Do Combination Forecasts Outperform the Historical Average? Economic and Statistical Evidence

Thomadakis, Apostolos

University of Wrawick

20 May 2016

Do Combination Forecasts Outperform the Historical Average? Economic and Statistical Evidence

Apostolos Thomadakis*†

University of Warwick

May 20, 2016

Abstract

This paper examines the out-of-sample predictability of monthly German stock returns, and addresses the issue of whether combinations of individual model forecasts are able to provide significant out-of-sample gains relative to the historical average. Empirical analysis over the period from 1973 to 2012 implies that firstly, term spread has the in-sample ability to predict stock returns, secondly, and most importantly, this variable successfully delivers consistent out-of-sample forecast gains relative to the historical average, and thirdly, combination forecasts do not appear to offer a significant evidence of consistently beating the historical average forecasts of the stock returns. Results are robust using both statistical and economic criteria, and hold across different out-of-sample forecast evaluation periods.

JEL classification: C22, C32, C53, G11, G17

Keywords: Equity Premium, Forecast Combination, Out-of-Sample Forecast, Mean-Variance Investor

*Department of Economics, University of Warwick, Coventry CV4 7AL, UK. E-mail: a.thomadakis@warwick.ac.uk.

†I would like to thank Fabio Bussetti, Daniele Massacci, Andrew Patton, Allan Timmermann, as well as participants at the DIW Macroeconometric Workshop 2012, ISF 2013, OMI-SoFiE Financial Econometrics Summer School 2013, CIdE WEEE 2013, 14th IWH-CIREQ Macroeconometric Workshop 2013, and CFE 2013 for their valuable comments.

1 Introduction

Predictability of stock returns has been a fundamental concern of both market practitioners and academic researchers for many years. Nowadays there is an extensive literature on the predictability of stock returns by financial and macroeconomic variables: dividend yield (Fama and French, 1988; Cochrane, 1992; Ang and Bekaert, 2007), price-earnings ratio (Campbell and Shiller, 1988; Weigand and Irons, 2007), short-term interest rate (Campbell, 1987; Ang and Bekaert, 2002), term spread (Rapach *et al.*, 2005; Hjalmarsson, 2010), inflation rate (Fama and Schwert, 1977; Campbell and Vuolteenaho, 2004), unemployment rate (Boyd *et al.*, 2005; Chen and Zhang, 2009) and oil prices (Driesprong *et al.*, 2008; Casassus and Higuera, 2011) among others.

However, stock returns contain a large unpredictable component, so that a forecaster is able to explain only a small part of high-frequency stock returns. Therefore, the degree of return predictability is small. Someone can distinguish two types of predictability, one that arises from the in-sample fit of a model and another that arises from the out-of-sample fit obtained from a sequence of expanding or rolling regressions. In terms of in-sample tests of predictability, the main argument put forward by Granger (1990) and Rapach and Wohar (2006) is the so-called over fitting problem which may spuriously indicate predictability patterns where there is none.

In this context, what matters for a model is not its ability to generate an accurate in-sample fit, but rather its out-of-sample performance (Campbell, 2008). In light of this, the current paper undertakes an extensive analysis of both in-sample and out-of-sample tests of stock returns predictability in univariate and multivariate level. We assess the in-sample predictability of many financial and economic variables that have been extensively used in the empirical literature using the t-statistic of the slope coefficient, and the out-of-sample predictability using three statistical measures: the mean squared prediction error ($MSPE$), the squared correlation coefficient ($CORR^2$), and the out-of-sample R^2 (R^2_{OS}).

More importantly, in an effort to evaluate the forecast accuracy from an economic point of view, we analyse the stock return forecasts performance with utility-based loss-functions. As suggested by Leitch and Tanner (1991) and Pesaran and Timmermann (1995), forecast evaluation made on the basis of conventional error-magnitude criteria find little justification for profit-maximising investors and has no systematic relationship to profits. Based on that, we follow Giacomini and White (2006), and focus not only on the forecasting model, but rather on the forecasting method, which includes the model, the estimation procedure and the estimation window. In other words, we calculate utility gains from the perspective of a mean-variance investor who optimally allocates portfolio between a risky asset and a risk-free asset using equity risk premium forecasts based on financial and economic variables, relative to an investor who uses the historical average equity risk premium forecast.

While stock return predictability of the US and UK stock market has been the subject of intense research, less academic attention has been given to the predictability of other stock markets, such as the

German. Germany is the world's fifth largest economy[1] (after China, US, India and Japan), as well as the largest in Europe, accounting for about one-fifth of the European Union's (EU) GDP. In addition, the fact that Germany has a welcoming attitude towards foreign direct investment (FDI)[2], rank Germany among the world's leading FDI countries with more than $743,515 millions in inward FDI stocks in 2014 - this represents a growth of 45% from 2004 to 2014.[3] According to the World Federation of Exchange statistics, Deutsche Börse was the second largest domestic equity market in level of capitalisation from within the EU-27, with $1,762 billion at the end of 2014. Given the health and functioning of the German economy, as well as its approaches to international and economic policy issues as a driving force in European integration, it is of great importance to understand whether predictability patterns arise in German stock.

The development of the German financial system has been characterised by two key features, both of which have their origins in the country's pattern of industrialisation in the nineteenth century. The first is that external finance for non-financial firms in Germany has been supplied predominantly by banks - indeed, Germany provides one of the archetypal examples of a bank-based financial system. The second key feature is that, while a small number of big banks played a dominant role amongst the privately-owned commercial banks, the German financial system has also included two other sectors that are not primarily motivated by making profit, namely the publicly-owned saving banks and the cooperative banks.[4]

This paper has four main objectives. First, to present an extensive case study for the predictability of German stock returns using nine financial and economic predictor variables in which univariate and multivariate considerations may lead to improving forecasting performance relative to the historical average model. Second, to evaluate the stock returns predictability using statistical and economic measures. The former, tests the null hypothesis of equal predictive ability between the random walk benchmark and an alternative predictive model, while the latter evaluates the performance of a portfolio using mean-variance analysis. Third, to verify whether forecast combinations - such as equally weighted, trimmed mean, median and discounted mean square prediction error methods - outperform the historical benchmark model. Fourth, to investigate the forecasting performance near business cycle peaks and troughs.

Results can be summarised in the following way. First, in-sample analysis over the entire sample period 1973:02-2012:01 shows that two variables - the term spread and the inflation rate - are able to predict excess stock returns in univariate and multivariate level. This is consistent with findings by Fama and Schwert (1977), Fama and French (1989) and Campbell and Thompson (2008) who have found that these variables are capable to predict stock returns. Second, in terms of out-of-sample R^2 and $adjusted\text{-}MSPE$ statistic, the term spread appear to be a fairly robust predictor of German excess stock returns in all out-of-sample forecast periods considered. This result is in line with Hjalmarsson (2010) who reports an R^2_{OS} equal to

[1] Country GDP ranking based on purchasing power parity (PPP) according to The World Factbook, CIA: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html

[2] This is mainly due to the fact that German law makes no distinction between German and foreign nationals regarding investments or the establishment of companies. Also because the legal framework for FDI favors the principle of freedom of foreign trade and payment.

[3] http://unctad.org/en/Pages/DIAE/World%20Investment%20Report/Annex-Tables.aspx

[4] For more information about the development of the German financial system see Detzer *et al.* (2013).

0.512 for Germany over the period 1953 to 2004, compared to 0.642 that we found for the 1985 to 2012 period. This finding provides evidence that a significant in-sample relationship also tends to be associated with out-of-sample predictive power. Third, despite the success of some individual predictive regression model forecasts to outperform the historical average, the combination of individual model forecasts does not deliver statistically and economically significant out-of-sample gains relative to the historical average on a consistent basis over time. Fourth, the term spread detects the decline in the actual equity premium early in recessions, as well as the increase in the actual equity premium late in recessions.

The paper is organised as follows. Section 2 introduces the econometric methodology that we use to study the implications of economic and financial variables for forecasting stock market returns. Data are discussed in Section 3, while Section 4 presents the empirical results. Finally, Section 5 offers the concluding remarks.

2 Methodology

2.1 Predictive Regression Model

The method that typically is used in order to examine excess returns predictability, it is based on the simple regression model

$$r_{t+1} = a_i + \beta_i x_{i,t} + \varepsilon_{t+1}, \quad i = 1, \ldots, N, \tag{1}$$

where r_{t+1} is the return on the stock market index in excess of the risk-free interest rate; $x_{i,t}$ is a financial-macroeconomic variable that it can help predict future returns; ε_{t+1} is an error term; and N is the number of predictor variables. We divide the sample into an in-sample and out-of-sample period; given a sample of T observations, the in-sample portion is composed of the first k observations ($t = 1, \ldots, k$), while the out-of-sample portion is composed of the last m observations ($k = s, \ldots, T$), and therefore $m = T - s$. Here, following Pesaran and Timmermann (1995), Campbell and Thompson (2008), Welch and Goyal (2008), and Ferreira and Santa-Clara (2011), we employ an expanding (recursive) estimation window, where the first k observations will be used to construct an initial set of regression estimates which are then will be used for the first prediction. The last m observations of the sample will be used for forecast evaluation. We label the model in equation 1 as the "predictive model" (P) when it incorporates only one predictor variable at a time; and as the "kitchen sink" (KS), when a multiple regression forecasting model includes all potential predictors

$$r_{t+1} = a_i + \sum_{i=1}^{N} \beta_i x_{i,t} + \varepsilon_{t+1}. \tag{2}$$

The expanding estimation window is used to generate an out-of-sample forecasts of r_{t+1} based on equation 1 and information available at time t

$$\widehat{r}_{i,t+1} = \widehat{\alpha}_{i,t} + \widehat{\beta}_{i,t} x_{i,t}, \quad i = 1, \ldots, N, \tag{3}$$

where $\widehat{a}_{i,t}$ and $\widehat{\beta}_{i,t}$ are the OLS estimates of α_i and β_i respectively, computed by regressing $\{r_t\}_{t=2}^{k}$ on a constant and $\{x_{i,t}\}_{t=1}^{k-1}$. Repeating this process m times generates a series of out-of-sample forecasts, $\{\widehat{r}_{i,t+1}\}_{t=k}^{T-1}$, of the excess returns based on $x_{i,t}$.

The in-sample predictive ability of $x_{i,t}$ is typically assessed by examining the t-statistic corresponding to $\widehat{\beta}_{i,t}$ in equation 3. Under the null hypothesis of no predictability, $\beta = 0$, expected returns are constant. In other words, the predictive model is evaluated against the so-called "historical average benchmark model" (B), which is a simple random walk with drift in the log prices

$$\log p_{t+1} = a_i + \log p_t + \varepsilon_{t+1}, \tag{4}$$

since

$$r_{t+1} = a_i + \varepsilon_{t+1}, \tag{5}$$

and

$$r_{t+1} = \log p_{t+1} - \log p_t. \tag{6}$$

Using an expanding window, the return forecasts for the period $t+1$ which are produced at time t are given by $(t-1)^{-1}\sum_{t=1}^{k-1} r_{t+1}$. Under a recursive estimation this model assumes "no predictability" (Timmermann, 2008), since only the constant term is included. Therefore, it is a first test on the direction of the effect of $x_{i,t}$ on r_{t+1} under the alternative hypothesis.

2.2 Forecast Combination

Next, we consider two general classes of combining methods, those who do not take into account previous information - simple combination methods; and those who use historical information to compute the combination forecasts - discounting methods.

2.2.1 Simple Combination Methods

Since the seminal work of Bates and Granger (1969), the combination of individual forecasts of the same event has frequently been found to outperform the individual forecasts, in the sense that combining forecasts deliver a smaller $MSPE$. This is what Granger and Jeon (2004) descibed as "thick modeling", where the forecaster uses many alternative specifications and then combines or synthesises them in order to produce the best forecast possible. Studies by Newbold and Granger (1974) and Clemen (1989) found that simple rules of combining forecasts, such as averages, often outperform more complicated weighting schemes (Stock and Watson, 1999 and 2004). Furthermore, combining forecasts across individual models can lead to improved forecast accuracy (Hendry and Clements, 2004). It has been also shown that these forecasts are an effective tool for forecasting in the presence of structural breaks (Paye and Timmermann, 2006). For these reasons, in an effort to generate improved equity premium forecasts based on economic variables, we consider combination forecasts of the equity premium.

We estimate combination forecasts of excess returns, r_{t+1}, as weighted averages (linear combinations) of the N individual predictive regression model forecasts

$$\widehat{r}^{c}_{t+1} = \sum_{i=1}^{N} \omega^{c}_{i,t} \widehat{r}_{i,t+1}, \qquad i = 1, 2, \ldots, N \tag{7}$$

where c indicates the different combination schemes (c =mean, trimmed-mean and median); i denotes the number of forecasts combined; and $\left\{\omega^{c}_{i,t}\right\}_{i=1}^{N}$ are the combining weights corresponding to each specific scheme ($\sum_{i=1}^{N} \omega^{c}_{i,t} = 1$). In this paper we employ three averaging schemes will be employed. The simplest or naive combination forecast which sets $\omega^{mean}_{i,t} = 1/N$, the trimmed-mean combination forecast which sets $\omega^{trimmed}_{i,t} = 0$ for the forecasts with the smallest and largest values and $\omega^{trimmed}_{i,t} = 1/(N-2)$ for the remaining individual forecasts, and the median combination forecast which is the median of the forecasts $\{\widehat{r}_{i,t+1}\}_{i=1}^{N}$.

2.2.2 Discounting Method

Following the theory of combination forecasting which suggests that methods that weight forecasts more heavily will perform better than simple combination forecasts, we also apply the discounted mean squared prediction error ($DMSPE$) method. This combining method requires a holdout period in order to estimate the combining weights. As a holdout period we are using: i) the first q observations of the out-of-sample period (Rapach *et al.,* 2010), and ii) the last q observations of the in-sample period (Rapach and Zhou, 2012). The discounted $MSPE$ computes the combination forecast as a weighted average of the individual forecasts, where the weights depend inversely on the historical performance of each individual forecast

$$\omega^{DMSPE}_{i,t} = \frac{DMSPE^{-1}_{i,t}}{\sum\limits_{j=1}^{N} DMSPE^{-1}_{j,t}}, \tag{8}$$

and

$$DMSPE_{i,t} = \sum_{l=p}^{t-1} \theta^{t-1-l} (r_{i,l+1} - \widehat{r}_{i,l+1})^2, \tag{9}$$

where $p+1$ denotes the start of the holdout period and θ is the discount factor. As a result, the combining weights formed at time t are functions of the historical forecasting performance of the individual models over the holdout period.

With this method the individual predictive regression model which generates the lower $MSPE$ value (better out-of-sample performance) over the holdout period assigned greater weight. When $\theta = 1$, equation 9 produces the optimal combination forecasts derived by Bates and Granger (1969) for the case where the individual forecasts are uncorrelated. In other words, $\theta = 1$ ignores any correlation in the errors of the individual forecasts. On the other hand, setting values of $\theta < 1$ allows for higher (lower) weights to be assigned to more recent (distant) forecast errors in the calculation of the combination weights. Following Sarno *et al.* (2005), Rapach and Strauss (2008), Rapach *et al.* (2010) and Della Corte and Tsiakas (2012), we consider the values $\theta = 1$, $\theta = 0.9$ and $\theta = 0.75$.

2.3 Forecast Evaluation

A common problem that many forecasters face is how to evaluate the performance of two or more forecast alternatives. A forecaster should focus not only on the forecasting model, but rather on the forecasting method. The latter, according to Giacomini and White (2006) includes the model, the estimation procedure, as well as the possible choice of estimation window. Following that, Section 2.3.1 evaluates the accuracy of the forecasting model, while Section 2.3.2 I evaluates the accuracy of the forecasting method.

2.3.1 Statistical Evaluation

In order to measure and compare the accuracy of two out-of-sample forecasts, from the predictive and the benchmark model, we will consider three well known metrics. The first one is the squared value of the correlation coefficient, $CORR^2$ (Pesaran and Timmermann, 1995 and 2000). This is a measure of how well the predicted values from a forecast model, $\widehat{r}_{i,t+1}$, fit the actual values, r_{t+1}. The $CORR^2$ ranges between 0 and 1 ($0 \leq CORR^2 \leq 1$), where higher values strengthen the relationship between predicted and actual values.

The second metric is the Campbell and Thompson (2008) out-of-sample R^2, which compares the predictive ability of the predictive model with the historical benchmark model

$$R^2_{OS} = 1 - \frac{MSPE_m}{MSPE_B}, \quad m = KS, P, c, DMSPE, \tag{10}$$

where $MSPE_m$ is the mean square prediction error of the predictive/forecasting model

$$MSPE_m = \frac{1}{T-s}\sum_{k=s}^{T-1}(r_{k+1} - \widehat{r}^m_{k+1})^2, \quad m = KS, P, c, DMSPE, \tag{11}$$

and $MSPE_B$ is the mean square prediction error of the historical average benchmark model

$$MSPE_B = \frac{1}{T-s}\sum_{k=s}^{T-1}(r_{k+1} - \widehat{r}^B_{k+1})^2. \tag{12}$$

By construction, when the competing forecast outperforms the historical average benchmark in terms of $MSPE$, which means that $MSPE_m \leq MSPE_B$, the out-of-sample R^2 is positive, $R^2_{OS} > 0$. Contrary to that, $R^2_{OS} \leq 0$ when the historical average benchmark is at least as good as the forecasting model.

Furthermore, we evaluate the significance of the R^2_{OS} using the *adjusted*-$MSPE$ proposed by Clark and West (2007) and applied by Rapach and Wohar (2006), Welch and Goyal (2008) among others. The *adjusted*-$MSPE$ is based on Diebold and Mariano (1995) and West (1996) (DMW) idea of testing for significant differences in loss-function. The DMW statistic tests the null hypothesis that the $MSPE_m$ of the predictive model is greater than or equal to the $MSPE_B$ of the benchmark model against the one-sided alternative hypothesis that the the predictive model has lower mean square prediction error. Equally, $H_0 : R^2_{OS} \leq 0$ against $R^2_{OS} > 0$.

