Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrics

Short interest and aggregate stock return: A new post COVID perspective

Joël Dwars (621002)

Supervisor:	Terri van der Zwan
Second assessor:	Mikhail Zhelonkin
Date final version:	28th June 2024

Abstract

I show that short interest is not necessarily the best predictor of stock return as claimed by Rapach et al. (2016). I include new important variables that were excluded in their research like for example the output gap (Cooper & Priestley (2009)) and the yearly growth rate in personal consumption expenditures (Møller & Rangvid (2015)), and investigate predictive power using several tests for an updated sample period. The results show that SII can be beaten both in-sample and out-of-sample. While short interest generates an in-sample annual R^2 statistic of 5.65%, the output gap can improve this by generating an annual R^2 statistic of 17.46%. Results also show that SII's predictive power disappears when we exclude the Global Financial crisis from our sample period, while other predictors are still able to predict stock return. Rapach et al. (2016) also claim that short interest's predictive power predominantly stems from a cash flow channel. I show that this is still true, however, estimates are not significant anymore.

1 Introduction

Understanding and predicting stock returns has always been an important point of research in the financial economics due to its significant implications for investors, portfolio managers and policymakers. Accordingly, numerous papers have been published attempting to predict changes in future aggregate excess stock returns. Campbell & Shiller (1988b) introduced the dividend-price ratio as a predictor of long-term stock returns, which showed the importance of fundamental economic indicators. Bansal & Yaron (2004) expanded this by integrating consumption-based asset pricing models, which also showed that aggregate consumption growth played a role in predicting equity returns. However, Welch & Goyal (2008)'s assessment of the predictive power of a set of financial and economic-based predictor variables that are meant to track business conditions resulted in questioning the idea that stock returns changes with business conditions. More recently, investor sentiment-based variables have been shown to play a important role in predicting stock returns as well. For example, Baker & Wurgler (2006) highlight investor sentiment's influence on stock prices, indicating periods of high sentiment preceding market reversals.

Continuing on this idea that investors sentiment can play a role in predicting stock returns, Rapach et al. (2016) introduce a new variable using the short interest. The idea behind this variable is that when there is a high level of short selling in the market, investors believe that the market is over-valued and future returns will be lower. They provide empirical evidence suggesting that short interest contains strong predictive power for future stock returns. They even state that "short interest is arguably the strongest known predictor of aggregate stock returns". Rapach et al. (2016) compare the performance of his short interest variable to that of the set of variables used in Welch & Goyal (2008), as these variables have become widely recognized in the literature and serve as benchmarks for assessing the predictive power of new factors.

However, a recent paper by Priestley (2019) questions this statement that short interest is the strongest predictor of aggregate stock returns. Priestley (2019) underscores the need to consider additional theoretically motivated predictor variables that may offer valuable insights into market dynamics by including other well-performing predictors into his research, as Rapach et al. (2016) only include the Welch & Goyal (2008)'s variables. This paper also shows that excluding data on the financial crisis of 2008 can actually make the predictive power of the short interest disappear. A recent paper by Goyal et al. (2023) compares a lot of new predictor variables obtained from the literature and conclude that short interest is not the only variable with predictive power.

In this paper, I further analyze the claim that short interest is the best predictor for stock returns. I reexamine the findings of Rapach et al. (2016) by extending the time period beyond the COVID crisis. Additionally, I incorporate new predictor variables, including the output gap of Cooper & Priestley (2009) (later referred to as "OGAP"), Neely et al. (2014)'s technical indicators (TCHI), Pollet & Wilson (2010)'s average correlation measure (AVGCOR), and Møller & Rangvid (2015)'s (GPCE) growth rate in personal consumption expenditures. I have chosen for these variables, as Goyal et al. (2023) provides empirical evidence that these variables are also strong predictors of future stock returns. These variables are absent in Rapach et al. (2016), and can therefore offer an opportunity to gain fresh insight into the assertion that short interest stands as premier predictor.

In-sample tests show that when the time period is extended and these variables are added, short interest no longer beats all the variables in performance. Short interest produces predictive regression R^2 statistics of 0.67% at monthly horizon and 5.65% at annual horizon, while this is 2.10% at monthly horizon and 17.46% at annual horizon for OGAP. When we regress the returns on short interest and one of the other five best predictors, this reveals that all those other variables are able to predict future returns in the presence of short interest. When we regress returns on short interest only, all the other five best predictors can explain the residuals of this regression. So, short interest and those other predictors do not contain the same information about future returns and those other five predictors also play a (different) role in predicting stock return. When we also consider different time periods excluding first the COVID crisis and then the COVID crisis and the financial crisis of 2008, I show that the performance of short interest drops even further and actually disappears in the latter case with a R^2 statistic of just 0.13%, while other predictors retain their significance

Consequently, I will also examine the out-of-sample performance of short interest. Welch & Goyal (2008) show that despite significance evidence of in-sample predictability, variables still can fail to predict the equity risk premium based on out-of-sample tests. For example, I show that our new variable OGAP, which generates the best results in-sample, actually preforms as one of the worst using the out-of-sample tests, with out-of-sample R^2 statistics (introduced by Campbell & Thompson (2008)) of -2.39%, -5.98%, -6.67% and -1.97% for the different horizons. With these tests, short interest generates the highest performance with respectively out-of-sample R^2 statistics of 1.17%, 3.73% and 5.91% for the monthly, quarterly and semi-annual horizon, respectively. Only for the annual horizon, inflation, TCHI and GPCE preform better with out-of-sample R^2 statistics of 3.55%, 2.85% and 6.02%, respectively. Using the encompassing tests, the evidence that forecasts based on SII have superior information content relative to forecasts based on popular predictors is not so clear anymore, as for example the encompassing test for annual horizon with GPCE generates a $\hat{\lambda}$ of only 0.38 and is not significant. Equivalent to the in-sample tests, I will also consider some out-of-sample tests on different time periods excluding crises. When both crises are excluded, short interest's out-of-sample predictive power completely disappears. In an attempt to improve the out-of-sample R^2, we combine single forecasts of the best preforming predictors, as suggested by Rapach et al. (2010). This does not improve performance for the entire sample compared to short interest's performance.

I will also analyse the economic significance of short interests predictive ability via an asset allocation analysis for a monthly horizon. With the extended time period and compared to the new variables as well, short interest still generates the highest utility gain for the complete sample, with a CER gain of 3.06. Around the financial crisis, the utility gains accruing to short interest are particularly large, while if we take the sample until 2006, short interest actually returns a negative utility gain. Short interest's Sharpe ratio is also the highest, but it doesn't clearly outperform TCHI, as both predictors generate the same Sharpe ratio of 0.69.

Rapach et al. (2016) explain why short interest can predict stock returns and present evidence that short interest's predictability primarily operates via a cash flow channel. The Campbell (1991) and Campbell & Ammer (1993) vector auto regression (VAR) approach and the information contained in popular predictors are used to decompose aggregate stock return into expected returns, discount rate news and cash flow news components. In my analysis, the claim that short interest's predictability primarily operates via a cash flow channel is still true, but the coefficients are not significant any more. If we treat the set of predictors as a proxy for the market information set, the results suggest that most part of the predictability can be found in the cash flow component, as these β's are biggest values, even though not significant. Therefore, the claim of Rapach et al. (2016) might not hold anymore

The remainder of this research is organized as follows. section 2 will describe our data including the construction of short interest, the Welch & Goyal (2008) variables and our four newly added variables. section 3 reports all our in-sample analysis and tests including the results for all the predictors. section 4 reports all of our out-of-sample tests and their results for all the predictors. section 5 reports all the results for the asset allocation analysis and section 6 reports results for the VAR decomposition and reexamination of the claim that short interest's predictive ability primarily operates through a cash flow channel

2 Data

In my research, I use two different datasets for the short interest. One dataset contains data from 1973:01-2014:12 and this dataset can be accessed from Guofo Zhou's web-page. This datasets utilize Compustat's firm-level short interest data, normalized by dividing short interest by each firm's shares outstanding from CRSP. Filtering out assets with stock prices below $5 per share and those below the fifth percentile breakpoint of NYSE market capitalization results in the equally weighted short interest series (EWSI). In figure 1 can be seen that EWSI exhibits a strong upward trend. Therefore, detrending is needed, as is explained in Rapach et al. (2016). After this the series is standardized to have a standard deviation of one. This processed series serves as the short interest index (SII), visualized in figure 2. This dataset can then be used for replication of the results obtained by Rapach et al. (2016). However, in my research, the main focus lays on a different time period from 1973:01-2021:12. For this time period, another dataset is used, which can be accessed on Amit Goyal's website. This dataset contains already detrended SII data until end 2021 and is constructed in the same way as the other dataset mentioned. This dataset also contains detrended data for different dates, until that specific date. This is needed for the out-of-sample tests and can not be accessed from Guofo Zhou's web-page for the updated sample period. This dataset is visualized in figure 5 in appendix A

The predictive ability of the short interest index will be compared to that of 14 monthly predictor variables from Welch & Goyal (2008). Specifically, we include the following predictors:

- Log dividend-price ratio (DP): log of a 12-month moving sum of dividends paid on the S&P 500 index minus the log of stock prices (S&P 500 index)
- Log dividend yield (DY): log of a 12-month moving sum of dividends minus the log of lagged stock prices
- Log earnings-price ratio (EP): log of a 12-month moving sum of earnings on the S&P 500 index minus the log of stock prices.
- Log dividend-payout ratio (DE): log of a 12-month moving sum of dividends paid on the S&P 500 index minus the log of a 12-month moving sum of earnings on the S&P 500 index.
- Excess stock return volatility (RVOL): computed using a 12-month moving standard deviation estimator, as in Mele (2007)

- Book-to-market ratio (BM): book-to-market value ratio for the Dow Jones Industrial Average
- . Net equity expansion (NTIS): ratio of a 12-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks.
- Treasury bill rate (TBL): interest rate on a three-month Treasury bill (secondary market)
- Long-term yield (LTY): long-term government bond yield
- Long-term return (LTR): return on long-term government bonds.
- Term spread (TMS): long-term yield minus the Treasury bill rate.
- Default yield spread (DFY): difference between Moody's BAA- and AAA-rated corporate bond yields.
- Default return spread (DFR): long-term corporate bond return minus the long-term government bond return.
- Inflation (INFL): calculated from the Consumer Price Index (CPI) for all urban consumers.

Updated data for these 14 variables can all be obtained from Amit Goyal's web-page. The data is updated until end of 2023, but we only utilize the data until the end of 2021. This data is also used for the analysis from 1973:01-2014:12. Because the data is updated, the results for this time period will not be exactly the same as Rapach et al. (2016), but can be slightly different.

So far I have only included variables that were used in Rapach et al. (2016), where was concluded that SII is the best known predictor of excess stock return. However, Priestley (2019) argue that there are some more theoretically motivated predictor variables, that Rapach et al. (2016) omit from their analysis. Therefore, I will consider some other predictor variables as well in my analysis:

1. the output gap of Cooper & Priestley (2009) (*OGAP*): This is the deviation of the log of industrial production from a trend that incorporates both a linear and a quadratic term.
   Industrial production data can be obtained from FRED. The gap is estimated from regressing the following linear model:
   $$y_t = b_0 + b_1 t + b_2 t^2 + v_t \tag{1}$$
   where y_t is the log of industrial production, t is a time trend, and v_t is the error term, which is the output gap.

