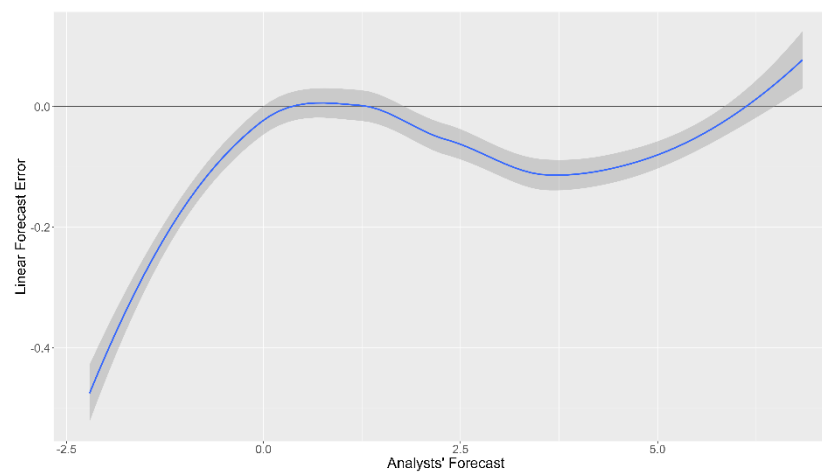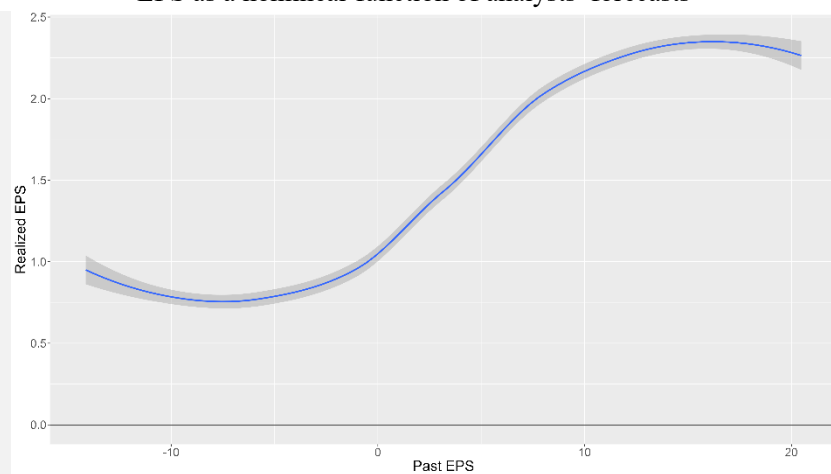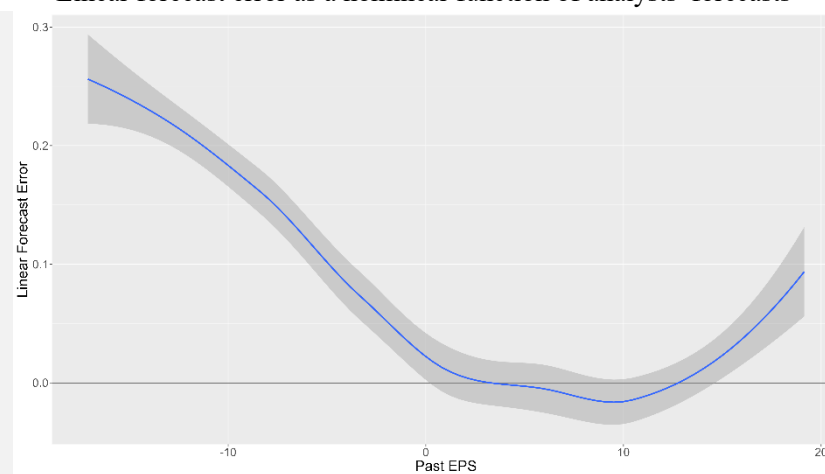EPS as a nonlinear function of analysts' forecasts

Linear forecast error as a nonlinear function of analysts' forecasts
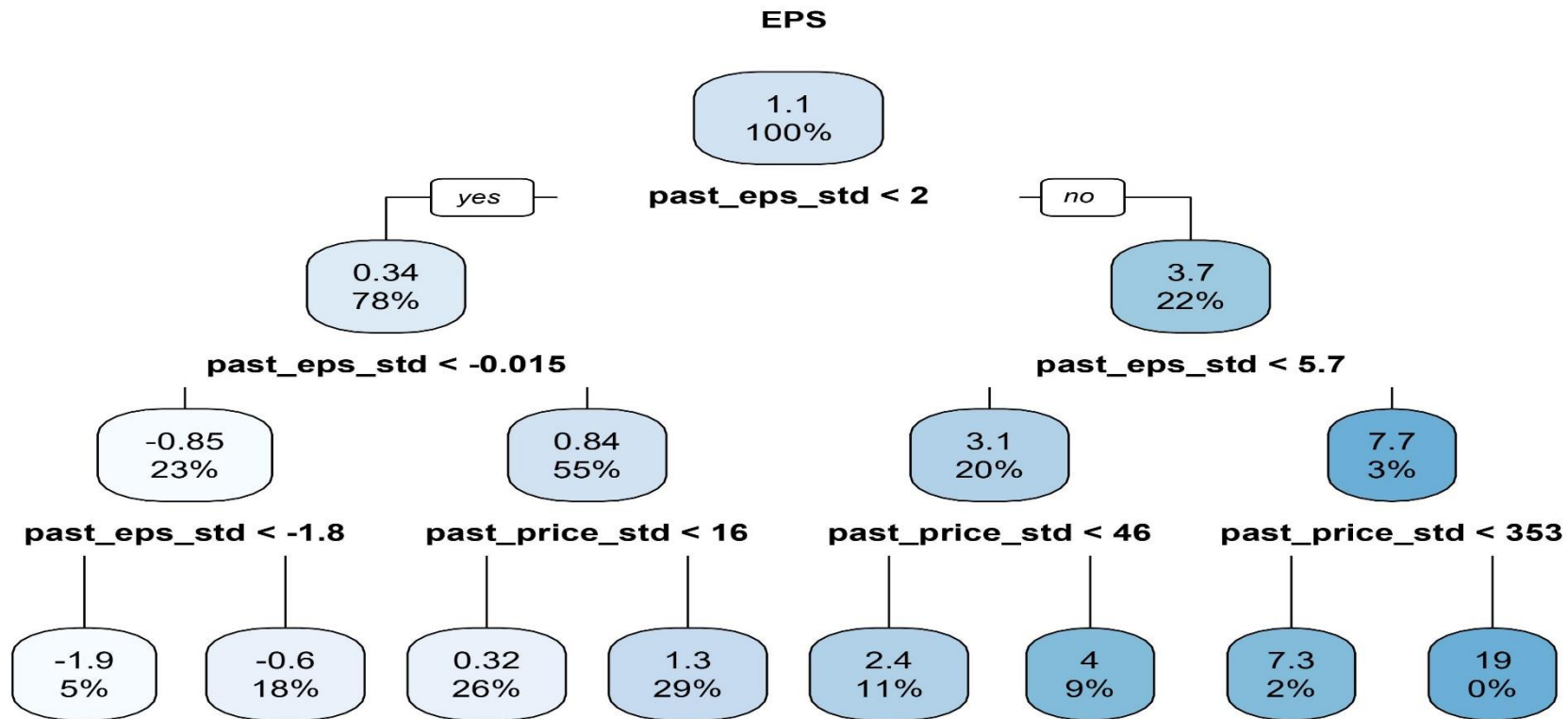
EPS as nonlinear function of past EPS

