Group X7
Predicting Aggregate Stock Returns with Short Interest and Forecast Combinations

Abbaan Nassar 413272
Hien Ngo 431495
Roland van Veen 430476
Veerle de Boer 411746

June 16, 2017

Abstract

In this paper, we look at the forecasting performance of short interest, as used by Rapach et al. (2016). We compare it to predictors introduced by Goyal and Welch (2008) for predicting aggregate stock returns in the full sample and when splitting the sample into half. To do this, we look at the in-sample and out-of-sample forecasting performance. When comparing the results, we look at the R^2 statistics and use the forecast encompassing test and the Diebold and Mariano (1995) test. Next, we construct a combined forecast model and again compare the results with the other predictors. Finally, we look Sharpe ratios and CER gains to investigate the economic value of the forecasts. Our results show that the short interest index performs better economically and statistically than the other predictors and our combined forecast model using the full sample.

1 Introduction

Stock return predictability has been a subject of much debate. It is an interesting issue for its theoretical value and is also naturally of great interest to investors in practice. The central debate in this area of finance is whether predictors can be found that are better at predicting future excess stock returns than simply using the historical mean. The increased computing power of past decades has been of great help in modeling and anticipating excess stock returns, which motivates the search for such predictors.

Past attempts at this problem have provided different results. Goyal and Welch (2008), for example, find that no single predictor can outperform the historical mean return when it comes to forecast accuracy. On the other hand, Rapach et al. (2016) interestingly finds a such single predictor: short interest. Moreover, Rapach et al. (2010) discovers that combining different predictive regression models results in better forecast accuracy than individual models.

We will look into this by testing the performance of several predictors in predicting monthly excess returns on the S&P 500 index. Also, an important predictor we shall look into is the short interest. Then, we will also be combining multiple models in a weighted average of forecasts. Our research question, therefore, consists out of two parts: firstly, whether short-interest performs significantly better in predicting aggregate stock returns than other predictors. Indeed, short interest outperforms all of 14 popular predictors. Moreover, we will show that utility gains and Sharpe ratios of short interest exceed those of other popular predictors. Secondly, whether better forecasting performance can be obtained from combining multiple forecast models.

First, in §2 the data is described and analyzed. Then, in §3 we formulate the base model using only one predictor and expand this model using combinations. The results and interpretation of the models are given in §4 and finally a conclusion follows in §5.

2 Data

Our data is taken from January 1973 to December 2014. We want to predict monthly excess return on the Standard & Poor's (S&P) 500 index and the main predictor variable will be the short-interest as used by Rapach et al. (2016) which we will explain shortly. We use the same dataset as in their paper, which includes the 14 monthly predictors from Goyal & Welch (2008).

We show some summary statistics in the Appendix, table 2 shows the full sample from 1973:01-2014:12 and gives a small explanation of the predictors. We have also split this sample in two smaller samples, namely 1973:01-1993:12 and 1994:01-2014:12. These summary statistics are given in the Appendix in table 3. Something which is striking is that the mean and median from EWSI (equal-weighted short interest) differ a lot in these two periods. The mean in the period 1973-1993 is 0.64% and in 1994-2014 is 3.66%. Rapach et al. (2016) gives some reasons for this upward trend. The first explanation is that the equity lending market has expanded significantly over the last few decades, which results in likely reducing short sale constraints. The growth of the hedge fund industry is also another reason. These increases are unrelated to the information set of short sellers. To adjust for the trend in the (log of the) EWSI, a detrended measure of aggregate short interest is created and the series is standardized to have a standard deviation of 1. A new predictor is the SII (short interest index) and can be interpreted as a measure of market pessimism based on the short interest data. In the remainder of the report, we will focus mainly on this predictor.

3 Methodology

3.1 Base Model

We use a linear regression model to compare the predictive performance of short interest with the 14 Goyal & Welch predictors. We consider the regression model:

$$r_{t:t+h} = \alpha + \beta x_t + \epsilon_{t:t+h} \ for \ t = 1, ..., T - h \tag{1}$$

In which we declare the following variables:
r_t: the S&P500 log excess return for month t: $r_{t:t+h} = (1/h)(r_{t+1} + ... + r_{t+h})$
x_t: the predictor variable for period t

This model consists of one predictor variable x and this can be either short interest or one of the Goyal & Welch predictors.

We use different horizons (h), namely monthly (h=1), bimonthly (h=2), quarterly, (h=3), semi-annual(h=6) and annual (h=12). Our total dataset consists of 504 observations. Therefore, we have 503, 502, 501, 498 and 492 usable observations respectively. Next, we also split our sample equally in 1973:01-1993:12 (sample I) and 1994:01-2014:12 (sample II). Both samples contain 252 observations.

In-Sample Tests
Our models contain one explanatory variable and we are interested in its predictive value, therefore we test the significance of β. We test $H_0 : \beta = 0$ against $H_A : \beta > 0$ in equation (1). We use a one-sided alternative hypothesis, because it is more powerful (Inoue and Kilian (2005)). To make the alternative hypothesis $H_A : \beta > 0$ relevant for all predictors, we take the negative of NTIS, TBL, LTY, INFL, and SII, because we expect a negative correlation with our dependent variable. We finally estimate equation (1) using OLS. We do this for the full sample, sample I and sample II and use the different horizons.

Out-of-Sample Tests
To examine the robustness of the in-sample results, we also pay attention in the results for out-of-sample tests of return predictability. Such tests are important in light of Goyal and Welch (2008), who show that the in-sample predictive ability of a variety of plausible return predictors generally does not hold up in out-of-sample tests. We compute a predictive regression forecast as:

$$\hat{r}_{t:t+h} = \hat{\alpha_t} + \hat{\beta}_t x_t \tag{2}$$

Where $\hat{\alpha_t}$ and $\hat{\beta}_t$ are the OLS estimates of α and β respectively. We compute the R^2_{OS} using equation (3) which is the proportional reduction in mean squared forecast error (MSFE) at the h-month horizon compared to the prevailing mean benchmark forecast. This benchmark is the average excess return from the beginning of the sample through month t. This forecast, hence, corresponds to the constant expected excess return model, that is equation (1) with $\beta = 0$. This model implies that the returns can not be predicted.

$$R^2_{OS} = 1 - \frac{\sum_{t=1}^{T}(r_t - \hat{r}_t)^2}{\sum_{t=1}^{T}(r_t - \bar{r}_t)^2} \tag{3}$$

We use the Clark and West (2007) statistic to test if our model has a better predictive ability than the constant expected return model. A positive R^2_{OS} implies that one of our predictors outperforms the prevailing mean benchmark in terms of MSFE. We test this using the full sample, sample I and sample II. For the full sample we use the forecast evaluation period from 1990:01 to 2014:12, for sample I we use the forecast evaluation period from 1982:01-1993:12 and for sample II we use the forecast evaluation period from 2003:01-2014:12. We test this for the different horizons 1, 2, 3, 6 and 12 months.

Next, we want to compare the predictive regression forecast using SII to that of the individual predictive regression forecasts based on the 14 popular predictors. We use two tests, namely to test if the G&W predictor encompasses SII and if the SII predictor encompasses a G&W predictor. This is done by forming an optimal combination forecast as a combination of a GW predictor and the SII predictor as seen in equation (4) and (5).

$$\hat{r^*}_{t:t+h} = (1-\lambda)\hat{r}^{i}_{t:t+h} + \lambda\hat{r}^{SII}_{t:t+h} \tag{4}$$

$$\hat{r^*}_{t:t+h} = (1-\lambda)\hat{r}^{SII}_{t:t+h} + \lambda\hat{r}^{I}_{t:t+h} \tag{5}$$

Where $0 \leq \lambda \leq 1$ and $\hat{r}^{SII}_{t:t+h}$ is the predictive regression forecast based on the short interest and $\hat{r}^{i}$ on one of the G&W predictors. In equation (4) if $\lambda = 1$ the SII predictor encompasses the G&W predictor, because the optimal combination forecast excludes the forecast based on the popular G&W predictor. In equation (5) if $\lambda = 1$ the G&W predictor encompasses the SII predictor. If $\lambda > 0$ both the SII and the G&W provides information that is useful for forecasting excess returns and is not contained in the other. To test if the estimate of λ is significant we will use the approach of Harvey, Leybourne and Newbold (1998).