2. Neely et al. (2014)'s *TCHI* is the first principal component of 14 technical indicators, themselves principally versions of moving price averages, momentum, and ("on-balance") dollartrading volume.
   For constructing this variable we first obtain S&P500 index prices and volume from Yahoo finance. The first six binary moving-average price indicators are based on moving average crossovers which each equal to 1 if the short-term moving average (based on the 1, 2 or 3 month average) crosses above the long-term moving average (based on the 9 or 12 month average), and 0 otherwise. The two binary momentum price indicators are based on whether the current index price exceeds price observed 9 or 12 months ago (they equal 1 if so, and 0 otherwise). The VOL indicators are based

on "on-balance" volume for a given period, which is the cumulative sum of the volume for the period if the index price exceeds the previous period's, and the negative of the same sum if not. The final six indicators are then computed in the same way as the moving-average indicators. Taking the first principal component of these 14 indicators results in our predictor variable

3. Pollet & Wilson (2010)'s *AVGCOR* is the average correlation among the 500 largest stocks (by capitalization).
   Data for this variable can be obtained from Amit Goyal's web-page.

4. Møller & Rangvid (2015)'s *GPCE* is the yearly growth rate in personal consumption expenditures. For constructing this variable we first obtain personal consumption expenditures data from FRED. We then construct the data as the yearly growth rate:

$$GPCE_t = \frac{PCE_t - PCE_{t-12}}{PCE_{t-12}} \tag{2}$$

Goyal et al. (2023) show us that these four variables are all doing well in predicting excess stock return both in-sample as out-of-sample. Graphs of the data for these four variables can be found in appendix B

Consistent with prior research, I focus on predicting the excess return on a value-weighted market portfolio, measured as the log return on the S&P 500 index minus the log return on a one-month Treasury bill.

All variables are monthly observed. All data ranges from January 1973 until December 2021 and will also be used, except SII, for the analysis for the sample period of January 1973 until December 2014. The main focus of the rest of this paper will be on the sample period until December 2021, but the results for the other sample period until December 2014 will also be available. Summary statistics for all the variables for both sample periods can be found in appendix C. Appendix D displays Pearson correlation coefficients for all the variables for both sample periods. SII remains largely unrelated to the 14 popular predictors, but also is largely unrelated to the four newly added variables. The strongest correlation in table 13 of SII with one of the other variables occurs with BM, which has a correlation of only -0.30. SII still contains substantially different information from the other predictors.

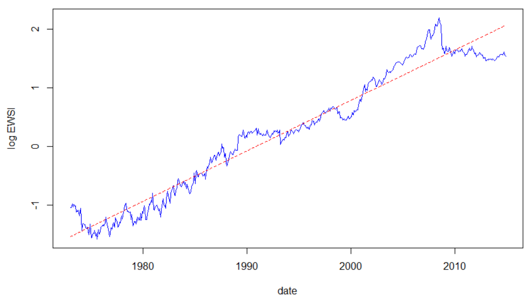

Figure 1: log equally weighted short interest for 1973:01-2014:12.
The blue line corresponds with the log of equally weighted short interest constructed as stated in section 2. The dashed red line is the lineair trend in the series

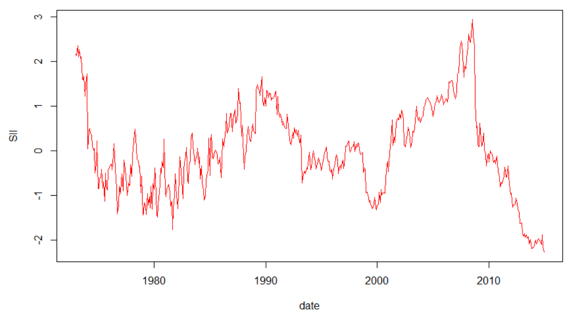

Figure 2: The short interest index (SII) for 1973:01-2014:12.
The detrended EWSI is standardized to have a standard deviation of one. This resulting series is here shown. The graph of SII for 1973:01-2021:12 can be found in appendix A

3 In-sample testing

3.1 Standard in-sample tests

For analyzing the aggregate stock return predictability, I use the following predictive regression model:

$$r_{t:t+h} = \alpha + \beta x_t + \epsilon_{t:t+h} \quad \text{for } t = 1, ..., T - h \tag{3}$$

where $r_{t:t+h} = (1/h)(r_{t+1} + ... + r_{t+h})$, r_t is the S&P 500 log excess return for month t, and x_t is one of the predictor variables. β is estimated using ordinary least squares (OLS). If a variable is able to predict the stock return, its own β should be significant. For testing significance of β, I follow Rapach et al. (2016), where they use a more powerful test with a one-sided alternative hypothesis, recommended by Inoue & Kilian (2005). The well-known Stambaugh (1999) bias and the use of overlapping observations when $h > 1$ can cause complications. Therefore, I will use a heteroskedasticity- and autocorrelation-robust t-statistic and compute a wild bootstrapped p-value to test $H_0 : \beta = 0$ against $H_a : \beta > 0$. To facilitate comparisons across predictors, I standardize each predictor to have a standard deviation of one. The negative is taken for NTIS, TBL, LTY, INFL, SII, OGAP, TCHI and GPCE before estimating, indicated in table 1 by the negative sign, so that $H_a : \beta > 0$ is the relevant alternative hypothesis for all the variables. For the sample period 1973:01-2021:12, we have $587, 585, 582$ and 576 usable observations at monthly ($h = 1$), quarterly ($h = 3$), semi-annual ($h = 6$) and annual ($h = 12$) horizons, respectively.

Table 1 reports the results for the predictive regression on sample period 1973:01-2021:12, while in appendix E the results for the sample period 1973:01-2014:12 can be found in table 14. At monthly horizon, table 1 shows that five of the Welch & Goyal (2008) variables (RVOL, TBL, LTY, LTR and DFR) combined with SII and the four new variables display significant predictive ability. OGAP has the largest $\hat{\beta}$ estimate (0.64). For the monthly horizon, Rapach et al. (2016) show that SII is clearly on par with the best individual predictors in their analysis. If we now look at the results for the extended time period and with the new variables, we can see in table 1 that this is not completely true anymore. Not only do the four new variables generate a higher or equal R^2 statistic than SII (2.10% (OGAP), 0.78% (TCHI), 0.84% (AVGCOR) and 0.67% (GPCE) compared to 0.67% of SII), but also DFR now delivers a higher R^2 of 0.88%.

Also for the other sample period in table 14, SII is not delivering the highest R^2 statistic, with OGAP and TCHI having a R^2 of 1.73% and 1.30% compared to 1.26%.

At the quarterly, semi-annual and annual horizon this remains the same for the updates sample period. SII is still generating significant estimates (at 5% for monthly horizon and at 10% for quarterly, semi-annual and annual horizon), but is not displays the strongest predictive ability, with OGAP and AVGCOR displaying stronger predictive ability on all horizons. The quarterly $\hat{\beta}$ for SII is 0.37, while this is 0.65 for OGAP and 0.48 for AVGCOR. The estimated slope coefficient for these two predictors are above the estimated slope coefficient of SII and while SII's estimate is significant at 10%, their estimates are significant at 1%. The $\hat{\beta}$ estimates implies that a one-standard-deviation increase in SII corresponds to a 4.44 percentage point decrease in future annualized excess return at monthly horizon. For those other two predictors, this implies that a one-standard-deviation increase corresponds to a 7.8 percentage point decrease in future annualized excess return for OGAP and a 5.76 percentage point increase for AVGCOR (as the negative is not taken for AVGCOR). For all horizons OGAP's R^2 is around three times as big as

SII's R^2.

Table 1: in-sample regression estimation results for sample period 1973:01-2021:12.
The table reports the ordinary least squares estimate of β and R^2 statistic for the predictive regression in equation 3 explained in section 3.1. *,** and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values. The last two rows corresponds to a multiple predictive regression that includes an intercept, SII, and the first three principal components of the Welch & Goyal (2008) variables (denoted with "GW") and first three components of all the variables excluding SII (denoted with "all"). For these two rows, the table displays estimated slope coefficients and partial R^2 statistics corresponding to SII.

	$h = 1$		$h = 3$		$h = 6$		$h = 12$	
Predictor	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)
DP	0.08	0.03	0.09	0.11	0.11	0.32	0.13	0.93
	[0.40]		[0.53]		[0.64]		[0.79]	
DY	0.09	0.04	0.09	0.12	0.11	0.34	0.14	0.99
	[0.47]		[0.55]		[0.66]		[0.81]	
EP	0.01	0.00	-0.03	0.01	-0.03	0.03	0.01	0.00
	[0.04]		[-0.13]		[-0.16]		[0.03]	
DE	0.09	0.04	0.15	0.36	0.19	1.00	0.16	1.47
	[0.35]		[0.81]		[1.24]		[1.67]	
RVOL	0.39	0.76	0.35	1.86	0.28	2.30	0.18	1.85
	[2.30]***		[2.54]***		[2.27]**		[1.48]	
BM	-0.04	0.01	-0.03	0.01	-0.01	0.00	0.00	0.00
	[-0.21]		[-0.17]		[-0.04]		[0.02]	
NTIS (-)	0.11	0.06	0.06	0.06	0.07	0.12	0.10	0.50
	[0.52]		[0.29]		[0.29]		[0.45]	
TBL (-)	0.34	0.57	0.29	1.24	0.26	1.89	0.23	2.95
	[1.82]**		[1.84]*		[1.60]		[1.60]	
LTY (-)	0.27	0.37	0.23	0.80	0.20	1.16	0.14	1.09
	[1.46]*		[1.51]		[1.29]		[0.94]	
LTR	0.30	0.45	0.15	0.34	0.21	1.32	0.14	1.10
	[1.55]*		[1.11]		[2.53]***		[3.10]***	
TMS	0.23	0.28	0.21	0.64	0.20	1.09	0.26	3.77
	[1.22]		[1.23]		[1.18]		[1.73]*	
DFY	0.14	0.09	0.14	0.29	0.20	1.17	0.16	1.47
	[0.54]		[0.62]		[1.15]		[1.19]	
DFR	0.42	0.88	0.10	0.15	0.11	0.33	0.03	0.06
	[1.37]*		[0.58]		[0.92]		[0.50]	
INFL (-)	0.00	0.00	0.17	0.45	0.27	2.07	0.25	3.44
	[0.01]		[0.99]		[1.91]*		[2.31]**	
SII (-)	0.37	0.67	0.37	1.95	0.37	3.60	0.35	5.65
	[1.79]**		[1.89]*		[1.78]*		[1.78]*	
OGAP (-)	0.64	2.10	0.65	6.30	0.61	10.68	0.56	17.46
	[3.75]***		[4.35]***		[3.98]***		[3.45]***	
TCHI (-)	0.39	0.78	0.34	1.72	0.35	3.45	0.22	2.54
	[1.71]*		[1.62]*		[1.69]*		[1.31]	
AVGCOR	0.41	0.84	0.48	3.51	0.37	3.97	0.28	4.33
	[2.64]**		[3.02]***		[2.79]***		[2.73]**	
GPCE (-)	0.36	0.67	0.37	2.06	0.36	3.67	0.44	8.91
	[1.87]**		[2.25]**		[2.56]**		[3.24]***	
SII (-) ‖PC(GW)	0.38	0.69	0.40	2.14	0.40	3.99	0.38	6.36
	[1.84]**		[1.99]*		[1.81]*		[1.84]	
SII (-) ‖PC(all)	0.36	0.62	0.38	1.99	0.38	3.65	0.36	5.90
	[1.80]**		[2.03]*		[1.85]*		[1.89]*	

The last two rows of table 1 report the OLS estimate of β_{SII} and the corresponding t-statistic for the following predictive regression:

$$r_{t:t+h} = \alpha + \beta_{SII} SII_t + \sum_{j=1}^{3} \beta_{f,j} \hat{f}_{j,t} + \epsilon_{t:t+h} \tag{4}$$

where $\hat{f}_{1,t}$, $\hat{f}_{2,t}$ and $\hat{f}_{3,t}$ are the first three principal components extracted from the entire set of Welch & Goyal (2008) variables with and without the new four variables added (referred to as "GW" and "all", respectively). Ludvigson & Ng (2007) show that incorporating principal components provide an effective strategy for incorporating the information from a large number of variables in regression models for aggregate stock return. Table 1 also provides partial R^2 statistics corresponding to SII for both equations in 4.

