Online Appendix for
"Do Cross-Sectional Predictors Contain Systematic Information?"[1]

This Online Appendix provides additional evidence to supplement the analyses provided in the published text. Section A contains a generalization of the theoretical relation between cross-sectional return predictability and time-series return predictability, Section B provides additional information about our implementation of multiple testing methods, and Section C contains several supplemental tables. Below, we briefly discuss each table.

- Table AI.I shows Benjamini–Yekutieli (2001) p-values for in-sample regressions. It is similar to Table 4 in the main text except the multiple testing procedure controls the false discovery rate instead of the family wise error rate.
- Table AI.II shows Benjamini–Yekutieli (2001) p-values for out-of-sample regressions. It is similar to Table 7 in the main text except the multiple testing procedure controls the false discovery rate instead of the family wise error rate.
- Table IA.III shows the raw and Romano and Wolf adjusted p-values for in-sample regressions using the 269 variables in the sample of "All Possible Predictors." It is similar to Table 3 in the main text, except Table 3 in the main text considers the 253 variables in the sample of "Raw Predictors + First Diff" predictors. While this expanded sample shows that several other variables are highly significant before adjusting for multiple hypothesis testing, including *Short Interest*, the conclusion is similar after applying the Romano and Wolf correction – few of the predictors are significant out-of-sample.
- Table IA.IV shows the raw and Romano and Wolf adjusted p-values for out-of-sample regressions using the 269 variables in the sample of "All Possible Predictors." It is similar to Table 5 in the main text, except Table 5 in the main text considers the 253 variables in the sample of "Raw Predictors + First Diff" predictors. While this expanded sample shows that several other variables are highly significant before

[1] Citation format: Engelberg, Joseph E., R. David McLean, Jeffrey Pontiff, and Matthew C. Ringgenberg, Online Appendix for "Do Cross-Sectional Predictors Contain Systematic Information?"

adjusting for multiple hypothesis testing, including *Short Interest*, the conclusion is the same after applying the Romano and Wolf correction – none of the predictors are significant out-of-sample.

- Table IA.V provides an overview of the predictor variables we use in the main text and discusses the construction methodology for each variable. In our main specification "Raw Predictors + First Diff." we drop eight predictors using a manual filter because these variables should not aggregate to form an economically meaningful time-series variable. These eight variables are: (i) Analyst Coverage; (ii) Forecast Dispersion Change; (iii) Idiosyncratic Risk, Distant; (iv) Number of Analysts, (v) Change; Book to Assets, Change; (vi) Misvalued Innovation; (vii) Sales Growth; and (vii) Agnostic Value. In one of our alternate specifications, "*All Possible Predictors*", we include these eight variables and our conclusions are unchanged.
- Table IA.VI lists the 26 cross-sectional predictors used in the existing time-series literature as of June 2021, along with the 23 time-series papers that each was featured in.

A. The Information in Non-Systematic Predictors

In Section I.C of the main text we examine the theoretical relation between cross-sectional return predictability and time-series return predictability using the market model. In this Section, we generalize this relation and show that it holds using *any* factor model. While it may seem natural that cross-sectional return predictors should aggregate across assets to generate time-series return predictors, it is possible to have one without the other. Specifically, note that cross-sectional predictor variables could predict returns because they forecast either the non-systematic component of returns or the systematic component of returns. As such, cross-sectional return predictors do not necessarily aggregate to form good time series predictors. To see this, define a variable $X_{i,t}^{Non-Syst}$ as an non-systematic predictor if it forecasts the *non-systematic* portion of stock returns for asset *i* on date *t+1*. Assuming a factor model for returns, we can write the return on asset *i* as:

$$R_{i,t} = \sum_{k=1}^{K} \beta_{i,k} F_k + \varepsilon_{i,t}, \quad (1)$$

where $R_{i,t}$ is the return on stock *i* on date *t*, F_k is the risk-premium on factor *k*, $\beta_{i,k}$ is the loading of stock *i* on factor *k*, and $\varepsilon_{i,t}$, is the non-systematic portion of stock *i*'s return. We define a non-systematic predictor as a variable $X_{i,t-1}^{Non-Syst}$ that satisfies $\gamma_1 \neq 0$ in a linear regression of the form:[2]

$$\widehat{\varepsilon_{i,t}} = \gamma_0 + \gamma_1 X_{i,t-1}^{Non-Syst} + \omega_{i,t}, \quad (2)$$

where $\widehat{\varepsilon_{i,t}}$ is the abnormal return. In other words, an non-systematic predictor, by definition, forecasts the portion of asset *i*'s return that is *not* explained by the systematic component of returns. However, while $X_{i,t-1}^{Non-Syst}$ contains information about individual stock returns, it will not aggregate to generate market return predictability. To see this, calculate the weighted sum of equation (2) across all *N* stocks in the economy:

$$\sum_{i=1}^{N} [w_{i,t} R_{i,t} - \sum_{k=1}^{K} w_{i,t} \beta_{i,k} F_k] = \overline{\gamma_0} + \overline{\gamma_1 X_{t-1}^{Non-Syst}}, \quad (3)$$

where the weight $w_{i,t} = \frac{mc_{i,t}}{\sum_i^N mc_{i,t}}$, $mc_{i,t}$ is the market capitalization of stock *i* on date *t*, and the bar above a variable denotes the value-weighted mean. It is simple to show that the left-hand side of equation (3) is equal to zero.[3] Thus, the value-weighted non-systematic predictor variable $\overline{X_{t-1}^{idio}}$ contains *no* information about aggregate market returns.

While non-systematic predictors contain no information about aggregate market returns, it is possible to have a variable that predicts information in the cross-section that contains information about the aggregate risk-premium. Define a variable $X_{i,t-1}^{syst}$ as a systematic predictor if it forecasts the *systematic* portion of stock returns for asset *i* on date *t+1*. Thus, define a systematic predictor as a variable $X_{i,t-1}^{syst}$ that satisfies $\gamma_1 \neq 0$ in a linear regression of the form:

[2] For simplicity, we ignore the sign of the abnormal return and define an anomaly as any variable that predicts abnormal returns in either direction.
[3] Note: the aggregate market return is, by definition, the weighted sum across all assets of the factors multiplied by their loading: $R_m = \sum_{i=1}^{N} \sum_{k=1}^{K} w_{i,t} \beta_{i,k} F_k$.

$$\sum_{k=1}^{K} \beta_{i,k} F_k = \gamma_0 + \gamma_1 X_{i,t-1}^{syst} + \omega_{i,t}. \quad (4)$$

Because the market return is the weighted sum across all assets of the factors multiplied by their loading, equation (4) implies a direct linear relation between the predictor variable and the market return.

B. Multiple Hypothesis Testing Procedures

In Section III.B of the main text we apply the Romano and Wolf (2016) procedure to our in-sample results. Formally, we proceed as follows for the case of predicting the 1-month ahead equity risk premium (the procedures are analogous for the 3-month, 6-month, and 12-month cases):

1) For each predictor, find the OLS coefficient ($\hat{\beta}$) and its associated t-statistic from Equation (6). Rank these t-statistics by their absolute value from 1 to 268. For the n^{th} ranked predictor, call the t-statistic $|t|_n$.
2) For each predictor, calculate an adjusted dependent variable $r^*_{t:t+1}$ where $r^*_{t:t+1} = r_{t:t+1} - \hat{\beta} x_t$. Now, for each predictor x, there is no relation between r^* and x by construction (i.e. if we regressed r^* on x in the full sample the coefficient and associated t-stat for x would be zero). We use r^* (rather than r) when building bootstrapped samples so that we operate *under the null of no predictability from x*.[4]
3) Create a $T \times 268$ matrix for the dependent r^* values where rows index time (T) and columns index the predictors. Call this matrix R^*. Also create a $T \times 268$ matrix for the independent x values where rows index time and columns index the predictors. Call this matrix X.
4) Bootstrap row samples from the R^* and X matrices using the stationary bootstrap (Politis and Romano, 1994).[5] Bootstrapping by row preserves the cross-sectional

[4] The procedure in Romano and Wolf (2005) does not impose any null hypothesis in the bootstrap data generating process. As such, our procedure is not a direct application of Romano and Wolf (2005) but rather an extension similar to Bakshi, Panayotov, and Skoulakis (2011). We thank Michael Wolf for helpful conversations on this point.

[5] We use the bootstrap procedure in Politis and Romano (1994) which preserves the underlying time-series properties of the data. Specifically, we resample the original data using a stationary block bootstrap with a

dependence in the data. Create 1000 bootstrapped datasets of dependent and independent variables.

5) For each bootstrap dataset and each predictor, regress the dependent variable on the independent variable and save the corresponding coefficient and t-statistic.
6) Finally, for each bootstrap sample, calculate the maximum of the absolute value of the t-statistic across all predictors.
    a. Then, for the best predictor in the data compute the Romano and Wolf (2016) adjusted p-value as the fraction of times the maximum bootstrap t-statistics is greater than or equal to the actual t-statistic.
    b. Then remove the best predictor, and repeat step 6 (i.e., for each bootstrap sample, re-calculate the maximum of the absolute value of the t-statistic across all remaining predictors and calculate the adjusted p-value for the new best predictor subject to a monotonicity condition that the p-value for the nth ranked predictor must be less than or equal to the p-value for the (n+1)th ranked predictor. Repeat until all predictors have adjusted p-values.

B.1. Out-of-Sample Multiple Hypothesis Testing

We also apply the Romano and Wolf (2016) procedure to our out-of-sample results. The procedure follows a similar process as outlined above, except we estimate rolling regressions and compare the prediction from these regressions to a rolling average market risk premium. Formally, we proceed as follows for the case of predicting the 1-month ahead equity risk premium (the procedures are analogous for the 3-month, 6-month, and 12-month cases):

mean block size of 5, however our conclusions are robust to alternate block sizes of 10, 25, and 50. Consistent with this, Sullivan, Timmerman, and White (1999) show the White (2000) procedure is robust to various block size choices.

1) At each point in time, forecast the equity risk premium next period using the benchmark model (i.e., using the prevailing mean equity risk premium from the start of the sample until the current date).