However, the DMW statistic has a non-standard distribution when comparing forecasts from nested linear models (Clark and McCracken, 2001; McCracken, 2007), which is the case when comparing predictive regression model forecasts of the excess returns to the historical average benchmark model. The reason is that when the data are generated from the more parsimonious model, the forecasts are identical when the parameters are known and as a result the asymptotic distribution theory does not holds. Clark and West (2007) modified the DMW with the *adjusted*-$MSPE$ statistic that generates asymptotic confidence intervals that can be calculated by a normal distribution when comparing nested models. Given the forecasts of r_{t+1} from the historical average benchmark model, $\widehat{r}^B_{k+1}$, and the predictive model, $\widehat{r}^m_{k+1}$, as well as the corresponding forecasts errors, $\widehat{u}^B_{k+1} = r_{k+1} - \widehat{r}^B_{k+1}$ and $\widehat{u}^m_{k+1} = r_{k+1} - \widehat{r}^m_{k+1}$ respectively, the *adjusted*-$MSPE$ statistic can be defined as

$$\widehat{f}_{k+1} = \left(\widehat{u}^B_{k+1}\right)^2 - \left[\left(\widehat{u}^m_{k+1}\right)^2 - \left(\widehat{r}^B_{k+1} - \widehat{r}^m_{k+1}\right)^2\right], \qquad m = KS, P, c, DMSPE. \tag{13}$$

The additional term, $(\widehat{r}^B_{k+1} - \widehat{r}^m_{k+1})^2$, simply adjusts for the upward bias in $MSPE$ produced by estimation of parameters that are zero under the null, since $[(\widehat{u}^m_{k+1})^2 - adjustment)] < (\widehat{u}^m_{k+1})^2$. By regressing $\widehat{f}_{k+1}$ on a constant, the *adjusted*-$MSPE$ is the t-statistic corresponding to a zero constant, while the p-value of a one-sided test is obtained using the standard normal distribution.

2.3.2 Economic Evaluation

The previous section described how to evaluate the out-of-sample performance of a forecasting model, using statistical measures. However, these measures are sometimes found to be unsatisfactory from an economic point of view (Granger and Machina, 2006). As an alternative to the statistical criteria, forecast performance can be measured by utility-based loss-functions. Leitch and Tanner (1991), using profit measures for interest rate forecast conclude that forecast evaluations made on the basis of conventional error-magnitude criteria often find little justification for profit-maximising investors and have no systematic relationship to profits. Similar conclusion has been drawn by Pesaran and Timmermann (1995 and 2000) and Granger and Pesaran (2000) where they argue that predictability of stock returns in itself does not guarantee that an investor can earn profits from a trading strategy based on such forecasts.

Based on these arguments, we consider a classic portfolio choice problem where two assets - a risk-free asset and a risky asset - are available to an investor at time t. The investor has a single-period horizon and prefers a high mean and a low variance of portfolio returns (mean-variance preferences). We also assume that the investor trades off mean and variance in a linear way. That is, he maximises a mean-variance utility function, with a positive weight on mean and a negative weight on variance[5]

$$\max_{\omega_{mt}} \left[E_{mt}\left(y_{p,m,t+1}\right) - \frac{1}{2}\gamma Var_{mt}\left(y_{p,m,t+1}\right) \right], \qquad m = KS, P, B, c, DMSPE, \tag{14}$$

where m denotes the forecasting method; E_{mt} and Var_{mt} are the expected excess return and variance computed under forecasting method m and conditional upon the information available at time t; and $y_{p,m,t+1}$ is the return on the investor's portfolio defined as

$$y_{p,m,t+1} = y_{ft} + \omega_{mt}\left(y_{t+1} - y_{ft}\right), \qquad m = KS, P, B, c, DMSPE, \tag{15}$$

or

$$y_{p,m,t+1} = \left(1 - \omega_{mt}\right) y_{ft} + \omega_{mt} y_{t+1}, \qquad m = KS, P, B, c, DMSPE, \tag{16}$$

where ω_{mt} is the proportion of the portfolio allocated to the risky asset with return y_{t+1}; $(1 - \omega_{mt})$ is the proportion of the portfolio allocated to the risk-free asset with return y_{ft}; and γ is the coefficient of relative risk aversion (RRA), representing the investor's degree of risk aversion. The solution of the maximisation problem which leads to the optimal portfolio weight for the investor indicates that the portfolio share in the risky asset should equal the expected excess return (risk premium) divided by conditional variance times the coefficient γ

$$\omega^*_{mt} = \frac{E_{mt}(y_{t+1}) - y_{ft}}{\gamma Var_{mt}\left(y_{t+1}\right)}, \qquad m = KS, P, B, c, DMSPE, \tag{17}$$

where the expected value $[E_{mt}(y_{t+1}) - y_{ft}]$ is predicted by $\widehat{r}^{m}_{t+1}$ for each forecasting method, as described earlier.

However, an increase in the average return does not necessarily incorporates "pure" gain for a risk-averse investor (Campbell and Thompson, 2008). Based on that, we calculate the welfare benefits generated by optimally trading on each predictor variable, as well as collectively, for an investor where the relative risk aversion coefficient is set $\gamma = 5$.[6] Furthermore, we impose portfolio constraints in order to prevent the

[5] The mean-variance analysis may involve three rules for optimal asset allocation: maximum expected utility, maximum expected return and minimum variance. Following Han (2006) and Della Corte *et al.* (2009) we focus on the maximum expected utility strategy.

[6] As a robustness check we estimate optimal portfolio weights using different levels of risk aversion coefficient ($\gamma = 0.5$, $\gamma = 1$, and $\gamma = 2$). Notice that for a simple mean-variance asset allocation exercise, the optimal weight for a single risky asset is $\omega_t = SR_p/\gamma$. Knowing that the Sharpe ratio on the German equity premium over the period 1973-2012 is 0.086, a coefficient of $\gamma = 0.5$ implies a weight of 0.172, i.e. 17.2% in German equity premium. That represents a rather moderate portofolio position.

investor from shorting stocks or taking more than 50% leverage. The portfolio weights are defined as:

$$\begin{aligned} \omega_{mt} &= 0 \quad if \quad \omega^*_{mt} < 0 \\ &= \omega^*_{mt} \quad if \quad 0 \leq \omega^*_{mt} \leq 1.5, \qquad m = KS, P, B, c, DMSPE, \\ &= 1.5 \quad if \quad \omega^*_{mt} > 1.5 \end{aligned} \tag{18}$$

where the investor estimates the stock return variance $Var_{mt}(y_{t+1})$ at each point in time under the historical average benchmark model using a rolling five-year window of monthly data. This allows for a time-varying variance.

In order to measure the economic forecast evaluation of the mean-variance analysis we use two measures, the Sharpe ratio (SH) and the Sortino ratio (SO). The realized SH can be defined as the ratio of the average excess returns of a portfolio and the standard deviation of the portfolio returns

$$SH_{p,m} = \frac{\overline{y}_{p,m} - \overline{y}_f}{\widehat{\sigma}^{SH}_{p,m}}, \qquad m = KS, P, B, c, DMSPE, \tag{19}$$

where $\overline{y}_{p,m}$ is the realised return on the portfolio from the forecasting method m; $\overline{y}_f$ is the realised return on the risk-free asset (also known as the minimum acceptable return, MAR); and $\widehat{\sigma}^{SH}_{p,m}$ is the realised standard deviation of the portfolio

$$\overline{y}_{p,m} = \frac{1}{T-s} \sum_{k=s}^{T-1} y_{p,m,t+1}, \tag{20}$$

$$\overline{y}_f = \frac{1}{T-s} \sum_{k=s}^{T-1} y_{ft}, \tag{21}$$

$$\widehat{\sigma}^{SH}_{p,m} = \left[\frac{1}{T-s} \sum_{k=s}^{T-1} \left(y_{p,m,t+1} - \overline{y}_{p,m} \right)^2 \right]^{1/2}. \tag{22}$$

respectively. However, because the Sharpe ratio by construction uses the sample standard deviation of the realized portfolio returns, it overestimates the conditional risk that an investor faces at each point in time. As a result, the SH underestimates the performance of dynamic asset allocation strategies (Marquering and Verbeek, 2004; Han. 2006).

For that reason, we also compute the Sortino ratio, SO which penalises only those returns falling below a minimum acceptable return (the risk-free rate). The SO ratio compares returns on a portfolio to downside risk - the risk of under-performing the benchmark - by differentiating between volatility due to up and down movements in portfolio returns. It is equal to the actual rate of return in excess of the investor's target rate or return, per unit of downside risk

$$SO_{p,m} = \frac{\overline{y}_{p,m} - \overline{y}_f}{\widehat{\sigma}^{SO}_{p,m}}, \qquad m = KS, P, B, c, DMSPE, \tag{23}$$

where

$$\widehat{\sigma}_{p,m}^{SO} = \left[\frac{1}{T-s} \sum_{k=s}^{T-1} \left(y_{p,m,t+1} - \overline{y}_{p,m} \right)^2 I \left(y_{p,m,t+1} \leq y_{ft} \right) \right]^{1/2}, \tag{24}$$

and $I(\cdot)$ is the indicator function. Larger values of SO ratio indicate a low risk of large losses.

For an investor that uses the historical benchmark model to make portfolio decisions, the average utility level over the out-of-sample period is given by

$$\overline{U}_B = \overline{y}_{p,B} - \frac{1}{2}\gamma \widehat{\sigma}_{p,B}^2, \tag{25}$$

where $\overline{y}_{p,B}$ and $\widehat{\sigma}_{p,B}^2$ correspond to the sample mean and variance, respectively, of the returns from the historical average benchmark portfolio over the out-of-sample period. On the other hand, if the investor uses the forecasting method m to make his portfolio decision, the average utility level over the out-of-sample period is given by

$$\overline{U}_m = \overline{y}_{p,m} - \frac{1}{2}\gamma \widehat{\sigma}_{p,m}^2, \quad m = KS, P, c, DMSPE, \tag{26}$$

where $\overline{y}_{p,m}$ and $\widehat{\sigma}_{p,m}^2$ correspond to the sample mean and variance, respectively, of the returns from the individual predictive portfolio over the out-of-sample period.

Finally, we measure the utility gain of using a particular forecasting method as the difference between equations 25 and 26

$$\Delta_m = \overline{U}_m - \overline{U}_B, \quad m = KS, P, c, DMSPE. \tag{27}$$

This utility gain or certainty equivalent return can be viewed as the portfolio management fee that an investor with mean-variance preferences would be willing to pay to access a particular forecasting method. Previous studies (Marquering and Verbeek, 2004; Campbell and Thompson, 2008; Rapach *et al.,* 2010; Rapach and Zhou, 2012) have found that there are substantial utility gains for a mean-variance investor who bases equity premium forecasts on economic variables.

3 Data

We use monthly data on German stock returns, as well as a large set of predictor variables for the period 1973:01-2012:01, a total of 469 observations.[7] All data are obtained form Datastream, except the industrial production which is collected from the FRED database of St. Louis Fed. The dependent variable is the equity premium, $r_{t+1} = y_{t+1} - y_{ft}$, that is the total rate of return on the stock market, $y_{t+1} = (\ln SR_{t+1} - \ln SR_t) * 100$, minus the one-month T-bill rate, calculated from the annualised three-month interbank rate $3MIR$ as $y_{ft} = [(1 + 3MIR_t)^{1/12} - 1]$.

We use the same set of independent variables as in Guidolin *et al.* (2014): dividend yield (DY); price-earning ratio (PER); short term interest rate (3MIR); term spread (TMS); CPI inflation rate (INFL);

[7] The beginning of our sample period is dictated by data availability of all variables.

industrial production (IP); exchange rate (ER); unemployment rate (UR); and oil prices (OP).[8] Table 1 gives full description on the exact sources used in the paper and on their mnemonics. Table 2 provides summary statistics of the variables. Data on equity premium returns display typical features well-known in the literature. In annualised terms, mean equity premium returns are 5.5% with a volatility of 18.4%. Noticeable, the equity premium series exhibits significant autocorrelation in squares, but not in levels. Figures 1 and 2 plot the monthly return series for the equity premium and the predictor variables over time, respectively, while the vertical grey bars depict OECD periods of peaks and troughs. [9] During the 1973-2012 period the German stock market return index reached an all time high of 2323.54 in January of 2008 and a record low of 73.49 in October of 1974. With regard to predictor variables, all of them exhibit significant deviations from normality as highlighted by the rejection of the null of zero skewness and zero excess kurtosis underlying the Jarque-Bera's test.

4 Empirical Results

4.1 In-Sample Analysis

Table 3 reports in-sample estimation results for individual (Panel A) and multiple (Panel B) variables respectively, over the entire sample period 1973:02 to 2012:01. In Panel A each row uses a different predictor variable. The in-sample R-square - expressed in percentage points - shows that most of the variables have modest predictive power for equity returns. The most successful variables are the term spread and the inflation with an R-square of 0.99% and 1.12% respectively. Both of them are statistically significant at the 5% level.[10]

Most of the variables have the expected sign. As Fama and French (1988) document the dividend yield can positively forecast future returns on equity risk premium, while on the other hand, increases in the price-earnings ratio, short term interest rate, inflation rate, and exchange rate will decrease future returns. The latter, the negative relation between inflation and the equity premium is particularly highlighted by Fama and Schwert (1977). The authors also report evidence that higher level of industrial production is expected to yield positive returns on equity premium. Finally, changes in oil prices are not good news for future returns.

Results from the in-sample regression analysis for multiple variables are exposed in Panel B. All coefficients are at the same magnitude and sign as before, except the dividend yield which now has a negative effect on excess returns. In this model the null hypothesis that there is no serial correlation in the residuals, as well as the null of no heteroskedasticity are always failed to be rejected. On the other hand, we reject the null hypothesis of normality on the standardised residuals.

[8] Inflation rate, industrial production and unemployment rate are seasonally adjusted.

[9] The OECD peak and trough dates are available at:
http://www.oecd.org/std/leadingindicatorsandtendencysurveys/germany-clicomponentseriesturningpoints.htm The German economy was in recession for approximately 44% of the period during 1973:01-2012:01.

[10] For all variables the null hypotheses of no serial correlation and no heteroskedasticity are always rejected.

4.2 Statistical Evaluation

4.2.1 Individual Variables

We consider three different out-of-sample forecast evaluation periods, from 1985:01 to 2012:01, from 1992:01 to 2012:01 and from 2007:01 to 2012:01. These periods correspond to the last ≈70%, ≈50% and ≈10% observations of our sample. The consideration of multiple out-of-sample periods will help us assess the robustness of the out-of-sample forecasting results and especially, whether or not any predictability pattern that may be arise over the whole out-of-sample period is persistent or not over the subperiods. Results for the three out-of-sample periods are presented at Panel A, Panel B and Panel C, respectively, of Table 4.

From Panel A, the term spread, TMS, generates the lowest $MSPE$, while the dividend yield, DY, has the highest $CORR^2$. Regarding the R^2_{OS}, for seven out of nine predictor variables the statistic is negative, indicating that predictive regression model performs worse than the historical benchmark in terms of mean square prediction errors ($MSPE_P > MSPE_B$). For the two predictors with positive R^2_{OS}, TMS (0.64%) and UR (0.09%), the Clark and West (2007) *adjusted*-$MSPE$ statistic shows significant evidence of predictability only from the term spread, which is also statistically significant at 5% level. Moreover, the large negative values (in absolute terms) of the R^2_{OS} statistic, indicate that the historical average benchmark model outperforms the individual predictive models by a substantial margin. For example the R^2_{OS} statistic for DY and IP reaches 0.95% and 0.86% in absolute terms, signaling that these predictors are outperformed by the historical average benchmark model. This result is consistent with findings at Welch and Goyal (2008) and Rapach *et al.* (2010).

Figure 3 presents differences between the cumulative square prediction error for the historical benchmark model forecasts and the cumulative square prediction error for the predictive regression model forecasts on each individual variable separately. The graphs provide just a informative visual impression about the consistency of the out-of-sample forecasting performance of the predictive model over time and should be treated with caution (Welch and Goyal, 2008). These figures can be used in order to determine which model performs better in terms of $MSPE$, by simply comparing the height of the curve at the beginning and end points of the segment corresponding to the period of interest. Positive values - when the curve is higher at the end of the segment relative to the beginning - indicate that the predictive model has outperformed the historical average benchmark model. Put it differently, a positive slope implies that the predictive regression forecast has lower forecasting error than the historical average in a given period. By looking in this figure it is difficulty to draw a solid conclusion and identify a single model that consistently outperforms the random walk for an entire out-of-sample period.

In order to illustrate how stock return forecasts vary over time, Figure 4 graphs individual predictive regression model forecasts along with the historical average benchmark for the 1985:01-2012:01 out-of-sample period. Overall, these plots confirm that predictive regression forecasts are often highly volatile and that there are many false signals and quite substantial "noise" in the individual forecasts (Rapach *et* *all.*, 2010). Therefore, there is no clear pattern regrading the performance of the models, and it is very difficult to identify individual economic variables capable to generate reliable out-of-sample forecasts of the equity premium.

Moving to the out-of-sample period 1992:01 to 2012:01, Panel B shows that most of the models have lower mean square prediction error and higher squared correlation coefficient than over the full out-of-sample period. The R^2_{OS} takes positive values for TMS, ER and UR, while the *adjusted-*$MSPE$ statistic detects significant evidence of predictability only from the term spread. Similarly, Panel C which reports results for the 2007:01-2012:01 out-of-sample period confirms the predictive power of term spread.

To conclude, in this section we found evidence of German equity premium predictability, not only over the full out-of-sample period, but over the subsample periods as well. To be more precise, predictability from the term spread, which describes the difference between the ten-year government bond and the three-month interbank rate, appears to be long-lived and very persistent.[11] This is consistent with arguments put forward in Campbell and Yogo (2006), Ang and Bekaert (2007), Hjalmarsson (2010) and Rapach and Zhou (2012), that the term spread is a fairly robust predictor of equity premium. On the other hand, this result contradicts the evidence put forward by Timmermann (2008) about short-lived periods of return predictability, what the author calls "elusive predictability".