Comparing the row "SII (-)" with the row "SII (-)‖PC (GW)" and with the row "SII (-)‖PC (all), this shows that including both principal components into the regression had vary little influence on the SII's predictive power. The estimated slope coefficients remain nearly the same and the R^2 statistics remain sizeable. So, we can argue that SII contains information that is different from the information contained in the other predictors, and this different information is still useful, however not that useful in-sample as is claimed by Rapach et al. (2016).

3.2 Deeper in-sample analysis

From the previous section is clear that the statement made by Rapach et al. (2016) that "SII outperforms a host of popular return predictors" is still debatable. Variables as OGAP and AVGCOR have a higher predictive power in-sample, looking at their stand-alone performance. In an attempt to assess the relative predictive power of SII and other important variables, we report the results of various regressions. For this, we will only look at the monthly horizon and only the five strongest predictors will be included: TMS, OGAP, TCHI, AVGCOR and GPCE. Goyal et al. (2023) show that these five predictors including SII are one of the strongest predictors out there at the moment. In the first regression, we regress aggregate stock return on both SII and one of the other variables:

$$r_{t+1} = \alpha + \beta_{SII} SII_t + \beta_i Z_{i,t} + \epsilon_{t+1} \tag{5}$$

where Z_i is one of the other variables. This equation could reveal information about the relative performance of the two variables. t-statistics and p-values are computed in the same way as the regression in equation 3. The results for this regression are reported in table 2. The results in table 2 show that all the other variables retain the significance, even in the presence of SII. Including one of the other variables with SII increases the R^2 from 0.67% (second column of table 1 to between 1.22% and 2.30%. SII retains its significance in nearly all the regression as well, except one. It is interesting to see that when we combine the short interest index and the output gap into one regression, SII's estimated slope coefficient drops from 0.37 to 0.21 and actually loses its significance level to be not significant at all anymore. With this result OGAP clearly overshadows SII in terms of predictive ability.

Table 2: in-sample estimation results for regression with two variables for sample period 1973:01-2021:12. The table reports least squares estimates of β and R^2 statistics for the predictive regression explained in equation 5. t-statistics and p-values are computed in the same way as explained in 3.1. *,** and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values.

Variables	$\hat{\beta}_{SII}$	$\hat{\beta}_{var}$	R^2 (%)
SII & TMS	0.45	0.34	1.22
	[2.25]***	[1.86]**	
SII & OGAP	0.21	0.59	2.30
	[0.97]	[3.14]***	
SII & TCHI	0.34	0.37	1.36
	[1.72]**	[1.64]*	
SII & AVGCOR	0.42	0.46	1.72
	[2.13]**	[2.23]**	
SII & GPCE	0.51	0.50	1.85
	[2.52]***	[2.83]***	

It is possible that the correlation between SII and the other variables, although small, makes relative comparisons in multivariate regressions difficult. To guard against this, we undertake two further regressions for each of the variables. First we regress excess return on SII and regress the residuals of this regression on one of the other variables:

$$r_{t+1} = \alpha + \beta_{SII} SII_t + e_{SII,t+1} \tag{6}$$

$$e_{SII,t+1} = \alpha + \beta_i Z_{i,t} + v_{t+1} \tag{7}$$

Next we regress excess return on the other variables first and then regress the residuals of these regressions on SII:

$$r_{t+1} = \alpha + \beta_i Z_{i,t} + e_{i,t+1} \tag{8}$$

$$e_{i,t+1} = \alpha + \beta_{SII} SII_t + v_{t+1} \tag{9}$$

The results for both regressions can be found in table 3. In the left panel of this table we can see that SII is still significant in all regressions, except one. This indicates that OGAP leaves no return left for SII to explain. Thus, all information of aggregate stock return that can be explained by SII can also be explained by OGAP. In the right panel of the table we can see that all variables still display significant estimated slope coefficients and the R^2 statistics remain sizeable, as Campbell & Thompson (2008) argue that a monthly R^2 statistic of approximately 0.5% represents an economically meaningful degree of return predictability. So, these variable still play a role in predicting stock return, even after first allowing for predictability with SII. All these results provide evidence that over the entire sample is not the best predictor of stock return in-sample, as is claimed by Rapach et al. (2016).

Table 3: in-sample estimation results for regression explained in equations 6 and 8 for sample period 1973:01-2021:12.
The table reports least squares estimates of β and R^2 statistics for the predictive regression explained in equations 6 (left panel) and 8 (right panel). t-statistics and p-values are computed in the same way as explained in 3.1. *,** and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values.

Regression	$\hat{\beta}$	R^2 (%)	Regression	$\hat{\beta}$	R^2 (%)
res_{TMS} on SII	0.42	0.89	res_{SII} on TMS	0.32	0.52
	[2.06]**			[1.68]**	
res_{OGAP} on SII	0.20	0.20	res_{SII} on OGAP	0.55	1.53
	[0.97]			[3.18]***	
res_{TCHI} on SII	0.34	0.58	res_{SII} on TCHI	0.37	0.69
	[1.70]**			[1.62]*	
res_{AVGCOR} on SII	0.42	0.87	res_{SII} on AVGCOR	0.45	1.04
	[2.04]**			[2.13]**	
res_{GPCE} on SII	0.47	1.10	res_{SII} on GPCE	0.46	1.10
	[2.30]***			[2.51]***	

3.3 In-sample predictive power on different time periods

To raise additional questions about Rapach et al. (2016)'s statement that SII is the best predictor, Priestley (2019) shows us that SII's predictive power is completely conditional on the inclusion of the financial crisis in 2008. To better evaluate the predictive power of the other variables, I will conduct the regression described in section 3.1 for two different sample periods: one excluding the COVID crisis and the other excluding both the COVID crisis and the financial crisis of 2008. The results for both sample periods can be found in table 4.

If we only exclude the COVID crisis, we can see in the left panel that SII is only significant at 10% for monthly and quarterly and not significant anymore for semi-annual and annual horizon. For the other variables, nearly all the variables are significant at any level for all the horizons, with exception of TMS at semi-annual horizon and TCHI at annual horizon. In the left panel, we can see that SII's significance levels disappear for all the horizons when we also exclude the Global Financial crisis from the sample period. GPCE also loses his significance level for all the horizons, with the exception of the quarterly horizon (significant at 10%). For the other four variables, the significance levels stay relatively the same. This concludes that SII only shows predictive power if we include the financial crisis into our sample period, while a variable as OGAP has a strong predictive power for different sub sample periods.

Table 4: in-sample regression estimation results for sample period 1973:01-2007:12 (left panel) and sample period 1973:01-2019:12 (right panel).
The table reports the ordinary least squares estimate of β and R^2 statistic for the predictive regression in equation 3 explained in section 3.1. *,** and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values.

	1973:01-2007:12								1973:01-2019:12							
	$h=1$		$h=3$		$h=6$		$h=12$		$h=1$		$h=3$		$h=6$		$h=12$	
Predictor	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)
TMS	0.36	0.67	0.31	1.53	0.27	2.38	0.30	5.32	0.26	0.35	0.25	0.95	0.25	1.76	0.34	6.10
	[1.69]**		[1.69]*		[1.46]		[1.65]*		[1.37]*		[1.49]*		[1.49]		[2.23]**	
SII	0.16	0.13	0.11	0.19	0.09	0.24	0.14	1.04	0.32	0.51	0.32	1.51	0.32	2.80	0.31	4.82
	[0.73]		[0.56]		[0.41]		[0.48]		[1.56]*		[1.60]*		[1.51]		[1.58]	
OGAP	0.45	1.07	0.44	3.02	0.39	4.64	0.32	6.14	0.58	1.74	0.59	5.36	0.57	9.44	0.53	15.85
	[2.18]**		[2.52]**		[2.36]**		[1.87]*		[3.28]***		[3.88]***		[3.67]***		[3.23]***	
TCHI	0.40	0.84	0.35	1.97	0.34	3.67	0.22	2.83	0.43	0.95	0.37	2.03	0.36	3.64	0.22	2.74
	[1.67]*		[1.66]*		[1.64]*		[1.13]		[1.86]**		[1.73]*		[1.70]*		[1.33]	
AVGCOR	0.47	1.13	0.48	3.59	0.33	3.28	0.22	2.76	0.31	0.50	0.41	2.59	0.34	3.26	0.22	2.79
	[2.04]**		[3.09]***		[2.18]**		[1.61]		[1.42]*		[2.48]**		[2.70]**		[1.75]*	
GPCE	0.26	0.34	0.27	1.17	0.24	1.87	0.27	4.46	0.32	0.52	0.34	1.72	0.33	3.16	0.32	5.76
	[1.18]		[1.58]*		[1.45]		[1.49]		[1.52]*		[2.00]*		[2.28]**		[2.40]**	

4 Out-of-sample tests

4.1 Standard out-of-sample tests

It is also important to analyze out-of-sample performance, as Welch & Goyal (2008) show that the in-sample predictive ability of a variable does not always hold up in out-of-sample tests. I compute a predictive regression forecast as

$$r_{t:t+h} = \hat{\alpha}_t + \hat{\beta}_t x_t \tag{10}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the OLS estimates in equation 3 based on data from the beginning of the sample through month t. For examining performance, I will use the R^2_{OS} as suggested by Campbell & Thompson (2008). This R^2_{OS} is equal to:

$$R^2_{OS} = 1 - \frac{\sum_{t=1}^{T}(r_{t+1} - \hat{r}_{t+1})^2}{\sum_{t=1}^{T}(r_{t+1} - \bar{r}_{t+1})^2} \tag{11}$$

where $\hat{r}_t$ is the fitted value of a predictive regression estimated in equation 10 and $\bar{r}_{t+1}$ is the prevailing mean forecast, the average excess return from the beginning of the sample through month t. This prevailing mean forecast serves therefore as the benchmark. The forecast evaluation period covers 1990:01-2021:12. To ascertain whether the predictive regression forecast delivers a statistically significant improvement in MSFE, I will use the Clark & West (2007) statistics as proposed in their paper, where the null hypothesis states $h_0 : R^2_{OS} \leqslant 0$ against the alternative $h_a : R^2_{OS} > 0$.

Next, I use encompassing tests to compare the information content of the forecast based on SII with teh forecast based on one of the other predictor variables. A combination of the two can be specified as follow:

$$\hat{r}^*_{t:t+h} = (1-\lambda)\hat{r}^i_{t:t+h} + \lambda \hat{r}^{SII}_{t:t+h} \tag{12}$$

where $\hat{r}^i_{t:t+h}$ and $\hat{r}^{SII}_{t:t+h}$ are the forecasts based on one of the variables and SII and $0 \leqslant \lambda \leqslant 1$. Intuitively, if $\lambda > 0$, SII contains information that is useful in forecasting beyond the information of the predictor. λ

is estimated using the approach of Harvey et al. (1998).