2) At each point in time, forecast the equity risk premium next period using out-of-sample regressions for each of the 253 candidate predictor variables on each date.

3) For each date and predictor, calculate a loss function as in White (2000) by comparing the mean square forecast error of the predictor to the mean square forecast error of the benchmark model. Call these values $\hat{f}_{x,t+1}$, where x indexes the predictor variable (in our case, x ranges from 1 to 253). Call the $T \times 253$ matrix of $\hat{f}_{x,t+1}$ values matrix F.

    a. For each predictor, calculate the time-series mean of $\hat{f}_{x,t+1}$ and call it $\bar{f}_x$.

4) Bootstrap rows of loss function values ($\hat{f}_{x,t+1}$) from matrix F using a stationary block bootstrap (Politis and Romano, 1994) to ensure that the bootstrap accurately estimates the sampling distribution.[6]

5) Calculate studentized test statistics based on the actual sample: $V_x = \frac{n^{1/2}\bar{f}_x}{\hat{\omega}}$, where $\hat{\omega}$ is an estimate of the standard deviation of $\bar{f}_x$.

6) For each bootstrap sample i, calculate the maximum test statistic across the models considered $T_i = max\ \bar{V}^*_{x,i} = \frac{n^{1/2}(\bar{f}^*_{x,i} - \bar{f}_x)}{\hat{\omega}^*}$, where $\hat{\omega}^*$ is an estimate of the standard deviation of $(\bar{f}^*_{x,i} - \bar{f}_x)$.

[6] As with the in-sample calculations, we bootstrap by row, which preserves the cross-sectional dependence in the data and we use the bootstrap procedure in Politis and Romano (1994) with 1000 replications and a mean block size of 5. Our results are robust to alternate block size choices.

a. Sort the predictors based on their realized test statistics (V_x). Starting with the best predictor, compute the p-value for predictor x as the percent of T_i observations that are greater than or equal to V_x.[7]

b. If the p-value exceeds the desired confidence level, stop. Otherwise, remove the current predictor from the sample and repeat step 6.

[7] Formally, we calculate the p-value as $100 \times (M+1)/(N+1)$, where N is the number of observations and M is the number of observations that are greater than or equal to the test statistic.

Table IA.I
Summary of In-Sample Performance using Benjamini and Yekutieli p-values

The table displays a count of the number of predictive variables that are statistically significant at the 10% level or better, as a fraction of the total number of variables examined. We calculate statistical significance using Benjamini and Yekutieli (2001) adjusted p-values controlling the false discovery rate at the 5% level. For each anomaly, we estimate an in-sample predictive regression of the form:

$$r_{t:t+h} = \alpha + \beta x_t + \varepsilon_{t:t+h} \text{ for } t = 1, \ldots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the continuously compounded S&P 500 return for month t from CRSP including dividends and excess of the monthly risk-free rate from Welch and Goyal (2008), h indicates the forecast horizon in months, and x_t is one of the 140 predictor variables. To construct time-series predictors out of cross-sectional predictors, we calculate the value-weighted and equal-weighted mean across all stocks on each date resulting in 280 possible predictors. In Panel A, we consider four different definitions: (1) *Predictors from Existing Papers* uses only those variables that are used in the existing literature on time-series return predictability. (2) *Raw Predictors* examines every possible variable for which we reject the null that the raw variable is non-stationary. (3) *Raw Predictors + First Diff.* examines every possible variable however if a variable is not stationary in raw form, we then examine whether it is non-stationary in first-differenced form. If we fail to reject the null that the first differenced variable is non-stationary, we drop the variable. (4) All Possible Predictors examines every possible variable. If we fail to reject the null that a variable is non-stationary, we calculate deviations from a linear trend model. If we fail to reject the null that the linearly detrended variable is non-stationary, we calculate the first-difference. If we fail to reject the null that the first differenced variable is non-stationary, we drop the variable. In Panel B, we examine subsamples of the variables in *Raw Predictors + First Diff.* formed on the ten most statistically significant cross-sectional predictors (*Best Cross-Sectional*), and two different groupings based on the categories in McLean and Pontiff (2016): (1) *Opinion* and (2) *Valuation*. In Panel C we examine equal-weighted vs. value-weighted predictors for the variables in *Raw Predictors + First Diff.*

(1)	**(2)**	**(3)**	**(4)**	**(5)**
	Return Horizon (*h*)			
Predictive Variable	***h*=1**	***h*=3**	***h*=6**	***h*=12**
Panel A: Candidate Predictors (number significant / total examined)				
Predictors from Existing Papers	0/51	0/51	3/51	4/51
Raw Predictors	1/137	2/137	6/137	8/137
Raw Predictors + First Diff.	1/253	2/253	6/253	8/253
All Possible Predictors	1/269	5/269	4/269	8/269
Panel B: By Subcategory (number significant / total examined)				
Best Cross-sectional	0/20	0/20	1/20	1/20
Opinion	0/38	0/38	0/38	0/38
Valuation	0/24	0/24	0/24	0/24
Panel C: By Aggregation Method (number significant / total examined)				
Equal-Weighted Predictors	0/125	1/125	1/125	2/125
Value-Weighted Predictors	1/128	1/128	5/128	6/128

Table IA.II

Summary of Out-of-Sample Performance using Benjamini and Yekutieli p-values

The table displays a count of the number of predictive variables that are statistically significant at the 10% level or better, as a fraction of the total number of variables examined. We calculate statistical significance using Benjamini and Yekutieli (2001) adjusted p-values controlling the false discovery rate at the 5% level. For each anomaly, we estimate an out-of-sample predictive regression of the form:

$$r_{t:t+h} = \alpha + \beta x_t + \varepsilon_{t:t+h} \text{ for } t = 1, \ldots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the continuously compounded S&P 500 return for month t from CRSP including dividends and excess of the monthly risk-free rate from Welch and Goyal (2008), h indicates the forecast horizon in months, and x_t is one of the 140 predictor variables. We estimate expanding rolling window regressions using only information available on each date. To construct time-series predictors out of cross-sectional predictors, we calculate the value-weighted and equal-weighted mean across all stocks on each date resulting in 280 possible predictors. In Panel A, we consider four different definitions: (1) *Predictors from Existing Papers* uses only those variables that are used in the existing literature on time-series return predictability. (2) *Raw Predictors* examines every possible variable for which we reject the null that the raw variable is non-stationary. (3) *Raw Predictors + First Diff.* examines every possible variable however if a variable is not stationary in raw form, we then examine whether it is non-stationary in first-differenced form. If we fail to reject the null that the first differenced variable is non-stationary, we drop the variable. (4) All Possible Predictors examines every possible variable. If we fail to reject the null that a variable is non-stationary, we calculate deviations from a linear trend model. If we fail to reject the null that the linearly detrended variable is non-stationary, we calculate the first-difference. If we fail to reject the null that the first differenced variable is non-stationary, we drop the variable. In Panel B, we examine subsamples of the variables in *Raw Predictors + First Diff.* formed on the ten most statistically significant cross-sectional predictors (*Best Cross-Sectional*), and two different groupings based on the categories in McLean and Pontiff (2016): (1) *Opinion* and (2) *Valuation*. In Panel C we examine equal-weighted vs. value-weighted predictors for the variables in *Raw Predictors + First Diff.*

(1)	(2)	(3)	(4)	(5)
	Return Horizon (*h*)			
Predictive Variable	***h*=1**	***h*=3**	***h*=6**	***h*=12**
Panel A: Candidate Predictors (number significant / total examined)				
Predictors from Existing Papers	0/51	0/51	2/51	2/51
Raw Predictors	0/137	0/137	0/137	0/137
Raw Predictors + First Diff.	0/253	0/253	0/253	0/253
All Possible Predictors	0/269	0/269	0/269	0/269
Panel B: By Subcategory (number significant / total examined)				
Best Cross-sectional	0/20	0/20	2/20	0/20
Opinion	0/38	0/38	0/38	0/38
Valuation	0/24	0/24	0/24	0/24
Panel C: By Aggregation Method (number significant / total examined)				
Equal-Weighted Predictors	0/125	0/125	0/125	1/125
Value-Weighted Predictors	0/128	0/128	0/128	0/128

Table IA.III
Best In-Sample Predictive Regression Results using Romano and Wolf p-values
– Alternate Set of Predictors –

The table reports the ordinary least squares estimate of β, p-values, and the R^2 statistic from in-sample predictive regression models of the form:

$$r_{t:t+h} = \alpha + \beta x_t + \varepsilon_{t:t+h} \text{ for } t = 1, \ldots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the continuously compounded S&P 500 return for month t from CRSP including dividends and excess of the monthly risk-free rate from Goyal and Welch (2008), h indicates the forecast horizon in months, and x_t is the predictor variable shown in columns (2) and (9). For each horizon, we run 269 regressions using the variables in the *All Possible Predictors* set of predictors. Panel A displays results for the 1-month horizon, Panel B shows the 3-month horizon, Panel C shows the 6-month horizon, and Panel D shows the 12-month horizon. Within each panel, predictors are sorted by their Romano and Wolf p-value and then their unadjusted p-value. We report all predictors that have unadjusted p-values less than 10% for a given horizon. Unadjusted p-values are shown in columns (6) and (13) and Romano and Wolf (2016) adjusted p-values are shown in columns (7) and (14).