4.2.2 Multiple Variables

Table 5 presents results for monthly equity premium forecasts based on eleven different models: the historical average benchmark, the kitchen sink, three simple averaging technics (*mean*, *trimmed-mean* and *median*), and six different specifications of discounting methods ($DMSPE$). As expected (Welch and Goyal, 2008; Rapach *et al.,* 2010; Rapach and Zhou, 2012), the "kitchen sink" model performs poorly in terms of R^2_{OS} statistic - values range from -3.19% for the full out-of-sample period to -10.30% for the short 2007:01-2012:01 out-of-sample period - with an insignificant *adjusted-*$MSPE$. In the following rows, even though the simplest combining scheme ($Mean$) and the $DMSPE$ deliver positive values of R^2_{OS}, the Clack and West (2007) *adjusted-*$MSPE$ is insignificant. Panels B and C suggest that for all the combining methods that have been considered in this exercise, the out-of-sample R^2 is always negative and insignificant. A more careful look points out that the $DMSPE$ combination forecasts select weights relatively close to the naive $1/N$ rule. In particular this happens only for the case where the combining weights computed over an out-of-sample holdout period (1985:01-1998:06).

Figure 5 plots the cumulative differences between squared forecast errors from the historical average benchmark model and the predictive regression model based on the combination methods, while Figure 6 graphs the forecasts themselves. By looking in panel A of Figure 5, there is no clear conclusion (either negative or positive trend in the slope) to be drawn whether the kitchen sink forecast underperforms or overperforrms the historical average benchmark model in terms of $MSPE$. This is also clear if we look at

[11] This finding is fairly robust under different forecast horizons; 3-, 6- and 12-step ahead. Results are reported at Appendix C.

panel A of Figure 6, where the kitchen sink forecasts are highly volatile, more so than any of the individual bivariate predictive regression forecasts in Figure 4. Particularly, someone has to note the difference in the vertical axis scaling between figures 4 and 6. The kitchen sink model forecasts nearly 12% in monthly expected equity premium at the beginning of 1991, while two years later falls to -4%.[12] Such extreme values are clearly reflected to the $MSPE$ which delivers the largest value, 35.1187 (Panel A of Table 5), among the combination methods. This is due to the "in-sample over fitting" problem that causes highly parameterised models to produce large forecast errors. Finally, panels B-D of Figure 6 show that the mean, the trimmed-mean and the median combining methods shrink the forecast toward the historical average. The observed stabilisation of these three forecasts compared to many of the individual bivariate predictive regression forecasts of Figure 4, is simply accomplished by weighting the forecasts. These methods seem to be necessary in order to accommodate the uncertainty and instability describing stock returns.

To conclude, in this section we found that prediction methods based on multiple variables do not provide any statistical significant predictive pattern. By incorporating information from multiple predictor variables do not produce forecasts that are statistically plausible. Next section will examine whether there is any economic evidence supporting the general argument that combination methods overperforrm the historical average.

4.3 Economic Evaluation

Results from the economic evaluation of out-of-sample predictability from individual and multiple variables are reported in tables 6 and 7. At a first glance, only one variable, the term spread (TMS) maximise the Sharpe ratio (SH) and the Sortino ratio (SO), over the three out-of-sample periods. On the other hand in terms of utility gains, none of the predictor variables in panels A and B is able to generate positive gains for a mean-variance investor. The only case where there are positive utility gains, relative to the historical average benchmark model, is for TMS and UR over the shorter out-of-sample period 2007:01-2012:01. In annualised terms, the utility gain for the term spread is greater than 1% (1.11%) when $\gamma = 5$. This implies that an investor would be willing to pay more than 100 basis points to have access to the information from the predictive regression forecast compared to the historical average forecast. This is coming to add to the fact that the average utility gains are typically higher during recessions than expansions. Consider, for example, TMS, which generates a negative utility gain of -12.6% in annualised terms during the 1985:01-2012:01 forecast evaluation period when the risk aversion coefficient γ is 0.5, while during the 2007:01-2012:01 period the annualised out-of-sample gain is 11.5%. These results are in line with findings reported in Table 4, indicating that only one variable, the term spread, is able to identify predictability patterns.

Table 7 shows results from the combining methods. It is very difficult to identify forecast combination methods with multiple variables capable of generating reliable and consistent out-of-sample forecasts of

[12] These findings are in line with the results of Welch and Goyal (2008) and Rapach and Zhou (2012), in which the kitchen sink forecast implies an annualised expected equity premium of nearly 48% and -50%.

the equity premium. These findings indicate that combining forecasts fail to outperform the historical average benchmark model, both statistically and economically for a variety of out-of-sample periods.

4.4 Forecast Near Peaks and Troughs

In order to better understand if out-of-sample gains are concentrated in recessions or expansions, we examine the behavior of the actual equity premium forecasts around business cycle peaks and troughs. Following Neely *et al.* (2014), we first estimate a regression model around peaks (beginnings of recessions)

$$r_t - \widehat{r}_t^B = a_P + \sum_{l=-2}^{4} b_{P,l} I_{l,t}^P + \varepsilon_{P,t}, \tag{28}$$

where $I_{l,t}^P$ is an indicator variable that takes a value of unity l months after an OECD-dated peak and zero otherwise, and $b_{P,l}$ is the coefficient which measures the change in the average difference between the forecasts from the realised equity risk premium and the historical average benchmark model l months after a peak. Next we estimate the difference between a forecast based on a predictor variable relative to the historical average benchmark forecast l months after a peak

$$\widehat{r}_t^m - \widehat{r}_t^B = a_P + \sum_{l=-2}^{4} b_{P,l} I_{l,t}^P + \varepsilon_{P,t}, \quad m = P, i, c. \tag{29}$$

We follow similar procedure around troughs (ends of recessions)

$$r_t - \widehat{r}_t^B = a_T + \sum_{l=-2}^{4} b_{T,l} I_{l,t}^T + \varepsilon_{T,t}, \tag{30}$$

$$\widehat{r}_t^m - \widehat{r}_t^B = a_T + \sum_{l=-2}^{4} b_{T,l} I_{l,t}^T + \varepsilon_{T,t}, \quad m = P, i, c, \tag{31}$$

where $I_{l,t}^T$ is an indicator variable equal to unity l months after an OECD-dated trough and zero otherwise.

The top-left panel of Figure 7 depicts estimates of the slope coefficients in equation 28, while the remaining panels depict estimates for equation 29 based on individual predictor variables. The first panel shows that the actual equity premium moves below the historical average one month before and two to three months after a peak, while on the other hand it moves above the historical average two months before the peak, during the peak, as well as one and four months after the peak. Most of the predictor variables fail to pick up these fluctuations in the equity risk premium early in recessions. The PER, the $3MIR$ and the OP forecasts are above the historical average during the month of a peak and also one month after a peak, matching the higher-than-average actual equity premium for those months. However, the OP forecast is also significantly higher than the historical average for the full period of two months before to four months after a peak, unlike the actual equity risk premium. What is more, the TMS and the UR, are always below the historical average and do detect the decline in the actual equity premium one month before and two to three months after a peak.

Figure 9 plots the slope coefficient estimates for equation 30 - top-left panel - and equation 31 - remaining panels - based on individual economic variables. The first panel shows that the actual equity premium tends to move significantly below the historical average benchmark forecast four to two months before a business-cycle trough, while it tends to move above the historical average benchmark forecast one month before through one month after a trough. The remaining panels show that the majority of the economic variables fail to pick up these movements in the equity premium late in recessions. In particular, forecasts based on variables such as the $3MIR$, the TMS and the UR are significantly above the historical average benchmark forecast for any of the months in the late stages of recessions when the equity premium itself is higher than average, although the size of the increase in the TMS forecast is very small. None of the nine economic variables is able to match the lower-than-average actual equity premium for the four to two months before a trough. However, the $3MIR$ and the UR forecasts are also higher than the historical average benchmark forecast for the whole period under investigation, unlike the actual equity premium.

In conclusion, the analysis of the forecast behavior near beginnings and ends of recessions, highlights the out-of-sample gains from the term spread and the unemployment rate. These two predictors can detect the decline in the actual equity premium early in recessions, as well as the increase in the actual equity premium late in recessions.

5 Conclusion

This paper examines whether Welch and Goyal's (2008) argument - in sample predictability of stock returns from some economic predictor variables fails to deliver consistent out-of-sample forecasting gains relative to the historical average benchmark forecast in terms of MSPE - holds or not for German stock returns. Furthermore, it explores whether Bates and Granger's (1969) finding - combining forecasts across models often produces a forecast that performs better than the best individual model - can be validated by statistical and economic criteria.

The results show that only one variable, the term spread, has in-sample and out-of-sample forecasting power and consistent outperform the historical average benchmark model. In addition, this variable recognises the typical drop in the equity premium near business cycle peaks, as well as the typical increase in the equity premium near business cycle troughs. Finally, there is evidence that combinations of individual model forecasts do not deliver any statistical and economic significant out-of-sample gains relative to the historical average on a consistent basis over time.

This paper provide evidence that there is a predictable component in stock returns, which is captured by the term spread – yield curve. This finding has important implications for monetary policy and investor expectations. Monetary policy can influence the slope of the yield curve. A tightening of monetary policy usually means a rise in short-term interest rates. Other things being equal, that would tend to flatten or invert the yield curve at the same time that it chokes the supply of credit to the economy and produces and economic slowdown. On the other hand, long-term interest rates are determined more directly by investor behaviour. If investors are risk averse, they tend to flock toward the safety of bonds when they sense that an economic downturn is on the horizon. That tends to push long-term interest rates below short-term interest rates when a recession is approaching. Therefore, the yield curve is a simple tool for reading the collective mind of the stock market both near peaks and troughs.

We conclude by suggesting avenues for future research. The literature on stock return forecasting primarily relies on popular economic variables as predictors. However, other variables that potentially contain relevant information for forecasting stock returns have received less attention. Such variables include options, features and other derivative prices; microstructure measures of liquidity; and institutional trading variables such as trading volumes and money flows for mutual and hedge funds. In addition, learning appears to play an important role in stock return predictability (Timmermann, 1993; Pastor and Veronesi, 2009). Theoretical models that explain how investors form return forecasts in light of available information and respond to their forecasting errors serves as a promising building blocks for forecasting models based on learning. Finally, recent studies find significant in-sample evidence of a positive relationship between expected returns and risk (Guo and Whitelaw, 2006; Lundblad, 2007; Bali, 2008). It would be interesting to examine whether these approaches could be used to generate reliable out-of-sample stock return forecasts based on the expected risk-return relationship (Ludvigson and Ng, 2007).

REFERENCES

[1] Ang, A. and G. Bekaert (2002), "Regime Switches in Interest Rates", *Journal of Business and Economic Statistics*, 20(2): 163-182.

[2] Ang, A. and G. Bekaert (2007), "Stock Return Predictability: Is It There?", *Review of Financial Studies*, 20(3): 651-707.

[3] Bali, G. (2008), "The Intertemporal Relation between Expected Returns and Risk", *Journal of Financial Economics*, 87(1): 101-131.

[4] Bates, J. M. and C. W. J. Granger (1969), "The Combination of Forecasts", *Operational Research Quarterly*, 20(4): 451-468.

[5] Boyd, J. H., J. Hu and R. Jagannathan (2004), "The Stock Market's Reaction to Unemployment News: Why Bad News Is usually Good for Stocks", *Journal of Finance*, 60(2): 649-672.

[6] Campbell, J. Y. (1987), "Stock Returns and the Term Structure", *Journal of Financial Economics*, 18(2): 373-399.

[7] Campbell, J. Y. (2008), "Viewpoint: Estimating the Equity Premium", *Canadian Journal of Economics*, 41(1): 1-21.

[8] Campbell, J. Y. and R. K. Shiller (1988), "Stock Prices, Earnings, and Expected Dividends", *Journal of Finance*, 43(3): 661-676

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

[10] Campbell, J. Y. and T. Vuolteenaho (2004), "Inflation Illusion and Stock Prices", *American Economic Review*, 94(2): 19-23.

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

[12] Casassus, J. and F. Higuera (2011), "Stock Return Predictability and Oil Prices", PUC Economics Institute Working Paper No. 406.

[13] Chen, L. and L. Zhang (2009), "The Stock Market and Aggregate Employment", National Bureau of Economic Research Working Paper No. 15219.

[14] Clark, T. E. and M. W. McCraken (2001), "Tests of Equal Forecast Accuracy and Encompassing for Nested Models", *Journal of Econometrics*, 105(1): 85-110.

[15] Clark, T. E. and K. D. West (2007), "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models", *Journal of Econometrics*, 138(1): 291-311.

[16] Clemen, R. T. (1989), "Combining Forecasts: A Review and Annotated Bibliography.", *International Journal of Forecasting*, 50(4): 559-583.

[17] Cochrane, J. H. (1992), "Explaining the Variance of Price-Dividend Ratios", *Review of Financial Studies*, 5(2): 243-280.

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

[19] Della Corte, P. and I. Tsiakas (2012), "Statistical and Economic Methods for Evaluating Exchange Rate Predictability", In: James, J., I. W. Marsh and L. Sarno, (Eds), Handbook of Exchange Rates, Wiley-Blackwell, United Kingdom.

[20] Detzer, D., N. Dodig, T. Evans, E. Hein, and H. Herr (2013), "The German Financial System", *FESSUD Studies in Financial Systems* No. 3.

[21] Diebold, R. T. and R. S. Mariano (1995), "Comparing Predictive Accuracy", *Journal of Business & Economic Statistics*, 13(3): 253-263.

[22] Driesprong, G., B. Jacobsen and B. Maat (2008), "Striking Oil: Another Puzzle?", *Journal of Financial Economics*, 89(2): 307-327.

[23] Fama, E. and K. French (1988), "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, 22(1): 3-25.

[24] Fama, E. and W. Schwert (1977), "Asset Returns and Inflation", *Journal of Financial Economics*, 5(2): 115-146.

[25] Ferreira, M. A. and P. Santa-Clara (2011), "Forecasting Stock Market Returns: The Sum of the Parts Is More than the Whole", *Journal of Financial Economics*, 100(3): 514-537.

[26] Giacomini, R. and H. White (2006), "Tests of Conditional Predictive Ability", *Econometrica*, 74(6): 1545-1578.

[27] Granger, C. W. J. (1990), "*Modelling Economic Series: Readings in Econometric Methodology*", Oxford University Press, Oxford, UK.

[28] Granger, C. W. J. and Y. Jeon (2004), "Thick Modeling", *Economic Modelling*, 21(2):323-343.

[29] Granger, C. W. J. and M. J. Machina (2006), "Forecasting and Decision Theory", In: Elliott, G. and A. Timmermann (Eds.), Handbook of Economic Forecasting, Vol. 1. Elsevier, Amsterdam.

[30] Granger, C. W. J. and M. H. Pesaran (2000), "Economic and Statistical Measures of Forecast Accuracy", *Journal of Forecasting*, 19(7): 537-560.

[31] Guidolin, M., S. Hyde, D. McMillan, and S. Ono (2014), "Does the Macroeconomy Predict U.K. Asset Returns in a Nonlinear Fashion? Comprehensive Out-of-Sample Evidence", *Oxford Bulletin of Economics and Statistics*, 76(4): 510-535.

[32] Guo, H. and F. Whitelaw (2006), "Uncovering the Risk-Return Relation in the Stock Market", *Journal of Finance*, 61(3): 1433-1463.

[33] Han, Y. (2006), "Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model", *Review of Financial Studies*, 19(1): 237-271.

[34] Hansen, P. R. and A. Timmermann (2012), "Choice of Sample Split in Out-of-Sample Forecast Evaluation", European University Institute Working Paper ECO 2012/10.

[35] Hendry, D. F. and M. P. Clements (2004), "Pooling of Forecasts", *Econometrics Journal*, 7(1): 1-31.

[36] Hjalmarsson, E. (2010), "Predicting Global Stock Returns", *Journal of Financial and Quantitative Analysis*, 45(1): 49-80.

[37] Leitich, G. and J. E. Tanner (1991), "Economic Forecast Evaluation: Profits Versus the Conventional Error Measures", *American Economic Review*, 81(3): 580-590.

[38] Ludvigson, C. and S. Ng (2007), "The Empirical Risk-Return Relation: A Factor Analysis Approach", *Journal of Financial Economics*, 83(1): 171-222.

[39] Lundblad, C. (2007), "The Risk Return Trade-off in the Long Run: 1836-2003", *Journal of Financial Economics*, 85(1): 123-150.

[40] Marquering, W. and M. Verbeek (2004), "The Economic Value of Predicting Stock Index Returns and Volatility", *Journal of Financial and Quantitative Analysis*, 39(2): 407-429.

[41] McCracken, M. W. (2007), "Asymptotics for Out Of Sample Tests of Granger Causality", *Journal of Econometrics*, 140(2): 719-752.

[42] Neely, C. J., D. E. Rapach, J. Tu and G. Zhou (2014), "Forecasting the Equity Risk Premium: The Role of Technical Indicators", *Management Science*, 60(7): 1772-1791.

[43] Newbold, P. and C. W. J. Granger (1974), "Experience with Forecasting Univariate Time Series and the Combination of Forecasts", *Journal of the Royal Statistical Society*, 137(2): 131-165.

[44] Pastor, L. and P. Veronesi (2009), "Learning in Financial Markets", *Annual Review of Financial Economics*, 1(1): 361-381.

[45] Paye, B. S. and A. Timmermann (2006), "Instability of Return Prediction Models", *Journal of Empirical Finance*, 13(3): 274-315.

[46] Pesaran, M. H. and A. Timmermann (1995), "Predictability of Stock Returns: Robustness and Economic Significance", *Journal of Finance*, 50(4): 1201-1228.

[47] Pesaran, M. H. and A. Timmermann (2000), "A Recursive Modelling Approach to Predicting UK Stock Returns", *Economic Journal*, 110(460): 159-191.

[48] Rapach, D. E. and J. K. Strauss (2008), "Forecasting US Employment Growth Using Forecast Combining Methods", *Journal of Forecasting*, 27(1): 75-93.

