. Out-of-sample performance - Japanese stock returns

Panel A. Individual ARDL models								
Predictor/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
HML	**6.153**	**2.683**	**5.965**	**2.332**	**1.408**	-0.306	0.073	0.668
SMB	0.321	-0.195	-0.409	0.275	0.814	0.453	2.747	**2.441**
MOM	-0.239	-1.635	0.356	-0.726	-0.345	-0.156	1.197	-0.013
HML_b	**5.937**	**1.914**	**1.856**	0.206	0.167	-0.161	-0.894	-0.447
HML_s	**3.925**	0.929	4.210	**3.117**	**2.488**	-0.424	1.367	**1.204**
SMB_g	2.192	-0.456	-0.893	0.348	0.835	0.115	**1.846**	**0.976**
SMB_n	-0.373	0.003	-0.089	0.352	**1.552**	**1.329**	**3.504**	**3.195**
SMB_v	0.108	-0.376	0.150	-0.592	-0.057	-0.158	0.718	**1.092**
MOM_b	-2.637	-1.291	1.444	-0.493	-0.453	-0.379	0.458	0.591
MOM_s	-0.431	-1.288	1.884	-1.077	-1.332	-1.119	-0.030	-2.298
Panel B. Combination forecasts -Empirical factors								
Method/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	**3.335**	0.267	**2.343**	0.858	**1.035**	0.245	**1.690**	**1.164**
Median	1.499	0.290	0.752	0.358	0.012	-0.579	0.465	**0.783**
Trimmed mean	1.499	0.597	0.752	0.358	0.012	-0.579	0.465	**0.783**
DMSE(1)	**3.352**	0.599	**2.371**	0.863	**1.028**	0.258	**1.724**	**1.168**
DMSE(0.9)	**3.362**	0.259	**2.387**	0.871	**1.038**	0.260	**1.712**	**1.173**
CL(2)	**2.981**	-0.175	**1.018**	**1.024**	**1.526**	**0.745**	**2.356**	**1.473**
CL(3)	**4.628**	**2.719**	0.827	0.474	**1.306**	0.327	**2.187**	**1.352**
Panel C. Combination forecasts- Decomposed factors								
Method/Horizon	**1**	**2**	**3**	**6**	**9**	**12**	**18**	**24**
Mean	**2.181**	0.204	**1.618**	**0.768**	**1.056**	**0.755**	**1.422**	**0.859**
Median	1.199	**0.303**	**1.000**	0.167	**0.499**	**0.836**	**1.558**	**0.892**
Trimmed mean	**1.745**	0.080	**1.139**	0.468	**0.835**	0.359	**1.387**	**0.810**
DMSE(1)	**2.171**	0.214	**1.607**	**0.756**	**1.032**	0.664	**1.437**	**0.870**
DMSE(0.9)	**2.195**	0.178	**1.602**	**0.747**	**1.099**	0.613	**1.434**	**0.879**
CL(2)	**3.398**	-0.142	0.870	0.080	**1.173**	1.052	**2.285**	**1.341**
CL(3)	**4.132**	-0.265	1.075	0.100	**2.733**	1.027	0.878	**1.792**

Notes: The table reports the out-of-sample R^2_{os} of the individual ARDL models and combination schemes relative to the AR benchmark. Bold entries indicate significance at the 10% significance level according to the $CW-t$ statistic, which tests the null hypothesis: $R^2_{os}=0$ against the alternative: $R^2_{os}>0$.

Table 12. Asset allocation benefits - Japanese Investor

| | Panel A. Horizon 1 | | | | Panel B. Horizon 3 | | | |
| | Three Factors | | Seven Factors | | Three Factors | | Seven Factors | |
Method	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ
Mean	2.106	1.567	1.822	1.235	5.822	2.732	4.303	1.794
Median	0.796	0.623	0.207	0.185	0.915	0.171	2.492	1.100
Tr.Mean	0.796	0.623	1.532	1.121	0.915	0.171	3.034	1.280
DMSE(1)	2.106	1.571	1.821	1.237	5.925	2.773	4.286	1.788
DMSE(0.9)	2.100	1.572	1.810	1.230	5.990	2.800	4.321	1.792
CL(2)	3.067	1.946	2.292	1.477	3.495	1.045	2.641	0.550
CL(3)	2.621	1.504	3.982	2.719	5.304	0.816	2.946	0.737

| | Panel C. Horizon 6 | | | | Panel D. Horizon 12 | | | |
| | Three Factors | | Seven Factors | | Three Factors | | Seven Factors | |
Method	ΔU	Θ	ΔU	Θ	ΔU	Θ	ΔU	Θ
Mean	9.012	2.068	3.083	0.546	9.748	0.805	>20.0	4.258
Median	6.860	1.626	2.247	0.517	-2.525	-0.833	18.505	2.106
Tr. Mean	6.860	1.626	2.252	0.550	-2.525	-0.833	13.006	1.482
DMSE(1)	9.085	2.084	3.146	0.567	10.479	0.887	>20.0	4.190
DMSE(0.9)	9.193	2.103	3.232	0.575	9.902	0.819	>20.0	4.067
CL(2)	16.943	3.799	10.651	2.452	9.554	0.834	>20.0	6.004
CL(3)	>20.0	3.631	16.188	3.667	6.156	-0.377	>20.0	9.978

Notes: The table reports the average utility gain (ΔU) and the difference between the manipulation-proof performance measure (Θ) of the proposed specification relative to the benchmark AR model. Figures are reported in annualized percentage points. Portfolio weights are constrained to lie between 0 and 1.5 and RRA is set equal to 3.