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
		EW or			Raw	RW			EW or			Raw	RW
Rank	**Predictor**	**VW**	$\hat{\beta}$	R^2	**P-Value**		**Rank**	**Predictor**	**VW**	$\hat{\beta}$	R^2	**P-Value**	
Panel A: 1 Month Horizon													
1	Z-score	VW	-0.0056	1.6%	0.00	0.82	18	Amihud's Measure	EW	0.0026	0.4%	0.02	1.00
2	ST reversal	VW	0.0050	0.8%	0.07	0.82	19	Spinoffs	VW	-0.0040	0.9%	0.03	1.00
3	Size	VW	-0.0048	0.8%	0.01	0.86	20	ΔRec + Accrual	EW	-0.0046	1.1%	0.03	1.00
4	ST reversal	EW	0.0049	0.8%	0.09	0.86	21	Δ# of Analysts	VW	-0.0043	1.1%	0.03	1.00
5	Asset Turnover	VW	0.0053	1.5%	0.00	0.88	22	Forecast Dispersion	VW	0.0032	0.6%	0.03	1.00
6	Z-score	EW	-0.0052	1.4%	0.01	0.91	23	Inventory Growth	EW	0.0033	0.6%	0.04	1.00
7	LT reversal	VW	-0.0046	0.7%	0.01	0.92	24	Secured / Total Debt	VW	0.0040	0.9%	0.04	1.00
8	Zero Trading Days	VW	0.0042	1.0%	0.00	0.96	25	Sustainable Growth	VW	-0.0031	0.5%	0.04	1.00
9	ΔNC Op. Assets	VW	-0.0036	0.7%	0.02	0.98	26	Moment LT Reverse	VW	0.0025	0.2%	0.05	1.00
10	SEO	VW	-0.0042	0.9%	0.02	0.99	27	Spreads	EW	0.0026	0.4%	0.05	1.00
11	ΔRec + Accrual	VW	-0.0049	1.3%	0.05	0.99	28	Gross Profitability	VW	-0.0030	0.5%	0.05	1.00
12	Coskewness	EW	0.0036	0.7%	0.03	0.99	29	Misvalue Innovation	VW	-0.0035	0.7%	0.07	1.00
13	Coskewness	VW	0.0037	0.8%	0.06	0.99	30	Target Price Return	EW	-0.0054	1.7%	0.07	1.00
14	NOA	VW	-0.0035	0.7%	0.02	1.00	31	Cash Flow Var.	EW	0.0034	0.6%	0.08	1.00
15	ΔTax to Assets	EW	0.0055	1.6%	0.00	1.00	32	Repurchases	EW	-0.0035	0.6%	0.09	1.00
16	Δ# of Analysts	EW	-0.0043	1.0%	0.01	1.00	33	Capex Growth	VW	-0.0032	0.5%	0.09	1.00
17	Asset Growth	VW	-0.0035	0.7%	0.02	1.00	34	Profit Margin	EW	0.0031	0.5%	0.10	1.00
Panel B: 3 Month Horizon													
1	Short Interest	EW	-0.0062	5.7%	0.03	0.25	23	Δ Rec + Accrual	VW	-0.0037	2.1%	0.03	1.00
2	Z-score	VW	-0.0055	4.6%	0.00	0.34	24	Cash Flow Variance	EW	0.0033	1.7%	0.05	1.00
3	Asset Turnover	VW	0.0056	4.7%	0.00	0.34	25	Misvalue Innovation	VW	-0.0032	1.6%	0.06	1.00
4	Size	VW	-0.0050	2.4%	0.00	0.35	26	Capex Growth	VW	-0.0032	1.6%	0.07	1.00
5	LT Reversal	VW	-0.0048	2.2%	0.01	0.45	27	Spreads	VW	0.0028	1.2%	0.07	1.00
6	Zero Trading Days	VW	0.0042	2.8%	0.00	0.63	28	Accruals	VW	0.0029	1.3%	0.08	1.00
7	Z-score	EW	-0.0046	3.1%	0.00	0.69	29	Analyst Intensity	EW	-0.0026	1.0%	0.01	1.00
8	SEO	VW	-0.0045	3.1%	0.00	0.71	30	Spreads	EW	0.0024	0.9%	0.04	1.00
9	NOA	VW	-0.0038	2.3%	0.01	0.80	31	Moment LT Reverse	VW	0.0022	0.4%	0.06	1.00
10	ΔNC Op. Assets	VW	-0.0033	1.8%	0.03	0.85	32	Secured / Total Debt	VW	0.0021	0.7%	0.06	1.00
11	Target Price Return	EW	-0.0062	6.0%	0.01	0.85	33	ΔBook to Assets	VW	-0.0013	0.2%	0.06	1.00
12	Asset Growth	VW	-0.0036	2.1%	0.01	0.86	34	Reverse Stock Split	VW	0.0028	1.3%	0.07	1.00
13	Coskewness	VW	0.0036	2.1%	0.04	0.86	35	Dividend Omission	EW	0.0026	1.0%	0.08	1.00
14	Inventory Growth	EW	0.0036	2.1%	0.01	0.88	36	Δ # of Analysts	VW	-0.0037	2.1%	0.08	1.00
15	% Operat. Accrual	VW	-0.0045	3.4%	0.04	0.93	37	Opportunistic Sells	VW	-0.0028	1.3%	0.09	1.00
16	Sustain Growth	VW	-0.0033	1.7%	0.03	0.95	38	Reverse Stock Split	EW	-0.0020	0.6%	0.09	1.00
17	Profit Margin	EW	0.0034	1.7%	0.04	0.97	39	Δ Rec + Accrual	EW	-0.0029	1.3%	0.09	1.00
18	Agnostic Value	VW	0.0040	2.5%	0.01	0.98	40	ΔNumber of Analyst	EW	-0.0034	1.9%	0.09	1.00
19	Gross Profitability	VW	-0.0029	1.3%	0.05	0.98	41	Low Recommend	EW	-0.0026	1.1%	0.10	1.00
20	Repurchases	EW	-0.0032	1.6%	0.07	0.99	42	Volume	VW	-0.0020	0.4%	0.10	1.00
21	Δ Tax to Assets	EW	0.0044	2.9%	0.00	1.00	43	ΔInstitutional owner	EW	-0.0022	0.7%	0.10	1.00
22	Agnostic Value	EW	-0.0037	2.2%	0.02	1.00							

Table IA.III - Continued

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
		EW or			Raw	RW			EW or			Raw	RW
Rank	Predictor	VW	$\hat{\beta}$	R^2	P-Value		Rank	Predictor	VW	$\hat{\beta}$	R^2	P-Value	
Panel C: 6 Month Horizon													
1	Short Interest	EW	-0.0063	10.8%	0.02	0.08	24	Repurchases	EW	-0.0030	2.6%	0.04	0.94
2	Asset Turnover	VW	0.0059	9.7%	0.00	0.09	25	Coskewness	VW	0.0027	2.2%	0.05	0.94
3	Z-score	VW	-0.0057	9.1%	0.00	0.10	26	Inventory Growth	VW	0.0028	2.3%	0.08	0.94
4	Size	VW	-0.0050	4.7%	0.00	0.11	27	Volume Trend	EW	0.0032	3.0%	0.09	0.95
5	LT Reversal	VW	-0.0049	4.6%	0.01	0.15	28	Spreads	EW	0.0026	2.0%	0.01	0.95
6	Zero Trading Days	VW	0.0043	5.5%	0.00	0.34	29	Momentum-reversal	VW	-0.0026	1.2%	0.06	0.95
7	Z-score	EW	-0.0044	5.2%	0.00	0.53	30	Dividends	EW	0.0026	2.0%	0.06	0.95
8	Inventory Growth	EW	0.0039	4.5%	0.01	0.56	31	Reverse Stock Split	VW	0.0029	2.6%	0.01	0.96
9	NOA	VW	-0.0037	4.0%	0.00	0.63	32	ΔSales-ΔInventory	EW	-0.0026	2.0%	0.05	0.97
10	Asset Growth	VW	-0.0037	3.9%	0.01	0.66	33	Share issues PW	VW	-0.0028	2.1%	0.10	0.98
11	Spreads	VW	0.0035	3.7%	0.00	0.67	34	R&D/MV	EW	0.0026	2.0%	0.04	0.98
12	Sustainable Growth	VW	-0.0036	3.8%	0.01	0.70	35	Cash Flow Variance	EW	0.0029	2.6%	0.05	0.98
13	ΔNC Op. Assets	VW	-0.0032	3.1%	0.04	0.70	36	Misvalue Innovation	VW	-0.0029	2.5%	0.07	0.99
14	% Operating Accrual	VW	-0.0047	6.7%	0.01	0.71	37	Org. Capital	EW	0.0023	1.5%	0.05	1.00
15	SEO	VW	-0.0038	4.0%	0.01	0.76	38	Share issues DT	VW	-0.0017	0.8%	0.06	1.00
16	Profit Margin	EW	0.0036	3.6%	0.00	0.80	39	Share issues DT	EW	-0.0021	1.2%	0.06	1.00
17	Gross Profitability	VW	-0.0030	2.7%	0.02	0.87	40	Earn Consistency	VW	-0.0016	0.7%	0.10	1.00
18	Target Price Return	EW	-0.0052	7.0%	0.02	0.91	41	Agnostic Value	VW	0.0017	0.9%	0.05	1.00
19	ΔCapex-ΔInd Capex	VW	-0.0029	2.5%	0.01	0.92	42	Analyst Intensity	EW	-0.0012	0.4%	0.04	1.00
20	Capex Growth	VW	-0.0031	3.0%	0.05	0.93	43	ΔBook to Assets	VW	-0.0010	0.3%	0.06	1.00
21	ΔTax to Assets	EW	0.0048	6.0%	0.00	0.93	44	Cash to Assets	VW	0.0014	0.6%	0.05	1.00
22	accruals	VW	0.0029	2.4%	0.03	0.94	45	Opportunistic Sells	VW	-0.0026	2.1%	0.02	1.00
23	ΔCapex-ΔInd Capex	EW	-0.0029	2.4%	0.00	0.94							
Panel D: 12 Month Horizon													
1	Z-score	VW	-0.0058	18.5%	0.00	0.03	28	Volume Trend	EW	0.0030	5.0%	0.05	0.91
2	Size	VW	-0.0054	10.0%	0.00	0.04	29	Share issues PW	VW	-0.0027	4.0%	0.05	0.92
3	Asset Turnover	VW	0.0058	17.8%	0.00	0.04	30	ΔAsset Turnover	VW	0.0028	4.2%	0.02	0.92
4	Short Interest	EW	-0.0052	14.2%	0.01	0.11	31	Dividends	EW	0.0023	3.2%	0.04	0.94
5	LT Reversal	VW	-0.0044	7.3%	0.00	0.13	32	Target Price Return	EW	-0.0043	8.7%	0.03	0.95
6	Zero Trading Days	VW	0.0040	9.5%	0.00	0.21	33	SEO	VW	-0.0026	3.6%	0.06	0.96
7	Z-score	EW	-0.0043	9.8%	0.00	0.30	34	Pension Funding	EW	-0.0028	4.4%	0.05	0.96
8	Sustainable Growth	VW	-0.0039	8.3%	0.00	0.33	35	Inventory Growth	VW	0.0023	2.9%	0.10	0.98
9	ΔNC Op. Assets	VW	-0.0035	6.8%	0.01	0.34	36	ΔSales-ΔInventory	EW	-0.0022	2.8%	0.02	0.98
10	Asset Growth	VW	-0.0037	7.7%	0.00	0.37	37	Momentum-reversal	VW	-0.0022	1.7%	0.04	0.98
11	Spreads	VW	0.0035	7.2%	0.00	0.40	38	Lagged Momentum	VW	-0.0021	1.6%	0.08	0.98
12	Inventory Growth	EW	0.0037	7.7%	0.01	0.40	39	R&D/MV	EW	0.0022	2.5%	0.04	0.99
13	Price	VW	-0.0036	4.1%	0.02	0.50	40	Insider Buy	VW	0.0025	3.8%	0.02	0.99
14	% Operating Accrual	VW	-0.0046	12.2%	0.01	0.52	41	Reverse Stock Split	VW	0.0022	2.8%	0.03	0.99
15	NOA	VW	-0.0033	6.3%	0.00	0.57	42	Coskewness	VW	0.0019	2.1%	0.06	0.99
16	Capex Growth	VW	-0.0034	7.1%	0.01	0.66	43	Cash Flow Variance	EW	0.0023	3.0%	0.06	0.99
17	Profit Margin	EW	0.0033	5.9%	0.00	0.72	44	Exchange Switch	VW	-0.0019	2.1%	0.07	0.99
18	ΔCapex-ΔInd Capex	EW	-0.0030	5.2%	0.00	0.73	45	Covariance Risk	VW	0.0018	2.0%	0.09	1.00
19	R&D Increases	EW	-0.0029	4.6%	0.05	0.82	46	M/B and Accruals	VW	-0.0020	2.3%	0.09	1.00
20	ΔCapex-ΔInd Capex	VW	-0.0028	4.6%	0.00	0.83	47	ΔTax to Assets	EW	0.0029	4.0%	0.01	1.00
21	Spreads	EW	0.0027	4.2%	0.00	0.83	48	Moment LT Reverse	VW	0.0015	0.8%	0.02	1.00
22	Gross Profitability	VW	-0.0027	4.3%	0.02	0.84	49	E/P	VW	0.0008	0.4%	0.04	1.00
23	Org. Capital	EW	0.0030	4.9%	0.00	0.86	50	Coskewness	EW	0.0010	0.6%	0.05	1.00
24	Repurchases	EW	-0.0031	4.9%	0.02	0.87	51	Share issues DT	EW	-0.0018	1.8%	0.05	1.00
25	Employee growth	VW	-0.0027	4.3%	0.05	0.87	52	E/P	EW	-0.0017	1.7%	0.05	1.00
26	Misvalue Innovation	VW	-0.0030	5.3%	0.03	0.91	53	Dividend Initiation	VW	0.0013	1.0%	0.08	1.00
27	LT Reversal	EW	-0.0025	2.3%	0.05	0.91	54	CF/MV	VW	0.0007	0.3%	0.08	1.00