3.2 Forecasting Performance

To compare the forecasting performance of the short interest with other predictors, we can compare the MSPE values of the different predictors. To find whether a predictor gives significantly more accurate forecasts than the other, we can compare the MSPE values using a Diebold and Mariano (1995) test, which asymptotically follows a standard normal distribution. The DM-statistic is defined as the following:

$$DM = \frac{\bar{d}}{\sqrt{V(\hat{d}_{t+1})/P}} \overset{\alpha}{\sim} N(0,1) \tag{6}$$

where $\bar{d}$ is the sample mean of d_{t+1}, $V(\hat{d}_{t+1})$ is an estimate of the variance of d_{t+1} and $d_{t+1} = e^2_{i,t+1|t} - e^2_{j,t+1|t}$ are the differences between the forecast errors of model i and j. The null hypothesis of this test is the assumption of equal forecast accuracy, so $E[d_{t+1}] = 0$.

3.3 Forecasting Combinations

Until now, we only considered predictive regression models with one predictor. However, it may also be useful to combine more predictors in one model. For this purpose, we use the paper of Elliot et al. (2013), in which a new method is proposed by combining forecasts based on complete subset regressions. We have 15 different linear models with one explanatory variable. However, with these 15 predictors, we can also create models with k = {1, ..., 15} predictor variables. As it is stated in the paper, for a given set of potential predictor variables we can combine forecasts from all possible linear regression models that keep the number of predictors fixed. For example, with K possible predictors, there are K unique univariate models and $n_{k,K} = K!/((K-k)!k!)$ different k-variate models for k $\leq$ K. The set of models for a fixed value of k is referred to as a complete subset. We decided to use $k = 2$ as this had one of the lowest out-of-sample mean squared error (MSE)-values in their paper. Therefore the number of linear model increases from 15 models with one predictive variable to 105 models with two predictive variables.

We will use these linear models to try to improve the predictive performance by constructing a weighted average of forecasts as in equation (7). Here $\hat{r}^{(c)}_t$ denotes the combined forecast, $\hat{r}^{(i)}_t$ denotes the forecast from model i and n is the number of models considered.

$$\hat{r}^{(c)}_t = \omega_{1,t}\hat{r}^{(1)}_t + \omega_{2,t}\hat{r}^{(2)}_t + ... + \omega_{n,t}\hat{r}^{(n)}_t \tag{7}$$

First we will use an equally-weighted average, however, we try to improve this using the method in Rapach et al. (2010). The method we propose is based on Stock and Watson (2004), where the combining weights formed at time t are functions of the historical forecasting performance of the individual models over a holdout out-of-sample period. This method assigns greater weights to individual predictive regression model forecasts that have lower MSPE values over the holdout out-of-sample period. The weights are calculated using:

$$\omega_{i,t} = \phi_{i,t}^{-1} / \sum_{j=1}^{N} \phi_{j,t}^{-1} \tag{8}$$

where

$$\phi_{i,t} = \sum_{s=m}^{t-1} \theta^{t-1-s} (r_{s+1} - \hat{r}_{i,s+1})^2 \tag{9}$$

We shall also re-estimate the weights after every time period after the holdout period as well. Using a θ below 1 (such as 0.9) makes sure that more recent predictor performance is valued more. Thus, the weights are continually updated after every time-period.

3.4 Moving Window

Instead of the expanding window we have been using so far, we can also estimate the coefficients using a moving window as we move through the out-of-sample period. This may be useful if our sample contains structural breaks in the relation between excess returns and predictor variables. Otherwise, an expanding window is more useful as it includes more data. To decide how many observations to include in this window we need to search for structural breaks. To do this, we will make a figure showing the changing nature of the relationships between the equity premium and the individual economic variables over the full sample. We will do this using the correlation between r_{t+1} and x_t calculated on the basis of ten-year moving windows of data as in Rapach et al. (2010).

3.5 Economic Value

We will interpret the economic value of SII's predictive performance from an asset allocation perspective as in Rapach et al. (2016). The values of R^2_{OS} are relatively small for predictive regression models. However, the limitation of R^2_{OS} does not explicitly account for the risk borne by an investor over the out-of-sample period. Specifically, we consider a mean-variance investor who allocates between equities and risk-free bills with relative risk aversion coefficient γ by using a predictive regression forecast of excess stock returns. At the end of period t, a mean-variance investor allocates the following share of his or her portfolio to equities in period t+1. It can be described by the equation below:

$$w_t = \frac{1}{\gamma} \frac{\hat{r}_{t+1}}{\hat{\sigma}^2_{t+1}} \tag{10}$$

where $\hat{r}_{t+1}$ is a predictive regression excess return forecast and $\hat{\sigma}^2_{t+1}$ is a forecast of the excess return variance.

The investor who allocates using equation (10) realizes an average utility or certain equivalent return (CER) of:

$$CER = \bar{R}_p - \frac{1}{2}\gamma\sigma_p^2 \tag{11}$$

where $\bar{R_p}$ and $\hat{\sigma_p}^2$ are the mean and variance respectively, of the portfolio return over the forecast period. The CER gain is the difference between using the predictive regression forecast and the prevailing benchmark forecast. Therefore, the CER gain shows if it has economic value to use a predictive regression forecast model. The interpretation of the CER gain can therefore be seen as the fee an investor would pay to have access to such a model.

We will also calculate Sharpe ratios for the portfolios, which allows us to compare portfolio performance without the need to take risk aversion into account. If R^2 is large relative to S^2, then an investor can use the information in the predictive regression to obtain a large proportional increase in portfolio return. We will compare the one predictor models with the weighted forecast combination model.

4 Results

4.1 Base model

In-Sample

Table 4 in the Appendix shows the OLS estimate of β from equation (1) for the full sample using horizons 1, 3, 6 and 12. Furthermore table 5 and 6 show the OLS estimate of β using sample I and sample II and adding an extra horizon 2. Looking at the monthly horizon for the full sample, only the predictors RVOL, LTR, TMS, DFR and SII are significant. DFR is the only predictor that competes with SII, regarding their influence. They both have an estimate of 0.50 and are significant at 1% level. If we compare this with the split samples, we see some differences. For sample I (1973:01-1993:12) DP, DY, RVOL, NTIS, TBL LTR, TMS, DFY, DFR and SII are significant. And the estimate for NTIS is very large in comparison to the full sample. We now get negative estimates for LTY, which implies that we should not have taken the minus for sample I. For the full sample all estimates were quite small, except for SII and DFR. However, for sample I the estimates are larger for the other predictors and SII is not larger compared to the estimates of the other predictors. But in sample II (1994:01-2014:12), see table 6, SII performs much better. Only DP, DY and SII show significant predictive ability and these predictors have almost the same $\hat{\beta}$ estimates, namely 0.48, 0.49 and 0.48 respectively.

Looking at the R^2 statistics, they may seem rather small especially for the monthly returns. However, Campbell and Thompson (2008) argue that a monthly R^2 statistic of 0.5% represents an economically meaningful degree of return predictability. For the full sample, the significant estimates are all above this threshold. The SII and DFR have the biggest monthly R^2, namely of 1.24. The "$SII(-)|PC$" estimated slope coefficient and partial R^2, is even a bit higher than a regression containing only the predictor SII. The "$SII(-)|PC$" is a multiple regression that includes an intercept, SII and the first three principal components.

For sample I, the R^2 of SII is not as high as some of the other predictors. DFR remains a strong predictor with an R^2 of 2.61. Nonetheless, in sample II, SII is by far the strongest predictor, even better than the multiple regression.

The In-Sample results show that the predictive power of SII is as good as or even better than the best individual G&W predictors for the full sample and sample II. On the contrary, it lacks predictive power in sample I (1973:01-1993:12).