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

[50] Rapach, D. E. and G. Zhou (2012), "Forecasting Stock Returns", In: Elliott, G. and A. Timmermann (Eds.), Handbook of Economic Forecasting, Vol. 2. Elsevier, Amsterdam.

[51] Rapach, D. E., J. K. Strauss 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.

[52] Rapach, D. E., M. E. Wohar and J. Rangvid (2005), "Macro Variables and International Stock Return Predictability", *International Journal of Forecasting*, 21(1): 137-166.

[53] Sarno, L., D. Thornton and G. Valente (2005),"Federal Funds Rate Prediction", *Journal of Money Credit and Banking*, 37(3): 449-471.

[54] Smith, J. and K. F. Wallis (2009), "Simple Explanation of the Forecast Combination Puzzle", *Oxford Bulletin of Economics and Statistics*, 71(3):331-355.

[55] Stock, J. H. and M. W. Watson (1999), "Forecasting Inflation", *Journal of Monetary Economics*, 44(2): 293-335.

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

[57] Stock, J. H. and M. W. Watson (2006), "Forecasting with Many Predictors", In: Elliott, G., C. Granger and A. Timmermann (Eds.), Handbook of Economic Forecasting, Vol. 1. Elsevier, Amsterdam.

[58] Timmermann, A. (1993), "How Learning in Financial Markets Generates Excess Volatility and Predictability of Stock Returns", *Quarterly Journal of Economics*, 108(4): 1135-1145.

[59] Timmermann, A. (2008), "Elusive Return Predictability", *International Journal of Forecasting*, 24(1): 1-18.

[60] Weigand, R. A. and R. Irons (2007), "The Market P/E Ration, Earnings Trends, and Stock Return Forecasts", *Journal of Portfolio Management*, 33(4): 87-101.

[61] Welch, I. and A. Goyal (2008), "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction", *Review of Financial Studies*, 21(4): 1455-1508.

[62] West, K. D. (1996), "Asymptotic Inference about Predictive Ability", *Econometrica*, 64(5): 1067-1084.

APPENDIX A

Table 1. Data

Variable	Source	Mnemonic/Code
Stock Return (SR) $100 \times [\ln(p_t) - \ln(p_{t-1})]$	Total Market Index, Datastream	TOTMKBD(IR)
Dividend Yield (DY) $\ln(\frac{DY_t}{100})$	Total Market Index, Datastream	TOTMKBD(DY)
Price-Earnings Ratio (PER) $\ln(PE_t)$	Total Market Index, Datastream	TOTMKBD(PE)
Change in Short-term Interest Rate (3MIR) $ir_t - ir_{t-1}$	3 Month Interbank Rate, Datastream	BDINTER3
Term Spread (TMS) $gb_t - ir_t$	10 Year Government Bond, Datastream	BDI61...
Inflation (INFL) $100 \times [\ln(p_t) - \ln(p_{t-1})]$	Consumer Price Index, Datastream	BDUUFA01F
Industrial Production (IP) $100 \times [\ln(p_t) - \ln(p_{t-1})]$	Production of Total Industry, Federal Reserve Bank of St. Louis	DEUPROINDMISMEI
Exchange Rate (ER) $100 \times [\ln(p_t) - \ln(p_{t-1})]$	Nominal Effective Exchange Rate, Datastream	BDI..NECE
Change in Unemployment Rate (UR) $un_t - un_{t-1}$	Unemployment Rate, Datastream	BDOUN013Q
Change in Oil Prices (OP) $100 \times [\ln(p_t) - \ln(p_{t-1})]$	World Crude Petroleum Price, Datastream	WDI76AADF

Table 2. Descriptive Statistics

Variables	Maximum	Minimum	Mean	Median	St. Dev.	Skewness	Kurtosis	JB statistic	LB(4)	LB(4) squares
EP	14.8865	-26.7898	0.4564	0.7389	5.3159	-0.7979	5.3640	158.6476***	7.1487	183.33***
DY	-2.9603	-4.3981	-3.6895	-3.7550	0.3406	0.0907	2.1097	16.0956***	1720.1***	97.245***
PER	3.2884	2.1282	2.6945	2.6810	0.2457	-0.0432	2.3273	8.9686**	1446.9***	93.553***
$3MIR$	2.9300	-1.4200	-0.0142	0.0000	0.3590	1.4471	16.5944	3767.114***	163.61***	180.66***
TMS	4.4100	-4.8100	0.9912	1.3450	1.5675	-0.8069	3.9972	70.1852***	1574.3***	86.154***
$INFL$	1.4847	-2.1323	0.2169	0.1984	0.2676	-0.6070	17.3195	4027.215***	135.95***	238.55***
IP	11.7075	-9.9931	0.1208	0.1498	1.7336	-0.1354	10.4162	1073.935***	44.512***	276.84***
ER	6.6065	-2.8992	0.1454	0.0364	1.0322	1.1157	7.2290	445.8560***	47.721***	148.73***
UR	1.9000	-0.5000	0.0128	0.0000	0.1439	5.0649	66.0051	79409.01***	136.28***	221.59***
OP	137.0546	-31.2159	0.7644	0.0000	10.1994	5.1224	71.3443	93130.31***	15.672***	180.43***

Note: The sample period is from February 1973 to January 2012, a total of 468 observations. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 3. In-Sample Analysis

Panel A: Individual Predictor Variables

	Coefficient	St. Error	OLS T-Ratio	P-Value	R^2_{IS} (%)
DY	0.1430	0.7230	0.1978	0.8433	0.0084
PER	-0.6269	1.0042	-0.6242	0.5328	0.0836
$3MIR$	-0.8506	0.6852	-1.2413	0.2151	0.3303
TMS	0.3383	0.1564	2.1632**	0.0310	0.9964
$INFL$	-2.1046	0.9152	-2.2294**	0.0219	1.1243
IP	0.0541	0.1425	0.3802	0.7040	0.0311
ER	-0.2143	0.2389	-0.8972	0.3700	0.1728
UR	2.1963	1.7081	1.2858	0.1991	0.3543
OP	-0.0220	0.0241	-0.9144	0.3609	0.1795

Panel B: Multiple Predictor Variables

	Coefficient	St. Error	OLS T-Ratio	P-Value
DY	-0.1603	1.0737	-0.1493	0.8814
PER	-1.4532	1.4305	-1.0158	0.3102
$3MIR$	-0.4107	0.7091	-0.5793	0.5627
TMS	0.2962	0.1671	1.7723*	0.0770
$INFL$	-1.9369	1.0286	-1.8829*	0.0603
IP	0.0831	0.1443	0.5757	0.5651
ER	-0.1995	0.2401	-0.8311	0.4063
UR	2.9308	1.7935	1.6342	0.1029
OP	-0.0128	0.0245	-0.5226	0.6015
R^2	0.0301			
$\overline{R}^2$	0.0110			

Diagnostic Tests

Serial Correlation	$\chi^2(12)$=12.9952 (0.3694)	F(12,444)=1.0614 (0.3913)
Normality	$\chi^2(2)$=190.1624 (0.0000)	Not applicable
Heteroscedasticity	$\chi^2(9)$=9.7587 (0.3704)	F(9,546)=1.0837 (0.3732)

Note: The table shows results from the in-sample regression analysis. Stock market return predictability is tested for each predictor variable separately (Panel A), and for all the predictor variables together (Panel B). In-sample R-squares are estimated over the full-sample period, following Ferreira and Santa-Clara (2011). ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively. $\overline{R}^2$ denotes the adjusted R^2.

Table 4. 1-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Method	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	34.0331	0.0152	-	-
DY	34.3586	0.0090	-0.9565	-1.3686*
PER	34.2120	0.0068	-0.5258	-0.6529
$3MIR$	34.1984	0.0051	-0.4785	-0.3771
TMS	33.8167	0.0028	0.6429	1.7799**
$INFL$	34.2687	0.0023	-0.6850	1.0942
IP	34.3257	0.0042	-0.8597	-0.6904
ER	34.0660	0.0003	-0.0895	0.2352
UR	34.0022	0.0003	0.0978	0.6752
OP	34.1389	0.0022	-0.3038	-0.1219

Panel B: January 1992 - January 2012

Method	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	32.3733	0.0067	-	-
DY	32.5407	0.0117	-0.5169	-1.6532**
PER	32.6837	0.0129	-0.9587	-1.1473
$3MIR$	32.6022	0.0110	-0.7092	-0.8098
TMS	31.9955	0.0152	1.1647	2.5182***
$INFL$	32.3827	0.0014	-0.0312	0.6301
IP	32.9109	0.0293	-1.6601	-2.3028**
ER	32.3509	0.0003	0.0671	0.4717
UR	32.2805	0.0005	0.2844	0.9499
OP	32.4444	0.0002	-0.2219	0.0745

Panel C: January 2007 - January 2012

Method	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	40.8333	0.0181	-	-
DY	41.0756	0.1684	-0.5933	-2.2685**
PER	41.5916	0.0595	-1.8571	-1.7311**
$3MIR$	41.6813	0.0921	-2.1064	-1.1605
TMS	39.6941	0.0831	2.7615	2.3124**
$INFL$	40.5605	0.0126	0.6392	0.5627
IP	41.0898	0.0151	-0.6380	-0.5855
ER	42.7886	0.1573	-4.8189	-2.7137***
UR	40.5573	0.0002	0.6469	0.9985
OP	42.4730	0.1090	-4.0457	-1.4795*

Note: This table shows results from the out-of-sample regression analysis from individual variables. $MSPE$ denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 5. 1-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Method	*MSPE*	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	34.0331	0.0152	-	-
Kitchen Sink	35.1187	0.0010	-3.1897	0.7063
Mean	34.0226	0.0016	0.0308	0.3303
Trimmed-Mean	34.0502	0.0058	-0.0502	-0.0782
Median	34.1085	0.0188	-0.2217	-1.0210
DMSPE (θ=1.0)[1]	34.0228	0.0017	0.0303	0.3269
DMSPE (θ=0.9)[1]	34.0202	0.0014	0.0379	0.3584
DMSPE (θ=0.75)[1]	34.0188	0.0012	0.0419	0.3771
DMSPE (θ=1.0)[2]	34.0109	0.0009	0.0651	0.4659
DMSPE (θ=0.9)[2]	34.0175	0.0012	0.0457	0.3956
DMSPE (θ=0.75)[2]	34.0241	0.0015	0.0263	0.3251

Panel B: January 1992 - January 2012

Method	*MSPE*	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	32.3733	0.0067	-	-
Kitchen Sink	33.5418	0.0003	-3.6093	-0.1587
Mean	32.3904	0.0033	-0.0526	-0.0721
Trimmed-Mean	32.3944	0.0046	-0.0652	-0.1418
Median	32.4353	0.0119	-0.1914	-0.6805
DMSPE (θ=1.0)[1]	32.3904	0.0034	-0.0528	-0.0735
DMSPE (θ=0.9)[1]	32.3877	0.0029	-0.0444	-0.0392
DMSPE (θ=0.75)[1]	32.3872	0.0028	-0.0428	-0.0296
DMSPE (θ=1.0)[2]	32.3855	0.0025	-0.0375	-0.0062
DMSPE (θ=0.9)[2]	32.3966	0.0035	-0.0720	-0.1287
DMSPE (θ=0.75)[2]	32.4071	0.0045	-0.1043	-0.2455

Panel C: January 2007 - January 2012

Method	*MSPE*	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	40.8333	0.0181	-	-
Kitchen Sink	45.0403	0.1020	-10.3029	-1.6545*
Mean	41.2061	0.1264	-0.9130	-1.3287*
Trimmed-Mean	41.1915	0.1303	-0.8771	-1.3175*
Median	41.0471	0.1015	-0.5235	-0.9257
DMSPE (θ=1.0)[1]	41.2092	0.1283	-0.9206	-1.3401*
DMSPE (θ=0.9)[1]	41.2120	0.1213	-0.9274	-1.3191*
DMSPE (θ=0.75)[1]	41.2115	0.1153	-0.9262	-1.2937
DMSPE (θ=1.0)[2]	41.1855	0.1026	-0.8624	-1.2190
DMSPE (θ=0.9)[2]	41.1851	0.1100	-0.8616	-1.2542
DMSPE (θ=0.75)[2]	41.1803	0.1148	-0.8498	-1.2806

Note: This table shows results from the out-of-sample regression analysis from individual variables. *MSPE* denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. As the initial holdout out-of-sample period when computing the *DMSPE* forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 6. 1-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0255	0.0331	-2.1423	-1.6875	-0.8750	-0.3496
PER	0.0347	0.0436	-0.2078	-0.1041	-0.0523	-0.0212
$3MIR$	0.0364	0.0449	-1.3056	-0.9879	-0.4951	-0.1995
TMS	0.0968	0.1246	-1.0533	-0.5242	-0.2597	-0.1009
$INFL$	0.0933	0.1248	-6.2761	-6.9708	-3.8543	-1.5424
IP	0.0444	0.0585	-1.6445	-1.0607	-0.5319	-0.2146
ER	0.0725	0.0942	-1.2671	-0.8263	-0.4132	-0.1653
UR	0.0664	0.0867	-0.0006	-0.0801	-0.0406	-0.0169
OP	0.0456	0.0560	-1.2344	-1.3344	-0.6670	-0.2666

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0334	0.0422	-0.9227	-0.4628	-0.2311	-0.0920
PER	0.0079	0.0098	-0.9151	-0.4567	-0.2276	-0.0900
$3MIR$	0.0299	0.0367	-2.0207	-1.2890	-0.6455	-0.2593
TMS	0.1276	0.1696	-0.3252	-0.1607	-0.0785	-0.0292
$INFL$	0.0854	0.1132	-5.3153	-3.5469	-1.7695	-0.7030
IP	-0.0085	-0.0104	-1.5797	-0.9070	-0.4535	-0.1814
ER	0.0850	0.1108	-1.3094	-0.9157	-0.4586	-0.1843
UR	0.0740	0.0984	0.1011	-0.0576	-0.0300	-0.0135
OP	0.0478	0.0588	-1.4997	-1.7188	-0.8593	-0.3437

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	-0.0852	-0.1055	-0.6429	-0.3203	-0.1590	-0.0622
PER	-0.1170	-0.1432	-4.2347	-2.1164	-1.0573	-0.4218
$3MIR$	-0.1164	-0.1378	-4.6357	-3.3482	-1.6703	-0.6636
TMS	0.1476	0.2088	0.9571	0.4769	0.2368	0.0928
$INFL$	0.0216	0.0282	-6.6243	-4.3989	-2.1919	-0.8662
IP	-0.0887	-0.1078	-0.4937	-0.2481	-0.1253	-0.0516
ER	-0.2077	-0.4867	-4.2032	-2.8496	-1.4196	-0.5618
UR	-0.0212	-0.0287	0.2400	0.1199	0.0599	0.0238
OP	-0.1611	-0.1789	-4.6353	-5.8044	-2.8948	-1.1491

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model.

Table 7. 1-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	0.0722	0.0994	-5.9175	-10.4872	-7.2449	-2.9774
Mean	0.0648	0.0833	-1.3138	-0.6695	-0.3349	-0.1342
Trimmed-Mean	0.0591	0.0753	-0.9626	-0.4812	-0.2405	-0.0960
Median	0.0495	0.0624	-0.6045	-0.3021	-0.1508	-0.0601
$DMSPE$ (θ=1.0)[1]	0.0647	0.0834	-1.3010	-0.6615	-0.3309	-0.1326
$DMSPE$ (θ=0.9)[1]	0.0653	0.0843	-1.3628	-0.6963	-0.3483	-0.1396
$DMSPE$ (θ=0.75)[1]	0.0657	0.0846	-1.4196	-0.7298	-0.3651	-0.1463
$DMSPE$ (θ=1.0)[2]	0.0670	0.0864	-1.4306	-0.7337	-0.3671	-0.1471
$DMSPE$ (θ=0.9)[2]	0.0660	0.0850	-1.3770	-0.7014	-0.3509	-0.1406
$DMSPE$ (θ=0.75)[2]	0.0650	0.0838	-1.3432	-0.6826	-0.3415	-0.1369