Table IA.IV
Best Out-of-Sample Predictive Regression Results using Romano and Wolf p-values – Alternate Set of Predictors –

The table reports the mean of the ordinary least squares estimate of β, p-values, and the Campbell and Thompson (2008) R^2_{OS} statistic from out-of-sample predictive regression models of the form:

$$r_{t:t+h} = \alpha + \beta x_t + \varepsilon_{t:t+h} \text{ for } t = 1, \ldots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the continuously compounded S&P 500 return for month t from CRSP including dividends and excess of the monthly risk-free rate from Goyal and Welch (2008), h indicates the forecast horizon in months, and x_t is the predictor variable in the first column. $\hat{\beta}$ (column (4)) is the time-series mean of the coefficient estimates for each predictor. The Campbell and Thompson R^2_{OS} statistic (columns (5) and (12)) is calculated as 1 minus the proportional reduction in mean squared forecast error (MSFE) at the h-month horizon for a predictive regression forecast of the S&P 500 log excess return based on the predictor variable in the first column vis-a-vis the prevailing mean benchmark forecast. For each horizon, we run 269 out-of-sample regressions using the variables in the *All Possible Predictors* set of predictors. Panel A displays results for the 1-month horizon, Panel B shows the 3-month horizon, Panel C shows the 6-month horizon, and Panel D shows the 12-month horizon. Within each panel, predictors are sorted by their Romano and Wolf p-value and then their unadjusted p-value. We report all predictors that have unadjusted p-values less than 10% for a given horizon. Unadjusted p-values are shown in columns (6) and (13) and Romano and Wolf (2016) adjusted p-values are shown in columns (7) and (14).

(1) Rank	(2) Predictor	(3) EW or VW	(4) $\hat{\beta}$	(5) R^2	(6) Raw P-Value	(7) RW P-Value	(8) Rank	(9) Predictor	(10) EW or VW	(11) $\hat{\beta}$	(12) R^2	(13) Raw P-Value	(14) RW P-Value
Panel A: 1 Month Horizon													
1	Asset Turnover	VW	0.0070	0.3%	0.04	1.00							
2	ΔRec + Accrual	VW	-0.0029	1.1%	0.04	1.00							
3	Z-score	EW	-0.0043	0.5%	0.04	1.00							
4	Z-score	VW	-0.0049	0.7%	0.05	1.00							
5	ΔTax to Assets	EW	0.0065	-0.9%	0.06	1.00							
6	Pension Funding	VW	-0.0043	-0.4%	0.08	1.00							
7	Coskewness	EW	0.0045	0.0%	0.09	1.00							
Panel B: 3 Month Horizon													
1	Misvalue Innovation	VW	-0.0030	1.3%	0.04	0.97	17	Short Interest	VW	0.0059	0.5%	0.04	1.00
2	Short Interest	EW	-0.0051	6.4%	0.02	0.99	18	Exchange Switch	VW	-0.0034	-0.9%	0.04	1.00
3	Idio Risk	VW	-0.0015	1.1%	0.00	1.00	19	Repurchases	EW	-0.0039	3.9%	0.04	1.00
4	Asset Turnover	VW	0.0073	3.8%	0.00	1.00	20	ΔInstitutional owner	VW	0.0036	-13.4%	0.04	1.00
5	ST Reversal	EW	-0.0002	0.4%	0.00	1.00	21	Cash Flow Variance	VW	0.0042	0.0%	0.05	1.00
6	ST Reversal	VW	-0.0007	0.3%	0.00	1.00	22	LT Reversal	VW	-0.0066	-0.6%	0.05	1.00
7	Z-score	VW	-0.0048	3.3%	0.01	1.00	23	Cash Flow Variance	EW	0.0020	1.1%	0.05	1.00
8	Lagged Momentum	VW	-0.0013	0.8%	0.01	1.00	24	Age	EW	0.0163	-106.8%	0.07	1.00
9	Max	VW	-0.0014	1.2%	0.01	1.00	25	Moment LT Reverse	VW	0.0021	0.3%	0.08	1.00
10	Analyst Intensity	EW	-0.0073	-0.1%	0.02	1.00	26	Age	VW	0.0019	-0.1%	0.08	1.00
11	Lagged Momentum	EW	-0.0010	0.6%	0.02	1.00	27	Reverse Stock Split	VW	0.0048	-0.1%	0.08	1.00
12	ΔNC Op. Assets	VW	-0.0013	1.6%	0.02	1.00	28	ΔTax to Assets	EW	0.0051	2.2%	0.08	1.00
13	Price	VW	-0.0040	1.0%	0.02	1.00	29	ΔRec + Accrual	VW	-0.0030	1.5%	0.09	1.00
14	Size	EW	0.0000	0.3%	0.03	1.00	30	Capex Growth	VW	-0.0024	1.3%	0.09	1.00
15	Coskewness	VW	0.0055	1.4%	0.03	1.00	31	Max	EW	-0.0001	0.4%	0.09	1.00
16	Z-score	EW	-0.0039	1.4%	0.04	1.00	32	Momentum-reversal	VW	-0.0009	0.5%	0.10	1.00