Out-of-Sample

For the Out-of-Sample tests, we will first look at the R^2_{OS} which compares the MSFE error at the h-month horizon with the prevailing mean benchmark forecast. In the full sample described via Appendix in table 7, we see that only the short index predictor is able to perform better than the historical mean average for all the different horizons. TMS outperforms the prevailing mean benchmark forecast annually and INFL semi-annually. However, the R^2_{OS} of SII is larger for both these horizons. The rest of the G&W predictors lack predictive value. In addition, if we look at the R^2_{OS} in the split samples in table 8 and 9 the R^2_{OS}'s for SII are all negative and most are not significant.

In order to evaluate whether one of the popular predictors encompasses the SII, we look at the second part of table 7, 8 and 9. If the $\hat{\lambda}$'s were negative we set them to 0 and if they were greater than 1 we set them to 1. In full sample, all $\hat{\lambda}$ estimates are significant and are very close to or equal to 1. If $\lambda = 1$, this means the total weight will go to SII in combination forecast and one of the popular predictors will get weight 0. Therefore, SII encompasses the popular G&W predictors for out-of-sample forecasts for the full sample. If we look at sample I, most of the $\hat{\lambda}$ estimates are insignificant, except for BM and NTIS. The $\hat{\lambda}$ of the Book-to-Market value is 1 for the monthly results, indicating that SII encompasses BM. Nonetheless, the net equity expansion seems to have an important predictive value and SII does not encompass this predictor in this sample. In sample II both SII and the G&W most $\lambda > 0$ but not 1, which implies that both the SII and the G&W

predictor provides useful information.

4.2 Forecasting Performance

To compare the relative accuracy of the SII predictor compared to other single-predictor models we ran the Diebold-Mariano test for the out-of-sample forecasting period starting from 1990. The results are shown in table 1 below. We see that the test statistic appears to be negative everywhere, indicating that the squared forecast errors of the SII are smaller than the forecast errors of any other single predictor variable. Many of the DM-statistics are significant at the 1% level, which shows that the relative accuracy of the SII predictor is significantly better than any other predictor, even the historical mean. The exceptions to this are the values for when h=1. There we see that many of the statistics are not, or barely, significant. For horizons larger than 1 we find that all the statistics are highly significant, so it appears to be that the relative predictive accuracy of the SII is high when forecasting over larger horizons.

Table 1: Diebold-Mariano statistics for out-of-sample forecast accuracy, 1990:01-2014:12

Predictor	h=1	h=3	h=6	h=12
DP	-2.81***	-5.19***	-7.33***	-10.80***
DY	-2.82***	-5.13***	-7.37***	-10.77***
EP	-1.90*	-3.86***	-5.60***	-7.91***
DE	-1.70*	-3.41***	-5.04***	-4.27***
RVOL	-1.76*	-3.35***	-4.47***	-4.57***
BM	-2.13**	-3.83***	-5.32***	-7.18***
NTIS	-2.55**	-4.34***	-5.73***	-6.21***
TBL	-1.63	-2.86***	-3.85***	-3.92***
LTY	-1.82*	-3.49***	-5.14***	-6.99***
LTR	-1.54	-2.99***	-3.65***	-3.74***
TMS	-1.67*	-2.94***	-3.39***	-2.18**
DFY	-2.79***	-4.69***	-5.29***	-5.13***
DFR	-1.07	-2.89***	-3.78***	-3.81***
INFL	-1.78*	-2.98***	-3.35***	-2.76***
Historical Mean	-1.55	-2.88***	-3.96***	-3.84***

This table shows the Diebold-Mariano statistic values for the comparisons of the accuracies between the SII predictor and the Goyal-Welch and historical mean predictors. * marks significant at the 10% level, ** marks significant at the 5% level and *** marks significant at the 1% level.

4.3 A Weighted Combination

For the out-of-sample evaluation of our weighted model, we use three different in-sample-ends (1989, 1999 and 2009) so that the forecast period is 25, 15 and 5 years respectively. In the equally weighted model, we combine the 105 linear regressions having two predictor variables and give them all the weight $\omega = 1/105$. The R^2_{OS}'s for the different horizons (monthly, quarterly, semi-annually and annually) are given in Table 10. The R^2_{OS} are negative for the out-of-sample period 1990:01-2014:12 and 2000:01-2014:12 (except for h=12) and this equally weighted model is hence not a better predictor than the constant expected return model. In contrast, if the out-of-sample period is only 5 years, we get positive R^2_{OS}'s for every horizon.

We tried to improve this by using the discount mean squared prediction error (DMSPE) combining method to create different weights. We used $\theta = 0.9$ as in Rapach et al. (2010). We use an intial holdout period of 1 year (first year of the out-of-sample period) to estimate the combining weights, after which we re-estimate these after every new observation. The R^2_{OS}'s for this combined weighted model are all positive, except for the monthly horizon as can be seen in table 11. The smaller the out-of-sample period the more accurate the model gets. This weighted combination model shows that a predictive regression model using SII and the G&W predictors outperforms the prevailing mean benchmark in terms of MSFE for h=3 or larger using different in-sample ends.

4.4 Moving Window

Figure 1 in the Appendix shows the correlation of the 14 G&W predictors and the SII with excess returns in the period 1990:01 to 2014:12, calculated on basis of a ten-year moving window of data. The correlations fluctuate and may indicate structural breaks. Therefore we decided to also use a ten-year moving window in our weighted combined model. However when using a moving window of ten years in our combined DMSPE model we get insignificant and negative R^2_{OS}'s. Next we tried different windows using trial and error and none of them gave significant results for the R^2_{OS}'s. We decided to continue with a recursively expanding estimation scheme, as in Rapach et al. (2010), Goyal and Welch (2008) and Elliot et all. (2013).

4.5 Asset Allocation

As in Rapach et al. (2016) we calculated the CER gains and Sharpe Ratios using three periods. The forecast evaluation period 1991:01-2014:12, a forecast evaluation period predating the Global Financial Crisis (1991:01-2006:12) and a period surrounding the recent crisis (2007:01-2014:12). We calculated it for the 15 predictor variables, the weighted combined forecast and for a Buy&Hold strategy. We deleted the first 12 months because this is the hold-out period for the combined forecast model. The results using a relative risk aversion coefficient of three are shown in table 12. The performance of SII is clearly better than that of the other predictors. Among the 14 popular predictors only TBL, TMS, DFR and INFL generate positive CER gains but not for every horizon and the gains are much lower than those of SII. We also included a buy-and-hold portfolio that passively holds the market portfolio. The SII predictor also outperforms this strategy. For the monthly horizon the SII provides 406 basis points in the full forecast period, while the buy-and-hold strategy only produces 209. In the period corresponding to the Global Financial Crisis, SII offers striking gains. The SII predictor is measure of market pessimism based on the short interest data and this table shows that the SII is especially very useful during acute macroeconomic stress. We also included the combined model using the DMSPE weights, however, this weighted model is not able to surpass the SII predictive regression.

Next, we will look at the Sharpe ratios in table 13. With these ratios, we can compare portfolio performance independently of risk aversion. The ratios for the prevailing mean benchmark are also included and we can compare this with the 14 G&W predictors, the SII , Buy & Hold and the combined forecast. Only some of the G&W predictors are able to outperform the prevailing mean for some horizons, TBL and DFR produce the highest Sharpe ratios. However, SII produces Sharpe ratios that are much higher than all the other predictors including prevailing mean and buy-and-hold for the full forecast period and the period surrounding the recent crisis. For the forecast predating the Global Financial Crisis, the Sharpe ratios are quite similar to those of the prevailing mean. And also this table shows the impressing performance of the SII during acute macroeconomic stress. Therefore, the information contained in SII is not only statistically significant, but very valuable economically as well. Also for the Sharpe ratios, our combined weighted model is not able to outperform the SII predictor.

5 Conclusion and Future Research

In this report, the focus has been on two key research questions. Firstly, whether the short interest is better in predicting aggregate stock returns compared to other predictors. And secondly, if forecasting combinations using short interest and other variables further improve the forecasting performance.

We compared the short index as a predictor to the 14 popular Goyal and Welch predictive variables. We first used a predictive regression model including one predictive variable and showed that the forecasting performance of short interest is significantly better than all the G&W predictors in predicting aggregate stock returns. The short index predictor is able to outperform the historical mean average for all the different horizons, while most of the G&W predictors lack predictive value. However, when we split the sample in sample I and sample II (1973:01-1993:12 and 1994:01-2014:12), the SII is not as strong of a predictor in sample I.