Linear forecast error as a nonlinear function of past EPS

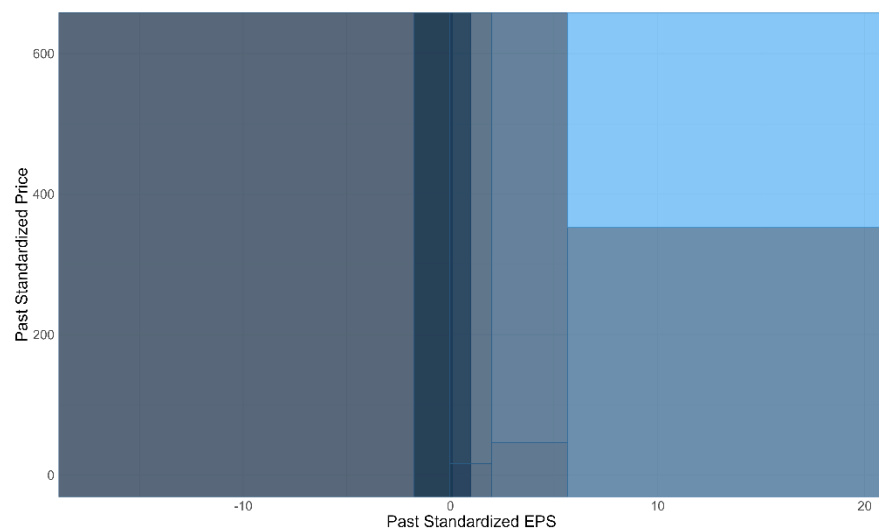**Figure 1: Partial dependence plot**

Table 1: Hyperparameters for the random forest regression

Number of trees	2000
Maximum depth	7
Sample fraction	5%
Minimum node size	5

Figure 2: Example decision tree
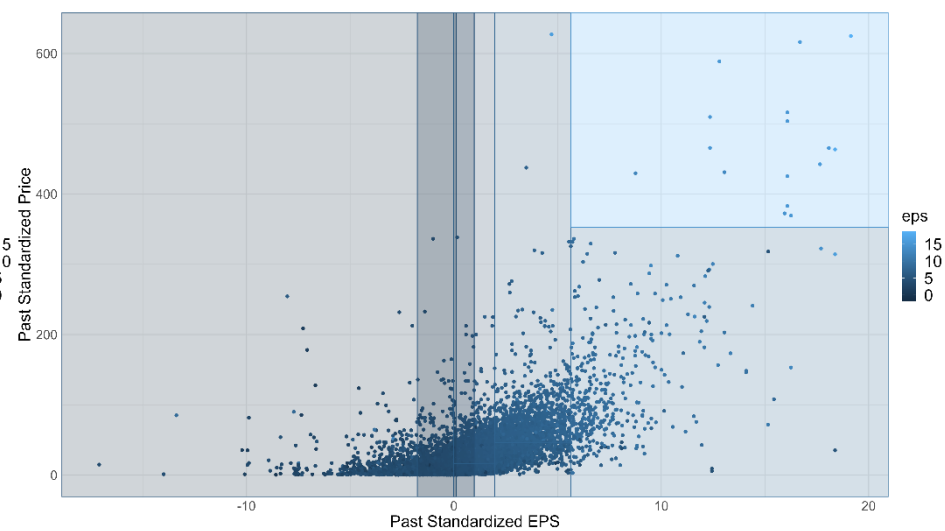
Decision tree prediction regions   Predictions and realized values

