Internet Appendix for "Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases"

Jules H. van Binsbergen, Xiao Han, and Alejandro Lopez-Lira

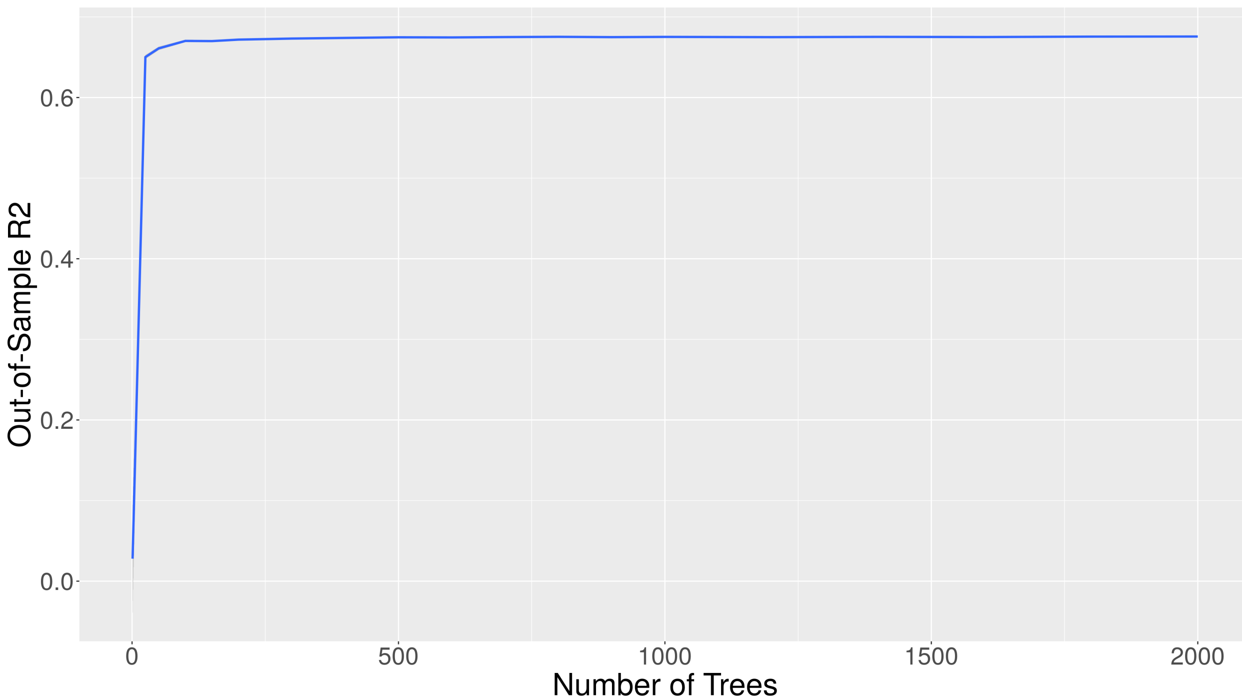
A1. Parameters in random forest

We choose the hyperparameters in a purely data-driven way using cross-validation. We use data up to (and including) 1986 by dividing the data into two partitions: training and testing (cross-validation). The training data contains the early part of the sample: from the beginning of the sample until January 1986. The testing data includes a single month: February 1986. The results are similar to other testing periods in 1986. We train the model using the training data for different configurations of the hyperparameters. We evaluate the results in the testing data and pick the parameters that result in the best performance. Notice that the testing data is not using information from future periods. We maintain the hyperparameters chosen in 1986 for the whole sample. The model is then trained using rolling windows, keeping the hyperparameters fixed.

We choose 2000 trees from the cross-validation procedure but remark that there is little difference after 500. We use the recommended minimum node size of 5. We find no significant differences in the out-of-sample R^2 and even a slight reduction after a depth of seven, so we choose that parameter. The result is explained in the following way: we train using a rolling window of 12 months for around 10,000 observations. Since each split divides the data into two and we use a minimum node of 5, the maximum number of splits is ten since $\frac{10^3}{2^{10}} = 9.77$. Figure A1, Figure A2, and Figure A3 show the cross-validation results for the first-period two-quarters-ahead, three-quarters-ahead, and two-years-ahead earnings forecasts.

The standard algorithm allows for the specification of the probability of a predictor being chosen at each step. We take advantage of that and implement a two-step procedure. First, we run a standard random forest regression, where every variable has the same probability of being chosen, and obtain the variable importance for each feature. We then run a different random forest where at each split, besides considering the strict random subset, we include

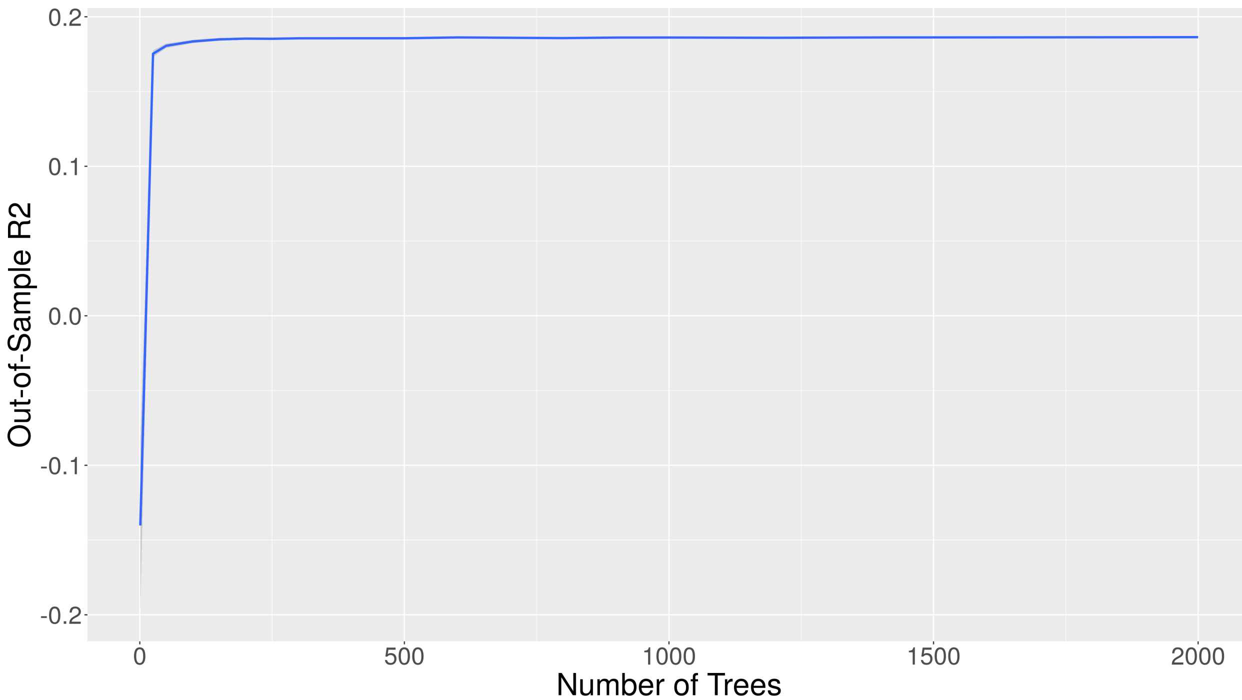
Figure A1: Cross-validation results of the number of trees in the two-quarters-ahead forecast



Notes: This figure plots the relationship between the number of decision trees used in the random forest for training up to 1986 January and the out-of-sample R^2 for the two-quarters-ahead earnings forecasts 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 we take the average of 100 runs to measure the out-of-sample R^2.

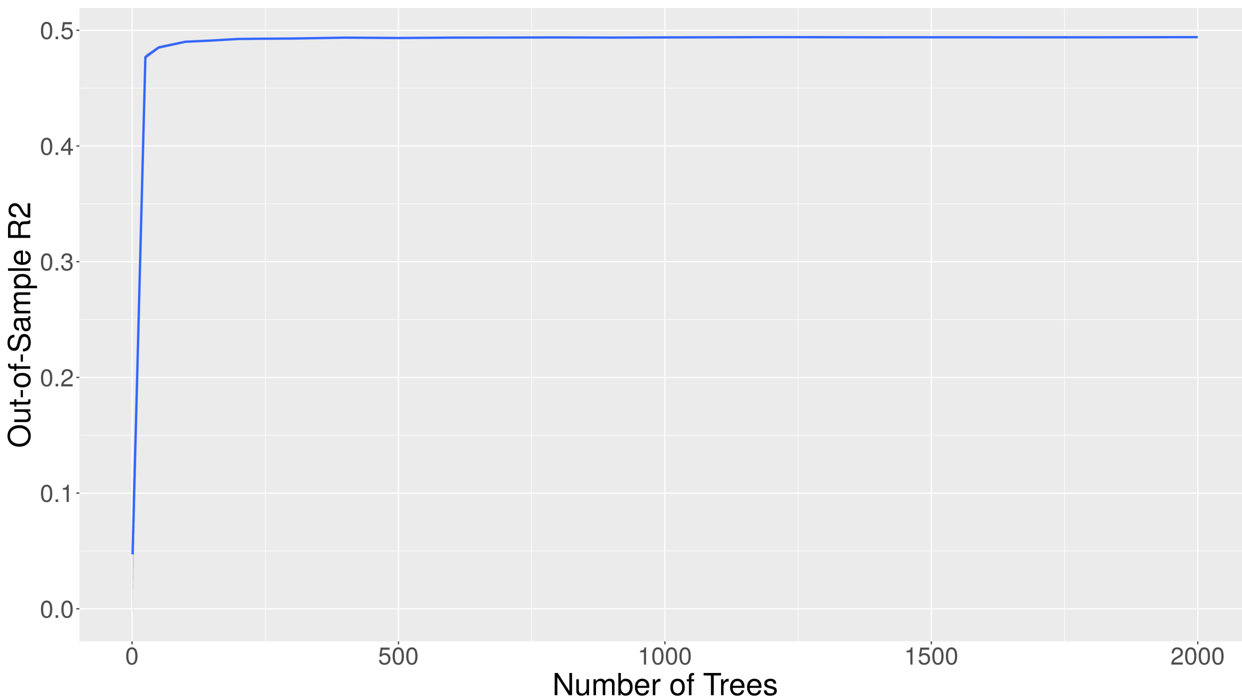
the top n features from the first step up until that point for consideration at each split. This method gives the algorithm the option, but not the obligation, of considering the best predictors from the first stage at each step. We find that adding this step increases the accuracy of the algorithm significantly. We choose $n = 5$ based on cross-validation.