In table 5, the results for the out-of-sample R^2 can be found in the left panel and the results for the encompassing tests can be found in the right panel. These results are on out-of-sample period 1990:01-2021:12. The results for out-of-sample period 1990:01-2014:12 can be found in appendix E in table 15.

In the left panel of table 5, results show that from the 14 original predictors no predictors achieve to outperform the prevailing mean forecast confirming the findings of Welch & Goyal (2008). In contrast, the monthly R^2_{OS} statistic for SII is positive (1.17%) and is significant following the Clark & West (2007) test statistic, so that, it outperforms the prevailing mean benchmark. But, like the rest of our analysis, I also considered the four newly added variables. However though, it is clear that these variables don't display a stronger predictive power than SII, which was the case for the in-sample tests. While OGAP was the best performing variable in-sample, it actually generates very poor results out-of-sample at monthly horizon with a R^2_{OS} of only -2.39%. Two variables, TCHI and GPCE, return positive R^2_{OS} statistics at monthly horizon (0.46 and 0.04, respectively), but these are not significant according to the Clark & West (2007) test statistic.

At quarterly and semi-annual horizon, SII still generates the highest R^2_{OS} statistic and both are still significant at level 1%. Still, SII is not the only variable to outperform the benchmark with TCHI and AVGCOR returning positive R^2_{OS} statistics for the quarterly horizon (only AVGCOR being significant) and INFL, TCHI, AVGCOR and GPCE are displaying positive R^2_{OS} statistics for the semi-annual horizon, while all being significant. Only at the yearly horizon, SII is not generating the highest R^2_{OS} statistic, with both INFL and GPCE having a higher R^2_{OS} (3.55% and 6.02% compared to SII's 2.58%. Together with SII, only TCHI generates positive R^2_{OS} statistics for all horizons. With these results, SII is preforming better out-of-sample than in-sample when compared to the other variables. However though, it is still not true that SII is the best preforming predictor at all horizons out-of-sample.

To further evaluate this statement, I can also look at the results for the encompassing tests, which can be found in the right panel of table 5. In contrast with Rapach et al. (2016), not all $\hat{\lambda}$ estimates are significant when we add those four variables. For the encompassing test with the TCHI based forecast at monthly horizon and with the GPCE based forecast at annual horizon, we cannot reject the null that the $\hat{\lambda}$ estimate is greater than 0 for the SII based forecast. All other $\hat{\lambda}$ estimates are significant. However though, the $\hat{\lambda}$ estimates are less sizeable, but more than 0.5 in nearly all cases. Also, when I follow the Harvey et al. (1998) procedure, I now can reject the null hypothesis, that the weight on $\hat{r}^i_{t:t+h}$ is 0, in some cases, which is in contrast with Rapach et al. (2016). Namely, at annual horizon, the weight on INFL, TCHI, AVGCOR and GPCE based forecasts are significant. This supports my claim that SII is not the best performing predictor at all horizons out-of-sample.

Table 5: Out-of-sample R^2 statistics (%) and Encompassing tests for 1990:01-2021:12.
In the left panel out-of-sample R^2 statistics can be found computed with equation 11 against the prevailing mean forecast, that serves as benchmark. Significance is tested with the Clark & West (2007) test statistic. In the right panel the results for the encompassing tests can be found, where forecasts are based on a convex combination of SII forecasts and forecasts based on one of the other predictor. Significance is based on the Harvey et al. (1998) test statistic where the null hypothesis states that the weight on SII's forecast is 0 against the alternative that this is greater than 0. *, ** and *** indicate significance at the 10%, 5%, and 1% levels, respectively

	Out-of-sample R^2 statistics (%)				Encompassing tests			
Predictor	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
DP	-1.78	-5.26	-10.50	-24.26	1.00***	1.00***	1.00***	1.00***
DY	-1.92	-5.20	-10.61	-23.82	1.00***	1.00***	1.00***	1.00***
EP	-0.98	-3.62	-7.90	-14.69	1.00***	1.00***	1.00***	0.92***
DE	-1.75	-4.93	-6.11	-2.00	1.00***	1.00***	1.00***	0.56***
RVOL	-0.06	-0.26	-0.62	-2.31	0.67***	0.72***	0.73***	0.52***
BM	-0.44	-1.41	-3.09	-8.42	1.00***	1.00***	1.00***	0.80***
NTIS	-2.44	-7.11	-15.98	-22.74	1.00***	1.00***	1.00***	0.97***
TBL	0.13	0.29	0.27	1.12	0.90**	1.00***	0.99***	0.52*
LTY	0.13	-0.30	-1.74	-9.13	0.99**	1.00***	1.00***	0.66***
LTR	-0.44	-1.07	-0.66	-0.32	0.79***	0.90***	0.76***	0.48**
TMS	-0.96	-2.38	-2.70	-0.61	0.90***	0.94***	0.87***	0.54**
DFY	-2.42	-6.04	-8.10	-6.32	1.00***	1.00***	1.00***	0.59***
DFR	-2.04	-1.84	-0.91	-0.55	0.89**	1.00***	0.85***	0.50**
INFL	-0.58	-0.21	2.44**	3.55*	0.99***	0.94***	0.71***	0.43*
SII	1.17***	3.73***	5.91***	2.58**	-	-	-	-
OGAP	-2.39	-5.98	-6.67	-1.97	1.00***	1.00***	1.00***	0.62**
TCHI	0.46	0.91	3.23*	2.85*	0.67	0.78**	0.61**	0.37*
AVGCOR	-0.27	0.67*	2.41*	1.54	0.73**	0.64***	0.62***	0.46**
GPCE	0.04	-0.03	0.44*	6.02*	0.73**	0.77**	0.74**	0.38

4.2 Out-of-sample predictive power on different time periods

In the same way as I did in section 3.3, I can further evaluate SII's out-of-sample predictive power by considering different sample periods, where in one I make out-of-sample predictions for 1990:01-2007:12 and one sample period where I make these predictions for 1990:01-2019:12. In the same way, I will also compare this performance to the other five best performing predictors used in section 3.2 and 3.3. The predictions are constructed in the same way as is explained in section 4, where I now only look at the R^2_{OS} statistic. The results can be found in table 6.

Looking at this table, the results that if I only exclude the COVID crisis in my analysis, SII remains significant at all levels. However though, comparing it to the results for the complete sample period (table 5, the R^2_{OS} statistics drops in value, with even the R^2_{OS} statistic being negative for the on SII based forecasts at annual horizon (-1.94%). Looking at the sample period where I exclude both the COVID crisis and the financial crisis of 2008, the results show that SII's out-of-sample predictive power drops completely with the R^2_{OS} statistics being negative at all horizons. This supports Priestley (2019)'s statement that SII's predictive power is fully conditional on the inclusion of the Global Financial crisis. OGAP and GPCE also drop in performance when we exclude the COVID crisis and even further drops when we exclude the COVID crisis and the financial crisis of 2008 (even though OGAP was not performing at all for the full sample period). Only TCHI is able to retain predictive power with displaying positive R^2_{OS} statistics at all horizons for all the sample periods.

Table 6: Out-of-sample R^2 statistics (%) for 1990:01-2007:12 (left panel) and 1990:01-2019:12 (right panel). In the table out-of-sample R^2 statistics can be found computed with equation 11 against the prevailing mean forecast, that serves as benchmark. Significance is tested with the Clark & West (2007) test statistic. *, ** and *** indicate significance at the 10%, 5%, and 1% levels, respectively

	Out-of-sample R^2 statistics (%)							
	1990:01-2007:12				1990:01-2019:12			
Predictor	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
TMS	-1.12	-3.08	-3.85	-2.50	-0.97	-2.15	-1.88	2.97*
SII	-1.07	-3.29	-7.33	-25.22	0.76**	2.32***	3.13***	-1.94*
OGAP	-4.01	-10.32	-15.05	-11.69	-2.83	-7.08	-8.05	-3.31
TCHI	0.61	2.02	6.21**	3.54*	0.74	1.42	3.54	3.07*
AVGCOR	0.27	0.51*	1.24	-0.86	-0.94	-1.17	1.07*	-1.20
GPCE	-0.58	-1.34	-3.73	-4.02	-0.21	-0.52	-0.46	0.75

In an attempt to further improve the accuracy of the predictions, I employ the method of combining forecasts. This approach is supported by the findings of Rapach et al. (2010), who demonstrated that combining forecasts significantly improves out-of-sample prediction accuracy for the equity premium. According to their research, combining forecasts reduces the volatility seen in predictions based on one predictor by incorporating information from a variety of economic variables. In my case, it was shown that for example SII is preforming well on the complete sample while preforming not so good on the sample period excluding the crises. On the other hand, TCHI is performing better when we exclude the crises compared to the complete sample. For this reason, combining forecasts could work, as this would lead to more stability in different time periods. In my analysis I use a simple equally weighted forecast combination, excluding the poor performing OGAP:

$$\hat{f}_{c,t} = 0.2 * \hat{f}_{TMS,t} + 0.2 * \hat{f}_{SII,t} + 0.2 * \hat{f}_{TCHI,t} + 0.2 * \hat{f}_{AVGCOR,t} + 0.2 * \hat{f}_{GPCE,t} \tag{13}$$

where $\hat{f}_{c,t}$ is the equal weight combination forecast and $\hat{f}_{i,t}$ is the forecast based on variable i. After I obtained the forecasts, out-of-sample R^2 statistics and significance are computed in the same way as explained in 4 for all three different sample periods mentioned previously. The results for this can be found in table 7.

Looking at the table, results show that for all different sample period positive R^2_{OS} statistics are generated at all horizons. Nearly all are significant as well according to the Clark & West (2007) test, with the exception of the combined forecast for sample period 1990:01-2007:12 at annual horizon. I can conclude out of this that combining forecasts for these five predictors is a great way of making stable predictions, which can all outperform the prevailing mean forecast. However, when I compare the results of the full sample period for the combining forecasts with the performance of SII alone, it shows that SII has a higher performance at all horizons, with the exception of the annual horizon. Also, for the 1990:01-2007:12 sample period, TCHI's standalone performance is better at all horizons. Combining forecasts is therefore not always generating the highest R^2_{OS} statistics, but is still a great way for making stable forecasts.