Table IA.IV - Continued

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
		EW or			Raw	RW			EW or			Raw	RW
Rank	Predictor	VW	$\widehat{\beta}$	R^2	P-Value		Rank	Predictor	VW	$\widehat{\beta}$	R^2	P-Value	
Panel C: 6 Month Horizon													
1	Misvalue Innovation	VW	-0.0027	2.9%	0.04	0.72	22	LT Reversal	EW	-0.0028	1.5%	0.03	1.00
2	Short Interest	EW	-0.0050	15.8%	0.01	0.87	23	Repurchases	EW	-0.0043	7.0%	0.03	1.00
3	Cash Flow Variance	EW	0.0018	2.4%	0.03	0.90	24	Exchange Switch	VW	-0.0041	-1.9%	0.04	1.00
4	ST Reversal	EW	0.0000	0.3%	0.00	0.95	25	Short Interest	VW	0.0047	4.2%	0.04	1.00
5	Z-score	VW	-0.0050	7.4%	0.01	0.97	26	Reverse Stock Split	VW	0.0043	-1.3%	0.04	1.00
6	Idio Risk	VW	-0.0020	1.6%	0.00	0.99	27	Spreads	VW	0.0080	-2.3%	0.04	1.00
7	Size	VW	0.0004	0.4%	0.00	1.00	28	Idio Risk	EW	-0.0001	1.3%	0.04	1.00
8	ST Reversal	VW	0.0003	0.3%	0.00	1.00	29	Momentum-reversal	EW	-0.0011	1.2%	0.04	1.00
9	Asset Turnover	VW	0.0075	9.6%	0.00	1.00	30	Age	VW	0.0019	-0.1%	0.05	1.00
10	Moment LT Reverse	VW	0.0008	1.2%	0.00	1.00	31	Coskewness	VW	0.0046	1.4%	0.05	1.00
11	Momentum-reversal	VW	-0.0024	2.2%	0.00	1.00	32	Dividends	EW	0.0000	1.0%	0.06	1.00
12	Price	VW	-0.0036	1.9%	0.01	1.00	33	Z-score	EW	-0.0041	1.9%	0.06	1.00
13	Max	EW	-0.0014	0.8%	0.01	1.00	34	ΔCapex-ΔInd Capex	VW	-0.0037	0.0%	0.06	1.00
14	Max	VW	-0.0024	1.5%	0.01	1.00	35	Price	EW	-0.0023	0.5%	0.06	1.00
15	Size	EW	0.0005	0.2%	0.01	1.00	36	Dividends	VW	0.0000	1.0%	0.06	1.00
16	Lagged Momentum	VW	-0.0015	1.9%	0.01	1.00	37	ΔCapex-ΔInd Capex	EW	-0.0038	-0.1%	0.06	1.00
17	LT Reversal	VW	-0.0064	1.2%	0.01	1.00	38	Capex Growth	VW	-0.0026	3.1%	0.07	1.00
18	Spreads	EW	0.0120	-1.6%	0.02	1.00	39	ΔSales-ΔInventory	EW	-0.0042	-4.3%	0.08	1.00
19	ΔNC Op. Assets	VW	-0.0009	3.3%	0.02	1.00	40	Asset Growth	VW	0.0003	1.0%	0.09	1.00
20	Lagged Momentum	EW	-0.0012	1.6%	0.02	1.00	41	Analyst Intensity	EW	-0.0043	-0.7%	0.09	1.00
21	Cash Flow Variance	VW	0.0057	0.3%	0.03	1.00							
Panel D: 12 Month Horizon													
1	Misvalue Innovation	VW	-0.0029	6.1%	0.04	0.62	24	Idio Risk	VW	0.0005	-0.8%	0.01	1.00
2	Short Interest	EW	-0.0049	20.6%	0.02	0.70	25	Capex Growth	VW	-0.0034	7.6%	0.02	1.00
3	Cash Flow Variance	EW	0.0013	4.4%	0.02	0.80	26	Max	EW	-0.0003	-0.3%	0.03	1.00
4	Z-score	VW	-0.0053	15.0%	0.03	0.94	27	Dividends	EW	0.0000	1.7%	0.03	1.00
5	Moment LT Reverse	VW	0.0013	4.0%	0.00	0.96	28	Dividends	VW	0.0000	1.7%	0.03	1.00
6	Moment LT Reverse	EW	0.0009	3.9%	0.00	0.99	29	ΔNC Op. Assets	VW	-0.0008	7.4%	0.03	1.00
7	Short Interest	VW	0.0026	5.6%	0.02	0.99	30	Age	VW	0.0021	-1.2%	0.03	1.00
8	Repurchases	EW	-0.0044	11.2%	0.04	0.99	31	Exchange Switch	VW	-0.0040	-2.3%	0.04	1.00
9	Lagged Momentum	VW	-0.0020	4.4%	0.00	1.00	32	Momentum-reversal	EW	-0.0006	1.1%	0.05	1.00
10	Size	VW	0.0009	-0.2%	0.00	1.00	33	Sustainable Growth	VW	-0.0021	2.3%	0.05	1.00
11	Reverse Stock Split	VW	0.0028	-2.5%	0.00	1.00	34	ΔCapex-ΔInd Capex	EW	-0.0025	2.5%	0.05	1.00
12	Lagged Momentum	EW	-0.0011	2.9%	0.00	1.00	35	ΔCapex-ΔInd Capex	VW	-0.0030	0.4%	0.07	1.00
13	Price	VW	-0.0047	3.0%	0.00	1.00	36	Cash Flow Variance	VW	0.0062	-0.8%	0.07	1.00
14	Momentum-reversal	VW	-0.0020	2.7%	0.00	1.00	37	Org. Capital	EW	0.0027	4.6%	0.07	1.00
15	Spreads	EW	0.0096	1.5%	0.00	1.00	38	Volume / MV	VW	-0.0012	0.0%	0.09	1.00
16	Max	VW	-0.0013	1.1%	0.00	1.00	39	Price	EW	-0.0030	-0.8%	0.09	1.00
17	Asset Turnover	VW	0.0079	17.5%	0.00	1.00	40	Volume Trend	EW	0.0031	2.6%	0.09	1.00
18	ST Reversal	VW	0.0011	-1.0%	0.01	1.00	41	Momentum	EW	0.0007	-1.6%	0.09	1.00
19	ST Reversal	EW	0.0010	-1.0%	0.01	1.00	42	Asset Growth	VW	0.0001	3.1%	0.09	1.00
20	Spreads	VW	0.0067	1.6%	0.01	1.00	43	ΔSales-ΔInventory	EW	-0.0024	-4.6%	0.09	1.00
21	Size	EW	0.0010	-1.3%	0.01	1.00	44	Coskewness	VW	0.0039	0.2%	0.09	1.00
22	LT Reversal	VW	-0.0052	6.1%	0.01	1.00	45	% Operating Accrual	VW	-0.0046	9.8%	0.10	1.00
23	LT Reversal	EW	-0.0025	3.6%	0.01	1.00							

Table IA.V
Predictor Database and Sample Construction Methodology

This table provides the author (s), original sample period, and construction of the 140 predictors that we consider in our study. Panel A contains the 97 return predictors studied in McLean and Pontiff (2016) and Engelberg, McLean, and Pontiff (2017). Panel B contains the predictors that are new in this study, and not used in the two earlier papers. The data used are from the CRSP, Compustat, IBES, 13F databases, and Thomson Reuters databases.

Predictor	Author(s) (Journal and Year of Publication)	Original Sample Period	Construction
Panel A: Predictors from McLean and Pontiff (2016) and Engelberg, McLean, and Pontiff (2018)			
52-Week High	George and Hwang (JF 2004)	1963-2001	Price scaled by the highest price or bid/ask average during the last 12 months. Updated monthly.
Accruals	Sloan (AR 1996)	1962-1991	Accruals = ((ΔCA - ΔCash) - (ΔCL - ΔSTD - ΔTP) – Dep) / Average Assets; ΔCA = Change in Current Assets; ΔCash = Change in Cash and Cash Equivalents; ΔCL = Change in Current Liabilities; ΔSTD = Change in Debt Included in Current Liabilities; ΔTP = Change in Income Taxes Payable; and Dep = Depreciation and Amortization Expense. Average Assets = (Assets + Assets (t-1))/2 Updated annually. Exclude if price<$5.
Advertising / Market Value of Equity	Chan et al. (2001)	1975-1996	Advertising expenses scaled by market value of equity. Updated annually.
Amihud's Measure (Illiquidity)	Amihud (JFM 2002)	1964-1997	Daily absolute value of return scaled by dollar trading volume on the same day. Averaged over the previous year. Updated monthly.
Analyst Value	Frankel and Lee (JAE 1998)	1975-1993	Using Compustat annual data, at each fiscal year-end we add half of the current book value per share to half of the book value per share from the previous year. For the first year with Compustat data, we simply use the current book value per share. The book value per-share measure is divided by Common Shares Used to Calculate Earnings Per Share (Basic), creating a per share measure. In non-fiscal year-ends, each month's value is the value from the last fiscal year-end. For each month, we calculate an expected return on equity by using the most recent I/B/E/S mean earnings estimate and dividing it by the book value measure from six months ago. Our analyst forecast measure is: $\left(1+\left(\frac{FROE_t-0.1}{1.1}\right)+\frac{FROE_t-0.1}{0.1X1.1}\right)\left(\frac{AvgBook_{t-6}}{P_t}\right)$

			FROE is expected return on equity, AvgBook is the average book value, and P is the current stock price.
Asset Growth	Cooper, Guylen, and Schill (JF 2008)	1968-2003	Yearly percentage change in total assets. Updated annually.
Asset Turnover	Soliman (AR 2008)	1984-2002	Asset Turnover (*t*) = Sales (*t*)/ ((Net Operating Assets (*t*) + Net Operating Assets (*t*-1)) / 2); Net Operating Assets = Receivables + Total Inventory + Other Current Assets + PP&E + Intangibles - Payables - Other Current Liabilities - Other Liabilities. Updated annually.
Beta	Fama and MacBeth (JPE 1973)	1926-1968	Beta with respect to the CRSP equal-weighted return index. Estimated over the past 60 months. Updated monthly.
Bid-Ask Spread	Amihud and Mendelsohn (1986)	1961-1980	Estimates from Shane Corwin's "Monthly High-Low Spread Estimates 1926-2013" dataset. See http://www3.nd.edu/~scorwin/ . We divide each estimate by the particular stock's month-end price.
Book Equity / Market Equity	FF (JF 1992)	1963-1990	The log of: book value of equity scaled by market value of equity. Updated annually.
Cash Flow / Market Value of Equity	LSV (JF 1994)	1968-1990	Net income plus depreciation and amortization, all scaled by market value of equity. Updated annually.
Cash Flow Variance	Haugen and Baker (JFE 1996)	1979-1993	Variance of the monthly ratio of cash flow to market value of equity. Cash flow is net income plus depreciation and amortization, all scaled by market value of equity. Measured over the last 60 months. Updated monthly.
Change in Asset Turnover	Soliman (AR 2008)	1984-2002	Asset Turnover (*t*) - Asset Turnover *(t*-1). Asset Turnover is defined above. Updated annually.
Change in Forecast + Accrual	Barth and Hutton (RAS 2004)	1981-1996	This variable is equal to one if the firm has an increase in the mean analysts' forecast and is and below the median for the accruals variable. The variable is equal to minus one if the firm has a decrease in the mean analysts' forecast and is above the median for the accruals variable. The analysts' forecasts and accruals variables are defined above.
Change in Profit Margin	Soliman (AR 2008)	1984-2002	Profit Margin (*t*) - Profit Margin (*t*-1). Updated annually. Data from year *t* are used to forecast returns for 12 months beginning in April of year *t*+1.
Change in Recommendation	Jegadeesh et al. (JF 2004)	1985-1998	Using I/B/E/S data, when an analyst assigns a new strong buy recommendation, we assign a value of one to that analyst's change variable. When the analyst assigns new