Figure A2: Cross-validation results of the number of trees in the three-quarters-ahead forecast



Notes: This figure plots the relationship between the number of decision trees used in the random forest for training up to 1986 January and the out-of-sample R^2 for the three-quarters-ahead earnings forecasts 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 we take the average of 100 runs to measure the out-of-sample R^2.

Figure A3: Cross-validation results of the number of trees in the two-years-ahead forecast



Notes: This figure plots the relationship between the number of decision trees used in the random forest for training up to 1986 January and the out-of-sample R^2 for the two-years-ahead earnings forecasts 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 we take the average of 100 runs to measure the out-of-sample R^2.

A2. WRDS financial ratios

In the random forest model, we use financial ratios obtained from the Financial Ratio Suit by Wharton Research Data Service (WRDS) as forecasting variables. According to WRDS, these variables are the most commonly used financial ratios by academic researchers and available at both quarterly and annual frequencies. The variables can be grouped into the following seven categories: Capitalization, Efficiency, Financial Soundness/Solvency, Liquidity, Profitability, Valuation, and others. Table A1 details the definitions of financial ratios.[1] Since our predicted variable is earnings per share, we also consider another twenty-six fundamental values per share derived from these financial ratios, such as book equity per share and current debt per share, to improve the forecasts.

We exclude PEG_1yrforward, PEG_ltgforward, pe_op_basic, and pe_op_dil from our forecast model, because these variables have too many missing observations. We replace the missing values of other variables as the industry medians. The industries are defined as in Fama-French 49 industry portfolios.

Table A1: WRDS financial ratios

Variable	Definition	Variable	Definition
accrual	Accruals/Average Assets	invt_act	Inventory/Current Assets
adv_sale	Avertising Expenses/Sales	lt_debt	Long-term Debt/Total Liabilities
aftret_eq	After-tax Return on Average Common Equity	lt_ppent	Total Liabilities/Total Tangible Assets
aftret_equity	After-tax Return on Total Stockholders Equity	npm	Net Profit Margin
aftret_invcapx	After-tax Return on Invested Capital	ocf_lct	Operating CF/Current Liabilities
at_turn	Asset Turnover	opmad	Operating Profit Margin After Depreciation
bm	Book/Market	opmbd	Operating Profit Margin Before Depreciation
capei	Shillers Cyclically Adjusted P/E Ratio	pay_turn	Payables Turnover
capital_ratio	Capitalization Ratio	pcf	Price/Cash flow
cash_conversion	Cash Conversion Cycle (Days)	pe_exi	P/E (Diluted, Excl. EI)
cash_debt	Cash Flow/Total Debt	pe_inc	P/E (Diluted, Incl. EI)
cash_lt	Cash Balance/Total Liabilities	pe_op_basic	Price/Operating Earnings (Basic, Excl. EI)
cash_ratio	Cash Ratio	pe_op_dil	Price/Operating Earnings (Diluted, Excl. EI)
cfm	Cash Flow Margin	PEG_1yrforward	Forward P/E to 1-year Growth (PEG) ratio
curr_debt	Current Liabilities/Total Liabilities	PEG_ltgforward	Forward P/E to Long-term Growth (PEG) ratio
curr_ratio	Current Ratio	PEG_trailing	Trailing P/E to Growth (PEG) ratio
de_ratio	Total Debt/Equity	pretret_earnat	Pre-tax Return on Total Earning Assets
debt_assets	Total Debt/Total Assets (1)	pretret_noa	Pre-tax return on Net Operating Assets
debt_at	Total Debt/Total Assets (2)	profit_lct	Profit Before Depreciation/Current Liabilities
debt_capital	Total Debt/Capital	ps	Price/Sales
debt_ebitda	Total Debt/EBITDA	ptb	Price/Book
debt_invcap	Long-term Debt/Invested Capital	ptpm	Pre-tax Profit Margin
divyield	Dividend Yield	quick_ratio	Quick Ratio (Acid Test)
dltt_be	Long-term Debt/Book Equity	RD_SALE	Research and Development/Sales
dpr	Dividend Payout Ratio	rect_act	Receivables/Current Assets
efftax	Effective Tax Rate	rect_turn	Receivables Turnover
equity_invcap	Common Equity/Invested Capital	roa	Return on Assets
evm	Enterprise Value Multiple	roce	Return on Capital Employed
fcf_ocf	Free Cash Flow/Operating Cash Flow	roe	Return on Equity
gpm	Gross Profit Margin	sale_equity	Sales/Stockholders Equity
GProf	Gross Profit/Total Assets	sale_invcap	Sales/Invested Capital
int_debt	Interest/Average Long-term Debt	sale_nwc	Sales/Working Capital

[1]The formulas to calculate these financial ratios are available at the WRDS website.

continued from previous page

int_totdebt	Interest/Average Total Debt	short_debt	Short-Term Debt/Total Debt
intcov	After-tax Interest Coverage	staff_sale	Labor Expenses/Sales
intcov_ratio	Interest Coverage Ratio	totdebt_invcap	Total Debt/Invested Capital
inv_turn	Inventory Turnover		

A3. Sample selection and machine learning forecasts

In this section, we detail the sample selection and the procedures of machine learning earnings forecasts. Our first step is to obtain actual realized earnings and analysts' earnings forecasts from the I/B/E/S database.[2] We keep firms that have both realized earnings and analysts' forecasts. We focus on 1-year- and 2-year-ahead forecasts for annual earnings (IBES FPI of 1 and 2), and one-quarter-, two-quarters-, and three-quarters-ahead forecasts for quarterly earnings (IBES FPI of 6, 7, and 8), because analysts' forecasts for other horizons have significantly fewer observations.

We then match the IBES actual file (actual realized earnings) with the summary file (analysts' consensus forecasts) using Ticker and fiscal end date.[3] As pointed out by Diether et al. (2002) and Bouchaud et al. (2019), mistakes occur when matching the I/B/E/S actual file with the I/B/E/S summary file because stock splits may occur between the earnings forecast day and the actual earnings announcement day. However, the I/B/E/S adjusted summary files round the forecast and actual earnings to the nearest penny for adjusting the splits. To circumvent these rounding errors, we obtain data from unadjusted actual and summary files. We use the cumulative adjustment factors (CFACSHR) from the CRSP monthly stock file to adjust the forecast and the actual EPS on the same share basis. For example, if forecasts are made at $t-1$ and the actual earnings are announced at t, we measure the adjusted actual earning as,

$$AdjustActual_t = Actual_t * CFACSHR_{t-1}/CFACSHR_t.$$

For matching /I/B/E/S with CRSP, we use the link table provided by the Wharton Research Data Service. We require firms' historical CUSIP to be same in both /I/B/E/S

[2]We do not get the actual earnings from Compustat, because I/B/E/S uses different accounting basis from Compustat to measure actual earnings. Since our primary goal is to construct a statistically optimal and unbiased benchmark for analysts' earnings forecasts, we obtain the realized earnings from the /I/B/E/S database.

[3]PENDS indicates the fiscal end date in the actual file and FPEDATS indicates the fiscal end date in the summary file.

and CRSP. We keep common stocks (share code 10 and 11) in stock exchanges of NYSE, AMEX, and NASDAQ (exchange code 1, 2, and 3).[4].

Our sample is in monthly frequency because analysts make earnings forecasts for firms' earnings every month (I/B/E/S estimate date is STATPERS). We therefore provide our statistically optimal forecast for every I/B/E/S estimate date (STATPERS). Specifically, we assume that we are making forecasts on the same date as when analysts make forecasts. We train the random forest model using the information available at the current time and then forecast earnings for the same fiscal end periods as analysts do. When matching the forecast variables such as firm characteristics and macroeconomic variables, we require announcement dates of these variables are before STATPERS. The forecasts are therefore out-of-sample and are not based on any future information. The resulting forecasting regression is:

$$E_t[eps_{i,t+\tau}] = RF[Fundamentals_{i,t}, Macro_t, AF_{i,t}],$$

where RF denotes the random forest model using data from the most recent periods. $Fundamentals_{i,t}$, $Macro_t$, and $AF_{i,t}$ denote firm i's fundamental variables, macroeconomic variables, and analysts' earnings forecasts respectively. The earnings per share of firm i in quarter t+ τ (τ=1 to 3) or year t+τ (τ=1 to 2) is $eps_{i,t+\tau}$.

For the quarterly earnings forecasts and 1-year ahead forecast, we train the random forest model using the data from the most recent year. We then forecast earnings in the following periods using information available at the current time. For the 2-year ahead forecasts, we train the model using the data from the two most recent years rather than from the most recent year because we do not have enough observations when using a 12-month window to train the model. Our forecasts remain consistent when using different windows to train the model. Our training data starts in January 1985, and our first forecast observations are in January 1986.

[4]We do not delete the smallest firms because smallest firms are not covered in /I/B/E/S, and the intersection of /I/B/E/S and CRSP heavily tilts towards stocks with large market-cap (Diether et al. (2002)).