Furthermore, we used forecasting combinations to try to improve the forecasting performance both statistically and economically. The weighted combination model, using the DMSPE method weights, was able to outperform the prevailing mean benchmark in terms of MSFE for h=3 and larger using different out-of-sample periods. The DMSPE weighted combination model gave better results than an equally weighted model and therefore we included this model in our economic evaluation. However, this combined model was not able to outperform the SII forecast regression model economically.

The SII forecast regression model generates positive CER gains and outperforms the buy-and-hold strategy. Also when comparing the portfolio performance using Sharpe ratios, the SII predictor was able to perform better than the 14 G&W predictors, the combined forecast and the buy-and-hold strategy. Especially for the forecasting period regarding the Global Financial Crisis, the results are impressing.

For future research we suggest focusing on combining the forecasting models, but not including all linear combinations as we did. Furthermore, there may be better weighting combinations than the DMSPE weights which we used. All this might still lead to a combined forecasting model that could outperform the SII.

References

Clark, T.E. and West, K.D. Approximately Normal Tests For Equal Predictive Accuracy In Nested Models. *Journal of Econometrics 138*, 291–311, 2007.

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

Elliott, G., Gargano, A., and Timmermann, A. Complete subset regressions. *Journal of Econometrics,* 177(2):357–373, 2013.

Goyal, A. and Welch, I. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies,* 21(4):1455–1508, 2008.

Inoue, A. and Kilian, L. In-Sample Or Out-of-Sample Tests of Predictability: Which One Would We Use? *Econometric Reviews 23*, 371–402, 2005.

Rapach, D., Ringgenberg, M., and Zhou, G. *Journal of the Operations Research Society of America"*, 2(3):339-341, 2016.

Rapach D., Strauss, J. and Zhou, G. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies,* 23(2):821–862, 2010.

Stock, J. H., and M. W. Watson Combination Forecasts of Output Growth in a Seven-Country Data Set. *Journal of Forecasting* 23:405–30, 2004.

Appendix

Table 2: Summary statistics 1973:01-2014:12

Variable	Mean	Median	1st percentile	99th percentile	Std. dev
DP	-3.62	-3.57	-4.47	-2.84	0.44
DY	-3.61	-3.57	- 4.47	-2.83	0.44
EP	-2.82	-2.83	-4.60	-1.97	0.49
DE	-0.80	-0.86	-1.24	1.03	0.35
RVOL (ann.)	0.15	0.14	0.06	0.31	0.05
BM	0.49	0.38	0.13	1.14	0.29
NTIS	0.01	0.01	-0.05	0.04	0.02
TBL (%, ann.)	5.05	5.05	0.02	14.98	3.44
LTY (%, ann.)	7.17	7.06	2.26	13.94	2.73
LTR (%)	0.74	0.82	-6.78	9.29	3.13
TMS (%, ann.)	2.11	2.35	-2.18	4.37	1.51
DFY (%, ann.)	1.10	0.96	0.56	2.86	0.47
DFR (%)	-0.01	0.05	-4.84	3.91	1.47
INFL (%)	0.34	0.29	-0.50	1.33	0.34
EWSI (%)	2.15	1.31	0.22	7.86	1.98
SII	0.00	-0.09	-2.13	2.45	1.00

The database contains 504 monthly observations for January 1973 to December 2014. The table displays summary statistics for 14 predictor variables from Goyal and Welch (2008) and aggregate short interest. DP is the log dividend-price ratio, DY is the log dividend yield, EP is the log earnings-price ratio, DE is the log dividend-payout ratio, RVOL is the volatility of excess stock returns, BM is the book-to-market value ratio for the Dow Jones Industrial Average, NTIS is net equity expansion, TBL is the interest rate on a three-month Treasury bill, LTY is the long-term government bond yield, LTR is the return on long-term government bonds, TMS is the long-term government bond yield minus the Treasury bill rate, DFY is the difference between Moody's BAA- and AAA-rated corporate bond yields, DFR is the long-term corporate bond return minus the long-term government bond return, and INFL is inflation calculated from the CPI for all urban consumers. EWSI is the equal-weighted mean across all firms of the number of shares held short in a given firm (from Compustat) normalized by each firm's shares outstanding. EWSI includes common equities, ADRs, ETFs, and REITs. SII is the detrended log of EWSI, constructed by removing a linear trend from the log of EWSI; SII is standardized to have a standard deviation of one.

Table 3: Summary statistics 1973:01-1993:12 and 1994:01-2014:12

	Variable	Mean	Median	1st perc	99th perc	Std. dev
1973:01-1993:12						
	DP	−3.24	−3.24	−3.62	−2.80	0.22
	DY	−3.23	−3.23	−3.62	−2.80	0.22
	EP	−2.51	−2.47	−3.24	−1.94	0.36
	DE	−0.73	−0.78	−0.97	−0.27	0.17
	RVOL (ann.)	0.15	0.14	0.06	0.31	0.05
	BM	0.72	0.73	0.31	1.19	0.25
	NTIS	0.01	0.01	−0.03	0.04	0.02
	TBL (%, ann.)	7.41	7.15	2.91	5.48	2.80
	LTY (%, ann.)	9.28	8.59	6.27	14.15	1.98
	LTR (%)	0.81	0.69	−7.09	11.40	3.24
	TMS (%, ann.)	1.87	2.31	−2.91	4.28	1.65
	DFY (%, ann.)	1.24	1.15	0.62	2.61	0.46
	DFR (%)	0.00	0.06	−3.92	3.70	1.15
	INFL (%)	0.49	0.41	−0.30	1.40	0.34
	EWSI (%)	0.64	0.48	0.21	1.34	0.38
	SII	0.00	−0.14	−1.61	3.09	1.00
1994:01-2014:12						
	DP	−4.00	−4.00	−4.48	−3.39	0.24
	DY	−4.00	−3.99	−4.48	−3.40	0.24
	EP	−3.13	−2.98	−4.81	−2.66	0.40
	DE	−0.87	−0.98	−1.24	1.28	0.45
	RVOL (ann.)	0.15	0.14	0.06	0.31	0.06
	BM	0.27	0.28	0.13	0.41	0.07
	NTIS	0.01	0.01	−0.05	0.04	0.02
	TBL (%, ann.)	2.69	2.75	0.01	6.09	2.17
	LTY (%, ann.)	5.05	4.98	2.19	8.00	1.41
	LTR (%)	0.67	0.92	−6.48	7.89	3.01
	TMS (%, ann.)	2.36	2.38	−0.26	4.43	1.32
	DFY (%, ann.)	0.96	0.86	0.55	3.09	0.44
	DFR (%)	−0.02	0.04	−6.07	5.92	1.73
	INFL (%)	0.19	0.19	−0.82	0.83	0.28
	EWSI (%)	3.66	4.08	1.24	8.22	1.78
	SII	0.00	−0.10	−1.70	2.46	1.00

The database contains 252 monthly observations for January 1973 to December 1993 and 252 monthly observations for January 1994 to December 2014. The table displays summary statistics for 14 predictor variables from Goyal and Welch (2008) and aggregate short interest. See the notes for table 2 for the variable definitions.