Figure 3: Example decision tree prediction regions

Figure 3 illustrates the forecast of the decision tree from Figure 2. The variable we wish to forecast is the earnings-per-share for a cross-section of firms. Left panel shows the prediction is constant within each color box and corresponds to the historical mean for each subspace. Right panel shows the realized values with different colors indicating different values.
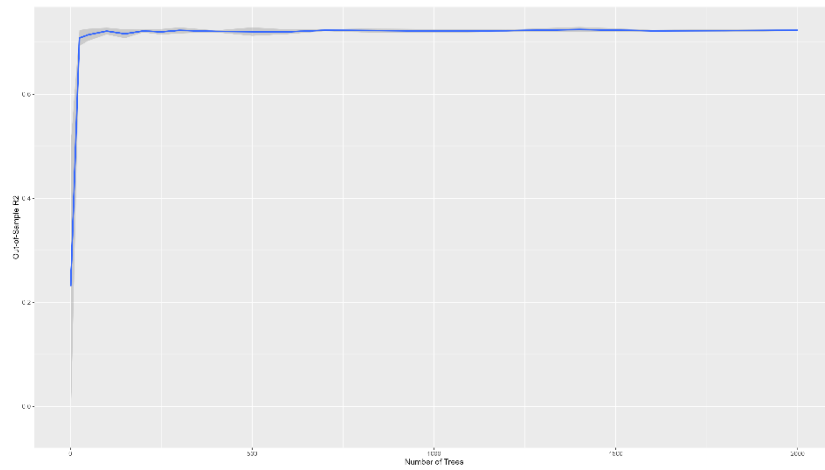
Number of trees in the one-quarter-ahead forecast
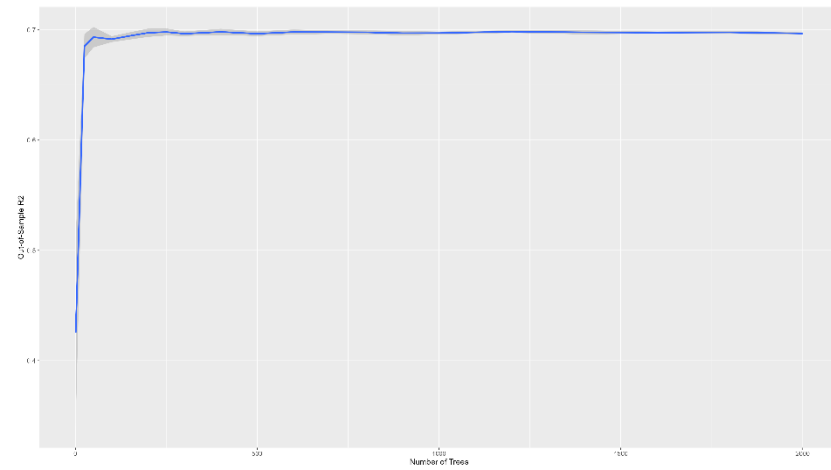
Number of trees in the one-year-ahead forecast
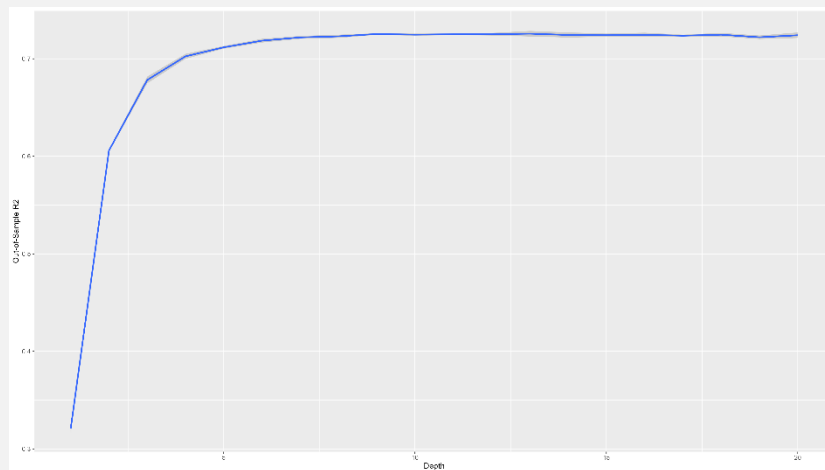
Maximum depth in the one-quarter-ahead forecast
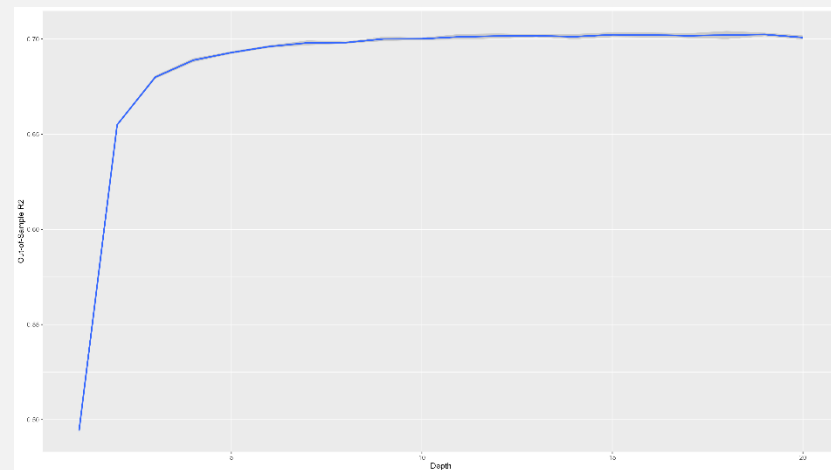
Maximum depth in the one-year-ahead forecast

Fraction of the sample in the one-quarter-ahead forecast   Fraction of the sample in the one-year-ahead forecast
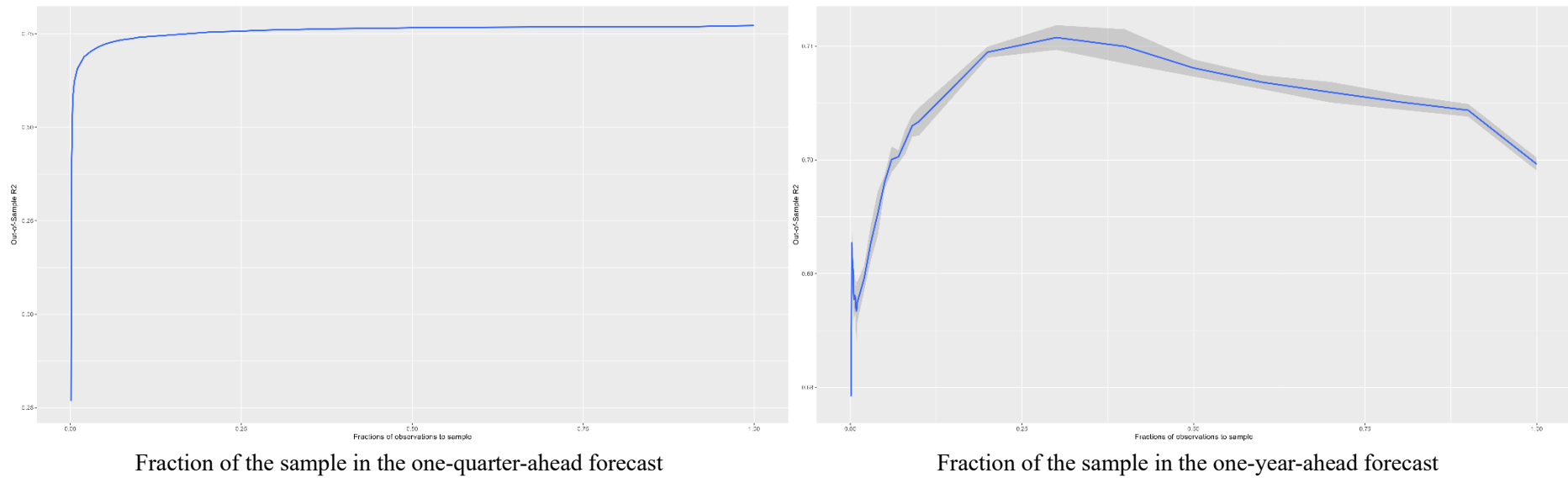
Figure 4: Cross-validation for hyperparameters