			recommendation that is not a strong buy, we assign a value of minus one to the analyst's change variable. Cases in which the analyst's change variable is neither one nor minus one, we assign it a value of zero. Each month we average the change variable over all analysts who follow the stock, and use this average as our change in recommendation variable.
Coskewness	Harvey and Siddique (JF 2000)	1963-1993	Calculated from a rolling window using data from 60 months ago until one month ago. For each stock, we regress the return of the stock on the return of the CRSP value-weighted index (VWRETD). Each month a variable is created by multiplying the residual from this regression by the deviation of the squared VWRETD minus the average squared VWRETD, and dividing by the product of the regression's root mean square error times the average of VWRETD. The average of the aforementioned variable is Coskewness.
Credit Rating Downgrade	Dichev and Piotroski (JF 2001)	1970-1997	Binary variable equal to one if S&P credit rating decreased during the previous year. Updated annually.
Debt Issuance	Spiess and Affleck-Graves (JFE 1999)	1975-1989	Binary variable equal to one if long-term debt issuance indicated in statement of cash flow. Updated annually.
Dividend Initiation	Michaely, Thaler, and Womack (1995)	1964-1988	Binary variable equal to one if a firm initiates a dividend during the previous 12 months, and did not pay a dividend during any of the 24 months preceding the month in which the dividend was initiated. NYSE firms only. Exclude if CRSP SHRCD > 11. Updated monthly.
Dividend Omission	Michaely, Thaler, and Womack (1995)	1964-1988	Using firms with at least 18 months of data, we identify stocks that paid dividends 3, 6, 9, 12, 15, and 18 months ago, but did not pay a dividend in the current month, or the two months before the current month. These firms are assigned a dividend ommitter indicator equal to one for the current month and the next 11 months. All other observations are assigned a value of zero.
Dividend Yield	Naranjo, Nimalendran, and Ryngaert (JF 1998)	1963-1994	Most recent dividend multiplied by four, all scaled by price. The firm had to pay a dividend during each of the last four quarters, and had to have not undergone an exchange or reorganization in the last year (CRSP distribution code = 3).

Dividends, Anticipated	Hartzmark and Solomon (JFE 2013)	1927-2011	We create an indicator variable that is equal to one if 11 months ago the difference between the monthly CRSP return with dividends (RET) is greater than the CRSP return without dividends (RETX), or if two months ago RET>RETX. The indicator variable is set to zero if this condition does not hold, or if the stock was not included in the CRSP data 11 months ago.
Down Forecast	Barber et al. (JF 2001)	1985-1997	Using I/B/E/S data, if the mean analyst earnings forecast decreases, we assign Down Forecast a value of one. If it stays the same or increases, we assign it a value of zero.
Earnings / Price	Basu (JF 1977)	1964-1971	Net income scaled by market value of equity. Sample is limited to NYSE firms. Updated annually.
Earnings Growth	Alwathainani (BAR 2009)	1971-2002	Geometric average of Earnings Growth from *t*-1 to *t*-5. Earnings growth is: Earnings Per Share (*t*) – Earnings Per Share (*t*-1) / ((Absolute Value of Earnings Per Share (*t*-1) + Absolute Value of Earnings Per Share (*t*-2) / 2). Exclude if: price < $5; absolute value of growth > 6; if growth is positive this year, but negative last year, or vice versa. Updated annually.
Earnings Surprise	Foster, Olsen, and Shevlin (AR 1984)	1974-1981	(Earnings per Share_t - Earnings per Share_{t-4} - Drift) all scaled by the standard deviation of this difference over the last eight quarters. Drift is the average yearly growth in earnings per share over the last eight quarters. Updated quarterly.
Enterprise Component of Book/Price	Penman, Richardson, and Tuna (JAR 2007)	1961-2001	The B/P ratio can be decomposed into an enterprise book-to-price (that pertains to operations and potentially reflects operating risk) and a leverage component (that reflects financing risk). EBP = (BV+ND)/(ND+MV); BV = Book Value of Equity; ND = Cash - Long-Term Debt - Debt in Current Liabilities - Preferred Stock - Preferred Dividends in Arrears + Preferred Treasury Stock; MV = Market Value of Equity. Updated annually.
Enterprise Multiple	Loughran and Wellman (JFQA 2011)	1963-2009	Enterprise Value / Operating Cash Flow. Enterprise Value = Market Value of Equity + Long-Term Debt + Debt in Current Liabilities + Preferred Stock – Cash and Short-Term Equivalents. Updated annually.
Exchange Switch	Dharan and Ikenberry (JF 1995)	1962-1990	Binary variable equal to one if during the past year a firm switched to AMEX from NASDAQ, or to NYSE from either AMEX or NASDAQ. Updated monthly.
Firm Age	Barry and Brown (JFE 1984)	1931-1982	The number of months that a firm has been listed in the CRSP database. Updated monthly.
Firm Age-Momentum	Zhang (JF 2004)	1983-2001	Buy-and-hold returns from *t*-6 through *t*-1. Exclude if age < 12. Firms are then sorted on age, and only firms in the bottom age quintile are included. Updated monthly.

Forecast Dispersion	Diether, Malloy, and Scherbina (JF 2002)	1976-2000	Computed by dividing the I/B/E/S standard deviation of analyst earnings forecasts by the absolute value of the average forecast. If the average forecast is zero, we do not use the observation.
G Index	Gompers, Ishi, and Metrick (2003)	1990-1998	The sum of 21 corporate governance variables. A higher value of the index suggests worse corporate governance. The governance data is from Investor Responsibility Research Center (IRRC). The index can be downloaded from Andrew Metrick's website: http://faculty.som.yale.edu/andrewmetrick/data.html.
Gross Profitability	Novy-Marx (JFE 2013)	1962-2010	(Revenue(*t*) - Cost of Goods Sold (*t*)) / Assets (*t*-1). Updated annually.
Inventory to Assets, Change	Thomas and Zhang (RAS 2002)	1970-1997	Growth in Inventory. (Inventory (t) - Inventory (t-1)) / (Assets (t) + Assets (t-1))/2. Inventory is total inventory (invt). Updated annually.
Growth in LTNOA	Fairfield, Whisenant, and Yohn (AR 2003)	1964-1993	Growth in Net Operating Assets minus Accruals. GRLTNOA=GRNOA-ACC; NOA=(RECT+INVT+ACO+PPENT+INTAN+AO-AP-LCO-LO)/AT; GRNOA=NOA-NOA(*t*-1) ACC=((RECT-RECT(*t*-1))+(INVT-INVT(*t*-1))+(ACO-ACO(*t*-1))-(AP-AP(*t*-1))-(LCO-LCO(*t*-1))-DP)/((AT+AT(*t*-1))/2); RECT = Receivables; INVT = Total Inventory; ACO = Current Assets; AP =Accounts Payable; LCO = Current Liabilities (Other); DP = Depreciation and Amortization; AT = Assets; PPENT = Property, Plant, and Equipment (net); INTAN = Intangible Assets; AO = Assets (Other); LO = Liabilities (Other). Updated annually.
F-Score	Piotroski (AR 2000)	1976-1996	Measure ranging from zero to nine based on the sum of the following dummy variables: one if net income > zero; one if cash flow from operations > zero; otherwise; one if return on assets (net income scaled by assets) increased during the previous year; one if cash flow from operations > net income; one if the ratio of long-term debt to total assets decreased during the previous year; one if the ratio of current assets to current liabilities increased during the previous year; one if the firm did not issue common shares; one if the ratio of EBIT to revenues increased during the previous year; and one if the ratio of revenues to assets increased during the previous year. Limit sample to firms in the highest book-to-market quintile. Updated annually.
G-Score	Mohanram (RAS 2005)	1978-2001	First, the sample is limited to firms in the lowest book-to-market quintile. Then, a measure ranging from zero to eight based on the sum of the following dummy variables: one if net income scaled by assets > industry (two-digit SIC code) median; one if cash flow scaled by assets > industry-median; one if cash flow from operations > net income; one if net income variability < median firm in the same industry; one if revenue variability is less that median firm in the same industry; one if capital

			expenditures scaled by assets > industry median; one if research and development expenditures scaled by assets > industry median; and one if advertising expenditures scaled by assets > industry median. Revenue and net income variability are both measured over the previous four quarters. Updated quarterly. Data from year *t* are used to forecast returns for 12 months beginning in April of year *t*+1.
Herfindahl Index	Hou and Robinson (JF 2006)	1963-2001	Herfindahl indices are constructed within each 3-digit SIC code, and then averaged over the last three years. Regulated industries, as defined by Barclay and Smith (1995), are not included. Index values from December are used to predict returns beginning in July of the subsequent year. Updated annually.
Idiosyncratic Risk	Ang et al. (JF 2006)	1963-2000	The standard deviation of the residual from a monthly, firm-level regression of daily stock returns on the daily innovations of the Fama and French three-factor model. Returns are market value-weighted.
Industry Momentum	Grinblatt and Moskwotiz (1999)	1963-1995	Value-weighted return from *t*-6 to *t*-1 within each industry. Industry is measured with two-digit SIC code. Updated monthly.
Initial Public Offering	Ritter (JF 1991)	1975-1984	We use Jay Ritter's data set "Founding dates for 9,902 IPOs from 1975-2016," which is available on his website, http://bear.warrington.ufl.edu/ritter/ipodata.htm. We consider IPO firms as firms that have had an IPO within the last 36 months.
Investment	Titman, Wei, and Xie (JFQA 2004)	1973-1996	CAPEX scaled by revenues, all scaled by the average CAPEX / revenues from the previous three years. Updated annually.
IPO and Age	Ritter (JF 1991)	1975-1984	We use Jay Ritter's dataset "Founding dates for 9,902 IPOs from 1975-2014," which is available on his website http://bear.warrington.ufl.edu/ritter/ipodata.htm. We consider IPO firms as firms that have had an IPO within the last 36 months. We sort all IPO stocks on age, and measure the difference in returns between the young and old IPO stocks. Young and old are defined as the highest and lowest age quintiles. We require at least 30 observations in a cross-section to estimate the effect.
IPO no R&D	Gou, Lev, and Shi (JBFA 2006)	1980-1995	This variable is equal to one if the firm is an IPO firm and had no R&D spending. We use Jay Ritter's dataset "Founding dates for 9,902 IPOs from 1975-2014," which is available on his website http://bear.warrington.ufl.edu/ritter/ipodata.htm. We consider IPO firms as firms that have had an IPO within the last 36 months. Updated monthly.
Lagged Momentum	Novy-Marx (JFE 2012)	1926-2010	Buy-and-hold returns from *t*-13 through *t*-8. Updated monthly.
Leverage	Bhandari (JFE 1988)	1946-1981	Log of long-term debt scaled by market value of equity. Updated annually.