Table 7: Out-of-sample R^2 statistics (%) for combined forecasts for sample period 1990:01-2007:12, 1990:01-2019:12 and 1990:01-2021:12.
In the table out-of-sample R^2 statistics can be found computed with equation 11 against the prevailing mean forecast, that serves as benchmark. Significance is tested with the Clark & West (2007) test statistic. *, ** and *** indicate significance at the 10%, 5%, and 1% levels, respectively

	Out-of-sample R^2 statistics (%)			
Sample period	$h = 1$	$h = 3$	$h = 6$	$h = 12$
1990:01-2021:12	0.95***	3.20***	5.55***	7.80***
1990:01-2019:12	0.67**	2.40***	4.39***	5.76**
1990:01-2007:12	0.54*	1.87*	2.45*	0.06

5 Asset allocation perspective

In this section, I compare the economic value of SII's predictive ability to that of the other variables from an asset allocation perspective. Like Rapach et al. (2016) and Campbell & Thompson (2008), I consider a mean-variance strategy, where wealth can be distributed between equities and risk-free bills. At the end of month t, the share of a portfolio allocated to equities during the subsequent month is equal to:

$$w_t = \frac{1}{\gamma}\frac{\hat{r}_{t+1}}{\hat{\sigma}^2_{t+1}} \tag{14}$$

where γ is the risk aversion coefficient. For $\hat{r}_{t+1}$, I use the excess return forecasts of the different predictor variables and $\hat{\sigma}^2_{t+1}$ is a forecast of the excess return variance. I also restrict w_t to lie between -0.5 and 1.5, as suggested by Rapach et al. (2016). They argue that this imposes realistic portfolio constraints and produces better-behaved portfolio weights. Like Campbell & Thompson (2008), the volatility forecast is generated using a ten-year moving window of the past returns.
To compare the obtained portfolios for the different variables, I compute the certainty equivalent return (CER) as

$$CER = \bar{R}_p - 0.5\gamma\sigma_p^2 \tag{15}$$

where $\bar{R}_p$ and σ_p^2 are the mean and variance of the portfolio return over the evaluation period. γ is the risk aversion coefficient, for which I assume in my research that it is equal to three.

I also compute the CER when the investor uses the prevailing excess return forecast instead of the forecasts based on one of the predictor variables. The CER gain is then the difference between CER based on one of the predictor variables minus the CER based on the prevailing mean excess return forecast. I annualize the CER gain, so that it can be interpreted as the annual portfolio management fee that a investor is willing to pay in exchange for information about the predictive regression forecast in place of the prevailing mean forecast. I also evaluate performance by comparing the Sharpe ratios, which allows me to compare portfolio performance independently of the risk aversion coefficient. The Sharpe ratio is also computed for the prevailing mean forecasts.

I generate results for the CER gain and the Sharpe ratio only at monthly horizon, where we assume that a investor rebalances his portfolio monthly. For this analysis, I consider three different sample periods. The complete sample period, a sub sample period excluding the Global Financial crisis, 1990:01-2006:12, and a sub sample period including the Global Financial crisis, 2007:01-2021:12 (instead of taking the sample period excluding the crisis until end 2007, I take it until end 2006, for replication of

Rapach et al. (2016) purposes).

The results for this can be found in table 8, where the left panel corresponds with the generated CER gains and the right panel corresponds with the Sharpe ratios. The replication of Rapach et al. (2016) can be found in appendix E in table 16.

For the CER gains, results show that SII is producing the highest value for the complete sample. SII generates a CER gain of 306 basispoints at monthly horizon. TCHI is also generating sizeable CER gains (257 basispoints), but this is lower than SII. However, both variables outperform the buy-and-hold strategy. Among the other variables, TBL, LTY and DFR generate positive CER gains ($112, 45$ and 148 basispoints, respectively). For the sub sample period excluding the Global Financial crisis, we actually see that SII is generating a negative CER gain (-7 basispoints), while TCHI is actually improving in performance (334 basispoints). SII is outperformed by several variables in this case. When I include the Global Financial crisis into a sub sample, the results show that SII is generating much higher CER gains (664 basispoints), while TCHI is performing worse (170 basispoints). SII is now again the highest performing predictor, while TCHI is outperformed by DFR (208 basispoint). This further supports the statement that SII's predictive power predominantly stems from the inclusion of the Global Financial crisis.

Looking at the Sharpe ratios in table 8 (right panel), the table shows that the results are relatively similar. Both SII and TCHI are displaying the highest Sharpe ratio (0.69) for the complete sample, where both predictors outperform the prevailing mean benchmark (0.52) and the buy-and-hold strategy (0.61). Besides these variables and the buy-and-hold strategy, only TBL, LTY and DFR ($0.57, 0.54$ and 0.62, respectively) are able to outperform the benchmark. For the sample until the end of 2006, SII is (again) performing much worse with having the same Sharpe ratio as the benchmark (0.39), while TCHI is now performing slightly worse (0.62). For the sample period including the financial crisis, SII is performing much better with a Sharpe ratio of 0.98, while TCHI is only performing slightly better (0.75), but still outperforming the benchmark (0.67).

Figure 3 and 4 provide additional perspective on the behavior of the monthly portfolio based on SII, like is done in Rapach et al. (2016). Figure 3 shows the equity weight for the monthly portfolios based on SII and the prevailing mean for period 1990:01-2021:12. The equity weight for the portfolio based on the prevailing mean is pretty stable and close to 0.75 for the complete out-of-sample period, while the equity weight for the portfolio based on SII has some big fluctuations. The biggest fluctuations are during the Global Financial crisis. In the beginning of this crisis, the portfolio based on SII takes a short position, while in the end, it takes an aggressive long position, and then going back to the short position for the remainder of the sample.

Figure 4 shows the log cumulative wealth for the two portfolios. Here the graph shows that the short position of the SII based portfolio in the beginning of the crisis enables it to make money, while the prevailing mean based portfolio actually loses money, as it doesn't has the information of SII about the Global Financial crisis. Before and after this crisis, the log cumulative wealth of both portfolios grow in the same way.

This shows that the main reason for the SII based portfolio to perform better than the prevailing mean based portfolio, comes from the fact that the Global Financial crisis is in the data. This also shows again that SII's predictive power is conditional on the inclusion of the Global Financial crisis.

Table 8: Out-of-sample CER gains and Sharpe ratios for sample periods 1990:01-2021:12, 1990:01-2006:12 and 2007:01-2021:12.
In the left panel annualized CER gains (in percentage) can be found as specified in section 5. In the right panel Sharpe ratios can be found as specified in section 5. Buy and hold corresponds to the investor passively holding the market portfolio.

	CER gain ($h = 1$)			Sharpe ratios ($h = 1$)		
Predictor	1990:01-2021:12	1990:01-2006:12	2007:01-2021:12	1990:01-2021:12	1990:01-2006:12	2007:01-2021:12
Prevailing mean	-	-	-	0.52	0.39	0.67
DP	-2.84	-4.51	-0.93	0.31	0.01	0.65
DY	-2.68	-4.38	-0.74	0.32	0.02	0.68
EP	-0.34	-0.96	0.36	0.51	0.31	0.73
DE	-0.79	-1.54	0.06	0.46	0.30	0.67
RVOL	-1.46	-2.97	0.27	0.42	0.21	0.66
BM	-0.62	-1.14	-0.03	0.48	0.31	0.67
NTIS	-1.86	-1.59	-2.16	0.40	0.26	0.51
TBL	1.12	1.11	1.11	0.57	0.47	0.68
LTY	0.45	-0.37	1.37	0.54	0.37	0.73
LTR	-0.80	-0.96	-0.63	0.46	0.33	0.60
TMS	-0.28	1.10	-1.84	0.50	0.46	0.53
DFY	-3.92	-4.49	-3.26	0.23	-0.08	0.45
DFR	1.48	0.96	2.08	0.62	0.45	0.81
INFL	-0.58	0.55	-1.85	0.48	0.43	0.54
SII	3.06	-0.07	6.64	0.69	0.39	0.98
OGAP	-2.18	-4.84	0.91	0.40	0.06	0.67
TCHI	2.57	3.34	1.70	0.69	0.62	0.75
AVGCOR	-0.21	0.95	-1.55	0.50	0.45	0.55
GPCE	-0.25	0.93	-1.60	0.51	0.47	0.55
Buy and hold	1.33	2.02	0.55	0.61	0.53	0.69

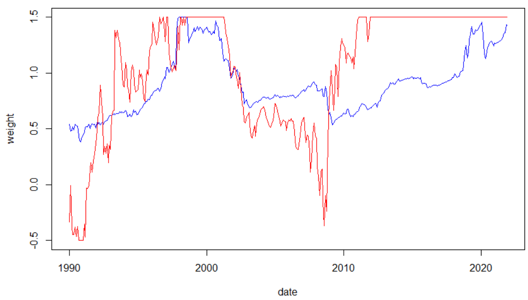

Figure 3: Equity weights for the asset allocation strategy explained in section 5.
The red line corresponds with the allocation using the SII based forecasts and the blue with the allocation using the prevailing mean forecasts for the period 1990:01-2021:12. The graph for 1990:01-2014:12 can be found in appendix F

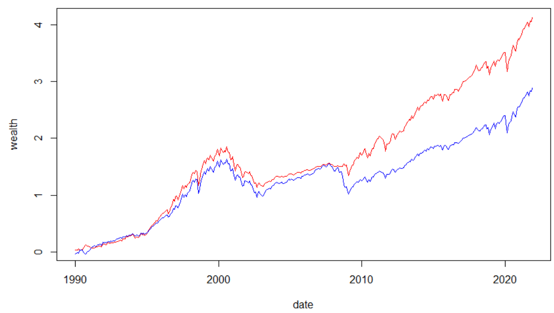

Figure 4: Log cumulative wealth for the investor assuming that the investor begins with $1 and reinvests all proceeds.
The red line corresponds with using SII based forecasts for the allocation and the blue line corresponds with using the prevailing mean forecasts for the period 1990:01-2021:12. The graph for 1990:01-2014:12 can be found in appendix F

6 Cash flow news and short interest

Rapach et al. (2016) argue that SII's ability to predict the stock return comes predominantly from a cash flow channel. I check whether this statement is still true nowadays with the new data and with incorporating the new variables. To test this, I use the Campbell & Shiller (1988a) log linearization of stock returns to decompose stock returns into news about future discount rates and news about future cash flows. The decomposition can be formulated as:

$$\begin{aligned} r_{t+1} &= E_t(r_{t+1}) + (E_{t+1} - E_t)\sum_{j=0}^{\infty}\rho^j d_{t+1+j} - (E_{t+1} - E_t)\sum_{i=1}^{\infty}\rho^i r_{t+1+i} \\ &= E_t(r_{t+1}) + N_{CF,t+1} + N_{DR,t+1} \end{aligned} \tag{16}$$

where N_{CF} and N_{DR} are news about cash flows and discount rates. To implement this, I use Campbell & Vuolteenaho (2004) who follow Campbell (1991) and estimate a VAR to obtain $E_t(r_{t+1})$ and N_{DR}. Assume that the data is generated following this framework:

$$y_{t+1} = Ay_t + u_{t+1} \tag{17}$$

where $y_t = (r_t, dp_t, z_t)$. z_t is a n-vector of predictor variables, dp_t is the dividend-price ratio, A is a $(n+2)$-by-$(n+2)$ matrix of VAR slope coefficients, and u_t is a $(n+2)$-vector of zero-mean innovations. In my analysis, z_t is each variable used in this paper separately, also including both first three principle components explained in section 3.1 separately. Campbell & Vuolteenaho (2004) show us then that the components in equation 16 can be computed as:

$$N_{DR,t+1} = e_1^{'}\rho A(I - \rho A)^{-1}u_{t+1} \tag{18}$$

$$N_{CF,t+1} = (e_1^{'} + e_1^{'}\rho A(I - \rho A)^{-1})u_{t+1} \tag{19}$$

$$E_t(r_{t+1}) = e_1^{'}Ay_t \tag{20}$$

wheree_1 denotes a $(n + 2)$-vector with one as first element and zeros for the remaining elements and ρ is:

$$\rho = \frac{1}{1 + \exp(\overline{d - p})} \tag{21}$$

where $\overline{d - p}$ is the mean of $d_t - p_t$. When the different components are obtained, I follow Rapach et al. (2016) and regress:

$$\hat{E}_t(r_{t+1}) = \alpha_{\hat{E}} + \beta_{\hat{E}}SII_t + \epsilon_{t+1}^{\hat{E}} \tag{22}$$

$$\hat{N}_{CF,t+1} = \beta_{\hat{CF}}SII_t + \epsilon_{t+1}^{\hat{CF}} \tag{23}$$

$$\hat{N}_{DR,t+1} = \beta_{\hat{DR}}SII_t + \epsilon_{t+1}^{\hat{DR}} \tag{24}$$

to assess whether SII capture information about the different components. Rapach et al. (2016) also show that this implies the following relation between the β of equation 3 when SII is the predictor and the β's of equations 22, 23 and 24:

$$\hat{\beta} = \hat{\beta}_{\hat{E}} + \hat{\beta}_{CF} - \hat{\beta}_{DR} \tag{25}$$

By comparing the different $\hat{\beta}$'s of the different components, I can ascertain to which extent SII's ability to predict returns comes from the SII's ability to predict one or more of the components. The results for equations 22, 23 and 24 can be found in table 9. the replication of Rapach et al. (2016) for this can be found in appendix E in table 17. The estimate for equation 3 with SII as predictor was equal to -0.37.