A4. Summary statistics of variables in Fama–MacBeth return regressions

Table A2 reports the summary statistics of conditional biases in analysts' earnings forecasts. "Average BE" is defined as the average of these conditional biases at multiple horizons. "BE score" is defined as the arithmetic average of the percentile rankings on each of the five conditional biases at different forecast horizons. We also report the summary statistics of control variables including the log of firm size (Lnsize), the log of book-to-market ratio (LNbeme), short-term reversal (Ret_1), medium-term momentum (Ret12_7), investment-to-asset (IA),idiosyncratic volatility (IVOL), return volatility (RetVol), and share turnover (Turnover).[5]

Table A2: Summary statistics

Variable	N	Mean	Std	P1	Q1	Median	Q3	Q99
Average BE	1268964	0.0167	0.0983	-0.0268	0.0007	0.0042	0.0137	0.2274
BE Score	1268964	51.1431	23.2862	7.0000	33.2000	48.2000	68.6667	98.6667
LNsize	1268633	13.1098	1.8849	9.3120	11.7508	12.9775	14.3328	17.9642
LNbeme	1153148	-0.7601	0.8561	-3.2706	-1.2255	-0.6680	-0.2000	1.0404
Ret12_7	1207915	0.0817	0.4535	-0.6842	-0.1329	0.0420	0.2228	1.5358
Ret1	1268389	0.0099	0.1574	-0.3819	-0.0625	0.0049	0.0731	0.4828
IA	1174640	0.3021	1.0069	-0.4184	0.0015	0.0893	0.2538	4.3283
IVOL	1268571	0.0247	0.0197	0.0049	0.0125	0.0195	0.0308	0.0954
RetVol	1268016	0.0297	0.0220	0.0067	0.0160	0.0240	0.0366	0.1094
Turnover	1266778	1.5416	12.7502	0.0748	0.4874	0.9836	1.8738	8.3893

A5. Fama–MacBeth regressions with conditional bias in each forecast horizon

Table A3 reports the Fama–MacBeth of monthly stock returns on conditional earnings forecast bias at each forecast horizon, including one-quarter-, two-quarters-, three-quarters-, 1-year-, and 2-year-ahead. Columns (1) and (2) in each panel report the regression results with and without control variables, respectively. We find that the two-quarters-, three-quarters-, and 2-year-ahead forecast bias negatively predict stock returns; the predictability remains robust after controlling for other return predictors.

Table A4 reports returns on the value-weighted portfolios sorted on conditional earnings forecast bias at each forecast horizon. We find consistent evidence that stocks with more

[5] We winsorize investment-to-asset at 1% to exclude outliers.

optimistic biases earn lower future returns. Table A5 shows that the return-predictability results from the cross-sectional regressions and portfolio sorts also hold in the time-series regressions against factor models such as the CAPM and the Fama-French five-factors model.

Table A3: Fama–Macbeth regressions

Notes: This table reports the Fama–MacBeth cross-sectional regressions of monthly stocks' excess returns on the conditional earnings forecast bias in each forecast horizon: one-quarter-, two-quarters-, three-quarters-, 1-year-, and 2-years-ahead. (1) and (2) report the regression results with and without control variables, respectively. The t-statistics are reported in parentheses. The sample period is 1986 to 2019.

	A: One-quarter		B: Two-quarters		C: Three-quarters		D: One-year		E: Two-years	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Bias	-0.029	0.045	-0.184	-0.219	-0.237	-0.278	-0.012	0.005	-0.030	-0.041
t-stat	-0.59	1.11	-3.42	-4.58	-5.12	-6.76	-0.83	0.40	-5.39	-7.80
LNsize		-0.041		-0.080		-0.093		-0.065		-0.118
t-stat		-1.15		-2.25		-2.62		-1.83		-3.29
LNbeme		0.090		0.130		0.130		0.060		0.099
t-stat		1.53		2.24		2.24		1.05		1.77
Ret1		-3.270		-2.559		-2.630		-2.976		-3.012
t-stat		-7.35		-6.12		-6.08		-7.08		-6.98
Ret12_7		0.509		0.357		0.329		0.476		0.338
t-stat		3.31		2.30		2.06		3.08		2.14
IA		-0.003		-0.003		-0.003		-0.003		-0.003
t-stat		-5.24		-5.03		-4.62		-5.46		-4.86
IVOL		-0.298		-0.266		-0.274		-0.197		-0.239
t-stat		-2.61		-2.30		-2.33		-1.81		-2.14
Retvol		0.184		0.186		0.242		0.097		0.214
t-stat		1.54		1.55		1.97		0.84		1.80
Turnover		-0.086		-0.065		-0.087		-0.063		-0.060
t-stat		-1.95		-1.46		-1.98		-1.39		-1.30
Intercept	0.955	1.806	1.048	2.348	1.114	2.483	0.981	2.102	1.166	2.799
t-stat	3.34	3.34	3.70	4.38	3.91	4.64	3.52	4.00	4.23	5.27
R^2 (%)	0.818	5.766	0.846	6.027	0.794	6.193	0.792	5.693	0.841	6.106

Table A4: Portfolios sorted on conditional bias

This table reports the time-series average of returns (in percent) on value-weighted portfolios sorted on the conditional earnings forecast bias at each forecast horizon. Panel A looks at the one-quarter-ahead conditional bias. Panel B looks at the two-quarters-ahead bias. Panel C looks at the three-quarters-ahead bias. Panel D looks at the 1-year-ahead bias. Panel E looks at the 2-year-ahead bias. The sample period is 1986 to 2019.

Quintile	1	2	3	4	5	5-1
Panel A: One-quarter-ahead BE						
Mean	1.04	0.88	0.92	0.97	0.83	-0.21
t-stat	4.51	4.01	3.81	3.52	2.14	-0.79
CAPM Beta	1.00	0.99	1.06	1.14	1.44	0.44
Panel B: Two-quarters-ahead BE						
Mean	1.14	0.91	0.92	0.61	0.21	-0.93
t-stat	5.29	4.26	3.92	2.19	0.52	-3.24
CAPM Beta	0.96	0.96	1.03	1.18	1.49	0.53
Panel C: Three-quarters-ahead BE						
Mean	1.33	1.03	0.80	0.50	-0.03	-1.36
t-stat	6.15	4.86	3.45	1.77	-0.09	-5.21
CAPM Beta	0.95	0.94	1.02	1.20	1.45	0.50
Panel D: One-year-ahead BE						
Mean	0.97	0.91	0.97	0.97	1.01	0.04
t-stat	4.61	4.18	3.96	3.47	2.73	0.14
CAPM Beta	0.94	0.98	1.08	1.18	1.40	0.46
Panel E: Two-years-ahead BE						
Mean	1.39	1.01	0.85	0.64	-0.32	-1.71
t-stat	6.65	4.85	3.59	2.28	-0.88	-6.65
CAPM Beta	0.91	0.93	1.06	1.19	1.42	0.51

Table A5: Time series tests with common asset-pricing models

This table reports the regression of stock returns (in percent) on the long-short portfolio sorted with the conditional earnings forecast bias at different horizons, in the CAPM, the Fama-French three-factors model (FF3), and the Fama-French five-factors model (FF5). Panel A looks at the one-quarter-ahead conditional bias. Panel B looks at the two-quarters-ahead bias. Panel C looks at the three-quarters-ahead bias. Panel D looks at the 1-year-ahead bias. Panel E looks at the 2-year-ahead bias. The sample period is 1986 to 2019. The t-statistics are adjusted by the heteroscedasticity robust standard errors (White (1980)).

	Panel A: CAPM		Panel B: FF3		Panel C: FF5	
	Coeffi	t-stat	*Coeffi*	t-stat	*Coeffi*	t-stat
			Panel A: One-quarter-ahead BE			
Intercept	-0.52	-2.03	-0.61	-2.74	-0.11	-0.50
Mkt_RF	0.44	5.75	0.40	5.70	0.22	3.55
SMB			0.78	8.56	0.57	5.61
HML			0.51	4.09	0.95	7.10
RMW					-0.78	-5.43
CMA					-0.66	-2.89
			Panel B: Two-quarters-ahead BE			
Intercept	-1.30	-4.92	-1.43	-6.08	-1.00	-4.02
Mkt_RF	0.53	6.38	0.51	6.70	0.36	4.94
SMB			0.76	7.48	0.56	4.86
HML			0.64	4.89	0.99	6.98
RMW					-0.74	-4.35
CMA					-0.46	-1.78
			Panel C: Three-quarters-ahead BE			
Intercept	-1.71	-7.18	-1.78	-8.34	-1.36	-5.98
Mkt_RF	0.50	7.26	0.46	7.21	0.32	5.06
SMB			0.67	6.78	0.46	4.23
HML			0.42	3.50	0.76	6.53
RMW					-0.74	-4.70
CMA					-0.43	-2.09
			Panel D: One-year-ahead BE			
Intercept	-0.28	-1.23	-0.35	-1.64	0.05	0.20
Mkt_RF	0.46	6.80	0.41	6.32	0.27	3.92
SMB			0.70	6.84	0.52	4.79
HML			0.38	3.35	0.72	5.24
RMW					-0.64	-3.97
CMA					-0.49	-1.83
			Panel E: Two-years-ahead BE			
Intercept	-2.06	-8.85	-2.17	-10.29	-1.86	-7.93
Mkt_RF	0.51	7.57	0.49	8.07	0.39	6.00
SMB			0.60	5.62	0.47	4.19
HML			0.51	4.71	0.78	5.99
RMW					-0.49	-3.36
CMA					-0.40	-1.65

A6. Cross-sectional return predictability: other robustness checks

In this section, we check the robustness of Fama–MacBeth regression results in Table 3 by omitting stocks whose prices are lower than \$5 and also by scaling the earnings forecast conditional biases with total asset (per share) from the last fiscal year. Total assets are obtained from Compustat (item AT). Table A6 and A7 report the two robustness checks results, respectively. Overall, we find robust return predictability of conditional biases.

Table A6: Fama–Macbeth regressions: omitting stocks with prices lower than \$5

Notes: This table reports the Fama–MacBeth regressions of monthly stocks' returns on the conditional earnings forecast bias. "Average BE" denotes the average of the conditional biases at different forecast horizons including one-quarter-, two-quarters-, three-quarters-, 1-year-, and 2-year-ahead. "BE score" denotes the arithmetic average of the percentile rankings on each of the five conditional biases at different forecast horizons. Columns (1) and (2) of each panel report the regression results with and without control variables, respectively. The t-statistics are reported in parentheses. The sample period is 1986 to 2019. We omit stocks whose closing prices in the previous month are smaller than \$5.