Leverage Component of Book/Price	Penman, Richardson, and Tuna (JAR 2007)	1961-2002	The B/P ratio can be decomposed into an enterprise book-to-price (that pertains to operations and potentially reflects operating risk) and a leverage component (that reflects financing risk). BPEBP=BP-EBP; EBP is defined above; BP =(BV+TSTKP-DVPA)/MV; BV = Book Value of Equity; TSTKP = Preferred Treasury Stock; DVPA = Preferred Dividends in Arrears; MC = Market Value of Equity. Updated annually.
Long-Term Reversal	Debondt and Thaler (JF 1985)	1926-1982	Buy and hold returns from *t*-60 to *t*-13. Updated monthly.
M/B and Accruals	Bartov and Kim (RQFA 2004)	1980-1998	Equal to one if both low book-to-market and high accrual quintiles; minus one if both high book-to-market and low accrual quintiles, and zero otherwise. Accruals are defined above by Sloan (1996). We exclude firms with negative book values of equity. Updated annually.
Max	Bali, Cakici, and Whitelaw (JF 2010)	1962-2005	Maximum daily return over the past month. Updated monthly.
Mergers	Langetieg (JFE 1978)	1929-1969	Binary variable equal to one if the CRSP variable *acperm* is less than 1000 in any of the last 12 months.
Momentum	Jegadeesh and Titman (JF 1993)	1964-1989	Buy-and-hold returns from *t*-6 to *t*-1. Updated monthly.
Momentum and LT Reversal	Chan and Kot (JOIM 2006)	1965-2001	Equal to one if both Momentum Winner and Reversal Loser, minus one if both Momentum Lose and Reversal Winner, and zero otherwise. Momentum and Long-Term Reversal are defined above (Jegadeesh and Titman (1993) and Debondt and Thaler (1985). Winners are in the top quintile, losers are in the bottom quintile. Updated monthly.
Momentum-Credit Ratings	Aramov et al. (JF 2007)	1985-2003	Buy-and-hold returns from *t*-6 through *t*-1. Sample is limited to firms with S&P credit ratings of BBB+ or lower. Updated monthly.
Momentum-Reversal	Jegadeesh and Titman (JF 1993)	1964-1989	Buy and hold returns from *t*-18 to *t*-13. Updated monthly.
Momentum-Volume	Lee and Swaminathan (JF 2000)	1965-1995	Buy- and- hold returns from *t*-6 through *t*-1. We limit the sample to high trading volume stocks, i.e., stocks in the highest quintile of average monthly trading volume measured over the past six months. 1. Updated monthly.
Net Operating Assets	Hirshleifer et al. (JAE 2004)	1964-2002	NOA_t = (Operating $Assets_t$ - Operating $Liabilities_t$) / Total $Assets_{t-1}$; Operating Assets = Total Assets - Cash and Short-Term Investment; Operating Liabilities = Total Assets – Current Portion of Long-Term Debt - Long-Term Debt - Minority Interest (Balance Sheet) - Preferred Stock - Common Equity. Updated annually.

Net Working Capital Changes	Soliman (AR 2008)	1984-2002	Yearly change in net working capital scaled by total assets. Net working capital is measured as current assets minus current liabilities. Current assets are measured as total current assets minus cash and cash equivalents. Current liabilities are measured as total current liabilities minus debt in current liabilities.
Noncurrent Operating Assets Changes	Soliman (AR 2008)	1984-2002	Yearly change in noncurrent operating assets scaled by total assets. Noncurrent operating assets are defined as noncurrent assets minus noncurrent liabilities. Noncurrent assets are total assets minus current assets and investment and advances. Noncurrent liabilities are total liabilities minus current liabilities and long-term debt. Updated annually.
Operating Leverage	Novy-Marx (ROF 2010)	1963-2008	SG&A + Cost of Goods Sold, all scaled by Total Assets.
Org. Capital	Eisfeldt and Papanikolaou (JF 2013)	1970-2008	Using Compustat annual data, we assign Selling, General and Administrative Expense (SG&A) values that are missing to zero. For the first year a company appears in Compustat, we assign a beginning value of organizational capital by using four times the original value of SG&A. At the end of each fiscal year, we multiply the last year's organizational capital by 0.85 and add SG&A. This value is divided by total assets.
O-Score (More Financial Distress)	Dichev (JFE 1998)	1981-1995	O-Score = –1.32 – 0.407*log(Total Assets/GNP Price-Level Index) + 6.03*(Total Liabilities/Total Assets) – 1.43*(Working Capital/Total Assets) + 0.076*(Current Liabilities/Current Assets) –1.72*(if Total Liabilities > Total Assets, else 0) – 2.37*(Net Income/Total Assets) – 1.83*(Funds from Operations/Total Liabilities) + 0.285*(1 if net loss for the last two years, else 0) – 0.521*(Net Income$_t$– Net Income$_{t-1}$)/(|Net Income$_t$| + |Net Income$_{t-1}$|). Updated annually. SIC codes 1 to 3999 and 5000 to 5999 only. Data from year *t* are used to forecast returns for 12 months beginning in April of year *t*+1.
Pension Funding Status	Franzoni and Marin (JF 2006)	1980-2002	FR=(FVPA-PBO)/Market Value of Equity; Used Compustat Items: pbnaa: Pension Benefits – Net Assets pbnvv: Pension Benefits –Present Value of vested interests pplao: Pension Plan Assets pplau: Pension Plan Assets (Underfunded) pbpro: Pension- Projected Benefit Obligation pbpru: Pension- Projected Benefit Obligation (Underfunded) 1980-1986: FVPA=pbnaa; PBO=pbnvv; 1987-1997: FVPA=pplao+pplau; PBO:=pbpro+pbpru; 1998 onwards: FVPA=pplao; PBO=pbpro. Exclude if CRSP SHRCD > 11 5. Updated annually.

Percent Operating Accrual	Hafzalla, Lundholm, and Van Winkle (AR 2011)	1989-2008	(Net Income - Cash Flow from Operations) / Absolute Value of Net Income. Updated annually.
Percent Total Accrual	Hafzalla, Lundholm, and Van Winkle (AR 2011)	1989-2008	Net Income - ((-Sale of Common and Preferred Stock + Purchase of Common and Preferred Stock + Total Dividends + Cash Flow from Operations + Cash Flow from Financing + Cash Flow from Investment) / Absolute Value of Net Income). Updated annually. Exclude if price<$5.
Price	Blume and Husic (JF 1972)	1932-1971	Log of stock price. Updated monthly.
Profit Margin	Soliman (AR 2008)	1984-2002	EBIT / Revenues. Updated annually.
Profitability	Karthik, Bartov, and Faurel (JAE 2010)	1976-2005	Quarterly Earnings per Share (*t*) x Number of Shares used to compute earnings per share (*t*), all scaled by Assets (*t*-1). Updated quarterly. Exclude if price<$1.
Seasoned Equity Offerings	Loughran and Ritter (JF 1995)	1975-1984	Binary variable equal to one if proceeds from seasoned equity offergins exceed 5% of assets, and zero otherwise.
R&D / Market Value of Equity	Chan et al. (2001)	1975-1995	R&D expenses scaled by market value of equity. Updated annually.
Return-on-Equity	Haugen and Baker (JFE 1996)	1979-1993	Net income scaled by book value of equity. Updated annually.
Revenue Surprises	Jegadeesh and Livnat (JAE 2006)	1987-2003	(Revenue per Share_t - Revenue per Share_{t-4} - Drift) all scaled by the standard deviation of this difference measured over the last 8 quarters. Drift is the average yearly growth in revenue per share over the last 8 quarters. Updated quarterly. Exclude if price<$5.
Sales Growth	LSV (JF 1994)	1968-1990	Average revenue growth rank over the past five years. Firms are ranked every year based on revenue growth. The average rank is then taken, giving growth in year *t*-1 a weight of 5, year *t*-2 a weight of 4, year *t*-3 a weight of 3, year *t*-4 a weight of 2, and year *t*-5 a weight of 1. NYSE and AMEX only. Updated annually.
Sales/Price	Barbee et al (FAJ - 1996)	1979-1991	Total revenues divided by stock price. Updated annually.
Seasonality	Heston and Sadka (JFE 2008)	1965-2002	Average monthly return in the same month over the last 20 years. As an example, the average return from prior Octobers is used to predict returns this October. The firm needs at least one year of data to be included in the sample. NYSE and AMEX only. Updated monthly.
Share Issuance (1-Year)	Pontiff and Woodgate (JF 2008)	1970-2003	Change in real number of shares outstanding from *t*-18 to *t*-6. Excludes changes in shares due to stock dividends and splits.