This holds for $\hat{\beta}$ in equation 25 and each set of the $\hat{\beta}_{\hat{E}}$, $\hat{\beta}_{CF}$ and $\hat{\beta}_{DR}$ estimates. Nearly all $\hat{\beta}_{\hat{E}}$ estimates are significant in table 9. However, they are all quite small in size and don't contribute much to the size of $\hat{\beta}$. The estimates for $\hat{\beta}_{DR}$ are mostly not significant. Most of them are also not big in size, such that they don't contribute much to the size of $\hat{\beta}$ as well. $\hat{\beta}_{CF}$ estimates are all much bigger in size, such that they contribute relative much to the size of $\hat{\beta}$ compared to the other components. However though, most of those estimates are not significant. Thus, it is still true that SII is particularly relevant for future aggregate cash flows. But, there is one exception. In table 9, the results show that when I include GPCE into the model, $\hat{\beta}_{DR}$ estimate is around as big as the $\hat{\beta}_{CF}$ estimate. In this case, SII contains around as much information about the cash flow as the discount rate.

Table 9: Predictive regression estimation results for market return components for sample period 1973:01–2021:12.
In the table, $\hat{\beta}_{\hat{E}}$, $\hat{\beta}_{CF}$ and $\hat{\beta}_{DR}$ estimates can be found for the framework explained in section 6 and equations 22, 23 and 24. PC indicates that the first three components are included in the VAR framework where one PC includes the Welch & Goyal (2008) variables and one PC includes all the variables excluding SII. The brackets report report heteroskedasticity- and autocorrelation-robust t-statistics. *, ** and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VAR variables	$\beta_{\hat{E}}$	β_{CF}	β_{DR}	VAR variables	$\beta_{\hat{E}}$	β_{CF}	β_{DR}
r, DP	-0.07	-0.26	0.07	r, DP, TMS	-0.04	-0.27	0.09
	[-3.63]***	[-1.62]	[1.53]		[-2.09]**	[-1.66]*	[1.89]*
r, DP, DY	-0.06	-0.26	0.07	r, DP, DFY	-0.06	-0.26	0.07
	[-2.92]***	[-1.63]	[1.76]*		[-3.31]***	[-1.63]	[1.58]
r, DP, EP	-0.07	-0.30	0.03	r, DP, DFR	-0.08	-0.26	0.05
	[-3.62]***	[-1.69]*	[0.87]		[-3.15]***	[-1.59]	[1.35]
r, DP, DE	-0.07	-0.30	0.03	r, DP, INFL	-0.07	-0.26	0.06
	[-3.62]***	[-1.69]*	[0.87]		[-3.78]***	[-1.63]	[1.46]
r, DP, RVOL	-0.12	-0.19	0.08	r, DP, OGAP	-0.17	-0.18	0.04
	[-4.34]***	[-1.34]	[1.10]		[-5.81]***	[-1.12]	[0.35]
r, DP, BM	0.00	-0.25	0.14	r, DP, TCHI	-0.10	-0.25	0.04
	[0.02]	[-1.72]*	[2.19]**		[-3.98]***	[-1.42]	[0.91]
r, DP, NTIS	-0.06	-0.25	0.08	r, DP, AVGCOR	-0.03	-0.25	0.12
	[-3.20]***	[-1.59]	[1.66]*		[-1.10]	[-1.73]*	[1.64]*
r, DP, TBL	-0.08	-0.21	0.10	r, DP, GPCE	0.01	-0.21	0.19
	[-3.80]***	[-1.44]	[1.17]		[0.18]	[-1.35]	[1.86]*
r, DP, LTY	-0.15	-0.16	0.09	r, DP, PC (GW)	-0.07	-0.21	0.12
	[-6.87]***	[-1.17]	[1.03]		[-2.36]**	[-1.45]	[1.61]
r, DP, LTR	-0.06	-0.26	0.07	r, DP, PC (all)	-0.08	-0.20	0.12
	[-2.63]***	[-1.63]	[1.60]		[-2.29]**	[-1.48]	[1.27]

7 Conclusion

In this paper, I show that the claim made by Rapach et al. (2016) that short interest is arguably the strongest known predictor of aggregate stock returns is not complete valid. By updating the sample period and incorporating some other important predictors of stock return into the analysis, I show that short interest actually can be beaten in terms of performance in several tests. Also, my research supports Priestley (2019)'s statement that SII's predictive power is fully conditional on the inclusion of the Global Financial crisis. When in-sample tests are applied, results show that SII is still generating good results, but actually can be beaten in terms of performance by OGAP for monthly, quarterly, semi-annual and annual horizon. When we further investigate this, results show that there is no return that cannot be explained by OGAP, that can be explained by SII. When the Global Financial crisis is excluded from the sample period, SII's predictive power completely disappears. For out-of-sample tests, results are better for SII, as it is the highest performing predictor at nearly all horizons. However, GPCE can outperform SII at annual horizon. Again, when the Global Financial crisis is excluded in the out-of-sample test, results show that SII's predictive power drops completely. In an attempt to improve the forecasts made, the concept of combining forecasts is implemented. This makes stable and pretty accurate forecasts. However, SII can outperform these forecasts at one, three and six months horizons. SII can also generate substantial utility gains for a mean-variance investor with a relative risk aversion coefficient of three, especially due to its strong performance during the Global Financial Crisis, where SII contains information about the crisis that other predictors do not have. SII is the best performing variable, but TCHI is close in second. For this, results also show that gains are especially large during the Global Financial crisis, while excluding this crisis drops the performance again. I also investigate whether SII's predictive power still stems predominantly from a cash flow channel. For this, estimates are still sizeable, but they are not significant anymore. Also, when I include GPCE into the framework, this statement is not so true anymore, as estimates for the discount rate component are just as big.

For further research, one can include for example more (in literature supported) variables and investigate different sample periods. For example, it is interesting to see what happens to SII's predictive power on the time period beginning in 2009 and ending in 2019, excluding the COVID crisis and the Global Financial crisis. Besides SII performing best in times of a crisis, this paper also shows that TCHI is best performing when there is no crisis. One could also consider a framework where SII predicts stock return when there is a crisis and TCHI predicts stock return when there is no crisis. This could improve performance.

References

Baker, M. & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645–1680.

Bansal, R. & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, *59*(4), 1481–1509.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The Economic Journal*, *101*(405), 157–179.

Campbell, J. Y. & Ammer, J. (1993). What moves the stock and bond markets? a variance decomposition for long-term asset returns. *The Journal of Finance*, *48*(1), 3–37.

Campbell, J. Y. & Shiller, R. J. (1988a). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, *1*(3), 195–228.

Campbell, J. Y. & Shiller, R. J. (1988b). Stock prices, earnings, and expected dividends. *The Journal of Finance*, *43*(3), 661–676.

Campbell, J. Y. & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, *21*(4), 1509–1531.

Campbell, J. Y. & Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review*, *94*(5), 1249–1275.

Clark, T. E. & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291–311.

Cooper, I. & Priestley, R. (2009). Time-varying risk premiums and the output gap. *The Review of Financial Studies*, *22*(7), 2801–2833.

Goyal, A., Welch, I. & Zafirov, A. (2023). A comprehensive 2022 look at the empirical performance of equity premium prediction. *Swiss Finance Institute Research Paper*(21-85).

Harvey, D. I., Leybourne, S. J. & Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics*, *16*(2), 254–259.

Inoue, A. & Kilian, L. (2005). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, *23*(4), 371–402.

Ludvigson, S. C. & Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics*, *83*(1), 171–222.

Mele, A. (2007). Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of Financial Economics*, *86*(2), 446–478.

Møller, S. V. & Rangvid, J. (2015). End-of-the-year economic growth and time-varying expected returns. *Journal of Financial Economics*, *115*(1), 136–154.

Neely, C. J., Rapach, D. E., Tu, J. & Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, *60*(7), 1772–1791.

Pollet, J. M. & Wilson, M. (2010). Average correlation and stock market returns. *Journal of Financial Economics*, *96*(3), 364–380.

Priestley, R. (2019). Short interest, macroeconomic variables and aggregate stock returns. *Macroeconomic Variables and Aggregate Stock Returns (May 8, 2019)*.

Rapach, D. E., Ringgenberg, M. C. & Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics*, *121*(1), 46–65.

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

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, *54*(3), 375–421.

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

A Graph for SII

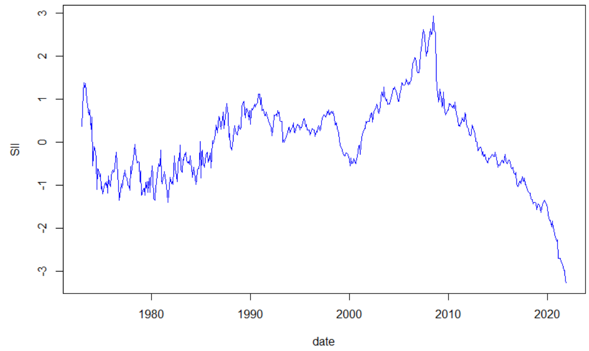

Figure 5: The short interest index (SII) for 1973:01-2021:12
The detrended EWSI is standardized to have a standard deviation of one. This resulting series is here shown.