	Panel A: Average BE		Panel B: BE Score	
	(1)	(2)	(1)	(2)
Bias	-0.383	-0.451	-0.027	-0.033
t-stat	-10.94	-14.46	-8.86	-13.57
LNsize		-0.112		-0.180
t-stat		-3.67		-6.00
LNbeme		0.153		0.193
t-stat		2.85		3.60
Ret1		-2.087		-2.163
t-stat		-5.29		-5.51
Ret12_7		0.416		0.328
t-stat		2.89		2.38
IA		-0.002		-0.002
t-stat		-4.35		-4.42
IVOL		-0.254		-0.228
t-stat		-2.32		-2.09
Retvol		0.159		0.153
t-stat		1.33		1.30
Turnover		-0.042		-0.025
t-stat		-0.99		-0.61
Intercept	1.197	3.004	2.208	5.099
t-stat	4.69	6.53	9.87	11.52
R^2 (%)	0.794	6.248	1.151	6.320

Table A7: Fama–Macbeth regressions: scaling conditional biases by total assets per share

Notes: This table reports the Fama-MacBeth regressions of monthly stocks' returns on the conditional bias, which is defined as the difference between analysts' earnings forecasts and machine learning forecasts, scaled by the total asset (per share) from the most recent fiscal period. "Average BE" denotes the average of the conditional biases at different forecast horizons including one-quarter-, two-quarters-, three-quarters-, 1-year-, and 2-year-ahead. "BE Score" denotes the arithmetic average of the percentile rankings on each of the five conditional biases at different forecast horizons. Columns (1) and (2) of each panel report the regression results with and without control variables, respectively. The t-statistics are reported in parentheses. The sample period is 1986 to 2019.

	Panel A: Average BE		Panel B: BE Score	
	(1)	(2)	(1)	(2)
Bias	-0.062	-0.079	-0.019	-0.025
t-stat	-4.15	-8.24	-4.36	-10.66
LNsize		-0.093		-0.178
t-stat		-2.63		-5.32
LNbeme		0.009		-0.050
t-stat		0.15		-0.90
Ret1		-2.860		-2.964
t-stat		-6.82		-7.13
Ret12_7		0.509		0.412
t-stat		3.25		2.68
IA		-0.003		-0.003
t-stat		-5.52		-5.79
IVOL		-0.219		-0.209
t-stat		-1.99		-1.91
Retvol		0.136		0.170
t-stat		1.18		1.49
Turnover		-0.056		-0.035
t-stat		-1.25		-0.79
Intercept	1.062	2.455	1.930	4.565
t-stat	3.85	4.67	8.73	9.29
R^2 (%)	0.539	5.448	1.473	5.697

A7. Downward revisions in analysts' earnings forecasts

To explain the intuition regarding return predictability, we note analysts revise their earnings forecasts every month. As the announcement dates approach, analysts should process new information and update their estimates to make better forecasts. Table A8 demonstrates that analysts revise their earnings forecasts.

We find that the average forecast error, defined as the difference between analysts' earnings forecasts per share and the realized earnings per share, is consistently positive for all horizons; the results suggest that analysts make over-optimistic forecasts. Further, the average error decreases as the earnings announcement dates approach; i.e., on average, a downward revision occurs in analysts' forecasts. As expected, the mean squared error also decreases. Analysts make more precise forecasts when the earning announcement dates approach.

For the 1-year-ahead forecast, the average forecast error decreases from 0.215, when analysts make the first forecast, to 0.081, when analysts make their last forecast for that fiscal year. This last forecast is usually made about one month after the fiscal year has ended, though it precedes the earnings announcement date for that fiscal year. The mean squared error declines from 1.197 (for the one-year-head forecast) to 0.365 for the last forecast the analysts make.

A downward revision also occurs in the one-quarter-ahead, the two-quarters-ahead, the three-quarters-ahead, and the 2-year-ahead forecasts. To the extent that investors follow analysts' forecasts and analysts make optimistic expectations, these downward updates may result in negative cross-sectional return predictability. Specifically, stocks with more optimistic expectations should earn lower subsequent returns than stocks with less optimistic expectations.

The realized values of earnings are not available when making the forecasts; therefore, the ex-post establishment of biases and their importance is not conducive to forming portfolios in real time. We cannot know which stocks have biased expectations when using the realized value as a benchmark until that realized value is revealed. In contrast, our statistically optimal benchmark allows us to study the effects of the bias before the earnings realization.

Table A8: Updates in analysts' beliefs

Notes: This table presents the time-series average of aggregate analysts' forecast errors, defined as the differences between analysts' earnings forecasts and the realized actual earnings. $Month-ahead$ denotes the number of months from the time when analysts make forecasts until the fiscal year/quarter end. N denotes the number of observations. FE and sqr_FE denote the average forecast error and the average square of the error respectively. The sample period is 1986 to 2019.

Panel A: One-quarter-ahead												
Month-ahead	1	0	-1									
N	356770	406433	185165									
FE	0.027	0.021	0.034									
Sqr_FE	0.078	0.070	0.085									
Panel B: Two-quarters-ahead												
Month-ahead	4	3	2									
N	322922	363517	323760									
FE	0.055	0.050	0.050									
Sqr_FE	0.107	0.100	0.097									
Panel C: Three-quarters-ahead												
Month-ahead	7	6	5									
N	290296	330900	296523									
FE	0.071	0.069	0.069									
Sqr_FE	0.133	0.126	0.121									
Panel D: One-year-ahead												
Month-ahead	10	9	8	7	6	5	4	3	2	1	0	-1
N	62905	91633	104809	106804	108226	109359	110562	111547	112710	113903	115354	80921
FE	0.215	0.223	0.218	0.207	0.192	0.171	0.151	0.133	0.111	0.089	0.073	0.081
Sqr_FE	1.197	1.075	1.017	0.888	0.806	0.730	0.603	0.545	0.490	0.388	0.341	0.365
Panel E: Two-years ahead												
Month-ahead	22	21	20	19	18	17	16	15	14	13	12	11
N	46977	70013	81557	85784	88049	89684	91417	92341	93389	94362	95217	92663
FE	0.434	0.442	0.427	0.425	0.414	0.397	0.382	0.370	0.348	0.316	0.302	0.305
Sqr_FE	2.743	2.475	2.310	2.154	2.066	1.964	1.826	1.754	1.651	1.491	1.423	1.381

A8. Cross-sectional return predictability: realized biases

As a placebo tests, we use the realized earnings forecasts biases, defined as the difference between analysts' forecasts and realized earnings scaled by share prices from the most recent month, to "predict" stock returns, though realized earnings are not available at real time. Table A9 reports the regressions with realized biases, and Table A10 and Table A11 report the mean return and alpha on the long-short portfolio strategy based on the realized bias. Overall, we find consistent results that stocks with more optimistic forecast biases earn lower future returns.

A9. Earnings forecasts errors of linear model

Table A12 reports the earnings forecasts via linear regressions. In sharp contrast to the random forest forecasts, linear forecasts have larger forecast errors, measured as the mean of squared difference between linear forecasts and realized earnings, than analysts' forecasts.

Table A9: Fama–Macbeth regressions: realized forecast bias

Notes: This table reports the unfeasible Fama–MacBeth regressions of monthly stocks' returns on the realized earnings forecast bias. We define the realized bias as the difference between analysts' earnings forecasts and actual realized values, scaled by stock prices from the most recent month. "Average BE" denotes the average of the realized biases at different forecast horizons including one-quarter-, two-quarters-, three-quarters-, 1-year-, and 2-year-ahead. "BE score" denotes the arithmetic average of the percentile rankings on each of the five realized biases at different forecast horizons. Columns (1) and (2) of each panel report the regression results with and without control variables, respectively. The t-statistics are reported in parentheses. The sample period is 1986 to 2019. It is important to remark that the realized bias are not available in real time, so this table is only presented for bench-marking purposes.

	Panel A: Average BE		Panel B: BE Score	
	(1)	(2)	(1)	(2)
Bias	-0.108	-0.132	-0.098	-0.110
t-stat	-14.92	-17.32	-38.37	-46.62
LNsize		-0.109		-0.264
t-stat		-2.99		-7.15
LNbeme		0.162		0.107
t-stat		2.80		1.89
Ret1		-3.215		-5.627
t-stat		-7.72		-13.11
Ret12_7		0.289		-0.196
t-stat		1.87		-1.31
IA		-0.003		-0.002
t-stat		-5.33		-3.47
IVOL		-0.177		-0.138
t-stat		-1.61		-1.26
Retvol		0.156		0.133
t-stat		1.35		1.16
Turnover		-0.056		-0.001
t-stat		-1.24		-0.03
Intercept	1.137	2.705	5.777	9.747
t-stat	3.98	5.05	20.73	17.48
R^2 (%)	0.988	6.133	3.368	8.770

Table A10: Portfolios sorted on realized bias

This table reports the time-series average of returns (in percent) on value-weighted portfolios formed on the realized earnings forecast bias. We define the realized bias as the difference between analysts' earnings forecasts and actual realized values, scaled by stock prices from the most recent month. Panel A looks at average conditional bias at different forecast horizons including one-quarter-, two-quarters-, three-quarters-, 1-year-, and 2-year-ahead. Panel B presents the sorts based on "BE score", defined as the arithmetic average of the percentile rankings on each of the five reazlied biases at different forecast horizons. The sample period is 1986 to 2019.

Quintile	1	2	3	4	5	1-5
			Panel A: Average BE			
Mean	3.21	1.59	0.24	-0.69	-1.73	-4.94
t-stat	13.01	7.53	1.13	-2.58	-5.04	-23.06
CAPM Beta	1.03	0.94	0.96	1.15	1.35	0.32
			Panel B: BE Score			
Mean	3.18	1.69	0.41	-0.73	-2.26	-5.44
t-stat	13.06	7.94	1.93	-2.94	-7.01	-26.36
CAPM Beta	1.03	0.94	0.95	1.08	1.29	0.26

Table A11: Time-series tests of long-short portfolios sorted on realized bias

This table reports the regressions of stock returns (in percent) on the long-short portfolio sorted with the realized earnings forecast bias, in the CAPM, the Fama-French three-factors model (FF3), and the Fama-French five-factors model (FF5). We define the realized bias as the difference between analysts' earnings forecasts and actual realized values, scaled by stock prices from the most recent month. Panel A looks at average conditional bias at different forecast horizons including one-quarter, two-quarters, three-quarters, 1-year, and 2-year-ahead. Panel B presents the sorts based on "BE score", defined as the arithmetic average of the percentile rankings on each of the five conditional biases at different forecast horizons. The sample period is 1986 to 2019. The t-statistics are adjusted by the heteroscedasticity robust standard errors (White (1980)).