The model is trained using data up to 1986 January and the out-of-sample R^2 for the 1-year-ahead earnings forecasts is calculated in 1986 February. The out-of-sample R^2 is defined as one minus the mean squared error implied by using the machine learning forecast divided by the mean squared error of using the realized average value as a forecast. The random forest algorithm is random by design, so I take the average of 5 runs to measure the out-of-sample R^2.
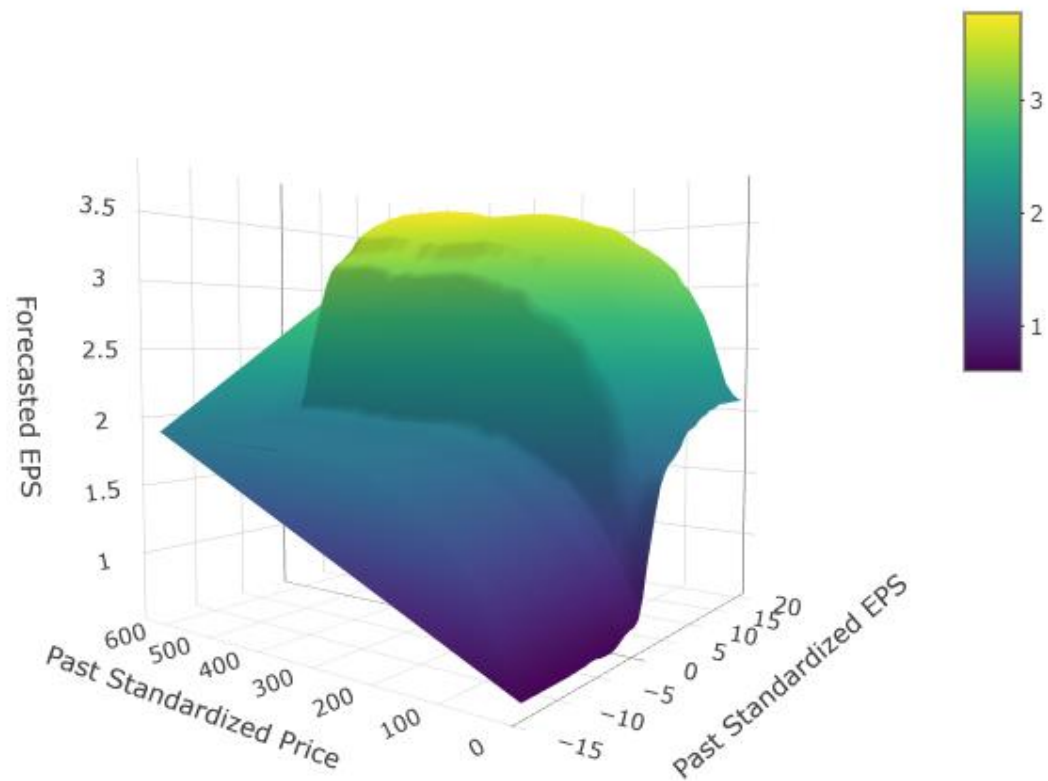
Figure 5: EPS as a nonlinear function of stock price and past EPS