Table 4: In-sample PR results 1973:01-2014:12

Predictor	$h=1$ $\hat{\beta}$	$R^2(\%)$	$h=3$ $\hat{\beta}$	$R^2(\%)$	$h=6$ $\hat{\beta}$	$R^2(\%)$	$h=12$ $\hat{\beta}$	$R^2(\%)$
DP	0.15	0.12	0.17	0.39	0.19	0.93	0.20	2.06
	[0.75]		[0.94]		[1.06]		[1.12]	
DY	0.18	0.15	0.18	0.44	0.19	1.00	0.21	2.18
	[0.86]		[0.99]		[1.10]		[1.15]	
EP	0.09	0.04	0.06	0.06	0.06	0.09	0.08	0.34
	[0.37]		[0.27]		[0.25]		[0.43]	
DE	0.06	0.02	0.12	0.22	0.16	0.68	0.14	1.03
	0.23		0.57		0.88		1.21	
RVOL	0.35	0.62	0.33	1.53	0.27	1.90	0.18	1.58
	[1.90]**		[2.12]**		[1.95]**		1.26	
BM	-0.01	0.00	0.01	0.00	0.04	0.05	0.05	0.14
	-0.05		0.06		0.23		0.29	
NTIS(-)	0.07	0.02	0.00	0.00	-0.01	0.00	0.02	0.02
	0.27		0.00		-0.02		0.07	
TBL (-)	0.26	0.34	0.22	0.67	0.19	0.91	0.17	1.38
	1.28		1.20		0.99		0.99	
LTY (-)	0.15	0.10	0.10	0.15	0.08	0.15	0.00	0.00
	0.71		0.57		0.40		0.03	
LTR	0.33	0.55	0.14	0.27	0.23	1.41	0.15	1.14
	[1.67]**		0.92		[2.48]***		[2.90]**	
TMS	0.33	0.56	0.31	1.33	0.28	2.14	0.36	6.52
	[1.66]**		[1.72]*		[1.62]*		[2.27]**	
DFY	0.15	0.11	0.16	0.37	0.24	1.51	0.19	1.75
	0.55		0.66		1.24		1.18	
DFR	0.50	1.24	0.23	0.72	0.16	0.71	0.06	0.17
	[1.58]*		1.26		[1.38]*		0.89	
INFL (-)	0.06	0.02	0.17	0.43	0.26	1.81	0.27	3.64
	0.21		0.90		[1.62]*		[2.04]*	
SII(-)	0.50	1.24	0.56	4.37	0.57	8.07	0.53	12.89
	[2.50]***		[2.88]***		[2.73]**		[2.70]**	
SII(-)|PC	0.51	1.27	0.58	4.54	0.59	8.64	0.55	13.69
	[2.65]***		[3.02]***		[2.79]**		[2.73]**	

This table shows the in-sample predictive regression estimate results for the period 1973:01-2014:12. It shows the OLS estimates of β and R^2 for the predictive regression using one variable. Where our explanatory variable is the S&P log excess return for month t. Each predictor variable is standardized. The regression uses a h-month horizon. We test the $\hat{\beta}$ estimates using a heteroskedasticity-and autocorrelation-robust t-statistic. The minus indicates that we take the negative of the variable. This is because we use a one-sides test, namely $H_0 = \beta = 0$ against $H_A = \beta > 0$. The significance levels are 10%, 5% and 1% shown using *, **, and ***. See the notes for table 2 for the variable definitions. The "$SII(-)|PC$" row corresponds to a multiple predictive regression that includes an intercept, SII, and the first three principal components extracted from the non-SII predictors in the first column. For this multiple predictive regression, the table reports the estimated slope coefficient and partial R2 statistic corresponding to SII.

Table 5: In-sample PR results 1973:01-1993:12

	$h=1$		$h=2$		$h=3$		$h=6$		$h=12$	
Predictor	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$
DP	0.48	1.04	0.48	2.07	0.47	3.02	0.47	5.87	0.45	11.11
	[1.47]*		[1.61]*		[1.68]*		[1.93]*		[2.07]*	
DY	0.49	1.09	0.47	1.99	0.46	2.91	0.47	5.77	0.45	10.80
	[1.46]*		[1.57]*		[1.69]*		[1.95]**		[2.03]*	
EP	0.14	0.08	0.15	0.21	0.16	0.34	0.19	0.91	0.20	2.21
	[0.45]		[0.56]		[0.63]		[0.85]		1.06	
DE	0.33	0.51	0.30	0.84	0.28	1.05	0.22	1.33	0.17	1.60
	[1.33]		[1.38]		[1.31]		[1.12]		[1.12]	
RVOL	0.54	1.33	0.51	2.38	0.49	3.32	0.42	4.57	0.29	4.58
	[1.90]**		[2.07]**		[2.11]**		[2.17]*		[1.61]	
BM	0.11	0.06	0.12	0.14	0.12	0.19	0.14	0.52	0.15	1.19
	[0.36]		[0.44]		[0.43]		[0.58]		[0.78]	
NTIS(-)	0.60	1.64	0.55	2.70	0.56	4.20	0.58	8.91	0.56	17.46
	[2.73]***		[2.86]***		[2.94]***		[3.05]***		[2.55]**	
TBL (-)	0.38	0.66	0.30	0.80	0.26	0.95	0.17	0.77	0.06	0.21
	[1.55]*		[1.54]*		[1.42]		[0.78]		[0.25]	
LTY (-)	-0.00	0.00	-0.05	0.02	-0.09	0.12	-0.14	0.55	-0.28	4.32
	[-0.01]		[-0.18]		[-0.37]		[-0.50]		[-0.99]	
LTR	0.48	1.08	0.48	2.14	0.33	1.49	0.37	3.81	0.25	3.92
	[1.65]*		[2.15]**		[1.57]*		[2.68]***		[3.66]***	
TMS	0.64	1.89	0.55	2.80	0.55	4.16	0.44	5.46	0.42	10.30
	[2.27]***		[2.34]**		[2.42]**		[1.96]*		[1.83]	
DFY	0.73	2.48	0.66	3.97	0.63	5.42	0.59	9.54	0.36	7.51
	[2.58]***		[2.39]**		[2.31]**		[2.32]**		[1.67]	
DFR	0.75	2.61	0.20	0.37	0.26	0.93	0.18	0.89	-0.03	0.07
	[2.26]**		[0.77]		[1.31]		[1.82]*		[-0.77]	
INFL (-)	0.26	0.31	0.33	1.03	0.14	0.26	0.19	1.04	0.21	2.64
	[0.87]		[1.38]		[0.60]		[0.82]		[1.22]	
SII(-)	0.48	1.09	0.50	2.29	0.46	2.94	0.43	5.16	0.50	15.03
	[1.77]**		[1.95]*		[1.90]*		[1.89]*		[2.19]*	
SII(-)|PC	0.47	0.80	0.46	1.51	0.40	1.73	0.32	2.29	0.43	8.44
	[1.69]**		[1.90]**		[1.78]*		[1.63]		[2.22]*	

This table shows the in-sample predictive regression estimate results for the period 1973:01-1993:12. It shows the OLS estimates of β and R^2 for the predictive regression using one variable. Where our explanatory variable is the S&P log excess return for month t. Each predictor variable is standardized. The regression uses a h-month horizon. We test the $\hat{\beta}$ estimates using a heteroskedasticity-and autocorrelation-robust t-statistic. The minus indicates that we take the negative of the variable. This is because we use a one-sides test, namely $H_0 = \beta = 0$ against $H_A = \beta > 0$. The significance levels are 10%, 5% and 1% shown using *, **, and ***. See the notes in table 2 for the variable definitions. The "$SII(-)|PC$" row corresponds to a multiple predictive regression that includes an intercept, SII, and the first three principal components extracted from the non-SII predictors in the first column. For this multiple predictive regression, the table reports the estimated slope coefficient and partial R2 statistic corresponding to SII.