B Graphs for other important variables

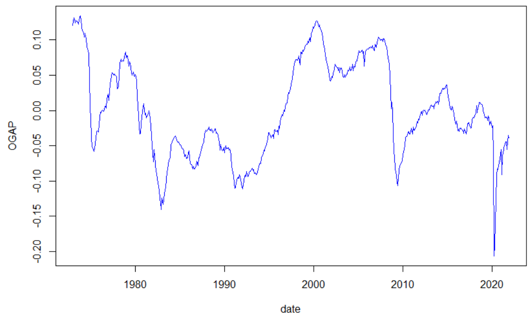

Figure 6: The output gap (OGAP) for 1973:01-2021:12

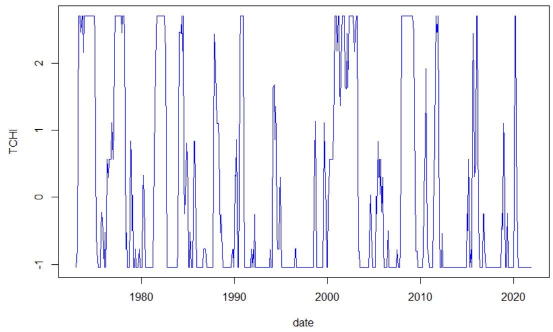

Figure 7: The first principal component of 14 technical indicators (TCHI) for 1973:01-2021:12

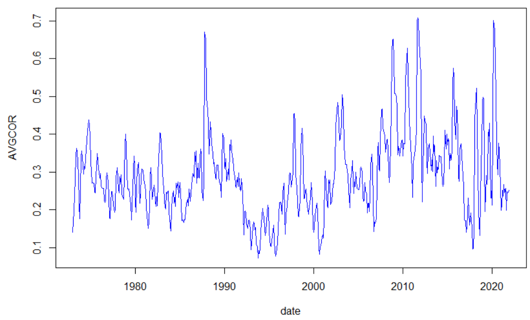

Figure 8: The average correlation among the 500 largest stocks (by capitalization) (AVGCOR) for 1973:01-2021:12

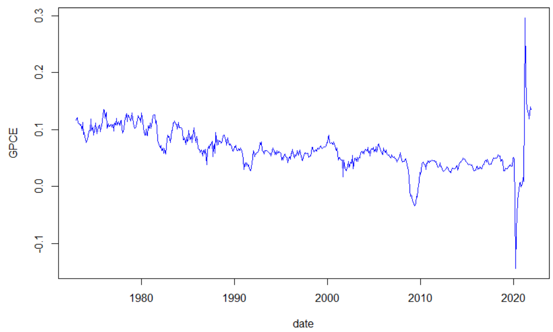

Figure 9: The yearly growth rate in personal consumption expenditures observed monthly (GPCE) for 1973:01-2021:12

C Summary statistics

Table 10: summary stats of all variables for sample period 1973:01-2014:12

Variable	Mean	Median	1st Percentile	99th Percentile	std. dev
DP	-3,62	-3,57	-4,47	-2,85	0,44
DY	-3,61	-3,57	-4,47	-2,84	0,44
EP	-2,82	-2,83	-4,54	-1,97	0,49
DE	-0,80	-0,86	-1,24	1,00	0,35
RVOL	0,15	0,14	0,06	0,30	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 (%)	5,05	5,05	0,02	14,95	3,44
LTY (%)	7,17	7,07	2,35	13,87	2,72
LTR (%)	0,74	0,78	-6,48	8,95	3,13
TMS (%)	2,12	2,36	-1,96	4,37	1,51
DFY (%)	1,10	0,96	0,56	2,81	0,47
DFR (%)	0,00	0,05	-4,85	3,96	1,48
INFL (%)	0,34	0,30	-0,54	1,27	0,38
SII	0,00	-0,09	-2,11	2,44	1,00
OGAP	0,00	0.00	-0,12	0,13	0,07
TCHI	0,00	-0,80	-1,04	2,70	1,47
AVGCOR	0,29	0,27	0,09	0,65	0,11
GPCE (%)	6,78	6,37	-1,86	12,83	3,00

Table 11: summary stats of all variables for sample period 1973:01-2021:12

Variable	Mean	Median	1st Percentile	99th Percentile	std. dev
DP	-3,67	-3,80	-4,46	-2,86	0,43
DY	-3,66	-3,79	-4,47	-2,86	0,43
EP	-2,87	-2,90	-4,45	-1,98	0,48
DE	-0,80	-0,85	-1,24	0,85	0,32
RVOL	0,15	0,14	0,06	0,30	0,05
BM	0,46	0,34	0,13	1,13	0,28
NTIS	0,01	0,01	-0,05	0,04	0,02
TBL (%)	4,45	4,82	0,02	14,81	3,53
LTY (%)	6,45	6,45	0,78	13,84	3,09
LTR (%)	0,69	0,67	-6,32	8,66	3,10
TMS (%)	2,00	2,10	-1,93	4,37	1,46
DFY (%)	1,08	0,94	0,57	2,71	0,45
DFR (%)	0,02	0,05	-4,91	4,02	1,54
INFL (%)	0,32	0,30	-0,55	1,27	0,38
SII	0,00	0,08	-2,78	2,52	1,00
OGAP	0,00	-0,01	-0,13	0,13	0,07
TCHI	0,00	-1,04	-1,04	2,70	1,43
AVGCOR	0,29	0,27	0,09	0,67	0,11
GPCE (%)	6,46	6,11	-2,94	13,50	3,50

D Correlation matrices

Table 12: correlation matrix of all variables for sample period 1973:01-2014:12

variables	DP	DY	EP	DE	RVOL	BM	NTIS	TBL	LTY	LTR	TMS	DFY	DFR	INFL	SII	OGAP	TCHI	AVGCOR	GPCE
DP	1,00																		
DY	0,99	1,00																	
EP	0,73	0,73	1,00																
DE	0,25	0,25	-0,48	1,00															
RVOL	0,01	0,02	-0,25	0,37	1,00														
BM	0,91	0,90	0,80	0,04	0,06	1,00													
NTIS	0,04	0,04	0,10	-0,09	-0,10	0,13	1,00												
TBL	0,67	0,67	0,66	-0,07	-0,06	0,68	0,08	1,00											
LTY	0,76	0,76	0,62	0,11	0,01	0,70	0,13	0,91	1,00										
LTR	0,03	0,04	0,02	0,00	0,01	0,01	-0,07	0,00	-0,02	1,00									
TMS	-0,15	-0,15	-0,38	0,35	0,15	-0,28	0,04	-0,64	-0,26	-0,04	1,00								
DFY	0,48	0,48	0,12	0,45	0,44	0,47	-0,32	0,25	0,36	0,10	0,08	1,00							
DFR	0,01	0,04	-0,08	0,14	0,13	0,02	0,03	-0,04	0,01	-0,44	0,12	0,10	1,00						
INFL	0,40	0,40	0,45	-0,13	-0,02	0,50	0,14	0,46	0,37	-0,08	-0,38	0,00	-0,05	1,00					
SII	-0,12	-0,13	-0,23	0,16	-0,13	-0,23	-0,28	-0,03	-0,06	-0,03	-0,05	-0,07	-0,07	0,00	1,00				
OGAP	-0,47	-0,48	-0,08	-0,48	-0,17	-0,18	0,02	-0,09	-0,37	-0,08	-0,46	-0,39	-0,08	0,14	0,14	1,00			
TCHI	0,16	0,13	0,04	0,16	0,11	0,24	-0,03	0,09	0,10	0,05	-0,04	0,26	-0,05	0,13	0,14	0,23	1,00		
AVGCOR	-0,01	-0,02	-0,09	0,11	0,41	-0,01	-0,33	-0,30	-0,30	0,10	0,15	0,33	-0,02	-0,14	0,10	0,01	0,26	1,00	
GPCE	0,57	0,56	0,75	-0,34	-0,17	0,69	0,33	0,70	0,65	-0,06	-0,43	-0,01	-0,05	0,49	-0,10	0,16	0,05	-0,32	1,00

Table 13: correlation matrix of all variables for sample period 1973:01-2021:12

variables	DP	DY	EP	DE	RVOL	BM	NTIS	TBL	LTY	LTR	TMS	DFY	DFR	INFL	SII	OGAP	TCHI	AVGCOR	GPCE
DP	1,00																		
DY	0,99	1,00																	
EP	0,75	0,75	1,00																
DE	0,24	0,24	-0,46	1,00															
RVOL	0,01	0,02	-0,23	0,35	1,00														
BM	0,91	0,91	0,81	0,04	0,06	1,00													
NTIS	0,12	0,11	0,17	-0,10	-0,02	0,20	1,00												
TBL	0,70	0,70	0,69	-0,07	-0,03	0,71	0,21	1,00											
LTY	0,76	0,76	0,64	0,08	0,02	0,71	0,29	0,91	1,00										
LTR	0,04	0,05	0,04	0,00	0,01	0,02	-0,05	0,02	0,01	1,00									
TMS	-0,08	-0,07	-0,29	0,33	0,11	-0,19	0,11	-0,49	-0,08	-0,04	1,00								
DFY	0,50	0,50	0,15	0,45	0,42	0,48	-0,25	0,27	0,36	0,11	0,10	1,00							
DFR	-0,01	0,02	-0,10	0,13	0,12	0,00	0,01	-0,05	-0,01	-0,45	0,10	0,08	1,00						
INFL	0,38	0,38	0,44	-0,13	0,01	0,48	0,20	0,45	0,36	-0,07	-0,31	-0,01	-0,07	1,00					
SII	-0,20	-0,20	-0,23	0,08	-0,15	-0,30	-0,10	-0,10	0,00	0,01	0,24	-0,06	-0,06	-0,16	1,00				
OGAP	-0,36	-0,38	-0,01	-0,48	-0,19	-0,11	0,07	0,00	-0,17	-0,05	-0,38	-0,35	-0,10	0,15	0,26	1,00			
TCHI	0,20	0,18	0,08	0,16	0,11	0,26	0,01	0,14	0,16	0,08	0,00	0,30	-0,08	0,10	0,07	0,23	1,00		
AVGCOR	-0,02	-0,03	-0,10	0,12	0,37	-0,03	-0,32	-0,29	-0,29	0,12	0,10	0,34	-0,06	-0,17	0,12	-0,03	0,29	1,00	
GPCE	0,45	0,44	0,63	-0,33	-0,13	0,57	0,40	0,59	0,55	-0,03	-0,26	-0,04	-0,06	0,49	-0,28	0,20	0,01	-0,36	1,00

E RRZ replication

Table 14: in-sample regression estimation results foor sample period 1973:01-2014:12.
The table reports the ordinary least squares estimate of β and R^2 statistic for the predictive regression in equation 3 explained in section 3.1. *,** and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values. The last two rows corresponds to a multiple predictive regression that includes an intercept, SII, and the first three principal components of the Welch & Goyal (2008) variables (denoted with "GW") and first three components of all the variables excluding SII (denoted with "all"). For these two rows, the table displays estimated slope coefficients and partial R^2 statistics corresponding to SII.