Panel B: January 1992 - January 2012

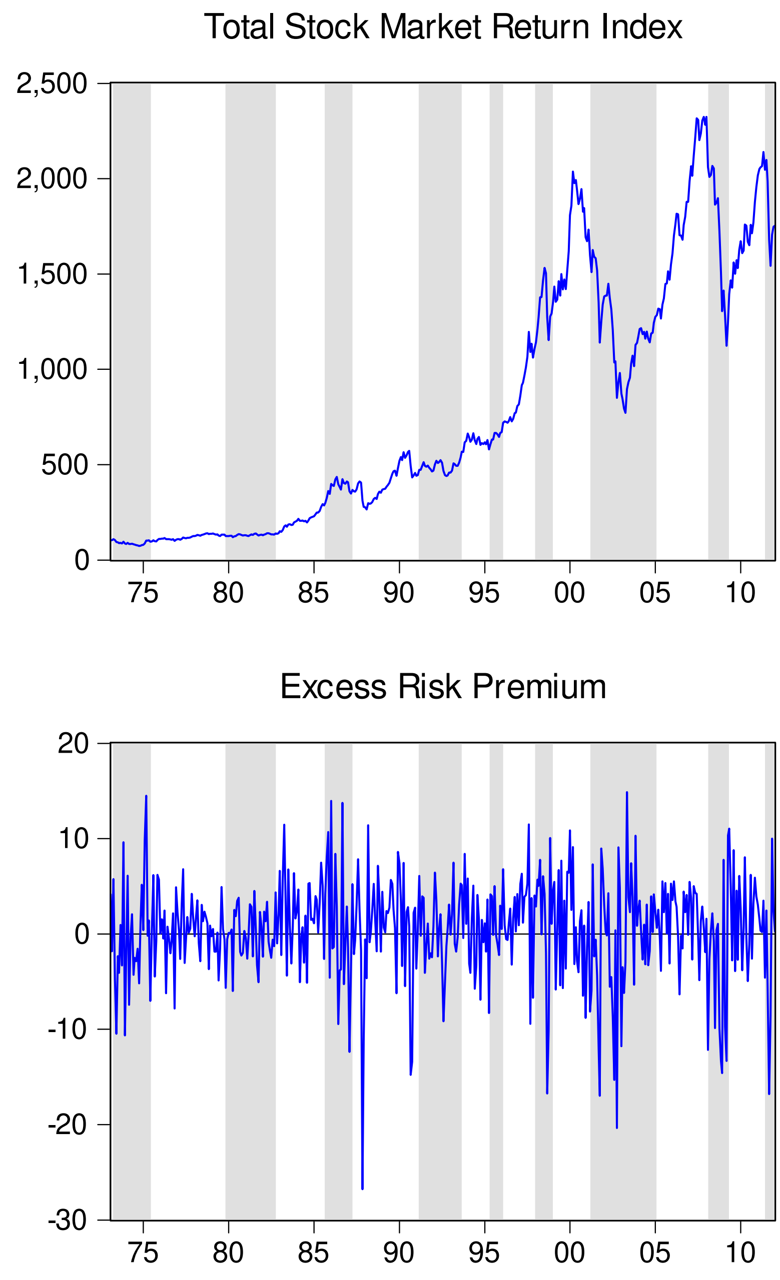
Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	0.0283	0.0366	-6.8329	-9.6590	-6.5410	-2.7159
Mean	0.0567	0.0715	-0.9452	-0.4721	-0.2355	-0.0936
Trimmed-Mean	0.0558	0.0703	-0.8143	-0.4067	-0.2029	-0.0806
Median	0.0493	0.0619	-0.5319	-0.2656	-0.1324	-0.0526
$DMSPE$ (θ=1.0)[1]	0.0567	0.0715	-0.9396	-0.4693	-0.2341	-0.0930
$DMSPE$ (θ=0.9)[1]	0.0573	0.0723	-0.9802	-0.4896	-0.2442	-0.0970
$DMSPE$ (θ=0.75)[1]	0.0575	0.0726	-1.0173	-0.5081	-0.2535	-0.1007
$DMSPE$ (θ=1.0)[2]	0.0580	0.0733	-1.0178	-0.5104	-0.2546	-0.1011
$DMSPE$ (θ=0.9)[2]	0.0562	0.0709	-1.0005	-0.5018	-0.2503	-0.0994
$DMSPE$ (θ=0.75)[2]	0.0546	0.0687	-0.9816	-0.4924	-0.2460	-0.0976

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	-0.1875	-0.2068	-10.1734	-13.1509	-10.3953	-4.1367
Mean	-0.0909	-0.1110	-2.2491	-1.1219	-0.5583	-0.2201
Trimmed-Mean	-0.0905	-0.1107	-2.1177	-1.0564	-0.5257	-0.2073
Median	-0.0753	-0.0933	-1.5665	-0.7813	-0.3886	-0.1531
$DMSPE$ (θ=1.0)[1]	-0.0914	-0.1115	-2.2484	-1.1215	-0.5581	-0.2200
$DMSPE$ (θ=0.9)[1]	-0.0911	-0.1111	-2.3064	-1.1504	-0.5725	-0.2257
$DMSPE$ (θ=0.75)[1]	-0.0904	-0.1103	-2.3570	-1.1757	-0.5850	-0.2307
$DMSPE$ (θ=1.0)[2]	-0.0872	-0.1066	-2.2920	-1.1432	-0.5689	-0.2242
$DMSPE$ (θ=0.9)[2]	-0.0879	-0.1074	-2.2199	-1.1073	-0.5510	-0.2172
$DMSPE$ (θ=0.75)[2]	-0.0880	-0.1076	-2.1308	-1.0628	-0.5289	-0.2085

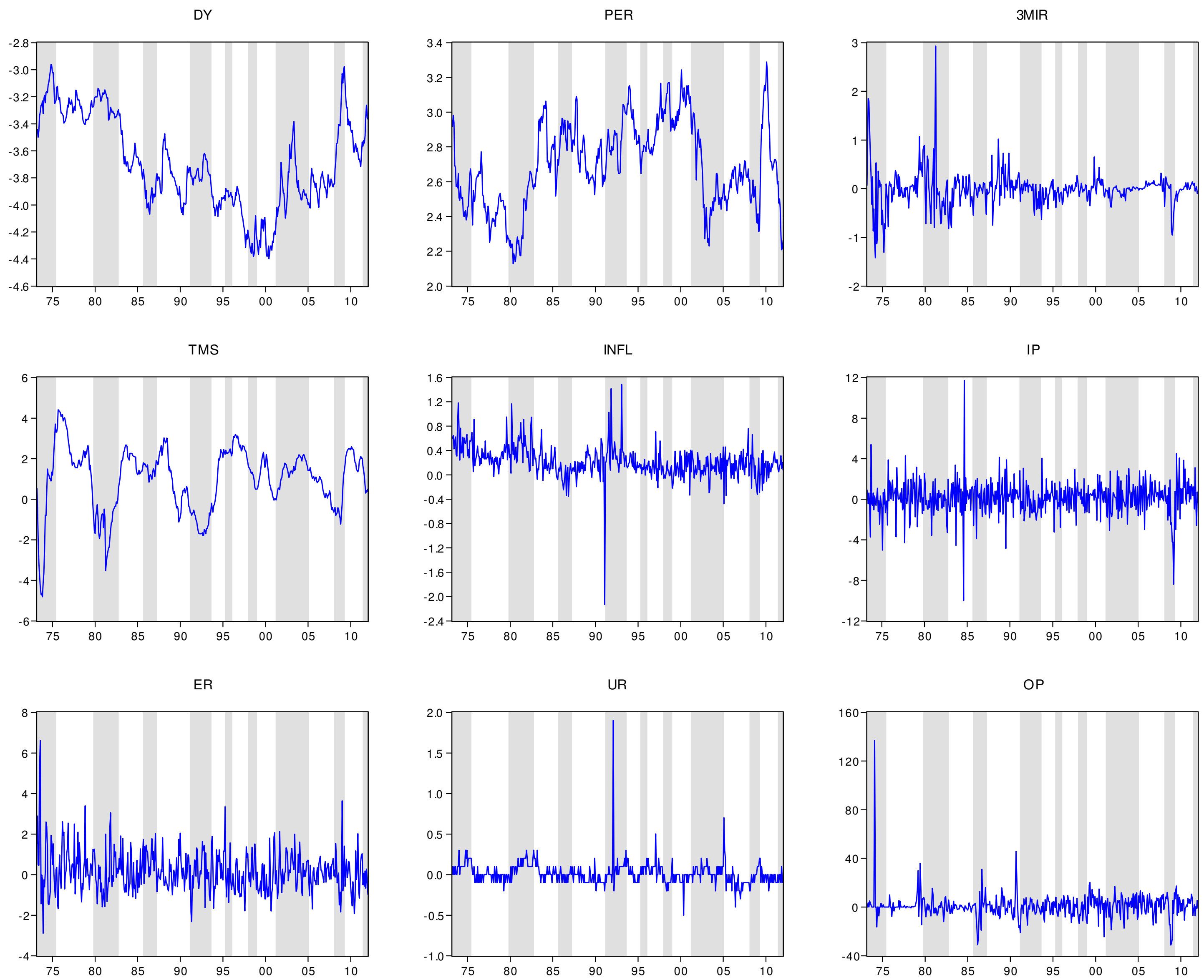
Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model. As the initial holdout out-of-sample period when computing the $DMSPE$ forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used.

Figure 1. Total Stock Market Return Index and Excess Risk Premium



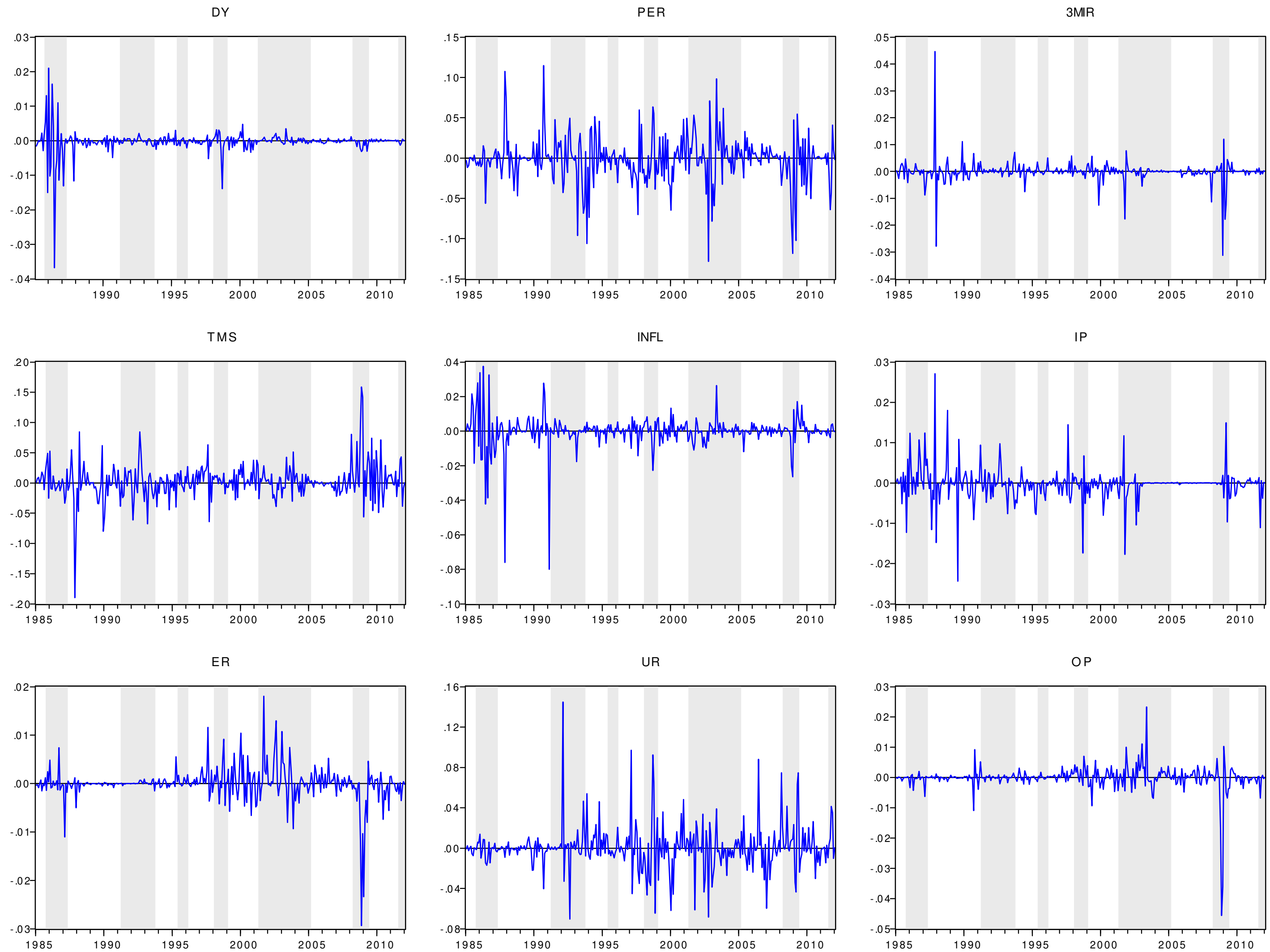
Note: Vertical grey areas depict OECD-dated peaks and troughs.

Figure 2. Return Series of Predictor Variables



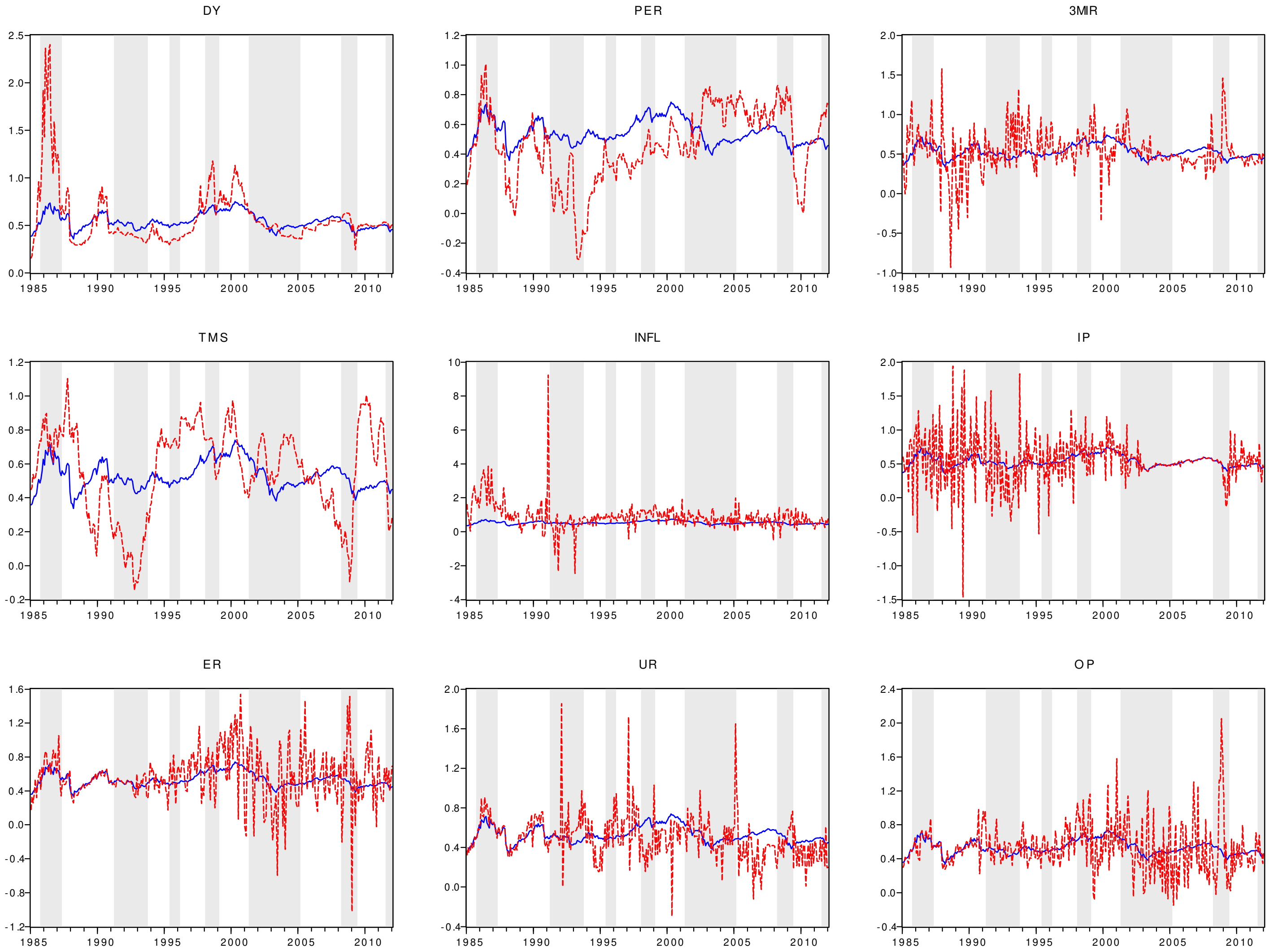
Note: Vertical grey areas depict OECD-dated peaks and troughs.

Figure 3. Cumulative Differences in Squared Forecast Errors, Out-of-Sample Forecasts Based on Individual Predictor Variables, 1985:01-2012:01



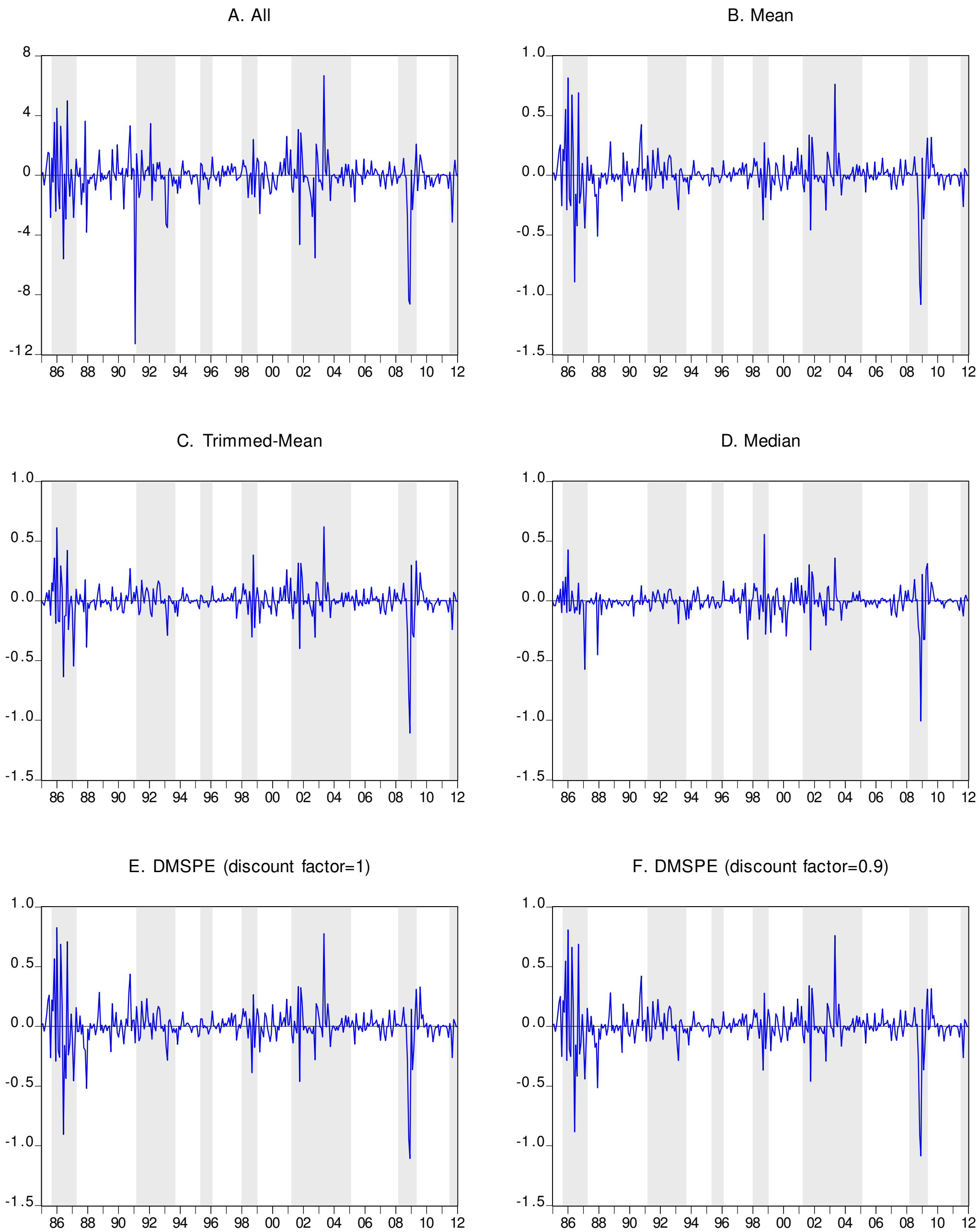
Note: Figures show cumulative differences between squared forecast errors from the historical average benchmark model and the predictive model based on the predictor variable given in the panel heading. Vertical grey areas depict OECD-dated peaks and troughs.