	CAPM		FF3		FF5	
	$Coeffi$	t-stat	$Coeffi$	t-stat	$Coeffi$	t-stat
			Panel A: Average BE			
Intercept	-5.17	-24.66	-5.21	-25.40	-4.88	-21.93
Mkt_RF	0.32	5.82	0.30	5.23	0.18	3.13
SMB			0.40	4.85	0.27	3.16
HML			0.23	2.13	0.53	4.68
RMW					-0.47	-3.86
CMA					-0.49	-2.27
			Panel B: BE Score			
Intercept	-5.62	-27.44	-5.64	-28.41	-5.42	-25.18
Mkt_RF	0.26	4.58	0.22	3.78	0.14	2.43
SMB			0.44	5.32	0.36	3.90
HML			0.16	1.49	0.36	3.23
RMW					-0.30	-2.76
CMA					-0.32	-1.69

Table A12: Earnings forecasts via linear regression

Notes: This table presents the time series average of OLS regression-based earnings per share forecasts (LF), analysts' earning forecasts (AF), and actual realized earnings (AE) —the difference as well as the squared difference between them. N denotes the number of the sample stocks. We report the Newey-West t-statistics of the 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 1986 January to 2019 December.

	LF	AF	AE	(LF-AE)	(AF-AE)	$(LF-AE)^2$	$(AF-AE)^2$	N
One-quarter-ahead	0.290	0.319	0.291	-0.001	0.028	0.081	0.081	1,022,661
t-stat				-0.26	6.59			
Two-quarters-ahead	0.316	0.376	0.323	-0.008	0.053	0.124	0.102	1,110,689
t-stat				-0.71	10.31			
One-year-ahead	1.150	1.320	1.167	-0.017	0.154	0.757	0.686	1,260,060
t-stat				-0.90	6.24			
Two-years-ahead	1.293	1.771	1.387	-0.094	0.384	2.573	2.009	1,097,098
t-stat				-1.22	8.33			

A10. Out-of-sample return predictability

In this section, we evaluate the return predictability resulting from analysts' earnings forecast biases when using OLS-regression-based earnings forecasts (as opposed to our ML forecasts) as a benchmark. The results are reported in Table A13.

We first document that when using the linear forecast bias from So (2013), the difference in mean returns of long-short portfolios between the pre-2000 period and the post-2000 period is large and significant with a value of -0.78%, consistent with the fact that the predictability from linear forecasts has largely disappeared. Moreover, we find that the average return of the long-short portfolio sorted on the forecast biased as measured in So (2013) equals 0.33% in the post-2000 period, which is not statistically significant. We also test the return predictability of predicted forecast errors as Frankel and Lee (1998) and find similar results.[6]

In contrast, when using the machine learning forecasts, the difference in mean returns between the pre- and post- 2000 periods is a mere -0.37%, which is statistically insignificant. In the post-2000 period the ML long-short portfolio earns a statistically significant mean return equal to -1.30%.

[6]We follow Frankel and Lee (1998) to predict forecast errors of analysts using the same variables including sale growth, book-to-price ratio, long-term earnings growth forecasts, and an optimism measure. Analysts' forecast errors are defined as the difference between analysts' earnings forecasts and realized earnings scaled by the closing stock price from most recent month. We apply these predicted errors to forecast monthly stock returns.

Table A13: Portfolios sorted on conditional bias: linear forecasts versus random forest versus predicted forecast error

Notes: This table reports the time-series average of returns (in percent) on portfolios sorted by various measures of the earnings forecast bias. Panel A uses random forest earnings predictions, whereas panel B uses linear earnings forecasts as the benchmark. Panel C reports the average returns on portfolios formed on the linear forecast bias measured in So (2013). Panel D reports the returns on portfolios sorted on predicted forecast errors as in Frankel and Lee (1998). The last row of each panel reports the p-value of testing (using an F-test) for the difference in mean returns between the two subperiods (pre and post-2000). The sample period is 1986 to 2019.

Quintile	1	2	3	4	5	5-1
Panel A: Random Forest						
Full Sample	1.32	0.98	0.79	0.47	-0.14	-1.46
t-stat	6.53	4.53	3.18	1.62	-0.35	-5.11
Pre-2000	1.42	1.35	1.17	0.51	-0.24	-1.67
t-stat	4.42	3.77	3.06	1.22	-0.49	-4.52
Post-2000	1.25	0.68	0.49	0.44	-0.05	-1.30
t-stat	4.81	2.61	1.51	1.09	-0.09	-3.09
Pre-Post	0.18	0.67	0.69	0.08	-0.19	-0.37
p-value	0.67	0.13	0.17	0.90	0.80	0.51
Panel B: Linear Forecast						
Full Sample	1.25	0.96	0.81	0.43	-0.01	-1.26
t-stat	5.86	4.43	3.34	1.39	-0.03	-4.05
Pre-2000	1.52	1.30	1.01	0.54	-0.28	-1.80
t-stat	4.35	3.86	2.82	1.27	-0.54	-4.84
Post-2000	1.04	0.69	0.66	0.35	0.20	-0.84
t-stat	3.93	2.46	1.99	0.78	0.32	-1.78
Pre-Post	0.49	0.61	0.35	0.20	-0.47	-0.96
p-value	0.27	0.17	0.47	0.75	0.56	0.11
Panel C: Linear Forecast of So (2013)						
Full Sample	1.63	1.23	0.99	0.79	0.95	-0.67
t-stat	4.74	5.49	3.97	2.65	2.67	-4.05
Pre-2000	1.89	1.30	1.10	0.74	0.78	-1.11
t-stat	3.98	3.90	3.00	1.77	1.52	-3.79
Post-2000	1.42	1.17	0.91	0.83	1.09	-0.33
t-stat	2.92	3.88	2.66	1.98	2.20	-1.79
Pre-Post	0.46	0.13	0.19	-0.09	-0.32	-0.78
p-value	0.50	0.77	0.70	0.89	0.66	0.03
Panel D: Predicted Forecast Error						
Full Sample	1.08	0.96	0.92	1.01	0.84	-0.24
t-stat	5.05	4.61	4.34	4.12	2.59	-1.21
Pre-2000	1.45	1.26	1.20	1.31	1.21	-0.24
t-stat	4.14	3.95	3.76	3.64	2.95	-0.97
Post-2000	0.78	0.72	0.69	0.77	0.54	-0.24
t-stat	2.99	2.64	2.47	2.31	1.14	-0.80
Pre-Post	0.67	0.54	0.51	0.55	0.66	0.00
p-value	0.13	0.20	0.23	0.27	0.29	0.99

Table A14 reports the return predictability resulting from various measures of earnings forecast biases in the pre- and post-2000 periods using Fama–MacBeth cross-sectional regressions. Panel A and B look at the average earnings forecast bias using random forest forecasts and linear forecasts as benchmarks, respectively. We find that the return predictability of random-forest-implied forecast biases remains robust when splitting the sample into the pre- and post-2000 periods. However, the return predictability resulting from linear forecast biases practically disappears after 2000.

Table A14: Return predictability: linear forecasts versus random forest versus predicted forecast error

Notes: This table reports the Fama–MacBeth cross-sectional regressions of monthly stocks' returns on the earnings forecast bias. Panel A and B look at the average of forecast bias at different horizons using random forest forecast and linear forecast as benchmarks, respectively. Panel C looks at the returns on the linear conditional bias computed as in So (2013). Panel D looks at the return predictability of predicted forecast errors as in Frankel and Lee (1998). The first and third rows of each panel report the regression results with and without control variables, respectively. The sample period is 1986 to 2019.

Sample Period	Full Sample	Pre-2000	Post-2000
Panel A: Random Forest			
FM without control	-0.054	-0.075	-0.037
t-stat	-3.94	-3.78	-1.99
FM with control	-0.064	-0.086	-0.047
t-stat	-5.08	-5.09	-2.58
Panel B: Linear Forecast			
FM without control	-0.029	-0.044	-0.017
t-stat	-2.39	-2.39	-1.05
FM with control	-0.021	-0.034	-0.011
t-stat	-2.06	-2.33	-0.78
Panel C: Linear Forecast of So (2013)			
FM without control	−0.015	−0.030	−0.002
t-stat	−3.79	−3.92	−0.64
FM with control	−0.020	−0.036	−0.004
t-stat	−5.45	−5.63	−1.28
Panel D: Predicted Forecast Error			
FM without control	-0.040	-0.110	0.015
t-stat	-0.75	-2.34	0.17
FM with control	-0.058	-080	-0.40
t-stat	-1.49	-2.35	-0.62

To further address the difference between the out-of-sample return predictability resulting from our ML conditional earnings forecast bias measure and the predictability resulting from linear forecast bias measures, we take the difference between our ML long-short portfolio returns and returns on the long-short portfolio formed on the linear bias. We test the average returns and alphas of this return difference during the post-2000 period. We find a significant difference in average returns relative to our ML numbers, taking the following three values: (1) -0.46%, (2) -0.97%, and (3) -1.06%, corresponding to three possible linear forecast benchmarks: (1) a linear forecast using our forecasting variables, (2) the linear predictions from So (2013), and (3) the predicted linear forecast error in Frankel and Lee

(1998). The results for alphas (as opposed to mean return differences) are very similar. As explained by our theoretical model, the significant decay in alpha and return predictability of linear forecasts during the post-2000 period is consistent with a forward-looking bias in forecasting variable selection, as well as a non-linear relationship between earnings and forecasting variables.