The figure plots the partial dependence plot of one-quarter-ahead realized EPS on past EPS and stock price. The partial dependence plot is calculated from a random forest regression of EPS on the variables mentioned in the original essay. The random forest regression for the figure uses 2,000 trees and a minimum node size of one. The data start in 1986 and end in 2021.
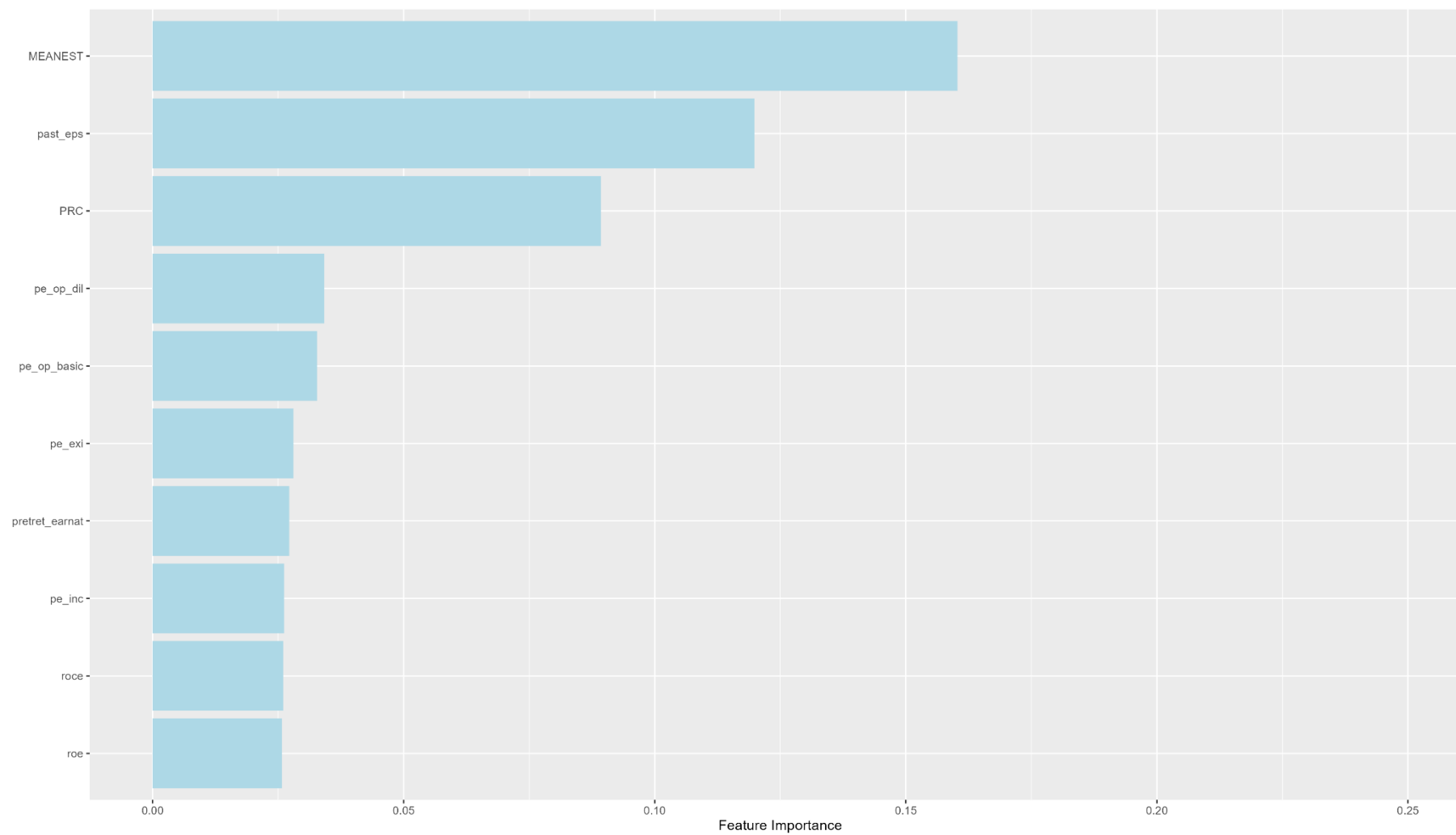
Feature importance of the one-quarter ahead forecast
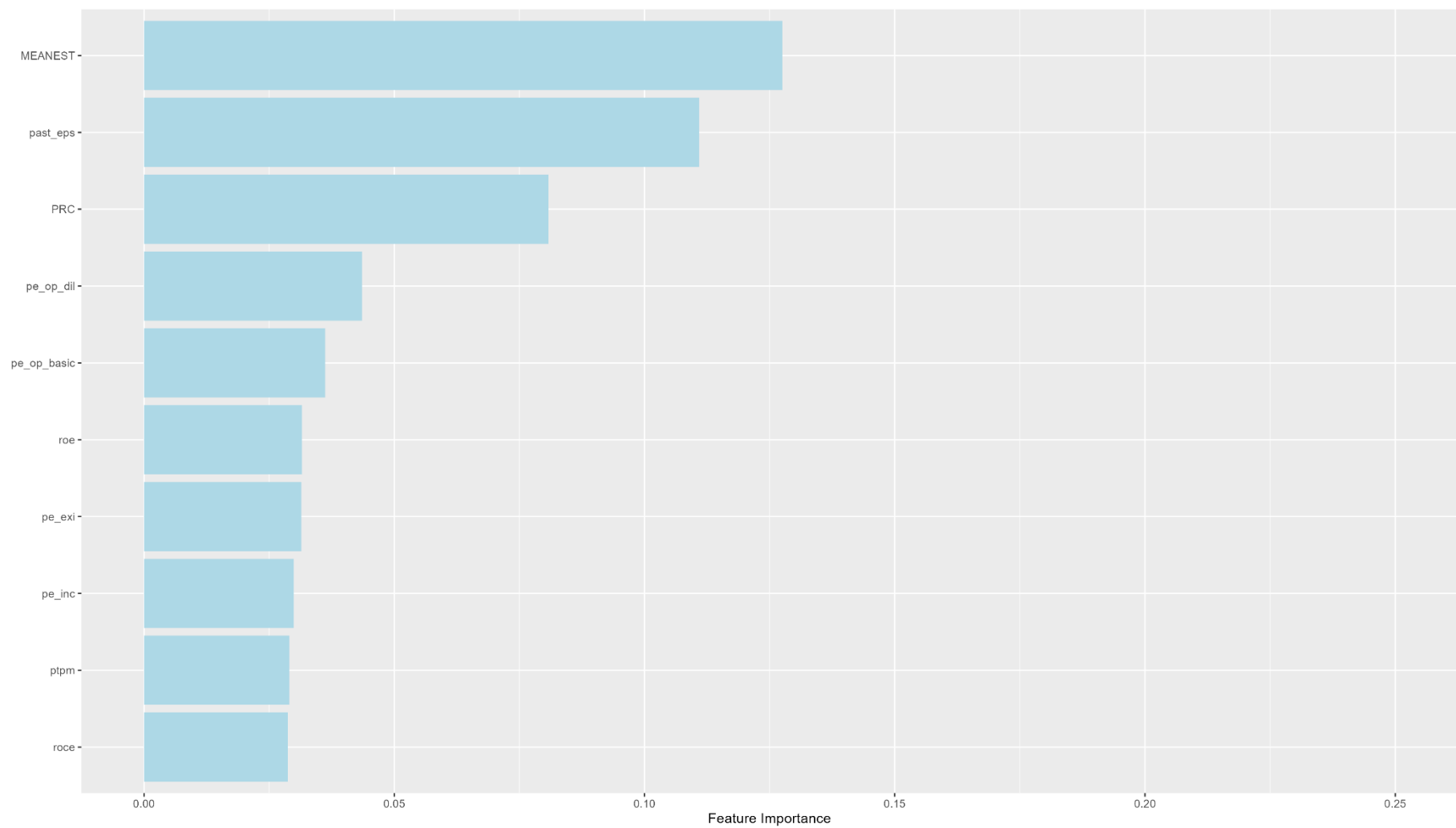
Feature importance of the one-year ahead forecast