Table 6: In-sample PR results 1994:01-2014:12

	$h = 1$		$h = 2$		$h = 3$		$h = 6$		$h = 12$	
Predictor	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$	$\hat{\beta}$	$R^2(\%)$
DP	0.53	1.46	0.57	3.15	0.59	4.98	0.66	10.98	0.71	22.74
	[1.47]*		[1.79]*		[1.93]*		[2.68]**		[3.74]**	
DY	0.60	1.86	0.59	3.40	0.63	5.69	0.68	11.79	0.73	24.25
	[1.80]**		[2.03]**		[2.32]**		[3.03]**		[4.20]***	
EP	0.31	0.51	0.26	0.67	0.21	0.62	0.16	0.68	0.20	1.88
	[0.75]		[0.66]		[0.54]		[0.43]		[0.65]	
DE	0.00	0.00	0.07	0.05	0.13	0.24	0.21	1.07	0.20	1.77
	[0.01]		[0.18]		[0.37]		[0.77]		[1.37]	
RVOL	0.21	0.24	0.21	0.43	0.19	0.51	0.15	0.53	0.09	0.34
	[0.84]		[0.91]		[0.85]		[0.69]		[0.39]	
BM	0.32	0.54	0.40	1.54	0.46	3.02	0.62	9.39	0.63	17.37
	[1.06]		[1.58]*		[2.04]**		[3.27]***		[2.98]**	
NTIS(-)	-0.48	1.21	-0.52	2.62	-0.60	5.13	-0.62	9.87	-0.53	12.78
	[-1.12]		[-1.29]		[-1.47]		[-1.64]		[-1.70]	
TBL(-)	0.07	0.03	0.08	0.06	0.09	0.12	0.14	0.45	0.24	2.36
	[0.27]		[0.33]		[0.37]		[0.56]		[0.97]	
LTY(-)	0.19	0.19	0.17	0.28	0.16	0.35	0.14	0.47	0.07	0.22
	[0.82]		[0.79]		[0.76]		[0.65]		[0.28]	
LTR	0.18	0.18	0.03	0.01	-0.06	0.04	0.09	0.20	0.05	0.09
	[0.67]		[0.11]		[-0.28]		[0.80]		[0.65]	
TMS	-0.08	0.04	-0.05	0.02	-0.02	0.00	0.08	0.14	0.29	3.89
	[-0.30]		[-0.18]		[-0.07]		[0.30]		[1.38]	
DFY	-0.35	0.62	-0.30	0.88	-0.22	0.68	-0.02	0.01	0.10	0.40
	[-0.70]		[-0.65]		[-0.48]		[-0.06]		[0.45]	
DFR	0.35	0.65	0.07	0.05	0.22	0.67	0.17	0.68	0.12	0.64
	[0.69]		[0.17]		[0.77]		[0.84]		[1.15]	
INFL(-)	-0.32	0.54	-0.22	0.45	0.13	0.25	0.29	2.15	0.30	4.00
	[-0.91]		[-0.60]		[0.41]		[1.66]		[2.28]**	
SII(-)	0.63	2.07	0.70	4.67	0.76	8.04	0.79	14.87	0.67	17.70
	[2.13]**		[2.42]**		[2.55]**		[2.42]**		[2.18]*	
SII(-)|PC	0.62	2.01	0.70	4.68	0.75	8.25	0.81	17.05	0.71	24.90
	[2.3]**		[2.63]**		[2.77]**		[2.57]**		[2.65]*	

This table shows the in-sample predictive regression estimate results for the period 1994:01-2014:12. It shows the OLS estimates of β and R^2 for the predictive regression using one variable. Where our explanatory variable is the S&P log excess return for month t. Each predictor variable is standardized. The regression uses a h-month horizon. We test the $\hat{\beta}$ estimates using a heteroskedasticity-and autocorrelation-robust t-statistic. The minus indicates that we take the negative of the variable. This is because we use a one-sides test, namely $H_0 = \beta = 0$ against $H_A = \beta > 0$. The significance levels are 10%, 5% and 1% shown using *, **, and ***. See the notes in table 2 for the variable definitions. The "$SII(-)|PC$" row corresponds to a multiple predictive regression that includes an intercept, SII, and the first three principal components extracted from the non-SII predictors in the first column. For this multiple predictive regression, the table reports the estimated slope coefficient and partial R2 statistic corresponding to SII.

Table 7: Out of sample test results, 1990:01-2014:12

Predictor	R^2_{OS} % h=1	h=3	h=6	h=12	λ_{SII} h=1	h=3	h=6	h=12
DP	-2.06	-5.80	-10.97	-26.39	1.00***	1.00***	1.00***	1.00***
DY	-2.20	-5.66	-11.04	-25.82	1.00***	1.00***	1.00***	1.00***
EP	-1.14	-4.24	-8.85	-16.39	1.00***	1.00***	1.00***	1.00***
DE	-2.27	-6.29	-7.85	-3.56	1.00**	1.00***	1.00***	0.96***
RVOL	-0.56	-1.40	-1.77	-3.40	0.98***	1.00***	1.00***	1.00***
BM	-0.56	-1.75	-3.55	-9.68	1.00***	1.00***	1.00***	1.00***
NTIS	-3.23	-8.88	-19.06	-27.82	1.00***	1.00***	1.00***	1.00***
TBL	-0.38	-0.92	-1.66	-1.75	1.00**	1.00***	1.00***	1.00**
LTY	-0.31	-1.53	-3.66	-11.85	1.00***	1.00***	1.00***	1.00***
LTR	-0.51	-1.59	-0.94	-0.91	1.00***	1.00***	1.00***	0.94***
TMS	-0.76	-1.72	-1.43	3.35**	1.00***	1.00***	1.00***	0.83**
DFY	-3.07	-7.15	-8.92	-7.35	1.00***	1.00***	1.00***	1.00***
DFR	-1.75	-1.11	-0.39	-0.78	0.95**	1.00***	1.00***	1.00***
INFL	-0.64	-0.32	1.95*	3.24	1.00***	1.00***	1.00***	0.89***
SII	1.94***	6.54***	11.70***	13.24**	-	-	-	-

This table shows the out of sample test results for the forecast period 1990:01-2014:12 using the in-sample period 1973:01-1989:12. The second through fifth columns are the out-of-sample R^2 statistics (%). They report the proportional reduction in mean squared forecast error at the h-month horizon for a predictive regression forecast of the S&P 500 log excess return. Clark & West (2007) is used to test H_0: the prevailing mean MSFE $\leq$ the predictive regression MSFE versus H_A: the prevailing mean MSFE $>$ the predictive regression MSFE. The sixth to ninth columns show the estimated weight on the predictive regression forecast based on SII using the encompassing tests. A combination forecast is used taking the form of a convex combination of a predictive regression forecast based on SII and a a predictive regression forecast based on one of the popular G&W variables shown in the first column. The statistical significance is based on the Harvey, Leybourne, and Newbold (1998) statistic for testing the null hypothesis that the weight on the SII-based forecast is equal to zero against the alternative that the weight on the SII-bases forecast is greater than zero. The significance levels are 10%, 5% and 1% shown using *, **, and ***.

Table 8: Out of sample test results, 1982:01-1993:12

Predictor	R^2_{OS} % h=1	h=2	h=3	h=6	h=12	λ_{SII} h=1	h=2	h=3	h=6	h=12
DP	-3.77	-7.04	-8.66	-13.56	-23.51	0.57	0.43	0.33	0.25	0.07
DY	-3.36	-5.74	-7.45	-11.41	-20.44	0.51	0.32	0.26	0.19	0.02
EP	-4.65	-9.04	-12.18	-22.37	-48.82	0.98	0.66	0.50	0.38	0.16
DE	-1.06	-2.46	-4.81	-13.04	-18.29	0.34	0.31	0.30	0.30	0.00
RVOL	-1.02	-1.91	-3.93	-7.92	-10.94	0.22	0.12	0.13	0.06	0.00
BM	-7.72	-14.72	-19.41	-38.56	-85.38	1.00**	0.97**	0.84**	0.75**	0.63**
NTIS	1.54***	2.26***	3.04***	4.63***	-14.26**	0.29*	0.27*	0.29**	0.34**	0.41***
TBL	0.57	0.71	0.77	0.73	5.67***	0.00	0.00	0.00	0.00	0.00
LTY	-0.52	-0.94	-1.22	-2.48	12.01*	0.00	0.00	0.00	0.00	0.00
LTR	0.45	3.18**	2.04*	7.51***	13.78***	0.00	0.00	0.00	0.00	0.00
TMS	1.26*	1.81	3.29*	6.59**	20.70***	0.05	0.00	0.00	0.00	0.00
DFY	-0.06	0.03	1.15	8.64*	15.79	0.05	0.18	0.00	0.00	0.00
DFR	2.44**	-0.76	-0.55	2.21**	7.89***	0.00	0.00	0.00	0.00	0.00
INFL	-0.00	0.78	-1.01	0.90	5.89**	0.00	0.00	0.00	0.00	0.00
SII	-3.25	-8.02	-12.16	-25.05	-69.75	-	-	-	-	

This table shows the out of sample test results for the forecast period 1982:01-1993:12 using the in-sample period 1973:01-1981:12. The second through sixth columns are the out-of-sample R^2 statistics (%). They report the proportional reduction in mean squared forecast error at the h-month horizon for a predictive regression forecast of the S&P 500 log excess return. Clark & West (2007) is used to test H_0: the prevailing mean MSFE $\leq$ the predictive regression MSFE versus H_A: the prevailing mean MSFE $>$ the predictive regression MSFE. The seventh through eleventh columns show the estimated weight on the predictive regression forecast based on SII using the encompassing tests. A combination forecast is used taking the form of a convex combination of a predictive regression forecast based on SII and a a predictive regression forecast based on one of the popular G&W variables shown in the first column. The statistical significance is based on the Harvey, Leybourne, and Newbold (1998) statistic for testing the null hypothesis that the weight on the SII-based forecast is equal to zero against the alternative that the weight on the SII-bases forecast is greater than zero. The significance levels are 10%, 5% and 1% shown using *, **, and ***.