Table A15: Return difference post-2000

Notes: This table evaluates the importance of various earnings forecast benchmark methods for generating return spreads in long-short portfolio sorts. First, we compute the average return on the long-short portfolio (in percentage points) that uses our machine-learning-implied earnings forecast bias as the sorting variable. Next, we compute that same average long-short portfolio return, but now using various other methods to construct an earnings forecast benchmark against which the earnings forecast bias (i.e., the sorting variable) is computed. The different forecasting methods include (1) a simple linear forecasting benchmark using the same variables as our ML framework, (2) the earnings forecast bias in So (2013), and (3) predicted forecast errors as in Frankel and Lee (1998). The table reports the difference between our ML average long-short return (alpha) and those resulting from these other benchmarks. The alphas are computed against several standard asset pricing models, including the CAPM and the Fama-French three- and five-factors models. The sample is the post-2000 period. Parentheses report the *t*-statistics adjusted by the heteroscedasticity robust standard errors (White (1980)).

	Linear Forecast	Forecast in So	Predicted Forecast Error
Mean Return	-0.46	-0.97	-1.06
	(-2.13)	(-2.19)	(-4.08)
CAPM Alpha	-0.41	-1.47	-1.16
	(-1.83)	(-4.00)	(-4.67)
FF3 Alpha	-0.39	-1.58	-1.21
	(-1.83)	(-4.70)	(-5.09)
FF5 Alpha	-0.34	-0.74	-1.01
	(-1.56)	(-2.07)	(-3.49)

A11. Anomaly variables

In this study, we follow Hou et al. (2015) as close as possible to define anomaly variables. Table A16 lists the significant anomalies documented in Hou et al. (2015). L-S ret (%) denotes the monthly average return (in percent) of each of the 27 long-short anomaly strategies. The sample period is July 1972 to December 2019, depending on data availability. Following Hou et al. (2015), for each firm characteristic obtained from Compustat database, we use NYSE breakpoints to sort stocks into deciles and assign decile rankings to them.

Table A16: List of significant anomalies

Anomalies	**Descriptions**	**Sample period**	L-S ret (%)
Sue-1	Earnings surprise (1-month holding period), Foster et al. (1984)	01/1974–12/2019	0.42

continued from previous page

Abr-1	Cumulative abnormal stock returns (1-month holding period), Chan et al. (1996)	07/1972−12/2019	0.89
R11-1	Price momentum (11-month prior returns, 1-month holding period), Fama and French (1996)	07/1972−12/2019	1.23
BM	Book-to-market equity, Rosenberg et al. (1985)	07/1972−12/2019	0.46
Dur	Equity duration, Dechow et al. (2004)	07/1972−12/2019	1.27
E/P	Earnings-to-price, Basu (1983)	07/1972−12/2019	0.39
CF/P	Cash flow-to-price, Lakonishok et al. (1994)	07/1972−12/2019	0.33
NO/P	Net payout yield Boudoukh et al. (2007)	07/1972−12/2019	0.56
I/A	Investment-to-assets, Cooper et al. (2008)	07/1972−12/2019	0.45
NOA	Net operating assets, Hirshleifer et al. (2004)	07/1972−12/2019	0.50
ΔPI/A	Changes in property, plant, and equipment plus changes in inventory scaled by assets Lyandres et al. (2007)	07/1972−12/2019	0.41
IG	Investment growth, Xing (2007)	07/1972−12/2019	0.34
CEI	Composite equity issues, Daniel and Titman (2006)	07/1972−12/2019	0.40
NSI	Net stock issues, Pontiff and Woodgate (2008)	07/1972−12/2019	0.59
IvC	Inventory changes, Thomas and Zhang (2002)	07/1972−12/2019	0.45
IvG	Inventory growth, Belo and Lin (2012)	07/1972−12/2019	0.34
OA	Operating accruals, Sloan (1996)	07/1972−12/2019	0.26
POA	Percent operating accruals, Hafzalla et al. (2011)	07/1972−12/2019	0.33
PTA	Percent total accruals, Hafzalla et al. (2011)	07/1972−12/2019	0.30
GP/A	Gross profits-to-assets, Novy-Marx (2013)	07/1972−12/2019	0.21
ROE	Return on equity, Haugen and Baker (1996)	07/1972−12/2019	0.72
ROA	Return on assets, Balakrishnan et al. (2010)	07/1972−12/2019	0.57
NEI	Number of consecutive quarters with earnings increases, Barth et al. (1999)	07/1972−12/2019	0.30
OC/A	Organizational capital-to-assets, Eisfeldt and Papanikolaou (2013)	07/1972−12/2019	0.26
Ad/M	Advertisement expense-to-market, Chan et al. (2001)	07/1972−12/2019	0.46
RD/M	R&D-to-market, Chan et al. (2001)	07/1972−12/2019	0.78
OL	Operating leverage, Novy-Marx (2010)	07/1972−12/2019	0.23

A12. Conditional biases and anomalies

In Table A17 we report estimated alphas relative to the Fama–French five-factors model for the portfolios formed on conditional earnings forecast bias and anomaly score. Consistent with the results in Table 8, we find that the anomaly long-short alpha is more significant across portfolios with a larger conditional earnings forecast bias. More importantly, the anomaly alpha becomes insignificant for portfolios with the smallest earnings forecast bias.

Table A17: Alphas on portfolios formed on conditional bias and anomaly score

Notes: This table reports the alphas (relative to the Fama–French five-factors model) on portfolios formed by sorting independently on the average conditional earnings forecast bias (BE) and the anomaly score. The anomaly score is defined as the equal-weighted average of the decile ranking on each of the 27 anomaly variables. The last two rows report the alphas and t-statistics on the ten decile portfolios formed on the anomaly score.

	Anomaly Decile										
BE Quintile	S	2	3	4	5	6	7	8	9	L	L-S
1	0.49	0.38	0.70	0.58	0.63	0.71	0.74	0.62	0.86	0.75	0.26
t-stat	1.75	2.08	4.40	3.96	4.26	5.90	6.04	5.17	5.79	5.00	0.84
2	-0.19	0.18	0.30	0.29	0.21	0.11	0.23	0.24	0.63	0.52	0.71
t-stat	-0.87	1.10	2.29	2.32	1.69	1.00	2.04	2.23	5.48	4.38	2.64
3	-0.69	-0.30	-0.06	-0.14	0.00	0.29	0.06	0.30	0.32	0.13	0.82
t-stat	-2.97	-1.59	-0.36	-0.90	0.03	1.86	0.41	2.11	2.15	0.92	3.11
4	-1.20	-1.05	-0.31	-0.40	-0.49	-0.31	-0.17	0.01	0.11	-0.12	1.08
t-stat	-4.58	-4.17	-1.64	-1.99	-2.33	-1.42	-0.82	0.07	0.55	-0.67	3.38
5	-1.90	-1.44	-1.11	-1.00	-0.95	-0.31	-0.66	-0.79	-0.05	-0.15	1.75
t-stat	-5.92	-3.90	-3.59	-3.18	-2.45	-0.87	-2.15	-2.78	-0.17	-0.53	5.13
5-1	-2.38	-1.83	-1.80	-1.57	-1.58	-1.03	-1.40	-1.41	-0.91	-0.90	1.48
t-stat	-6.20	-4.64	-5.11	-4.42	-4.07	-2.59	-4.08	-4.14	-2.63	-2.63	3.50
All Stocks	-0.58	-0.20	0.16	0.17	0.14	0.22	0.23	0.31	0.55	0.40	0.98
t-stat	-3.54	-1.79	1.74	2.27	1.58	3.00	3.79	4.22	7.37	4.82	4.81

We also explore our results related to anomalies (Table 8 in the paper) for the post-2000 period. We find that our results cannot be obtained if we use linear earnings forecasts as the benchmark to measure analysts' earnings forecast biases. First, we study the relationship between biases and anomalies.

Recall that we focus on the 27 significant and robust anomalies considered in Hou et al. (2015). We first define an anomaly score. To calculate the score, for each month, we assign decile ranks to each stock based on the 27 anomaly variables. The anomaly score for an individual stock is calculated as the arithmetic average of its ranking on each of the 27 anomalies. On average, stocks with a higher anomaly score have higher returns.

When using our conditional bias measure, we find that stocks with a lower anomaly score have a higher earnings forecast bias than stocks with a higher anomaly score, and the difference in the bias is statistically significant. As such, the short-leg portfolio is comprised of stocks with more over-optimistic expectations, suggestive of overpricing, consistent with the existing literature and basic behavioral intuition.

In contrast, when using the variable-selection-bias-free linear forecasts to measure analysts' biases, we find an insignificant difference in the earnings forecast bias between the stocks with higher anomaly scores and stocks with lower anomaly scores. Furthermore, when using the linear forecast bias measured as in So (2013), we even find the opposite (wrong) sign: stocks with a higher anomaly score (and higher returns) have a larger (linear) earnings forecast bias. As such, our results highlight the challenges of using linear forecast biases to provide a behavioral mispricing interpretation of anomalies.

When sorting portfolios on the conditional earnings forecast bias and the anomaly score independently for the post-2000 period, we find that the average monthly returns of the long-short anomaly portfolio in the quintile group with the smallest conditional earnings forecast bias is insignificant (0.66% with per month with t-statistic of 1.50). One way to interpret this result is that anomaly sorts are not delivering returns (in long-short portfolios) when the conditional bias in earnings forecasts is the lowest. In contrast, the average monthly long-short returns in the quintile group with the largest conditional earnings forecast bias is significant (1.91% with per month with t-statistic of 4.02). The difference in average long-short anomaly returns between these two extreme quintile portfolios is significantly positive. These findings are consistent with our conjecture that biased earnings expectations are an important driver of anomaly returns.

Using any of the linear forecasts as the benchmark, however, we find no strong relationship between conditional earnings forecast biases and the effectiveness of anomaly sorts. That is, in all of the quintile groups of the conditional earnings forecast bias, the long-short anomaly strategy earns significant returns.

Table A18: Returns on portfolios formed on conditional earnings forecast bias and anomaly score post-2000

Notes: This table reports the time-series average of returns on portfolios formed by sorting independently on the average conditional earnings forecast bias (using machine learning forecasts as the benchmark) and the anomaly score for the post-2000 periods. The anomaly score is defined as the equal-weighted average of the decile ranking on each of the 27 anomaly variables. The last two rows report the average earnings forecast bias (with Newey–West t-statistic) of each of the ten decile portfolios formed on the anomaly score.