Figure 6: Feature importance

The figure plots the time-series average of feature importance of the 10 most important variables for the one-quarter-ahead earnings forecasts in the first picture and for the 1-year-ahead in the second one. The feature importance for each variable is the normalized sum of the reduced mean squared error decrease when splitting on that variable using the method in Nembrini, König, and Wright (2018). The feature importance of each variable is normalized so that the features' importance sums to one.

For one-quarter ahead forecast, the 5 most important variables are: Analysts' forecasts(17%), past_eps(13%), stock_price(8%), P/E, P/E(dilluted), the result for one-year-ahead forecast is nearly the same.

Table 2: The term structure of earnings forecasts via machine learning

	RF	AF	AE	(RF-AE)	(AF-AE)	(RF-AE)^2	(AF-AE)^2	(AF-RF)/P	N
One-quarter-ahead	1.120	0.318	1.130	-0.010	-0.812	0.673	3.484	0.031	378,416
t-stat				-1.12	-21.53			4.35	
Two-quarters-ahead	1.302	0.398	1.321	-0.019	-0.923	0.793	3.726	0.009	299,945
t-stat				-1.65	-22.76			2.21	
Three-quarters-ahead	1.346	0.437	1.371	-0.025	-0.933	1.044	4.024	0.011	286,437
t-stat				-1.59	-21.10			2.28	
One-year-ahead	1.173	1.347	1.190	-0.017	0.157	0.660	0.803	0.035	1,372,431
t-stat				-1.64	6.35			6.96	
Two-years-ahead	1.300	1.779	1.328	-0.028	0.451	1.523	2.349	0.049	1,171,083
t-stat				-0.89	8.39			10.28	

This table presents the time-series average of machine learning earnings per share forecasts (RF), analysts' earning forecasts (AF), actual realized earnings (AE)—the difference, as well as the squared difference between them. N denotes the number of sample stocks. I report the Newey-West (Newey and West 1987) t-statistics of differences between earnings forecasts and realized earnings. Because the earning forecasts are made monthly, we adjust the quarterly forecasts with three lags and the annual forecasts with 12 lags when reporting the Newey-West t-statistics. The sample period is January 1986 to December 2021.

For forecasts at all horizons, analysts make overpessimistic forecasts for quarterly ones and overoptimistic forecasts for yearly ones. The realized analysts' forecasts errors, defined as the difference between the analysts' forecasts and the realized value, increase in the forecast horizon, ranging from -0.812 to 0.451 on average. All of

these are statistically significantly different from zero. In sharp contrast, the time-series averages of the differences between the machine-learning forecast and realized earnings are statistically indistinguishable from zero, with an average absolute value of around -0.0175 for the quarterly earnings forecasts, -0.017 for the 1-year-ahead forecast, and −0.028 for the 2-year-ahead forecast.

The mean squared errors of the machine learning forecast are smaller than the analysts' mean squared errors, demonstrating that machine learning forecasts are more accurate than the forecasts provided by analysts.

This essay defines the conditional expectation bias for every stock as the difference between the analysts' forecast and the machine-learning forecast, scaled by the closing stock price in the most recent month. The second-to-last column of Table 2 reports the time-series average of the real-time-biased expectations. The average conditional earnings forecast bias is statistically different from zero for all horizons. Furthermore, I find that analysts are more biased at longer horizons generally.

Average bias of analysts' earnings expectations relative to ML forecasts

Average realized bias of analysts' earnings expectations

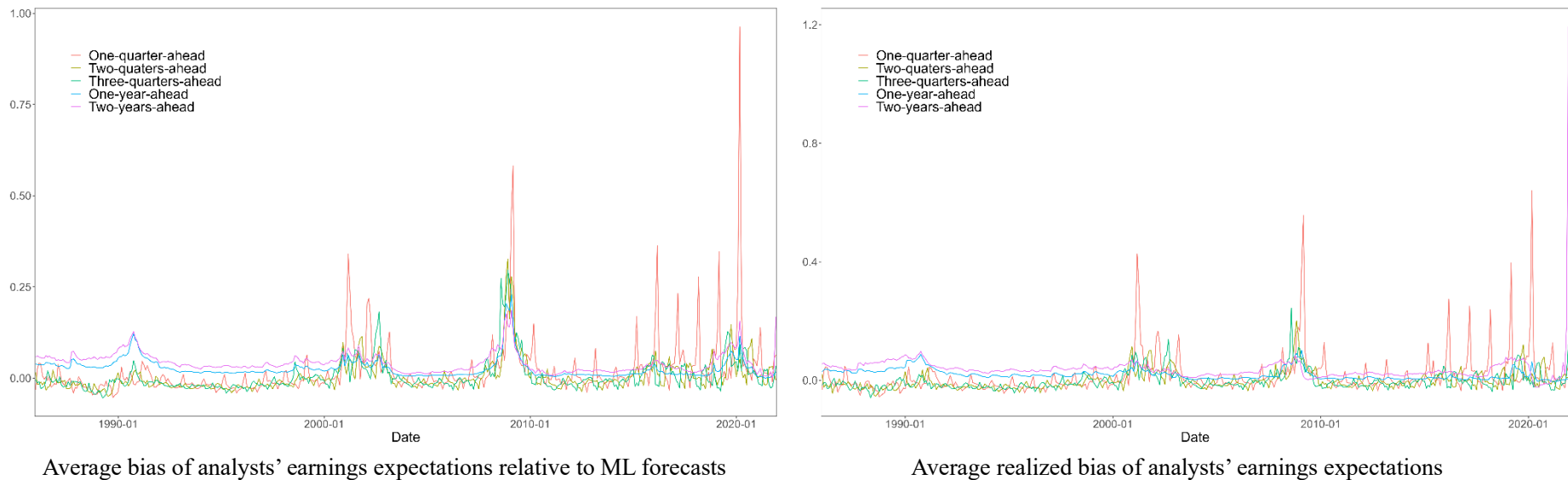**Figure 7: Average bias of analysts' earnings expectations**

Figure 7, left panel, shows the conditional aggregate bias, defined as the average of the individual stocks' expectations. I consider five different forecast horizons and consider the possibility that the aggregate bias is higher during historical bubbles. I find clear spikes during the internet bubble of the early 2000s (Griffin et al.