Figure 4. Out-of-Sample Forecasts Based on Individual Predictor Variables, 1985:01-2012:01



Note: Dotted (straight) lines display predictive regression forecasts based on the predictor variable given in the panel heading (historical average). Vertical grey areas depict OECD-dated peaks and troughs.

Figure 5. Cumulative Differences in Squared Forecast Errors, Out-of-Sample Forecasts Based on Multiple Predictor Variables, 1985:01-2012:01

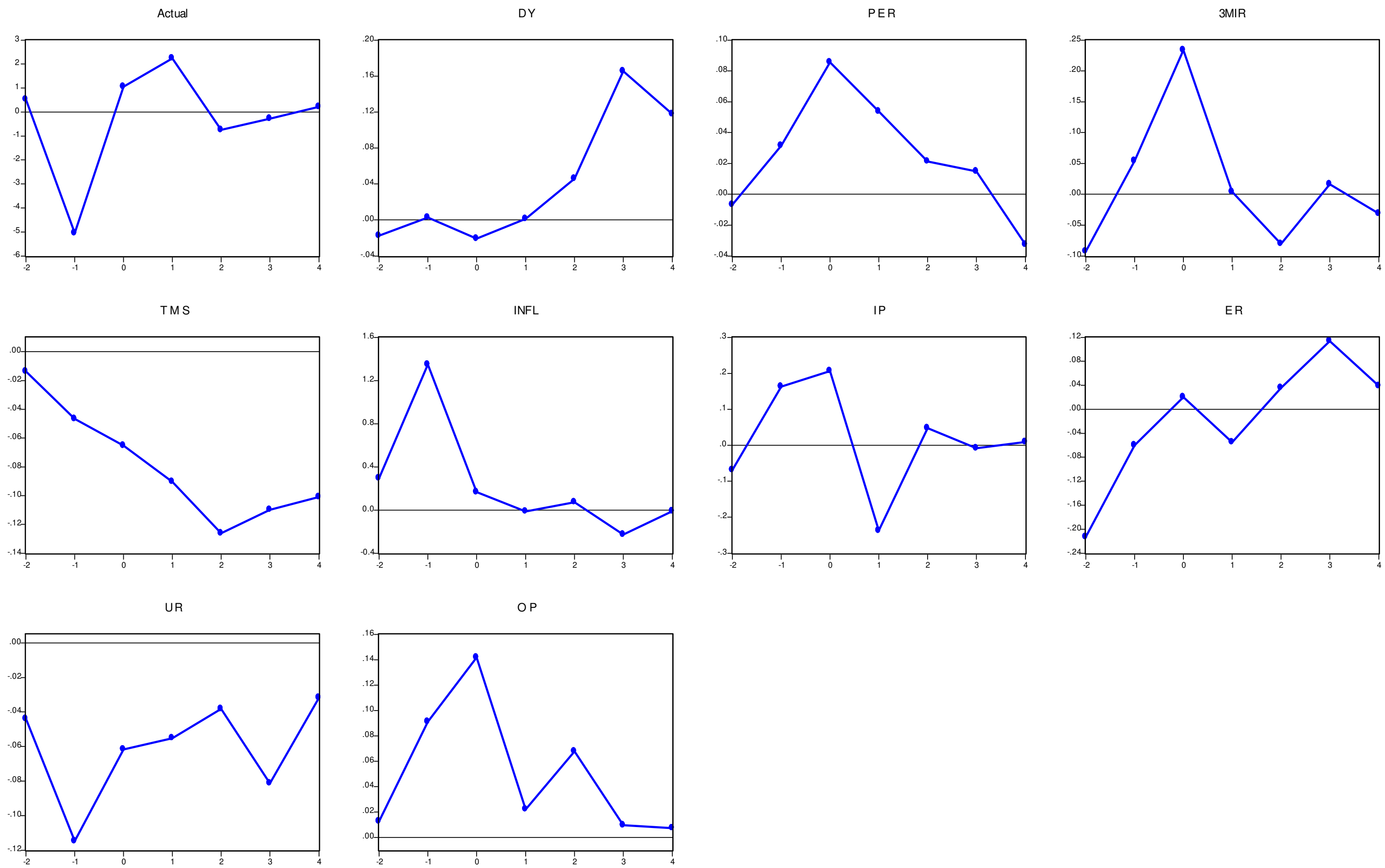


Note: Figures show cumulative differences between squared forecast errors from the historical average benchmark model and the predictive model based on the predictor variable given in the panel heading. Vertical grey areas depict OECD-dated peaks and troughs.

Figure 6. Out-of-Sample Forecasts Based on Multiple Predictor Variables, 1985:01-2012:01

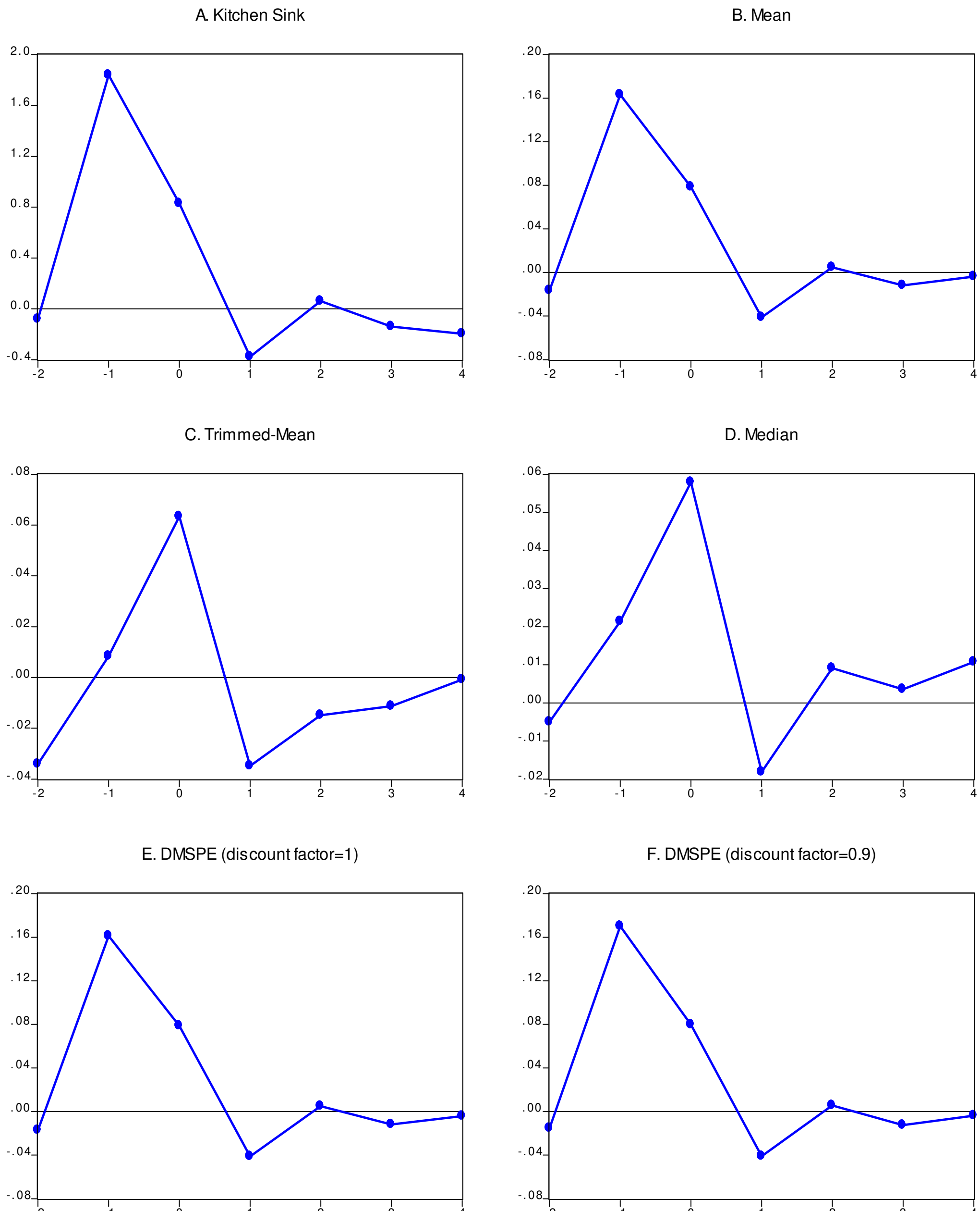
Note: Dotted (straight) lines display predictive regression forecasts based on the predictor variable given in the panel heading (historical average). Vertical grey areas depict OECD-dated peaks and troughs.

Figure 7. Actual Equity Premium and Equity Premium Forecasts Based on Individual Predictor Variables Near OECD Business-Cycle Peak, 1985:01-2012:01



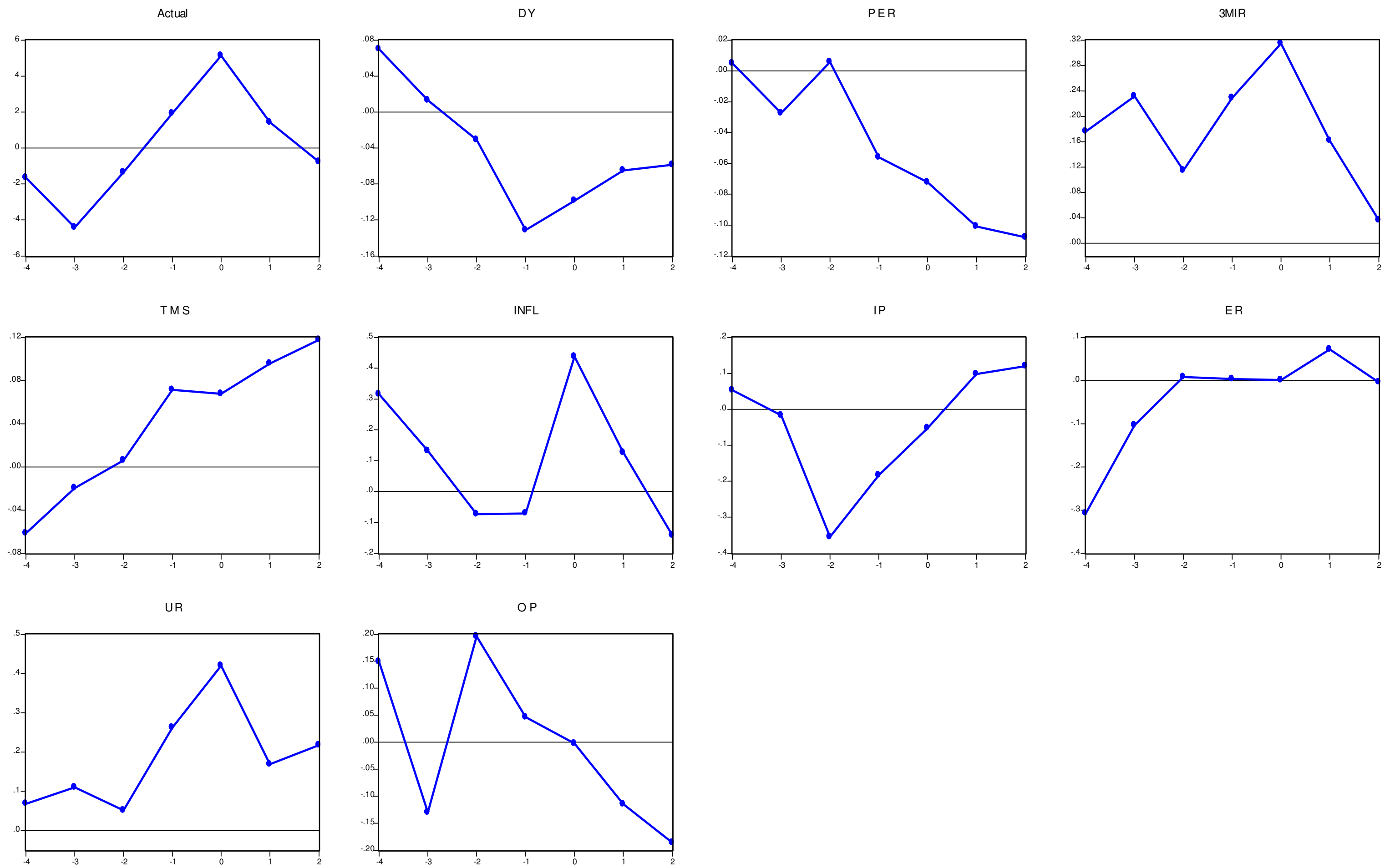
Note: The figures show the difference between the actual equity premium or the equity premium forecast based on the predictor variable given in the panel heading and the historical average benchmark model forecast two months before to four months after an OECD-dated peak. The dots indicate point estimates.

Figure 8. Actual Equity Premium and Equity Premium Forecasts Based on Multiple Predictor Variables Near OECD Business-Cycle Peak, 1985:01-2012:01



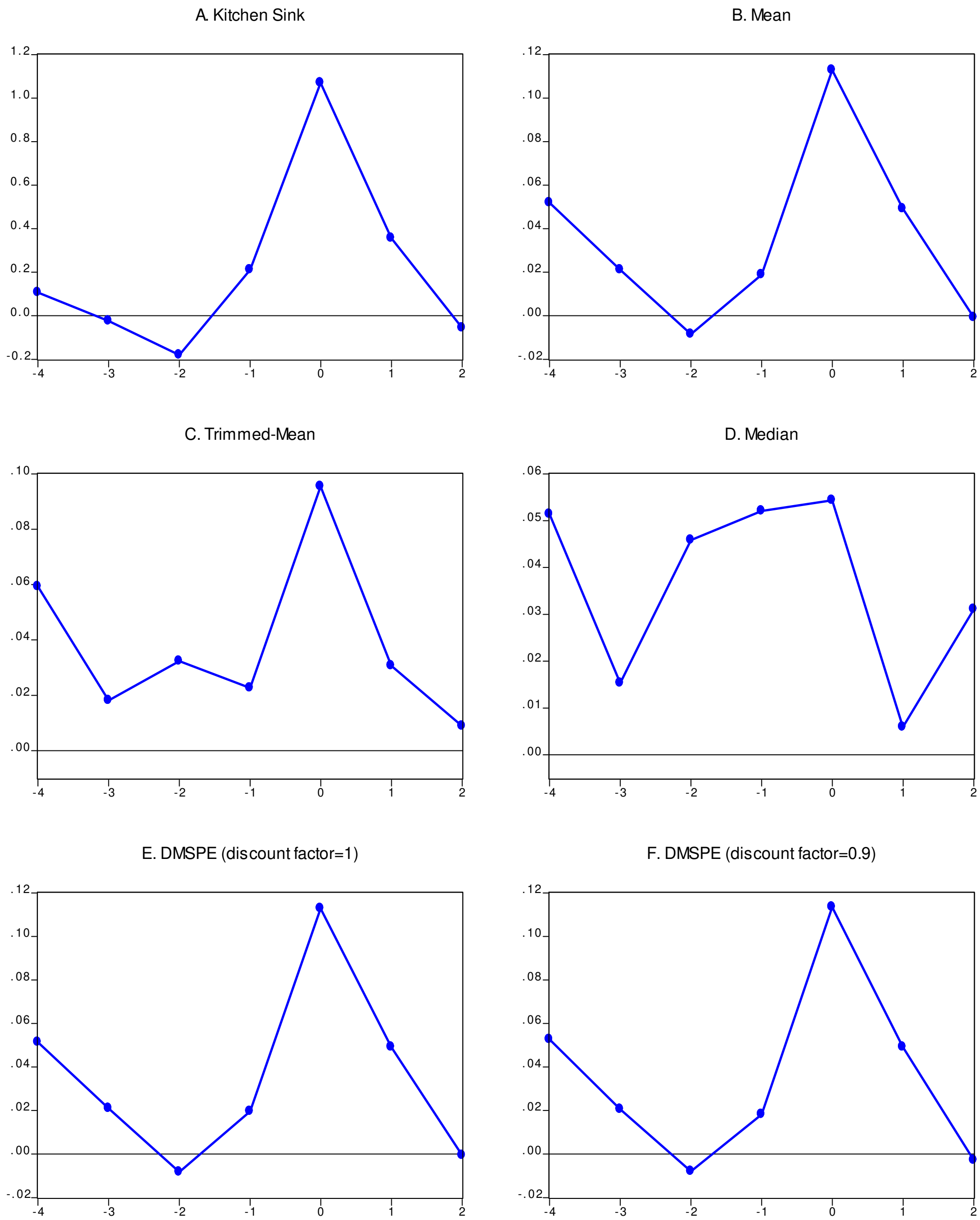
Note: The figures show the difference between the actual equity premium or the equity premium forecast based on the combination method given in the panel heading and the historical average benchmark model forecast two months before to four months after an OECD-dated peak. The dots indicate point estimates.