Table 9: Out of sample test results, 2003:01-2014:12

Predictor	R^2_{OS} % h=1	h=2	h=3	h=6	h=12	λ_{SII} h=1	h=2	h=3	h=6	h=12
DP	-1.63	2.77	-4.12	-1.97	-0.63	0.68	0.62*	0.60**	0.33**	0.11
DY	-0.28	-1.28	-1.54	0.36	2.49	0.43	0.49*	0.47*	0.26*	0.05
EP	-8.80	-19.45	-31.27	-56.51	-78.91	1.00	1.00**	1.00**	0.95**	0.99**
DE	-7.79	-21.98	-36.65	-50.85	1.32	0.82	0.95*	0.98**	0.89**	0.16
RVOL	-0.49	-2.07	-4.35	-10.16	-27.81	0.48	0.55**	0.58***	0.52**	0.49*
BM	-1.06	-1.06	-0.28	5.40**	-4.14**	0.59	0.47*	0.39*	0.00	0.04
NTIS	-0.09	-1.30	-1.30	-5.04	-18.34	0.46	0.50*	0.49**	0.47**	0.46**
TBL	-4.66	-9.65	-15.32	-25.49	-24.05	0.81**	0.81***	0.79***	0.71***	0.45
LTY	-2.63	-5.83	-10.08	-21.54	-49.73	0.71*	0.70**	0.69***	0.64***	0.63**
LTR	-0.99	-2.36	-2.41	-1.47	-0.30	0.55	0.56**	0.51**	0.36**	0.04
TMS	-1.95	-3.88	-5.44	-5.57	7.21**	0.65*	0.64**	0.61***	0.43**	0.00
DFY	-6.66	-19.59	-44.38	-125.23	-120.23	0.67*	0.73**	0.81**	0.96**	0.95**
DFR	-6.05	-13.04	-9.35	-7.56	-4.76	1.00*	1.00***	0.74***	0.47***	0.11
INFL	-1.10	-4.74	-7.17	1.47*	3.73*	0.55	0.70**	0.66**	0.30*	0.00
SII	-0.65	-1.36*	-2.11*	-9.08	-28.59	-	-	-	-	-

This table shows the out of sample test results for the forecast period 2003:01-2014:12 using the in-sample period 1994:01-2002:12. The second through sixth columns are the out-of-sample R^2 statistics (%). They report the proportional reduction in mean squared forecast error at the h-month horizon for a predictive regression forecast of the S&P 500 log excess return. Clark & West (2007) is used to test H_0: the prevailing mean MSFE $\leq$ the predictive regression MSFE versus H_A: the prevailing mean MSFE $>$ the predictive regression MSFE. The seventh through eleventh columns show the estimated weight on the predictive regression forecast based on SII using the encompassing tests. A combination forecast is used taking the form of a convex combination of a predictive regression forecast based on SII and a a predictive regression forecast based on one of the popular G&W variables shown in the first column. The statistical significance is based on the Harvey, Leybourne, and Newbold (1998) statistic for testing the null hypothesis that the weight on the SII-based forecast is equal to zero against the alternative that the weight on the SII-bases forecast is greater than zero. The significance levels are 10%, 5% and 1% shown using *, **, and ***.

Table 10: Out of sample test results, forecast combination with expanding window using equal weights

Out-Of-Sample Period	R^2_{OS} % h=1	h=3	h=6	h=12
1990:01-2014:12	-0.13**	-0.77***	-1.37***	-2.72**
2000:01-2014:12	-0.21*	-0.35**	-0.26**	0.19**
2010:01-2014:12	3.32***	9.80***	16.76***	37.44***

This table shows the out-of-sample R-squared results from using a weighted combination of forecast models. The combined model combines the forecasts of all possible linear regression models having two predictor variables. The weights are all equal. Statistical significance for the out-of-sample R-squared statistic is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic.

Table 11: Out of sample test results, forecast combination with expanding window using DMSPE weights

Out-Of-Sample Period	R^2_{OS} % h=1	h=3	h=6	h=12
1990:01-2014:12	-0.29**	-0.18***	0.26***	-2.77**
2000:01-2014:12	-0.01	0.72**	2.27**	1.79*
2010:01-2014:12	4.68***	13.04***	24.57***	49.25***

This table shows the out-of-sample R-squared results from using a weighted combination of forecast models. The combined model combines the forecasts of all possible linear regression models having two predictor variables. The weights follow from using the discount mean square prediction error (DMSPE) combining method. Statistical significance for the out-of-sample R-squared statistic is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic. Note that the year 1990 is used as an intial holdout period.

Table 12: Out-of-sample CER gains

	1991 : 01 − 2014 : 12 period				2007 : 01 − 2014 : 12 period				1991 : 01 − 2006 : 12 period			
Predictor	h=1	h=3	h=6	h=12	h=1	h=3	h=6	h=12	h=1	h=3	h=6	h=12
DP	−3.89	−3.16	−4.06	−4.07	−0.57	0.01	0.81	1.21	−5.26	−4.49	−6.14	−6.27
DY	−3.68	−2.99	−3.90	−3.94	−0.06	0.58	1.36	1.38	−5.18	−4.48	−6.13	−6.16
EP	−0.65	−0.75	−1.52	−1.77	1.15	−0.03	−0.33	−1.13	−1.39	−1.05	−2.02	−2.01
DE	−0.61	−1.05	−2.11	−0.44	−0.12	−2.14	−5.22	−0.91	−0.81	−0.57	−0.58	−0.41
RVOL	−2.19	−2.01	−1.60	−0.40	0.90	0.70	0.51	0.99	−3.47	−3.19	−2.62	−1.09
BM	−1.00	−0.89	−1.28	−1.73	0.00	−0.14	0.18	−0.23	−1.42	−1.20	−1.91	−2.43
NTIS	−2.41	−2.44	−2.57	−3.20	−5.25	−7.18	−6.43	−9.22	−1.23	−0.45	−0.95	−1.04
TBL	0.73	0.39	−0.04	−0.85	−0.33	−1.02	−1.26	−0.42	1.16	0.96	0.45	−1.15
LTY	−0.06	−0.63	−0.89	−2.05	0.70	−0.02	0.06	−1.16	−0.37	−0.87	−1.27	−2.33
LTR	−1.04	0.65	−0.20	1.19	−1.19	−1.30	−2.24	−0.44	−0.99	1.46	0.67	1.83
TMS	−0.12	0.56	0.50	1.27	−1.37	−1.99	−0.39	4.34	0.38	1.57	0.75	−0.20
DFY	−5.69	−5.14	−5.21	−2.20	−6.85	−5.58	−2.90	−0.31	−5.22	−5.01	−6.32	−3.25
DFR	1.65	0.68	0.93	0.54	1.43	2.25	3.46	1.02	1.73	0.00	−0.23	0.30
INFL	−0.45	−0.24	2.12	1.88	−3.73	−2.26	4.80	6.15	0.91	0.59	0.92	−0.19
SII	4.06	4.50	5.32	3.07	13.05	15.15	18.17	13.30	0.38	0.21	0.16	−1.14
Combined Forecast	0.46	−3.01	−4.68	−4.23	6.63	−0.60	0.41	−0.97	−2.08	−4.00	−6.81	−5.49
Buy and hold	2.09	3.04	2.65	2.45	1.01	1.90	3.43	1.09	2.52	3.46	2.18	2.73

This table reports certainty equivalent return (CER) gains (in percent) of a mean-variance investor with relative risk aversion coefficient who allocates between equities and risk-free bills using a predictive regression excess return forecast. The CER gain is the difference between predictive regression forecast and prevailing benchmark forecast. Buy and hold corresponds to the investor passively holding the market portfolio. Note that the year 1990 is used as an intial holdout period for the combined forecast model.