2011) and in the financial crisis. For comparison, Figure 7, right panel, displays the average realized bias. Both the realized and the conditional bias show similar patterns, albeit with different magnitudes, and both figures show spikes during the internet bubble and the financial crisis, as well as the covid-19 pandemic.

Table 3: Fama-Macbeth regressions

	A. Average BE		B. BE score	
	(1)	(2)	(1)	(2)
Bias	-0.010		-0.015	
t-stat	-1.34		-3.14	
LNsize				
t-stat				
LNbeme				
t-stat				
Ret1				
t-stat				
Ret12_7				
t-stat				
IA				
t-stat				
IVOL				
t-stat				
Retvol				
t-stat				
Turnover				
t-stat				
Intercept	1.068		1.864	
t-stat	3.82		8.10	
R-sqr (%)	1.373		1.831	

Table 3 shows the regression results. The first column in each panel of Table 3 reports the regression without control variables. Both the conditional bias and the bias score are associated with negative cross-sectional stock return predictability. However, the coefficient of condition bias is not significant at the 1% level, whereas that of bias score is significant. Because of the lack of data on other variables, I didn't add control variables.

Table 4: Correlations between conditional bias

Variable	Average BE	BE Score	BE_Q1	BE_Q2	BE_Q3	BE_A1	BE_A2
Average BE	1						
BE Score	0.432	1					
BE_Q1	0.934	0.484	1				
BE_Q2	0.939	0.529	0.785	1			
BE_Q3	0.937	0.542	0.722	0.801	1		
BE_A1	0.895	0.348	0.75	0.782	0.817	1	
BE_A2	0.866	0.452	0.638	0.621	0.607	0.725	1

Table 4 reports the correlations between the bias measures and the control variables. The conditional bias and the bias score are highly positively correlated.

Table 5: Portfolios sorted on conditional bias

Quintile	1	2	3	4	5	5-1
A. Average BE						
Mean	1.80	1.37	1.66	1.85	2.74	0.94
t-stat	8.74	6.82	6.63	5.77	5.54	2.40
CAPM Beta	0.86	0.89	1.08	1.32	1.74	0.88
B. BE score						
Mean	1.79	1.53	1.52	1.70	2.49	0.70
t-stat	9.20	7.11	6.19	5.38	4.98	1.74
CAPM Beta	0.86	0.97	1.08	1.32	1.70	0.84

I sort stocks into five quintile portfolios based on the conditional bias. Table 5 reports the portfolio sorts. Two interesting patterns emerge. First, the value-weighted returns increase in the conditional bias. A long-short portfolio of the extreme quintiles results in a return spread of 0.94% for the average bias and 0.70% for the bias

score. Second, the capital asset pricing model (CAPM) betas of these portfolios tend to increase with higher biased expectations, a finding that is consistent with the results of Antoniou, Doukas, and Subrahmanyam (2015) and Hong and Sraer (2016), who show that high-beta stocks are more susceptible to speculative overpricing.

Table 6: Time-series tests with common asset pricing models

	Coef(β)	t-stat	Coef(β)	t-stat	Coef(β)	t-stat
	CAPM		FF3		FF5	
A.Average BE						
Intercept	0.27	0.81	0.31	1.06	0.95	3.22
Mkt_RF	0.88	9.17	0.71	7.38	0.55	5.98
SMB			1.19	7.81	0.80	5.20
HML			0.07	0.47	0.48	2.73
RMW					-1.23	-6.19
CMA					-0.34	-1.14
B. BE score						
Intercept	0.05	0.15	0.05	0.17	0.60	1.92
Mkt_RF	0.84	9.44	0.70	7.64	0.57	6.40
SMB			1.19	6.31	0.84	4.83
HML			0.29	2.06	0.62	3.75
RMW					-1.11	-5.11
CMA					-0.19	-0.53

Table 6, panel A, reports the results of using the average conditional bias as the portfolio sorting variable. The long-short strategy has a unsignificant CAPM alpha of 0.27% per month, with a significantly positive market beta of 0.88. Columns 4 to 7 show the regression results with the Fama-French three-factor (Fama and French 1993) and five-factor models (Fama and French 2015). Neither model can explain the documented return spread. The alpha in the three-factor model is with a t-statistic of 1.06; the alpha in the five-factor model is −0.95%with a t-statistic of 3.22. Table 6, panel B, reports the long-short strategy using the bias score as the sorting variable, and I find consistent results.

I also plot the equal-weighted result of table 5 and 6, which perform much greated than the value-weighted one. Perhaps the reason is that I calculated the firm value with previous outstanding shares which may not be correct.

Table 5: Portfolios sorted on conditional bias (equal-weighted)

Quintile	1	2	3	4	5	5-1
A. Average BE						
Mean	1.74	1.24	1.06	0.76	0.61	-1.13
t-stat	7.35	5.21	3.87	2.34	1.34	-3.51
CAPM Beta	0.96	1.02	1.16	1.32	1.56	0.59
B. BE score						
Mean	1.67	1.32	1.07	0.75	0.58	-1.10
t-stat	7.62	5.35	3.82	2.29	1.31	-3.51
CAPM Beta	0.94	1.06	1.18	1.32	1.52	0.58

Compared to previous Table 5, this time the portfolios return decreases as the conditional bias increase, which is consistent with the original essay. Moreover, the long-short portfolios also earn a significantly return of around -1.10% when using Average BE or BE score.