Figure 9. Actual Equity Premium and Equity Premium Forecasts Based on Individual Predictor Variables Near OECD Business-Cycle Trough, 1985:01-2012:01



Note: The figures show the difference between the actual equity premium or the equity premium forecast based on the predictor variable given in the panel heading and the historical average benchmark model forecast four months before to two months after an OECD-dated trough. The dots indicate point estimates.

Figure 10. Actual Equity Premium and Equity Premium Forecasts Based on Multiple Predictor Variables Near OECD Business-Cycle Trough, 1985:01-2012:01



Note: The figures show the difference between the actual equity premium or the equity premium forecast based on the combination method given in the panel heading and the historical average benchmark model forecast four months before to two months after an OECD-dated trough. The dots indicate point estimates.

APPENDIX B

Table 1. 3-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-*$MSPE$
DY	34.5525	0.0104	-0.9355	-1.3535*
PER	34.4304	0.0075	-0.5787	-0.6580
$3MIR$	34.3703	0.0044	-0.3975	-0.2586
TMS	34.0316	0.0016	0.5922	1.6817**
$INFL$	34.5119	0.0018	-0.8113	0.9863
IP	34.5046	0.0043	-0.7903	-0.6071
ER	34.3030	0.0013	-0.2010	0.0597
UR	34.2083	0.0009	0.0756	0.6147
OP	34.4301	0.0055	-0.5724	-0.4151

Panel B: January 1992 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-*$MSPE$
DY	32.7112	0.0169	-0.5784	-1.7477**
PER	32.8416	0.0119	-0.9796	-1.0184
$3MIR$	32.7480	0.0113	-0.6961	-0.7498
TMS	32.1234	0.0159	1.2243	2.7072***
$INFL$	32.5630	0.0009	-0.1274	0.5327
IP	33.0276	0.0288	-1.5512	-2.1901**
ER	32.5605	1.284×10^{-5}	-0.1197	0.2477
UR	32.4788	0.0001	0.1312	0.6685
OP	32.6533	0.0011	-0.4050	-0.0700

Panel C: January 2007 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-*$MSPE$
DY	41.8481	0.0809	-0.4529	-1.2185
PER	42.4290	0.0461	-1.8472	-1.4716*
$3MIR$	42.4738	0.0630	-1.9913	-1.0126
TMS	40.4358	0.0998	2.9024	2.5044***
$INFL$	41.4321	0.0119	0.5101	0.4755
IP	41.7932	0.0073	-0.3330	-0.3186
ER	43.8880	0.1727	-5.3872	-2.6999***
UR	41.2195	0.0052	1.0205	1.5408*
OP	43.5892	0.1064	-4.6695	-1.3778*

Note: This table shows results from the out-of-sample regression analysis from individual variables. $MSPE$ denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 2. 6-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-*$MSPE$
DY	34.8774	0.0141	-1.0734	-1.7737**
PER	34.6975	0.0058	-0.5522	-0.4957
$3MIR$	34.6420	0.0043	-0.3881	-0.2443
TMS	34.3082	0.0014	0.5792	1.6520**
$INFL$	34.7636	0.0018	-0.7403	0.9333
IP	34.8165	0.0058	-0.8927	-0.7579
ER	34.6052	0.0021	-0.2814	-0.0645
UR	34.5022	0.0015	0.0171	0.4649
OP	34.6603	0.0040	-0.4411	-0.2593

Panel B: January 1992 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-*$MSPE$
DY	33.0470	0.0150	-0.5109	-1.5365*
PER	32.2054	0.0103	-0.9925	-0.9040
$3MIR$	33.1082	0.0113	-0.7015	-0.7424
TMS	32.4766	0.0161	1.2194	2.7251***
$INFL$	32.9036	0.0010	-0.0791	0.5810
IP	33.4461	0.0405	-1.7245	-2.5618***
ER	32.9622	0.0004	-0.2575	0.0886
UR	32.8488	0.0003	0.0874	0.5813
OP	33.0214	0.0015	-0.4377	-0.1039

Panel C: January 2007 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted-*$MSPE$
DY	43.4772	0.1022	-0.6343	-1.1171
PER	44.1290	0.0565	-2.1430	-1.5173*
$3MIR$	44.0406	0.0627	-1.9793	-0.9692
TMS	41.9012	0.1063	2.9745	2.6131***
$INFL$	42.9303	0.0128	0.5916	0.4999
IP	43.4835	0.0555	-0.6618	-1.0944
ER	45.6796	0.1841	-5.7746	-2.6765***
UR	42.7058	0.0019	1.1114	1.6984**
OP	45.2157	0.1097	-4.7004	-1.3685*

Note: This table shows results from the out-of-sample regression analysis from individual variables. $MSPE$ denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 3. 12-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted*-$MSPE$
DY	34.9234	0.0316	-1.5733	-2.6852***
PER	34.6156	0.0080	-0.6781	-0.6042
$3MIR$	34.5229	0.0049	-0.4191	-0.2623
TMS	34.2267	0.0004	0.4422	1.3587*
$INFL$	34.8932	0.0005	-1.4961	0.4614
IP	34.6994	0.0071	-0.9221	-0.8276
ER	34.5218	0.0004	-0.41600	-0.2748
UR	34.3295	0.0008	0.1432	0.7874
OP	34.5380	0.0046	-0.4631	-0.2720

Panel B: January 1992 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted*-$MSPE$
DY	33.3563	0.0214	-0.6786	-1.9364**
PER	33.5176	0.0102	-1.1656	-0.9398
$3MIR$	33.3651	0.0115	-0.7015	-0.7113
TMS	32.8122	0.0087	0.9670	2.3306**
$INFL$	33.0750	0.0015	0.1738	0.8054
IP	33.8204	0.0617	-2.0760	-3.1044***
ER	33.2534	0.0013	-0.3644	-0.0370
UR	33.0553	7.225×10^{-7}	0.2332	0.8664
OP	33.2648	0.0015	-0.3988	-0.0594

Panel C: January 2007 - January 2012

Variance	$MSPE$	$CORR^2$	R^2_{OS} (%)	*adjusted*-$MSPE$
DY	47.6093	0.0132	-0.3700	-0.5300
PER	48.5301	0.0516	-2.3113	-1.4872*
$3MIR$	48.3523	0.0414	-1.9796	-0.9001
TMS	46.0066	0.1376	2.9678	2.7249***
$INFL$	47.0271	0.0208	0.8153	0.5833
IP	47.7535	0.0745	-0.6878	-1.3516*
ER	50.2290	0.1770	-5.9377	-2.6396***
UR	46.8612	0.0048	1.1652	1.8087**
OP	49.6864	0.0989	-4.7932	-1.3616*

Note: This table shows results from the out-of-sample regression analysis from individual variables. $MSPE$ denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 4. 3-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0216	0.0284	-2.6152	-1.9651	-0.9876	-0.3944
PER	0.0279	0.0353	-0.7573	-0.3787	-0.1894	-0.0755
$3MIR$	0.0348	0.0431	-1.5118	-1.1395	-0.5708	-0.2297
TMS	0.0904	0.1158	-1.2893	-0.6421	-0.3185	-0.1231
$INFL$	0.0861	0.1152	-6.7762	-7.5803	-4.2099	-1.6847
IP	0.0422	0.0554	-1.9551	-1.1548	-0.5788	-0.2331
ER	0.0628	0.0808	-1.5246	-0.9687	-0.4845	-0.1939
UR	0.0582	0.0754	-0.2876	-0.1986	-0.1072	-0.0403
OP	0.0288	0.0349	-1.2001	-1.9667	-0.9826	-0.3922

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0217	0.0272	-1.3369	-0.6689	-0.3338	-0.1328
PER	0.0084	0.0102	-1.5912	-0.7948	-0.3965	-0.1632
$3MIR$	0.0237	0.0293	-2.2824	-1.5322	-0.7669	-0.3077
TMS	0.1253	0.1661	-0.4633	-0.2302	-0.1136	-0.0419
$INFL$	0.0762	0.1006	-5.7433	-3.9181	-1.9549	-0.7771
IP	-0.0100	-0.0121	-1.7684	-0.9949	-0.4973	-0.1988
ER	0.0710	0.0914	-1.6126	-1.0871	-0.5442	-0.2185
UR	0.0566	0.0736	-0.0897	-0.0778	-0.0391	-0.0158
OP	0.0332	0.0404	-1.6224	-2.4411	-1.2202	-0.4877

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	-0.0963	-0.1209	-1.5943	-0.7952	-0.3957	-0.1261
PER	-0.1311	-0.1633	-5.7381	-2.8668	-1.4312	-0.5699
$3MIR$	-0.1229	-0.1482	-5.1485	-4.0197	-2.0058	-0.7975
TMS	0.1308	0.1871	0.7896	0.3927	0.1942	0.0752
$INFL$	0.0013	0.0017	-6.9508	-5.2459	-2.6137	-1.0343
IP	-0.0941	-0.1151	-0.7153	-0.3585	-0.1801	-0.0731
ER	-0.2364	-0.2757	-5.0229	-3.4534	-1.7203	-0.6805
UR	-0.0247	-0.0338	-0.1370	-0.0692	-0.0354	-0.0150
OP	-0.1710	-0.1909	-4.7629	-8.4809	-4.2305	-1.6802

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model.

Table 5. 6-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	SR	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0109	0.0143	-2.2803	-1.6562	-0.8271	-0.3296
PER	0.0273	0.0347	-0.9264	-0.4684	-0.2341	-0.0936
$3MIR$	0.0302	0.0375	-1.7299	-1.2499	-0.6260	-0.2516
TMS	0.0850	0.1101	-1.2518	-0.6418	-0.3182	-0.1241
$INFL$	0.0826	0.1107	-7.0235	-7.7457	-4.2289	-1.6922
IP	0.0346	0.0452	-2.0489	-1.2522	-0.6275	-0.2527
ER	0.0577	0.0741	-1.5300	-0.9985	-0.4998	-0.2000
UR	0.0516	0.0670	-0.3226	-0.2104	-0.1055	-0.0426
OP	0.0305	0.0374	-1.5339	-1.9630	-0.9810	-0.3917

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0218	0.0275	-1.5419	-0.7702	-0.3844	-0.1529
PER	0.0091	0.0114	-2.1430	-1.0778	-0.5381	-0.2142
$3MIR$	0.0213	0.0263	-2.6385	-1.7312	-0.8663	-0.3474
TMS	0.1194	0.1577	-0.5903	-0.2938	-0.1456	-0.0567
$INFL$	0.0764	0.1015	-6.0758	-4.1738	-2.0828	-0.8282
IP	-0.0202	-0.0244	-2.1407	-1.1558	-0.5777	-0.2308
ER	0.0645	0.0824	-1.8578	-1.2478	-0.6246	-0.2507
UR	0.0504	0.0654	-0.3251	-0.1942	-0.0972	-0.0390
OP	0.0303	0.0369	-1.7957	-2.5223	-1.2610	-0.5042

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	-0.1230	-0.1547	-3.1944	-1.5945	-0.7945	-0.3145
PER	-0.1565	-0.1951	-7.4457	-3.7195	-1.8563	-0.7385
$3MIR$	-0.1377	-0.1668	-6.3837	-4.7771	-2.3837	-0.9477
TMS	0.1049	0.1497	0.4566	0.2255	0.1100	0.0406
$INFL$	-0.0123	-0.0162	-7.7674	-5.9749	-2.9950	-1.1864
IP	-0.1417	-0.1732	-1.5001	-0.7502	-0.3740	-0.1503
ER	-0.2552	-0.2985	-6.2071	-4.2119	-2.0984	-0.8302
UR	-0.0478	-0.0653	-0.7533	-0.3777	-0.1898	-0.0771
OP	-0.1857	-0.2093	-5.6752	-9.2291	-4.6041	-1.8291

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model.

Table 6. 12-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Individual Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	-0.0267	-0.0323	-2.7346	-1.7900	-0.8935	-0.3557
PER	0.0033	0.0042	-1.5159	-0.8022	-0.4010	-0.1603
$3MIR$	0.0135	0.0166	-1.6985	-1.4327	-0.7171	-0.2877
TMS	0.0627	0.0799	-1.4859	-0.7428	-0.3686	-0.1441
$INFL$	0.0541	0.0697	-7.0475	-8.1288	-4.3026	-1.7193
IP	0.0151	0.0194	-2.0902	-1.2885	-0.6456	-0.2599
ER	0.0355	0.0450	-1.6606	-1.0809	-0.5404	-0.2161
UR	0.0397	0.0513	-0.3550	-0.2438	-0.1222	-0.0493
OP	0.0151	0.0183	-1.5965	-2.0784	-1.0383	-0.4142

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	0.0228	0.0289	-1.6640	-0.8305	-0.4138	-0.1637
PER	0.0094	0.0119	-2.6361	-1.3769	-0.6867	-0.2726
$3MIR$	0.0283	0.0349	-2.7983	-2.0011	-1.0008	-0.4007
TMS	0.1156	0.1533	-0.8213	-0.4082	-0.2017	-0.0778
$INFL$	0.0898	0.1207	-6.2407	-4.5396	-2.2657	-0.9013
IP	-0.0252	-0.0304	-2.6677	-1.3967	-0.6965	-0.2764
ER	0.0663	0.0848	-2.0376	-1.3631	-0.6819	-0.2732
UR	0.0661	0.0865	-0.3592	-0.2149	-0.1077	-0.0434
OP	0.0371	0.0452	-1.8554	-2.6825	-1.3408	-0.5357

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
DY	-0.1091	-0.1427	-4.6053	-2.2983	-1.1448	-0.4527
PER	-0.1712	-0.2165	-9.3169	-4.6523	-2.3200	-0.9207
$3MIR$	-0.1375	-0.1691	-7.5486	-6.0942	-3.0402	-1.2078
TMS	0.1026	0.1481	0.3352	0.1627	0.0765	0.0248
$INFL$	-0.0073	-0.0099	-8.9430	-7.4501	-3.7130	-1.4707
IP	-0.1536	-0.1920	-2.1125	-1.0554	-0.5268	-0.2097
ER	-0.2688	-0.3201	-7.5366	-5.1424	-2.5589	-1.0089
UR	-0.0500	-0.0692	-1.2144	-0.6086	-0.3057	-0.1239
OP	-0.1968	-0.2239	-6.5658	-10.6251	-5.2968	-2.0998

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model.

Table 7. 3-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Multiple Predictor Variables

Method	*MSPE*	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Panel A: January 1985 - January 2012				
Benchmark	34.2342	0.0236	-	-
Kitchen Sink	35.4842	0.0006	-3.6512	0.5588
Mean	34.2369	0.0034	-0.0079	0.1917
Trimmed-Mean	34.2525	0.0084	-0.0535	-0.0749
Median	34.3244	0.0269	-0.2635	-1.1631
DMSPE (θ=1.0)[1]	34.2370	0.0034	-0.0083	0.1881
DMSPE (θ=0.9)[1]	34.2350	0.0031	-0.0023	0.2152
DMSPE (θ=0.75)[1]	34.2339	0.0028	0.0006	0.2328
DMSPE (θ=1.0)[2]	34.1823	0.0008	0.1515	0.7585
DMSPE (θ=0.9)[2]	34.1888	0.0011	0.1324	0.6979
DMSPE (θ=0.75)[2]	34.1954	0.0014	0.1131	0.6333
Panel B: January 1992 - January 2012				
Method	*MSPE*	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	32.5215	0.0121	-	-
Kitchen Sink	34.0905	0.0029	-4.8243	-0.5610
Mean	32.5565	0.0068	-0.1074	-0.2226
Trimmed-Mean	32.5394	0.0062	-0.0549	-0.0876
Median	32.6025	0.0192	-0.2490	-0.8492
DMSPE (θ=1.0)[1]	32.5566	0.0069	-0.1076	-0.2241
DMSPE (θ=0.9)[1]	32.5540	0.0063	-0.0997	-0.1923
DMSPE (θ=0.75)[1]	32.5536	0.0059	-0.0984	-0.1820
DMSPE (θ=1.0)[2]	32.5515	0.0055	-0.0921	-0.1584
DMSPE (θ=0.9)[2]	32.5613	0.0068	-0.1221	-0.2562
DMSPE (θ=0.75)[2]	32.5705	0.0081	-0.1505	-0.3512
Panel C: January 2007 - January 2012				
Method	*MSPE*	$CORR^2$	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	41.6445	0.0305	-	-
Kitchen Sink	46.4771	0.0836	-11.6043	-1.4201
Mean	42.0426	0.0917	-0.9558	-1.1193
Trimmed-Mean	41.9685	0.0694	-0.7778	-0.9718
Median	41.8990	0.0787	-0.6109	-0.8821
DMSPE (θ=1.0)[1]	42.0459	0.0933	-0.9637	-1.1293
DMSPE (θ=0.9)[1]	42.0514	0.0901	-0.9768	-1.1206
DMSPE (θ=0.75)[1]	42.0522	0.0861	-0.9789	-1.1038
DMSPE (θ=1.0)[2]	42.0229	0.0768	-0.9085	-1.0395
DMSPE (θ=0.9)[2]	42.0189	0.0802	-0.8988	-1.0550
DMSPE (θ=0.75)[2]	42.0106	0.0821	-0.8789	-1.0633

Note: This table shows results from the out-of-sample regression analysis from individual variables. *MSPE* denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. As the initial holdout out-of-sample period when computing the *DMSPE* forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 8. 6-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Method	*MSPE*	*CORR*^{2}	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	34.5081	0.0257	-	-
Kitchen Sink	35.7491	0.0004	-3.5963	0.4141
Mean	34.5208	0.0041	-0.0370	0.0929
Trimmed-Mean	34.5285	0.0089	-0.0592	-0.0954
Median	34.5635	0.0183	-0.1606	-0.5950
DMSPE (θ=1.0)[1]	34.5211	0.0042	-0.0376	0.0883
DMSPE (θ=0.9)[1]	34.5184	0.0038	-0.0299	0.1221
DMSPE (θ=0.75)[1]	34.5170	0.0035	-0.0259	0.1436
DMSPE (θ=1.0)[2]	34.5096	0.0029	-0.0044	0.2193
DMSPE (θ=0.9)[2]	34.5154	0.0034	-0.0212	0.1598
DMSPE (θ=0.75)[2]	34.5217	0.0039	-0.0395	0.0975