	Anomaly Decile										
BE Quintile	S	2	3	4	5	6	7	8	9	L	L-S
1	0.84	0.86	1.25	1.31	1.31	1.34	1.47	1.20	1.61	1.50	0.66
t-stat	1.60	2.17	3.64	4.04	3.94	4.42	5.23	4.48	6.05	5.06	1.50
2	0.49	0.69	0.74	0.90	0.71	0.63	0.75	0.71	0.96	0.93	0.44
t-stat	1.05	1.92	2.34	2.85	2.40	2.21	2.61	2.50	3.92	3.39	1.17
3	-0.25	0.30	0.38	0.42	0.39	0.54	0.53	0.67	0.95	0.60	0.85
t-stat	-0.50	0.70	0.99	1.15	1.07	1.58	1.64	2.02	2.81	1.79	2.34
4	-0.56	-0.29	0.41	0.57	0.40	0.53	0.70	0.47	1.06	0.66	1.22
t-stat	-0.98	-0.53	0.90	1.12	0.87	1.17	1.55	1.16	2.77	1.59	2.84
5	-1.19	-0.74	-0.30	0.03	0.24	0.42	0.36	0.09	1.16	0.72	1.91
t-stat	-1.64	-0.96	-0.48	0.04	0.36	0.64	0.60	0.16	1.76	1.25	4.02
5-1	-2.03	-1.60	-1.56	-1.28	-1.07	-0.92	-1.11	-1.11	-0.45	-0.78	1.25
t-stat	-3.63	-2.78	-3.20	-2.25	-1.98	-1.64	-2.24	-2.15	-0.76	-1.71	2.34
BE	0.006	0.005	0.003	0.003	0.003	0.002	0.002	0.002	0.003	0.003	-0.003
t-stat	3.04	2.64	3.40	3.48	3.00	3.26	3.36	3.31	2.60	2.45	-3.21

Table A19: Returns on portfolios formed on linear forecast bias and anomaly score post-2000

Notes: This table reports the time-series average of returns on portfolios formed by independently sorting on the average linear earnings forecast bias and the anomaly score for the post-2000 periods. The anomaly score is defined as the equal-weighted average of the decile ranking on each of the 27 anomaly variables. The last two rows report the average linear earnings forecast bias (with Newey–West t-statistic) of each of the ten decile portfolios formed on the anomaly score.

	Anomaly Decile										
BE Quintile	S	2	3	4	5	6	7	8	9	L	L-S
1	0.21	0.92	1.00	1.45	1.11	0.92	1.18	1.05	1.31	1.43	1.21
t-stat	0.44	2.42	3.20	4.41	3.42	3.07	4.02	3.66	4.85	4.61	2.92
2	0.17	0.43	0.74	0.66	0.64	0.75	0.75	0.75	1.06	0.84	0.67
t-stat	0.39	1.13	2.00	1.97	1.97	2.48	2.51	2.59	4.09	2.68	1.94
3	-0.27	0.36	0.52	0.62	0.66	0.83	0.67	1.00	0.92	0.88	1.15
t-stat	-0.53	0.85	1.27	1.51	1.78	2.41	1.86	3.00	3.21	2.51	2.84
4	-0.30	-0.29	0.31	0.59	0.84	0.33	0.51	0.48	0.76	0.64	0.94
t-stat	-0.51	-0.52	0.66	1.21	1.77	0.67	1.08	1.10	1.82	1.43	2.35
5	-0.60	-0.22	-0.14	0.19	0.13	0.61	0.80	0.13	1.56	0.93	1.53
t-stat	-0.75	-0.28	-0.23	0.30	0.20	0.88	1.30	0.22	2.25	1.57	2.78
5-1	-0.82	-1.14	-1.15	-1.26	-0.98	-0.31	-0.38	-0.92	0.25	-0.50	0.32
t-stat	-1.27	-1.79	-2.34	-2.30	-1.80	-0.52	-0.73	-1.72	0.41	-1.02	0.56
BE	-0.002	0.000	-0.002	-0.003	-0.005	-0.004	-0.003	-0.003	-0.004	-0.002	0.000
t-stat	-0.14	0.00	-0.30	-0.39	-0.54	-0.51	-0.45	-0.50	-0.47	-0.24	-0.02

Table A20: Returns on portfolios formed on linear forecast bias measured in So (2013) and anomaly score post-2000

Notes: This table reports the time-series average of returns on portfolios formed by sorting independently on the linear earnings forecast bias measured in So (2013) and the anomaly score for the post-2000 periods. The anomaly score is defined as the equal-weighted average of the decile ranking on each of the 27 anomaly variables. The last two rows report the average linear bias (with Newey–West t-statistic) of each of the ten decile portfolios formed on the anomaly score.

	Anomaly Decile										
BE Quintile	S	2	3	4	5	6	7	8	9	L	L-S
1	0.73	0.74	1.14	1.25	1.36	1.15	1.25	1.22	1.62	1.63	0.90
t-stat	1.37	1.37	2.68	2.93	3.46	3.21	3.48	3.65	4.49	4.09	2.13
2	-0.36	0.34	0.72	0.82	0.70	0.83	0.97	0.70	1.15	1.38	1.73
t-stat	-0.71	0.88	2.13	2.58	2.24	2.68	3.29	2.72	4.31	4.35	3.97
3	-0.83	0.18	0.24	0.60	0.27	0.66	0.62	0.65	0.56	1.03	1.86
t-stat	-1.49	0.40	0.64	1.44	0.76	2.03	1.89	2.18	1.87	3.14	4.00
4	-1.25	-0.17	-0.39	-0.02	0.21	0.19	0.26	0.49	0.63	0.83	2.09
t-stat	-2.12	-0.34	-0.80	-0.04	0.45	0.41	0.63	1.17	1.60	2.20	4.41
5	-1.70	-0.42	-0.34	0.30	1.00	0.75	1.07	0.32	1.35	1.21	2.91
t-stat	-2.54	-0.67	-0.63	0.55	1.73	1.38	1.96	0.60	2.27	2.16	5.79
5-1	-2.43	-1.17	-1.49	-0.95	-0.36	-0.40	-0.18	-0.90	-0.27	-0.43	2.00
t-stat	-5.40	-2.22	-3.27	-2.23	-0.79	-0.88	-0.39	-2.24	-0.56	-1.06	3.82
BE	-0.000	0.001	0.001	0.002	0.003	0.002	0.003	0.005	0.006	0.012	0.012
t-stat	-0.01	0.39	0.92	1.73	1.42	1.78	3.07	4.21	4.96	4.28	4.04

Moreover, we study the properties of the orthogonal component of our conditional earnings forecast bias measure relative to the anomaly score. We find similar results with the orthogonalized component with stocks with a larger value earning significantly lower returns than stocks with a lower (orthogonalized) earnings forecast bias. Further, common factor models cannot subsume the returns on the long-short portfolios sorted on this component measure. Within each of the ten anomaly deciles, the difference in return between the stocks with the largest bias and the stocks with the smallest bias is significantly negative. Furthermore, we consider the conditional bias to be the primitive economic object, as interpreting anomalies is notoriously difficult.

Table A21: Portfolios sorted on the orthogonal component of conditional bias

Notes: This table reports the time-series average of returns (in percentage points) on portfolios formed on the orthogonal component of the average conditional earnings forecast bias relative to the anomaly score. The anomaly score is defined as the equal-weighted average of the decile ranking on each of the 27 anomaly variables. Panel A look at the portfolio mean returns and Panel B look at alphas relative to the CAPM, the Fama–French three-factor model (FF3), and the Fama–French five-factor model (FF5), respectively. The sample period is 1986 to 2019.

$$LS\_Port_t = \alpha + \sum_{i=1}^{5} \beta_i F_{i,t} + \epsilon_t$$

Panel A: Portfolio Mean Returns

Quintile	1	2	3	4	5	5-1
Mean	1.24	1.00	0.93	0.61	-0.00	-1.24
t-stat	5.60	4.91	4.18	2.23	-0.01	-4.84

Panel B: Portfolio Alphas

	CAPM		FF3		FF5	
	Coef	t-stat	Coef	t-stat	Coef	t-stat
Intercept	-1.54	-6.41	-1.68	-8.35	-1.39	-6.22
Mkt_RF	0.44	6.24	0.43	7.21	0.33	5.36
SMB			0.74	8.44	0.58	5.89
HML			0.67	6.88	0.87	6.97
RMW					-0.57	-4.24
CMA					-0.20	-0.92

A13. Net stock issuance: robustness check

We check the robustness of results in Table 9 by matching average of conditional earnings forecast bias from the past 24-12 months to net stock issuance of the fiscal year ending in t. Table A22 reports this robustness check. Overall, we find consistent results that managers of those companies for which analysts' upward biases are larger tend to issue more stocks.

Table A22: Net stock sssuances and conditional bias

Panel A reports the time-series average of net stock issuances of value-weighted portfolios sorted on the conditional earnings forecast bias. "Average BE" denotes the average of the conditional bias at different forecast horizons."BE score" denotes the arithmetic average of the percentile rankings on each of the five conditional biases at different forecast horizons. Panel B reports the Fama–MacBeth regressions of firms' net stock issuances on the conditional bias and control variables include the log of firm size (LNsize), the log of book-to-market ratio (LNbeme), and earnings before interest, taxes, and depreciation divided by total assets (EBITDA). The sample period is 1986 to 2019. We report the time-series average of slope coefficients associated with Newey-West t-statistics.