Second, the capital asset pricing model (CAPM) betas of these portfolios tend to increase with higher biased expectations as well.

Table 6: Time-series tests with common asset pricing models (equal-weighted)

	Coef(β)	t-stat	Coef(β)	t-stat	Coef(β)	t-stat
	CAPM		FF3		FF5	
	A. Average BE					
Intercept	-1.58	-5.43	-1.45	-5.58	-0.87	-3.17
Mkt_RF	0.59	7.95	0.42	5.51	0.28	3.87
SMB			0.90	7.02	0.50	3.80
HML			-0.43	-3.38	-0.10	-0.75
RMW					-1.23	-7.65
CMA					-0.13	-0.51
	B. BE score					
Intercept	-1.54	-5.56	-1.47	-5.71	-1.00	-3.42
Mkt_RF	0.58	8.16	0.44	5.79	0.33	4.11
SMB			0.88	5.85	0.57	3.83
HML			-0.16	-1.26	0.11	0.77
RMW					-0.96	-5.21
CMA					-0.16	-0.54

Surpisingly, Table 6 using equal-weighted portfolios also generate a better result. No matter FF3 or FF5 factors, neither of them could explain the alpha. And alpha of each portfolios are significant as well, with a significant market beta. In coclusion, I would say the result is acceptable when using equal-weighted methods. The reason of the falilure in value-weighted ones may be the wrong shares outstanding data. Then, I would compare the performance of equal-weighted portfolio and the market.
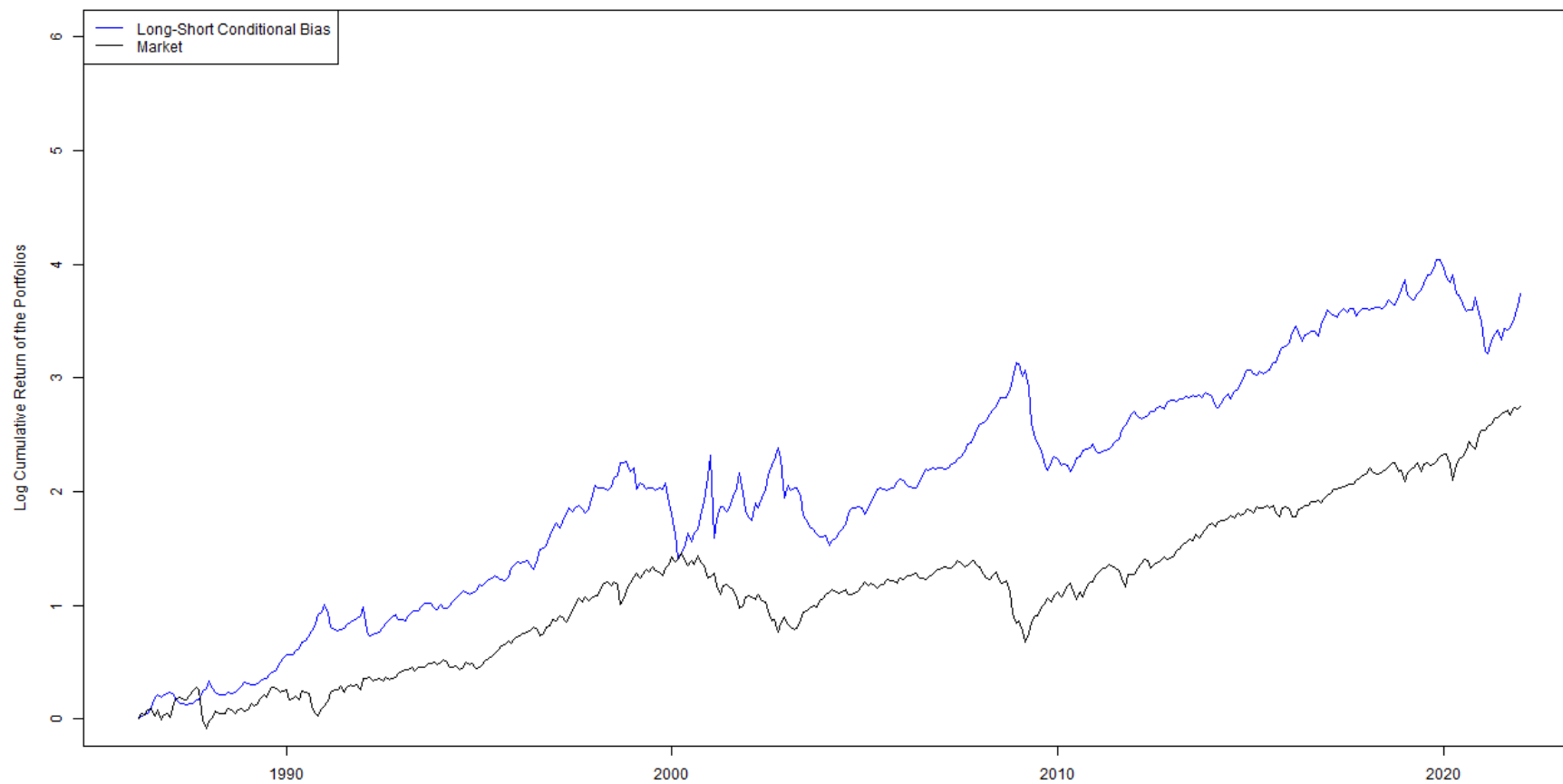
Figure 9: Cumulative Performance of the Portfolios sorted on Conditional Bias

This figure plots the (log) cumulative performance of the return of a **equal-weighted** long-short portfolio that is short on firms with the highest conditional earnings

forecast bias and long on the firms with the lowest. The figure also plots the market return for comparison. The market return data come from Kenneth R. French's website. The sample period is 1986 to 2021.

I didn't continue to the Anomoly part due to the lack of data on Anomoly portfolios.

Supplement: Correlation Heatmap