Panel B: January 1992 - January 2012

Method	*MSPE*	*CORR*^{2}	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	32.8775	0.0162	-	-
Kitchen Sink	34.6702	0.0044	-5.4526	-0.6877
Mean	32.9225	0.0093	-0.1368	-0.3000
Trimmed-Mean	32.8991	0.0082	-0.0657	-0.1181
Median	32.9326	0.0144	-0.1675	-0.4858
DMSPE (θ=1.0)[1]	32.9227	0.0093	-0.1374	-0.3027
DMSPE (θ=0.9)[1]	32.9200	0.0086	-0.1291	-0.2704
DMSPE (θ=0.75)[1]	32.9193	0.0082	-0.1271	-0.2577
DMSPE (θ=1.0)[2]	32.9182	0.0079	-0.1238	-0.2446
DMSPE (θ=0.9)[2]	32.9289	0.0095	-0.1561	-0.3470
DMSPE (θ=0.75)[2]	32.9389	0.0111	-0.1865	-0.4455

Panel C: January 2007 - January 2012

Method	*MSPE*	*CORR*^{2}	R^2_{OS} (%)	*adjusted-MSPE*
Benchmark	43.1858	0.0874	-	-
Kitchen Sink	48.8487	0.0948	-13.1129	-1.4865
Mean	43.6442	0.1135	-1.0613	-1.1598
Trimmed-Mean	43.5497	0.0818	-0.8426	-0.9660
Median	43.4306	0.0647	-0.5668	-0.7143
DMSPE (θ=1.0)[1]	43.6476	0.1152	-1.0694	-1.1694
DMSPE (θ=0.9)[1]	43.6527	0.1120	-1.0811	-1.1611
DMSPE (θ=0.75)[1]	43.6533	0.1076	-1.0824	-1.1449
DMSPE (θ=1.0)[2]	43.6223	0.1149	-1.0107	-1.0869
DMSPE (θ=0.9)[2]	43.6199	0.1036	-1.0052	-1.1045
DMSPE (θ=0.75)[2]	43.6138	0.1073	-0.9909	-1.1175

Note: This table shows results from the out-of-sample regression analysis from individual variables. *MSPE* denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. As the initial holdout out-of-sample period when computing the *DMSPE* forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 9. 12-Step Ahead Out-of-Sample Analysis, Statistical Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Method	$MSPE$	$CORR^2$	R^2_{OS} (%)	$adjusted\text{-}MSPE$
Benchmark	34.3824	0.0302	-	-
Kitchen Sink	36.0677	9.3758×10^{-6}	-4.9017	0.0336
Mean	34.4501	0.0095	-0.1969	-0.4455
Trimmed-Mean	34.4534	0.0169	-0.2064	-0.6677
Median	34.4790	0.0278	-0.2810	-1.0662
$DMSPE$ (θ=1.0)^1	34.4498	0.0096	-0.1959	-0.4470
$DMSPE$ (θ=0.9)^1	34.4487	0.0090	-0.1929	-0.4172
$DMSPE$ (θ=0.75)^1	34.4490	0.0085	-0.1936	-0.4000
$DMSPE$ (θ=1.0)^2	34.4400	0.0074	-0.1674	-0.3048
$DMSPE$ (θ=0.9)^2	34.4440	0.0082	-0.1793	-0.3613
$DMSPE$ (θ=0.75)^2	34.4498	0.0090	-0.1960	-0.4283
Panel B: January 1992 - January 2012				
Method	$MSPE$	$CORR^2$	R^2_{OS} (%)	$adjusted\text{-}MSPE$
Benchmark	33.1326	0.0214	-	-
Kitchen Sink	35.1277	0.0048	-6.0213	-0.6742
Mean	33.1969	0.0139	-0.1939	-0.4465
Trimmed-Mean	33.1821	0.0142	-0.1494	-0.3633
Median	33.2101	0.0204	-0.2339	-0.6645
$DMSPE$ (θ=1.0)^1	33.1972	0.0140	-0.1948	-0.4500
$DMSPE$ (θ=0.9)^1	33.1939	0.0131	-0.1850	-0.4134
$DMSPE$ (θ=0.75)^1	33.1926	0.0125	-0.1811	-0.3943
$DMSPE$ (θ=1.0)^2	33.1910	0.0122	-0.1761	-0.3778
$DMSPE$ (θ=0.9)^2	33.2041	0.0146	-0.2156	-0.4987
$DMSPE$ (θ=0.75)^2	33.2166	0.0169	-0.2535	-0.6178
Panel C: January 2007 - January 2012				
Method	$MSPE$	$CORR^2$	R^2_{OS} (%)	$adjusted\text{-}MSPE$
Benchmark	47.4137	0.0591	-	-
Kitchen Sink	53.5983	0.0710	-13.0440	-1.3313
Mean	47.9104	0.0679	-1.0475	-1.0891
Trimmed-Mean	47.8032	0.0442	-0.8215	-0.8847
Median	47.6515	0.0267	-0.5015	-0.5902
$DMSPE$ (θ=1.0)^1	47.9145	0.0692	-1.0562	-1.0992
$DMSPE$ (θ=0.9)^1	47.9194	0.0674	-1.0664	-1.0906
$DMSPE$ (θ=0.75)^1	47.9191	0.0647	-1.0658	-1.0738
$DMSPE$ (θ=1.0)^2	47.8865	0.0988	-0.9971	-1.0869
$DMSPE$ (θ=0.9)^2	47.8845	0.1036	-0.9928	-1.1045
$DMSPE$ (θ=0.75)^2	47.8778	0.1073	-0.9788	-1.1175

Note: This table shows results from the out-of-sample regression analysis from individual variables. $MSPE$ denotes the mean squared prediction error, $CORR^2$ denotes the squared correlation coefficient between the forecasts and the actual realisations of the excess returns, R^2_{OS} denotes the out-of-sample R^2 which compares the predictive ability of the model with the historical average benchmark model. The statistical significance of the R^2_{OS} is assessed with the Clark and West (2007) adjusted-MSPE statistic. As the initial holdout out-of-sample period when computing the $DMSPE$ forecast, the 1985:01-1998:06 (^{1}) and 1981:01-1984:12 (^{2}) periods have been used. ***, **, * correspond to the 1%, 5%, and 10% level of significance, respectively.

Table 10. 3-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	0.0671	0.0904	-8.5922	-11.5474	-8.7511	-3.5605
Mean	0.0557	0.0714	-1.7264	-0.8852	-0.4426	-0.1771
Trimmed-Mean	0.0516	0.0656	-1.3380	-0.6687	-0.3340	-0.1332
Median	0.0403	0.0504	-0.8989	-0.4492	-0.2243	-0.0894
$DMSPE$ (θ=1.0)[1]	0.0556	0.0713	-1.7113	-0.8766	-0.4383	-0.1754
$DMSPE$ (θ=0.9)[1]	0.0562	0.0720	-1.7805	-0.9140	-0.4570	-0.1863
$DMSPE$ (θ=0.75)[1]	0.0566	0.0726	-1.8464	-0.9503	-0.4752	-0.1936
$DMSPE$ (θ=1.0)[2]	0.0654	0.0843	-1.4402	-0.7386	-0.3695	-0.1480
$DMSPE$ (θ=0.9)[2]	0.0643	0.0829	-1.3862	-0.7061	-0.3533	-0.1451
$DMSPE$ (θ=0.75)[2]	0.0631	0.0815	-1.3162	0.6695	-0.3354	-0.1385

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	0.0049	0.0060	-6.8948	-10.8680	-8.0575	-3.2182
Mean	0.0453	0.0569	-1.3596	-0.6790	-0.3388	-0.1346
Trimmed-Mean	0.0475	0.0597	-1.1817	-0.5902	-0.2945	-0.1170
Median	0.0377	0.0471	-0.8789	-0.4390	-0.2190	-0.0870
$DMSPE$ (θ=1.0)[1]	0.0452	0.0568	-1.3535	-0.6760	-0.3372	-0.1340
$DMSPE$ (θ=0.9)[1]	0.0459	0.0576	-1.4357	-0.7170	-0.3576	-0.1419
$DMSPE$ (θ=0.75)[1]	0.0461	0.0579	-1.4769	-0.7376	-0.3679	-0.1460
$DMSPE$ (θ=1.0)[2]	0.0467	0.0587	-1.4378	-0.7181	-0.3582	-0.1423
$DMSPE$ (θ=0.9)[2]	0.0451	0.0567	-1.4159	-0.7072	-0.3528	-0.1437
$DMSPE$ (θ=0.75)[2]	0.0437	0.0548	-1.3925	-0.6955	-0.3469	-0.1414

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	-0.1719	-0.1902	-11.9672	-15.7179	-15.7767	-6.2824
Mean	-0.1085	-0.1331	-3.3302	-1.6618	-0.8275	-0.3270
Trimmed-Mean	-0.1016	-0.1258	-3.0194	-1.5067	-0.7504	-0.2966
Median	-0.0991	-0.1228	-2.4193	-1.2072	-0.6011	-0.2375
$DMSPE$ (θ=1.0)[1]	-0.1090	-0.1336	-3.3287	-1.6610	-0.8272	-0.3268
$DMSPE$ (θ=0.9)[1]	-0.1089	-0.1334	-3.4006	-1.6969	-0.8450	-0.3339
$DMSPE$ (θ=0.75)[1]	-0.1083	-0.1326	-3.4635	-1.7282	-0.8606	-0.3400
$DMSPE$ (θ=1.0)[2]	-0.1052	-0.1292	-3.3725	-1.6828	-0.8380	-0.3311
$DMSPE$ (θ=0.9)[2]	-0.1054	-0.1296	-3.2862	-1.6398	-0.8166	-0.3357
$DMSPE$ (θ=0.75)[2]	-0.1052	-0.1294	-3.1810	-1.5873	-0.7904	-0.3123

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model. As the initial holdout out-of-sample period when computing the $DMSPE$ forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used.

Table 11. 6-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	0.0538	0.0728	-6.3194	-12.2315	-9.8197	-3.9973
Mean	0.0509	0.0655	-1.7515	-0.8811	-0.4405	-0.1761
Trimmed-Mean	0.0477	0.0609	-1.3198	-0.6594	-0.3293	-0.1312
Median	0.0415	0.0526	-1.0214	-0.5103	-0.2547	-0.1014
$DMSPE$ (θ=1.0)[1]	0.0509	0.0654	-1.7357	-0.8729	-0.4364	-0.1745
$DMSPE$ (θ=0.9)[1]	0.0515	0.0662	-1.8108	-0.9113	-0.4556	-0.1909
$DMSPE$ (θ=0.75)[1]	0.0525	0.0678	-1.8790	-0.9464	-0.4731	-0.1979
$DMSPE$ (θ=1.0)[2]	0.0533	0.0686	-1.8883	-0.9513	-0.4756	-0.1902
$DMSPE$ (θ=0.9)[2]	0.0523	0.0673	-1.8259	-0.9189	-0.4594	-0.1925
$DMSPE$ (θ=0.75)[2]	0.0513	0.0660	-1.7861	-0.8985	-0.4492	-0.1884

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	-0.0052	-0.0064	-6.7453	-12.1847	-9.4667	-3.7836
Mean	0.0407	0.0511	-1.6562	-0.8273	-0.4129	-0.1643
Trimmed-Mean	0.0432	0.0545	-1.3506	-0.6761	-0.3388	-0.1365
Median	0.0384	0.0481	-1.1663	-0.5826	-0.2907	-0.1156
$DMSPE$ (θ=1.0)[1]	0.0407	0.0510	-1.6498	-0.8242	-0.4113	-0.1636
$DMSPE$ (θ=0.9)[1]	0.0413	0.0518	-1.7815	-0.8904	-0.4447	-0.1772
$DMSPE$ (θ=0.75)[1]	0.0415	0.0522	-1.8234	-0.9113	-0.4552	-0.1814
$DMSPE$ (θ=1.0)[2]	0.0420	0.0527	-1.7361	-0.8672	-0.4328	-0.1721
$DMSPE$ (θ=0.9)[2]	0.0403	0.0506	-1.7165	-0.8574	-0.4279	-0.1794
$DMSPE$ (θ=0.75)[2]	0.0388	0.0486	-1.6945	-0.8465	-0.4224	-0.1772

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	-0.1986	-0.2214	-13.0954	-20.0820	-19.2867	-7.6855
Mean	-0.1336	-0.1639	-4.4798	-2.2360	-1.1141	-0.4410
Trimmed-Mean	-0.1248	-0.1548	-4.1625	-2.0777	-1.0353	-0.4098
Median	-0.1159	-0.1446	-3.6049	-1.7994	-0.8966	-0.3550
$DMSPE$ (θ=1.0)[1]	-0.1341	-0.1644	-4.4770	-2.2346	-1.1134	-0.4407
$DMSPE$ (θ=0.9)[1]	-0.1338	-0.1641	-4.5485	-2.2703	-1.1312	-0.4477
$DMSPE$ (θ=0.75)[1]	-0.1331	-0.1632	-4.6146	-2.3032	-1.1476	-0.4542
$DMSPE$ (θ=1.0)[2]	-0.1303	-0.1601	-4.5068	-2.2495	-1.1208	-0.4436
$DMSPE$ (θ=0.9)[2]	-0.1308	-0.1607	-4.4263	-2.2093	-1.1008	-0.4660
$DMSPE$ (θ=0.75)[2]	-0.1310	-0.1610	-4.3267	-2.1596	-1.0761	-0.4260

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model. As the initial holdout out-of-sample period when computing the $DMSPE$ forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used.

Table 12. 12-Step Ahead Out-of-Sample Analysis, Economic Evaluation, Multiple Predictor Variables

Panel A: January 1985 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	0.0245	0.0321	-5.8618	-12.8293	-10.1952	-4.1642
Mean	0.0267	0.0334	-1.9741	-0.9868	-0.4930	-0.1967
Trimmed-Mean	0.0239	0.0298	-1.5865	-0.7925	-0.3955	-0.1573
Median	0.0187	0.0233	-1.2252	-0.6119	-0.3052	-0.1212
$DMSPE$ (θ=1.0)[1]	0.0266	0.0334	-1.9563	-0.9779	-0.4885	-0.1949
$DMSPE$ (θ=0.9)[1]	0.0272	0.0341	-2.0307	-1.0151	-0.5072	-0.2160
$DMSPE$ (θ=0.75)[1]	0.0275	0.0345	-2.1010	-1.0504	-0.5248	-0.2230
$DMSPE$ (θ=1.0)[2]	0.0289	0.0363	-2.1159	-1.0580	-0.5286	-0.2110
$DMSPE$ (θ=0.9)[2]	0.0281	0.0353	-2.0459	-1.0229	-0.5111	-0.2176
$DMSPE$ (θ=0.75)[2]	0.0272	0.0341	-2.0016	-1.0007	-0.5000	-0.2132

Panel B: January 1992 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	-0.0005	-0.0007	-7.1070	-13.7538	-11.0616	-4.4333
Mean	0.0461	0.0579	-1.9569	-0.9771	-0.4872	-0.1932
Trimmed-Mean	0.0480	0.0604	-1.7751	-0.8863	-0.4419	-0.1753
Median	0.0437	0.0548	-1.4427	-0.7202	-0.3590	-0.1423
$DMSPE$ (θ=1.0)[1]	0.0460	0.0578	-1.9503	-0.9738	-0.4856	-0.1926
$DMSPE$ (θ=0.9)[1]	0.0467	0.0587	-2.0836	-1.0430	-0.5226	-0.2100
$DMSPE$ (θ=0.75)[1]	0.0470	0.0591	-2.1277	-1.0650	-0.5336	-0.2144
$DMSPE$ (θ=1.0)[2]	0.0474	0.0597	-2.0451	-1.0211	-0.5091	-0.2019
$DMSPE$ (θ=0.9)[2]	0.0455	0.0572	-2.0283	-1.0127	-0.5049	-0.2131
$DMSPE$ (θ=0.75)[2]	0.0437	0.0549	-2.0073	-1.0022	-0.4997	-0.2110

Panel C: January 2007 - January 2012

Variance	SH	SO	$\Delta(\%)$, γ=0.5	$\Delta(\%)$, γ=1	$\Delta(\%)$, γ=2	$\Delta(\%)$, γ=5
Kitchen Sink	-0.1905	-0.2166	-14.8370	-24.7812	-23.7248	-9.5147
Mean	-0.1379	-0.1723	-5.7880	-2.8881	-1.4382	-0.5682
Trimmed-Mean	-0.1276	-0.1614	-5.5133	-2.7513	-1.3704	-0.5418
Median	-0.1156	-0.1482	-4.8616	-2.4264	-1.2087	-0.4781
$DMSPE$ (θ=1.0)[1]	-0.1384	-0.1729	-5.7838	-2.8860	-1.4371	-0.5678
$DMSPE$ (θ=0.9)[1]	-0.1381	-0.1724	-5.8684	-2.9282	-1.4581	-0.5760
$DMSPE$ (θ=0.75)[1]	-0.1373	-0.1714	-5.9474	-2.9676	-1.4777	-0.5838
$DMSPE$ (θ=1.0)[2]	-0.1347	-0.1684	-5.8162	-2.9022	-1.4451	-0.5709
$DMSPE$ (θ=0.9)[2]	-0.1352	-0.1692	-5.7248	-2.8566	-1.4225	-0.6380
$DMSPE$ (θ=0.75)[2]	-0.1354	-0.1695	-5.6093	-2.7990	-1.3938	-0.5507

Note: This table shows results from the economic evaluation of the out-of-sample analysis. SR denotes the Sharpe ration, SO denotes the Sortino ratio, $\Delta(\%)$ denotes the maximum fee (in monthly returns) that a mean-variance investor would be willing to pay to have access to the forecasting method based on the economic variable given in each row relative to the historical benchmark model. As the initial holdout out-of-sample period when computing the $DMSPE$ forecast, the 1985:01-1998:06 ([1]) and 1981:01-1984:12 ([2]) periods have been used.