Panel A: Net stock issuances of portfolios formed on BE						
Quintile	1	2	3	4	5	5-1
Average BE	0.007	0.010	0.019	0.021	0.071	0.064
t-stat	1.26	1.56	2.33	3.18	4.96	4.67
BE score	0.005	0.013	0.015	0.033	0.065	0.060
t-stat	0.78	1.78	2.81	4.90	5.14	5.35

Panel B: Fama–MacBeth regressions				
	A: Average BE		B: BE Score	
	(1)	(2)	(1)	(2)
Bias	0.514	0.398	0.091	0.063
t-stat	4.11	3.96	8.10	4.60
LNsize		-0.493		-0.307
t-stat		-2.78		-1.42
LNbeme		-1.689		-1.803
t-stat		-5.32		-5.68
EBITDA		-0.122		-0.121
t-stat		-5.18		-5.04
Intercept	0.032	0.096	-0.007	0.045
t-stat	8.15	3.36	-1.01	1.17
R^2 (%)	1.963	7.674	1.201	7.187

A14. Conditional biases and stock issuance

We test the predictability of net stock issuance (NSI) using linear forecasts and compare the predictability between pre and post-2000 periods in Table A23. Using our conditional earnings forecast bias measure, we do not find any decay in the NSI predictability in the post-2000 periods, consistent with the superior out-of-sample performance of machine learning earnings forecasts. In contrast, we observe a significant decline in the NSI predictability of the linear forecast bias as measured in So (2013). The variable-selection-bias-free linear forecast can predict both in- and out-of-sample net stock issuance.

Further, we report average net stock issuance for portfolios sorted independently on conditional bias and anomaly scores in Table A24. Given the independent sort on anomaly score, we find that stocks in the anomaly short-leg have more net stock issuance than stocks in the long leg. This finding is consistent with a market timing strategy where managers tend to issue more stocks if their firms are more overpriced. In addition, within 9 out of 10 anomaly deciles, we find a significantly positive difference in NSI between stocks with the largest conditional earnings forecast bias and stocks with the smallest conditional bias. Therefore, our conditional bias has an independent effect on issuance relative to other anomaly characteristics. Interestingly, we find that stocks in the short-leg portfolio with the largest earnings forecast bias have the most NSI, while stocks in the long-leg portfolio with the smallest bias have the least NSI. These patterns suggest an interaction between analysts' biases and anomaly characteristics driving managers' financing decisions.

Table A23: Net stock issuances of portfolios sorted on conditional bias

Notes: This table reports the time-series average of net stock issuances on portfolios formed on the conditional earnings forecast bias during the pre-2000 and post-2000 periods. Panel A uses random forest earnings predictions, whereas panel B uses a linear earnings forecasts as the benchmark. Panel C reports net stock issuances for portfolios formed on the linear forecast bias measured in So (2013). The last row of each panel reports the p-value of testing (using an F-test) for the difference in net stock issuances between the two subperiods (pre-2000 and post-2000). The sample period is 1986 to 2019.

Quintile	1	2	3	4	5	5-1
Panel A: Random Forest						
Pre-2000	0.019	0.031	0.036	0.046	0.070	0.051
t-stat	2.51	2.40	4.46	9.40	17.87	7.65
Post-2000	-0.005	-0.004	0.001	0.012	0.061	0.065
t-stat	-1.52	-1.20	0.29	1.75	2.47	2.63
p-value	0.01	0.01	0.00	0.00	0.71	0.58
Panel B: Linear Forecast						
Pre-2000	0.024	0.025	0.034	0.039	0.069	0.045
t-stat	2.34	2.63	10.99	17.73	12.83	5.62
Post-2000	0.000	-0.001	0.005	0.019	0.080	0.080
t-stat	-0.07	-0.21	1.03	2.47	3.82	3.58
p-value	0.04	0.02	0.00	0.02	0.62	0.14
Panel C: Linear Forecast in So (2013)						
Pre-2000	0.012	0.011	0.022	0.033	0.062	0.050
t-stat	2.52	1.92	3.45	4.91	6.38	5.69
Post-2000	0.028	0.003	0.007	0.013	0.040	0.012
t-stat	3.75	0.54	0.82	1.90	3.33	0.98
p-value	0.05	0.26	0.13	0.05	0.16	0.01

Table A24: NSI on portfolios formed on conditional bias and anomaly score

Notes: This table reports the net stock issuances on portfolios formed by sorting independently on the conditional earnings forecast bias (BE) and the anomaly score, defined as the equal-weighted average of the decile ranking on each of the 27 anomaly variables. The last two rows report the net stock issuances and t-statistics on the ten decile portfolios formed on the anomaly score. The sample period is 1986 to 2019.

	Anomaly Decile										
BE Quintile	S	2	3	4	5	6	7	8	9	L	L-S
1	0.066	0.037	0.032	0.023	0.009	0.002	0.000	0.004	0.001	-0.008	-0.074
t-stat	4.31	4.30	3.26	2.23	1.58	0.73	0.02	0.66	0.22	-1.34	-5.05
2	0.045	0.026	0.036	0.028	0.021	0.017	0.007	-0.005	0.001	-0.004	-0.050
t-stat	3.84	3.60	3.12	2.08	2.07	1.56	1.31	-0.80	0.15	-0.76	-5.55
3	0.056	0.043	0.025	0.026	0.036	0.020	0.005	0.003	0.004	0.005	-0.050
t-stat	3.74	6.78	3.36	2.79	2.24	2.85	0.84	0.44	0.69	0.60	-2.84
4	0.054	0.043	0.039	0.025	0.034	0.026	0.018	0.039	0.019	0.008	-0.046
t-stat	7.78	5.32	5.00	4.12	3.63	3.42	1.77	2.82	1.84	0.94	-4.87
5	0.099	0.080	0.069	0.067	0.086	0.046	0.069	0.044	0.055	0.061	-0.038
t-stat	5.46	4.66	5.98	5.62	2.95	5.25	3.42	3.05	3.41	4.97	-1.78
5-1	0.033	0.043	0.037	0.044	0.077	0.044	0.068	0.040	0.053	0.069	0.036
t-stat	1.21	2.05	2.45	2.84	2.58	4.97	3.30	3.61	3.65	5.90	1.32
All Stocks	0.065	0.034	0.032	0.026	0.018	0.010	0.003	0.002	0.004	-0.002	-0.066
t-stat	5.32	7.70	4.22	2.89	2.59	2.04	0.54	0.39	0.56	-0.34	-5.56

References

Balakrishnan, K., Bartov, E. and Faurel, L. (2010). Post loss/profit announcement drift, *Journal of Accounting and Economics* **50**(1): 20–41.

Barth, M. E., Elliott, J. A. and Finn, M. W. (1999). Market rewards associated with patterns of increasing earnings, *Journal of Accounting Research* **37**(2): 387–413.

Basu, S. (1983). The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence, *Journal of Financial Economics* **12**(1): 129–156.

Belo, F. and Lin, X. (2012). The inventory growth spread, *Review of Financial Studies* **25**(1): 278–313.

Bouchaud, J., Krueger, P., Landier, A. and Thesmar, D. (2019). Sticky expectations and the profitability anomaly, *Journal of Finance* **74**: 639–674.

Boudoukh, J., Michaely, R., Richardson, M. and Roberts, M. R. (2007). On the importance of measuring payout yield: Implications for empirical asset pricing, *Journal of Finance* **62**(2): 877–915.

Chan, L. K., Jegadeesh, N. and Lakonishok, J. (1996). Momentum strategies, *Journal of Finance* **51**(5): 1681–1713.

Chan, L. K., Lakonishok, J. and Sougiannis, T. (2001). The stock market valuation of research and development expenditures, *Journal of Finance* **56**(6): 2431–2456.

Cooper, M. J., Gulen, H. and Schill, M. J. (2008). Asset growth and the cross-section of stock returns, *Journal of Finance* **63**(4): 1609–1651.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information, *Journal of Finance* **61**(4): 1605–1643.

Dechow, P. M., Sloan, R. G. and Soliman, M. T. (2004). Implied equity duration: A new measure of equity risk, *Review of Accounting Studies* **9**(2-3): 197–228.

Diether, K. B., Malloy, C. J. and Scherbina, A. (2002). Differences of opinion and the cross section of stock returns, *Journal of Finance* **57**: 2113–41.

Eisfeldt, A. L. and Papanikolaou, D. (2013). Organization capital and the cross-section of expected returns, *Journal of Finance* **68**(4): 1365–1406.

Fama, E. F. and French, K. R. (1996). Multifactor explanations of asset pricing anomalies, *Journal of Finance* **51**(1): 55–84.

Foster, G., Olsen, C. and Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* pp. 574–603.

Frankel, R. and Lee, C. M. (1998). Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and Economics* **25**: 283–319.

Hafzalla, N., Lundholm, R. and Matthew Van Winkle, E. (2011). Percent accruals, *Accounting Review* **86**(1): 209–236.

Haugen, R. A. and Baker, N. L. (1996). Commonality in the determinants of expected stock returns, *Journal of Financial Economics* **41**(3): 401–439.

Hirshleifer, D., Hou, K., Teoh, S. H. and Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* **38**: 297–331.

Hou, K., Xue, C. and Zhang, L. (2015). Digesting anomalies: An investment approach, *Review of Financial Studies* **28**: 650–705.

Lakonishok, J., Shleifer, A. and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk, *Journal of Finance* **49**: 1541–1578.

Lyandres, E., Sun, L. and Zhang, L. (2007). The new issues puzzle: Testing the investment-based explanation, *Review of Financial Studies* **21**(6): 2825–2855.

Novy-Marx, R. (2010). Operating leverage, *Review of Finance* **15**(1): 103–134.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium, *Journal of Financial Economics* **108**(1): 1–28.

Pontiff, J. and Woodgate, A. (2008). Share issuance and cross-sectional returns, *Journal of Finance* **63**: 921–45.

Rosenberg, B., Reid, K. and Lanstein, R. (1985). Persuasive evidence of market inefficiency, *Journal of Portfolio Management* **11**(3): 9–16.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* pp. 289–315.

So, E. C. (2013). A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts?, *Journal of Financial Economics* **108**: 615–40.

Thomas, J. K. and Zhang, H. (2002). Inventory changes and future returns, *Review of Accounting Studies* **7**(2-3): 163–187.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* **48**: 817–38.

Xing, Y. (2007). Interpreting the value effect through the q-theory: An empirical investigation, *Review of Financial Studies* **21**(4): 1767–1795.