	$h = 1$		$h = 3$		$h = 6$		$h = 12$	
Predictor	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)
DP	0.16	0.13	0.17	0.43	0.20	1.01	0.21	2.22
	[0.78]		[0.98]		[1.10]		[1.17]	
DY	0.18	0.17	0.18	0.48	0.20	1.08	0.21	2.34
	[0.90]		[1.04]		[1.14]		[1.20]	
EP	0.10	0.05	0.07	0.07	0.06	0.11	0.09	0.40
	[0.39]		[0.30]		[0.28]		[0.47]	
DE	0.06	0.02	0.12	0.22	0.16	0.67	0.14	1.04
	[0.23]		[0.56]		[0.88]		[1.21]	
RVOL	0.35	0.60	0.32	1.49	0.27	1.87	0.17	1.57
	[1.88]**		[2.10]**		[1.93]**		[1.26]	
BM	0.03	0.00	0.04	0.02	0.07	0.12	0.07	0.27
	[0.13]		[0.22]		[0.35]		[0.39]	
NTIS (-)	0.07	0.02	0.00	0.00	-0.01	0.00	0.01	0.01
	[0.27]		[0.00]		[-0.03]		[0.06]	
TBL (-)	0.25	0.32	0.21	0.63	0.18	0.86	0.16	1.29
	[1.24]		[1.17]		[0.96]		[0.95]	
LTY (-)	0.14	0.09	0.10	0.13	0.07	0.12	0.00	0.00
	[0.66]		[0.52]		[0.35]		[-0.02]	
LTR	0.33	0.53	0.14	0.26	0.23	1.41	0.15	1.11
	[1.65]**		[0.91]		[2.48]***		[2.86]***	
TMS	0.33	0.55	0.31	1.34	0.29	2.18	0.36	6.59
	[1.65]**		[1.73]*		[1.63]*		[2.27]**	
DFY	0.15	0.12	0.17	0.39	0.24	1.56	0.19	1.82
	[0.56]		[0.68]		[1.26]		[1.21]	
DFR	0.48	1.15	0.22	0.69	0.16	0.66	0.06	0.17
	[1.52]*		[1.23]		[1.32]		[0.88]	
INFL (-)	-0.03	0.00	0.16	0.37	0.27	1.88	0.24	2.84
	[-0.11]		[0.83]		[1.72]*		[1.98]*	
SII (-)	0.51	1.26	0.56	4.43	0.57	8.19	0.53	13.12
	[2.52]***		[2.90]***		[2.74]**		[2.72]**	
OGAP (-)	0.59	1.73	0.61	5.30	0.59	9.18	0.55	15.36
	[3.06]***		[3.75]***		[3.68]***		[3.28]***	
TCHI (-)	0.51	1.30	0.42	2.45	0.39	4.04	0.24	3.06
	[2.11]**		[1.85]*		[1.76]*		[1.38]	
AVGCOR	0.30	0.46	0.42	2.54	0.37	3.59	0.24	2.89
	[1.24]*		[2.25]**		[2.62]**		[1.66]*	
GPCE (-)	0.30	0.45	0.33	1.55	0.32	2.76	0.31	5.00
	[1.32]*		[1.75]*		[1.98]*		[2.11]*	
SII (-)‖PC(GW)	0.51	1.27	0.58	4.55	0.59	8.71	0.55	13.72
	[2.64]***		[3.03]***		[2.79]**		[2.73]**	
SII (-)‖PC(all)	0.47	1.06	0.54	4.09	0.55	7.92	0.51	12.79
	[2.49]***		[2.99]***		[2.90]**		[2.98]**	

Table 15: Out-of-sample R^2 statistics (%) and Encompassing tests for 1990:01-2014:12.
In the left panel out-of-sample R^2 statistics can be found computed with equation 11 against the prevailing mean forecast, that serves as benchmark. Significance is tested with the Clark & West (2007) test statistic. In the right panel the results for the encompassing tests can be found, where forecasts are based on a convex combination of SII forecasts and forecasts based on one of the other predictor. Significance is based on the Harvey et al. (1998) test statistic where the null hypothesis states that the weight on SII's forecast is 0 against the alternative that this is greater than 0.
*, ** and *** indicate significance at the 10%, 5%, and 1% levels, respectively

	Out-of-sample R^2 statistics (%)				Encompassing tests			
Predictor	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
DP	-2.06	-5.79	-10.91	-26.24	1.00***	1.00***	1.00***	1.00***
DY	-2.20	-5.64	-10.96	-25.65	1.00***	1.00***	1.00***	1.00***
EP	-1.13	-4.25	-8.88	-16.42	1.00***	1.00***	1.00***	1.00***
DE	-2.27	-6.31	-7.85	-3.58	1.00***	1.00***	1.00***	0.95***
RVOL	-0.54	-1.34	-1.72	-3.46	0.98***	1.00***	1.00***	1.00***
BM	-0.56	-1.70	-3.42	-9.53	1.00***	1.00***	1.00***	1.00***
NTIS	-3.16	-8.67	-18.59	-27.27	1.00***	1.00***	1.00***	1.00***
TBL	-0.40	-0.99	-1.78	-1.99	1.00**	1.00***	1.00***	1.00***
LTY	-0.34	-1.60	-3.77	-11.96	1.00***	1.00***	1.00***	1.00***
LTR	-0.44	-1.52	-0.90	-0.95	1.00***	1.00***	1.00***	0.93***
TMS	-0.76	-1.72	-1.40	3.44**	1.00***	1.00***	1.00***	0.82**
DFY	-3.06	-7.18	-8.90	-7.31	1.00***	1.00***	1.00***	1.00***
DFR	-1.96	-1.21	-0.52	-0.85	0.97**	1.00***	1.00***	0.99***
INFL	-0.71	-0.51	2.02**	2.23	1.00***	1.00***	1.00***	0.92***
SII	1.95***	6.56***	11.71***	13.22**	-	-	-	-
OGAP	-3.43	-8.26	-9.54	-5.58	1.00***	1.00***	1.00***	0.95***
TCHI	1.38*	2.08	4.12*	3.29*	0.65	0.94***	0.95***	0.85***
AVGCOR	-1.18	-1.57	1.52	-1.39	1.00***	0.88***	0.93***	0.94**
GPCE	-0.42	-1.04	-1.41	-0.92	1.00**	1.00***	1.00***	0.83**

Table 16: Out-of-sample CER gains and Sharpe ratios for sample periods 1990:01-2014:12, 1990:01-2006:12 and 2007:01-2014:12.
In the left panel annualized CER gains (in percentage) can be found as specified in section 5. In the right panel Sharpe ratios can be found as specified in section 5. Buy and hold corresponds to the investor passively holding the market portfolio.

Predictor	CER gain ($h = 1$)			Sharpe ratios ($h = 1$)		
	1990:01-2014:12	1990:01-2006:12	2007:01-2014:12	1990:01-2014:12	1990:01-2006:12	2007:01-2014:12
Prevailing mean	-	-	-	0.39	0.39	0.40
DP	-3.22	-4.51	-0.51	0.11	0.01	0.36
DY	-2.99	-4.38	-0.02	0.13	0.02	0.42
EP	-0.27	-0.96	1.19	0.37	0.31	0.54
DE	-1.09	-1.54	-0.15	0.32	0.30	0.39
RVOL	-1.77	-2.97	0.75	0.29	0.21	0.45
BM	-0.76	-1.14	0.04	0.33	0.31	0.41
NTIS	-2.55	-1.59	-4.59	0.21	0.26	0.11
TBL	0.64	1.11	-0.37	0.45	0.47	0.39
LTY	-0.08	-0.37	0.51	0.39	0.37	0.45
LTR	-0.75	-0.96	-0.32	0.35	0.33	0.38
TMS	0.29	1.10	-1.45	0.42	0.46	0.35
DFY	-4.92	-4.49	-5.85	-0.02	-0.08	0.05
DFR	0.77	0.96	0.36	0.44	0.45	0.42
INFL	-0.66	0.55	-3.18	0.35	0.43	0.16
SII	4.20	0.91	11.23	0.66	0.45	1.05
OGAP	-3.51	-4.84	-0.70	0.22	0.06	0.44
TCHI	4.71	3.34	7.61	0.73	0.62	0.96
AVGCOR	-0.31	0.95	-3.04	0.38	0.45	0.30
GPCE	0.39	0.93	-0.78	0.44	0.47	0.38
Buy and hold	1.67	2.02	0.91	0.51	0.53	0.46

Table 17: Predictive regression estimation results for market return components for sample period 1973:01–2021:12.
In the table, $\hat{\beta}_{\hat{E}}$, $\hat{\beta}_{CF}$ and $\hat{\beta}_{DR}$ estimates can be found for the framework explained in section 6 and equations 22, 23 and 24. PC indicates that the first three components are included in the VAR framework where one PC inlcudes the Welch & Goyal (2008) variables and one PC includes all the variables excluding SII. The brackets report report heteroskedasticity- and autocorrelation-robust t-statistics. *, ** and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VAR variables	$\beta_{\hat{E}}$	β_{CF}	β_{DR}	VAR variables	$\beta_{\hat{E}}$	β_{CF}	β_{DR}
r, DP	-0.07	-0.35	0.09	r, DP, TMS	-0.08	-0.34	0.09
	[-3.10]***	[-2.27]**	[2.05]**		[-3.55]***	[-2.22]**	[1.63]
r, DP, DY	-0.06	-0.35	0.10	r, DP, DFY	-0.07	-0.36	0.09
	[-2.37]**	[-2.27]**	[2.39]**		[-3.12]***	[-2.27]**	[2.01]**
r, DP, EP	-0.08	-0.40	0.04	r, DP, DFR	-0.09	-0.34	0.08
	[-3.50]***	[-2.37]**	[0.98]		[-2.73]***	[-2.21]**	[1.95]*
r, DP, DE	-0.08	-0.40	0.04	r, DP, INFL	-0.07	-0.36	0.09
	[-3.50]***	[-2.37]**	[0.98]		[-3.03]***	[-2.28]**	[2.04]**
r, DP, RVOL	-0.11	-0.28	0.13	r, DP, OGAP	-0.14	-0.24	0.13
	[-3.62]***	[-1.96]**	[1.92]*		[-4.61]***	[-1.61]*	[1.75]*
r, DP, BM	0.00	-0.34	0.17	r, DP, TCHI	-0.13	-0.37	0.01
	[-0.01]	[-2.35]**	[2.80]***		[-3.92]***	[-2.11]**	[-0.17]
r, DP, NTIS	-0.05	-0.36	0.10	r, DP, AVGCOR	-0.04	-0.33	0.14
	[-2.38]**	[-2.33]**	[2.17]**		[-1.90]*	[-2.33]**	[2.06]**
r, DP, TBL	-0.09	-0.31	0.11	r, DP, GPCE	-0.05	-0.30	0.16
	[-3.64]***	[-2.09]**	[1.45]		[-1.92]*	[-2.15]**	[2.17]**
r, DP, LTY	-0.08	-0.32	0.11	r, DP, PC (GW)	-0.05	-0.31	0.15
	[-3.34]***	[-2.20]**	[1.62]		[-2.07]**	[-2.18]**	[2.31]**
r, DP, LTR	-0.07	-0.35	0.09	r, DP, PC (all)	-0.09	-0.28	0.14
	[-2.67]***	[-2.28]**	[1.93]*		[-2.85]***	[-2.11]**	[1.81]*

F Wealth and weight graph for RRZ replication

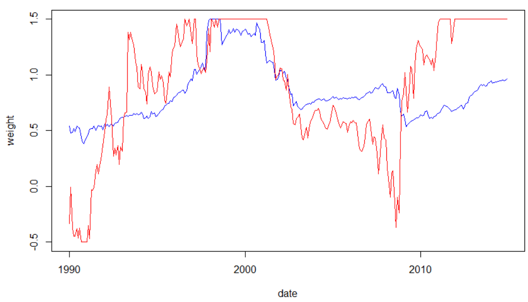

Figure 10: Equity weights for the asset allocation strategy explained in section 5 where the red line corresponds with the allocation using the SII based forecasts and the blue with the allocation using the prevailing mean forecasts for the period 1990:01-2014:12

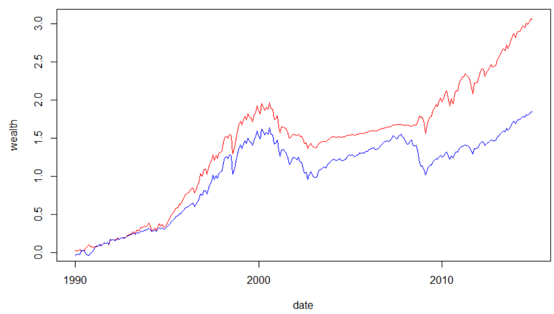

Figure 11: Log cumulative wealth for the investor assuming that the investor begins with $1 and reinvests all proceeds. The red line corresponds with using SII based forecasts for the allocation and the blue line corresponds with using the prevailing mean forecasts for the period 1990:01-2014:12