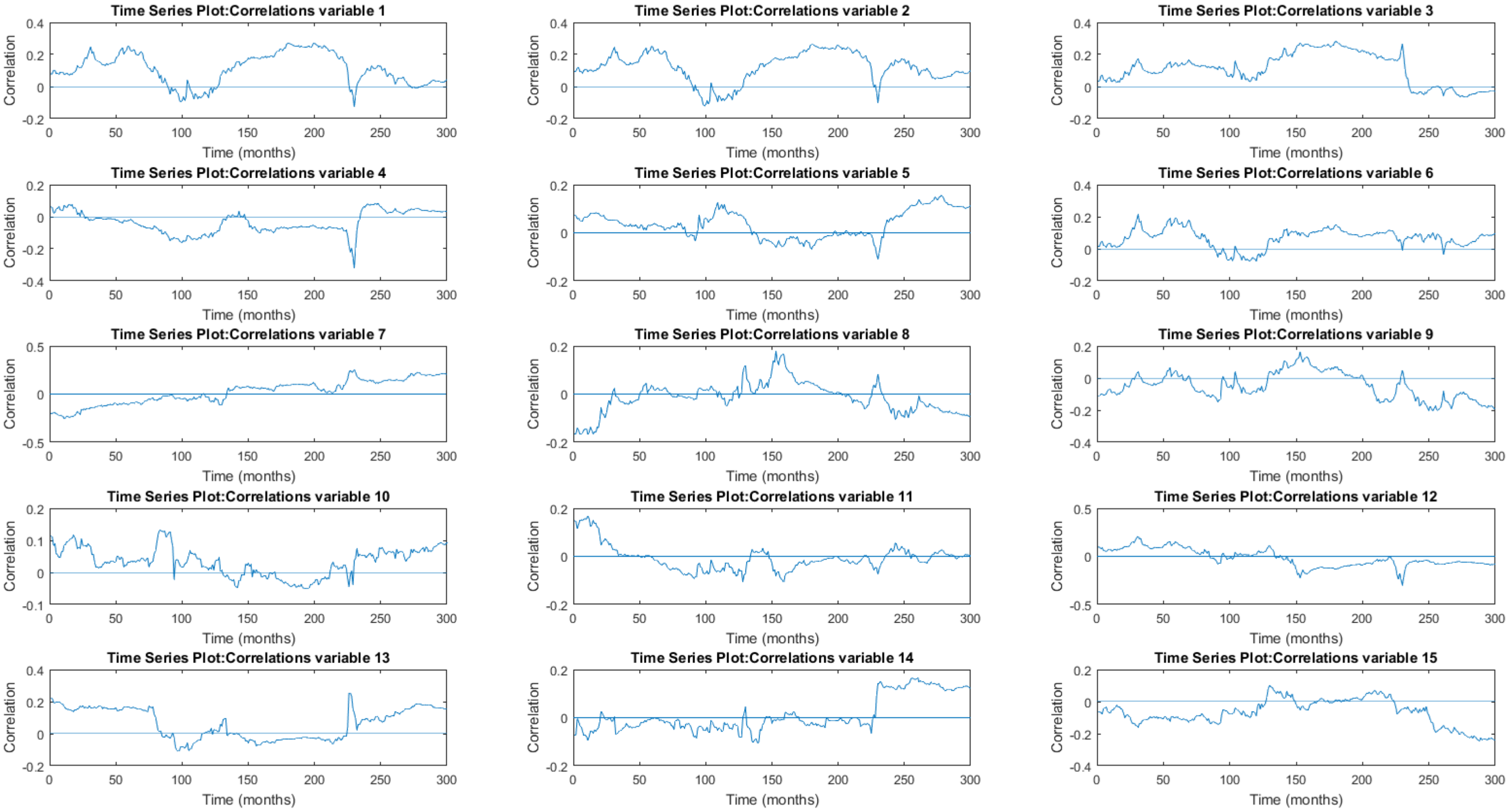
Table 13: Sharpe Ratios

	1991 : 01 − 2014 : 12 period				2007 : 01 − 2014 : 12 period				1991 : 01 − 2006 : 12 period			
Predictor	h=1	h=3	h=6	h=12	h=1	h=3	h=6	h=12	h=1	h=3	h=6	h=12
Prevailing mean	0.41	0.34	0.39	0.37	0.43	0.30	0.19	0.29	0.41	0.35	0.48	0.40
DP	0.08	0.07	−0.01	0.00	0.39	0.29	0.21	0.33	−0.03	0.01	−0.11	−0.14
DY	0.10	0.08	0.00	0.01	0.46	0.34	0.25	0.35	−0.02	0.01	−0.11	−0.14
EP	0.36	0.27	0.26	0.22	0.57	0.28	0.11	0.17	0.30	0.27	0.32	0.24
DE	0.37	0.28	0.25	0.36	0.43	0.14	−0.22	0.29	0.36	0.33	0.43	0.38
RVOL	0.29	0.25	0.28	0.34	0.49	0.39	0.28	0.34	0.20	0.19	0.28	0.33
BM	0.34	0.28	0.29	0.26	0.44	0.28	0.18	0.26	0.32	0.29	0.33	0.25
NTIS	0.24	0.18	0.23	0.23	0.10	−0.07	−0.03	0.11	0.31	0.30	0.41	0.32
TBL	0.47	0.39	0.40	0.35	0.42	0.28	0.18	0.30	0.49	0.43	0.50	0.37
LTY	0.41	0.29	0.32	0.18	0.49	0.28	0.14	0.11	0.39	0.30	0.38	0.19
LTR	0.35	0.38	0.38	0.44	0.36	0.24	0.11	0.30	0.35	0.44	0.53	0.49
TMS	0.41	0.39	0.43	0.44	0.38	0.27	0.27	0.54	0.44	0.44	0.53	0.40
DFY	−0.05	−0.03	0.01	0.22	0.04	0.05	0.13	0.33	−0.13	−0.08	−0.07	0.15
DFR	0.53	0.38	0.45	0.40	0.52	0.47	0.43	0.34	0.54	0.35	0.46	0.41
INFL	0.39	0.35	0.53	0.49	0.18	0.20	0.51	0.63	0.47	0.40	0.53	0.43
SII	0.67	0.60	0.73	0.53	1.14	1.16	1.35	1.06	0.44	0.36	0.49	0.36
Combined Forecast	0.45	0.18	0.17	0.07	0.85	0.34	0.31	0.15	0.27	0.13	0.11	0.04
Buy and hold	0.56	0.53	0.57	0.50	0.49	0.45	0.45	0.41	0.60	0.57	0.67	0.54

This table shows the Sharpe ratios of a mean-variance investor who allocates between equities and risk-free bills using a predictive regression excess return forecast. Buy and hold corresponds to the investor passively holding the market portfolio. Note that the year 1990 is used as an intial holdout period for the combined forecast model.

Figure 1: Correlations of the predictors with the excess returns, 1990:01-2014:12

Correlation graphs



These graphs show the correlations between the 15 dependent variables and the excess returns. The first 14 variables are the GW-variables (1:DP, 2:DY, 3:EP, 4:DE, 5:RVOL, 6:BM, 7:NTIS, 8:TBL, 9:LTY, 10:LTR, 11:TMS, 12:DFY, 13:DFR, 14:INFL) and the fifteenth variable is the SII. We obtained the resulting correlations for the 300 months by using a 10 year rolling window for the period 01-1990 through 12-2014.