Share Issuance (5-Year)	Daniel and Titman (JF 2006)	1968-2003	Five-year real change in number of shares outstanding. Excludes changes in shares due to stock dividends and splits. Updated monthly.
Share Repurchases	Ikenberry, Lakonishok, and Vermaelen (JFE 1995)	1980-1990	Binary variable equal to one if repurchase of common or preferred shares indicated in statement of cash flow. Updated annually.
Share Volume	Datair, Naik, and Radcliffe (JFM 1998)	1962-1991	Average number of shares traded over the previous three months scaled by shares outstanding. Drop observations for which shares outstanding changed over the last three months. NYSE Only. Updated monthly.
Short Interest	Dechow et al. (2001)	1976-1993	Shares Shorted / Shares Outstanding. Updated monthly.
Short-Term Reversal	Jegadeesh (1989)	1934-1987	Return in month t. Updated monthly.
Size	Banz (JFE 1981)	1926-1975	The log of market value of equity. Updated monthly.
Spinoffs	Cusatis, Miles, and Wooldridge (JFE 1993)	1965-1988	Spinoffs are identified by the variable ACPERM in CRSP. The variable indicates a spinoff during months *t*-1 to *t*-12, and predicts returns in month *t*+1.
Sustainable Growth	Lockwood and Prombutr (JFR 2010)	1964-2007	Growth in book value of equity. BE_t / BE_{t-1} - 1. Updated annually.
Tax	Lev and Nissim (AR 2004)	1973-2000	Income Tax scaled by Net Income. Following Lev and Nissim we make the following adjustments: Tax = Tax/0.48 if year>= 1973 & year<=1978 replace Tax = Tax/0.46 if year>= 1979 & year<=1986 replace Tax = Tax/0.40 if year==1987 replace Tax = Tax/0.34 if year>= 1988 & year<=1992 replace Tax = Tax/0.35 if year>= 1993. Updated annually.
Total XFIN	Bradshaw, Richardson, and Sloan (JAE 2006)	1971-2000	Total net external financing (Net Share Issuance + Net Debt Issuance - Cash Dividends) scaled by Total Assets. XFIN=(SSTK-DV-PRSTKC+DLTIS-DLTR)/AT; SSTK = Sale of Common and Preferred Stock; DV = Cash Dividends; PRSTKC = Purchase of Common and Preferred Stock; DLTIS = Sale of Long-Term Debt; DLTR = Purchase of Long-Term Debt. Updated annually.
Unexpected R&D Increases	Eberhart, Maxwell, and Siddique (JF 2004)	1974-2001	Binary variable equal to one if: both research and development scaled by revenue and R&D scaled by assets are greater than zero; the yearly percentage change in R&D expenditures is greater than 5%; and R&D scaled by assets increased by more than 5%. Data from year *t* are used to forecast returns for 12 months beginning in April of year *t*+1.

Up Forecast	Barber et al. (JF 2001)	1985-1996	Using I/B/E/S data, if the mean analyst earnings forecast increases, we assign Up Forecast a value of one. If it stays the same or decreases, we assign it a value of zero.
Volume / Market Value of Equity	Haugen and Baker (JFE 1996)	1979-1993	Monthly average dollar trading volume over the past 12 months scaled by shares outstanding. Updated monthly.
Volume Trend	Haugen and Baker (JFE 1996)	1979-1993	Five-year trend in monthly trading volume scaled by average trading volume during the same five-year period. Exclude NASDAQ, CRSP SHRCD > 11. Updated monthly.
Volume Variance	Chordia, Subranhmanyam, and Anshuman (JFE 2001)	1966-1995	Standard deviation of monthly trading volume over the last 36 months. NYSE only. Updated monthly.
Z-Score (Less Financial Distress)	Dichev (JFE 1998)	1981-1995	Z-Score = 1.2*(Working Capital / Assets) + 1.4*(Retained Earnings / Assets) + 3.3*(EBIT / Assets) + 0.6*(Market Value of Equity / Book Value of Total Liabilities) + (Revenues / Assets). NYSE only. Updated annually. SIC codes 1 to 3999 and 5000 to 5999 only. Data from year *t* are used to forecast returns for 12 months beginning in April of year *t*+1.
ΔCAPEX-ΔIndustry CAPEX	Abarbanell and Bushee (AR 1998)	1974-1988	ΔCAPEX = CAPEX at time *t* minus the average value of capex from *t*-1 and *t*-2, all scaled by the average value of CAPEX from t-1 and t-2. ΔIndustry CAPEX is the two-digit SIC industry average of ΔCAPEX. Updated annually.
ΔSales-ΔInventory	Abarbanell and Bushee (AR 1998)	1974-1988	ΔSales = Sales at time *t* minus the average value of sales from *t*-1 and *t*-2, all scaled by the average value of sales from *t*-1 and *t*-2. Inventory is computed using total inventories. Updated annually.
ΔSales-ΔSG&A	Abarbanell and Bushee (AR 1998)	1974-1988	ΔSales = Sales at time *t* minus the average value of sales from *t*-1 and *t*-2, all scaled by the average value of sales from *t*-1 and *t*-2. ΔSG&A is computed similarly. Updated annually.
Panel B: New Predictors not in McLean and Pontiff (2016) and Engelberg, McLean, and Pontiff (2018)			
Agnostic Value	Bartram and Grinblatt (RFS, 2017)	1987-2012	Agnostic value is the ratio of a cross-sectional, monthly regression residual to the firm's market value. The dependent variable for the regression is market value. The independent variables are quarterly Compustat variables. When quarterly variables are unavailable, we use annual variables. Before and including 1955, the right-hand side variables are Compustat annual variables: AT, ICAPT, PPENT, PSTK, DLTT, AO, LT, LO, CHE, ACO, DVP, SALE, IB, NI, XIDO, IBADJ, IBCOM, PI, TXT, and NOPI. For months after and including 1960 the regression also includes SEQ. For months after and including 1965 the regression also includes CEQ. For months

			after and including 1975, the regression also includes LCO and AP. For months after and including 1980, the regression also includes PSTKR. For months after and including 2010, the regression also includes TEQ.
Analyst Coverage	Barber, Lehavy, McNichols, and Trueman (JF, 2001)	1985-1996	An indicator variable equal to one if a firm is covered by one or more analysts. All other firms receive a value of zero.
Analyst Intensity, Abnormal	Lee and So (JFE, 2017)	1982-2007	The residual from the monthly cross-sectional regression on TOT on SIZE, MOMENTUM and TURNOVER. TOT is the natural log of one plus the number of unique earnings forecasts during the current month and the last two months. TOT includes all analysts and all earnings periods. Size is the natural log of market capitalization. MOMENTUM if the return of the stock over the last twelve months. TURNOVER is the trading volume over the last year, divided by shares outstanding. ATOT is calculated for all firms, even those without analyst coverage.
Book to Assets, Change	Richardson, Sloan, Solimon, and Tuna (JAE, 2005)	1962-2001	Book value of equity divided by total assets, minus the ratio of last year's book value of equity to last year's total assets.
Book Value to Assets, Change	Richardson, Sloan, Soliman, and Tuna (JAE, 2005)	1962-2001	Current year book value dividend by current year total assets, minus last year's book value dividend by last year's total assets.
Breadth of ownership, Change	Chen, Hong, and Stein (JFE, 2002)	1979-1998	The three-month change in the number of institutions that own shares in the firm.
Capex Growth	Anderson and Garcia-Feijoo (JF, 2006)	1976-1999	The difference between this year's capital expenditures and last year's capital expenditures, divided by the 2 times the sum of these measures.
Cash to Assets	Palazzo (JFE, 2012)	1975-1989	Cash and cash equivalents divided by total assets
Cash to Assets, Change	Palazzo (JFE, 2012)	1975-1989	Cash and cash equivalents divided by total assets minus the value of this ratio in the last year.
Complicated Firms	Cohen and Lou, (JFE, 2012)	1977-2009	Multi-segment firms are identified with Compustat Segment data as firms for which no individual segment comprises more than 80% of the sum of revenue from all segments. Using single segment returns we compute the SIC-two-digit average return over the last 6 months. The complicated return measure for each multi-segment firm is the sum of its segments' revenue-weighted industry 6-month return.

Continuing Overreaction	Byun, Lim, and Yun (JFQA, 2016)	1964-2009	We create a signed volume measure for each month that is equal to the number of shares trade if the return of the stock is positive, zero for months with zero returns, and negative volume for months with negative returns. Continuing Overreaction is computed by summing twelve months of signed volume, and dividing by the sum of twelve months of total volume.
Convertible Debt	Valta (JFQA, 2016)	1985-2012	Indicator variable equal to 1 of the firm has convertible debt, or zero otherwise.
Net Debt Issuance, Change	Bradshaw, Richardson, and Sloan (JAE, 2006)	1971-2000	
Earnings Forecast Anchoring	Cen, Hilary, and Wei (JFQA, 2013)	1983-2005	The difference between the average analyst's annual earnings forecast minus the median forecast of all firms in the same Fama-French 49 industry, divided by the median industry forecast
Earnings-to-Price, Forward Looking	Elgers, Lo, and Pfeiffer (TAR, 2001)	1982-1998	The average analyst forecasted annual earnings divided by current price.
Employee growth rate	Belo, Bazdresch, and Lin (JPE, 2014)	1965-2010	This year's number of employees minus last year's number of employees, divided by one half the sum of these quantities
Forecast Dispersion, Change	Cen, Wei, and Yang (MS, 2016)	1983-2009	The monthly change in our forecast dispersion variable.
Forecast Dispersion, Time Series	Kim and Na (JEF, 2016)	1986-2014	For each firm, we estimate a regression using the last 5 years of monthly returns. The mean annual earnings forecast is regressed on last month's mean annual earnings forecast. The residuals are modeled with a fourth-order moving average process. Time series dispersion is the standard deviation of the residuals. This measure is only computed if more than 24 months of data is available.
Hybrid Tail Covariance Risk	Bali, Cakici, and Whitelaw (RAPS, 2014)	1963-2012	Using the last 160 trading days, we determine the bottom 10th percentile of the stock's daily returns and the bottom 10th percentile of the returns of CRSP Equal-Weighted Index. Using data from the last 60 days for trading for which the stock return is in the bottom 10th percentile (as calculated above), we multiply the difference between the stock return and its 10th percentile level with the difference between the equal-weighted index return and its 10th percentile level. Hybrid Tail Covariance Risk is the sum of these products over the last 60 days.

Idiosyncratic Risk, Distant	Rachwalski and Wen (RAPS, 2016)	1966-2012	The standard deviation of the residual from a monthly, firm-level regression of daily stock returns on the daily innovations of the Fama and French three-factor model. Returns are market value-weighted. An average of individual months of is calculated for 11 months, after skipping the current month.
Insider Buy	Seyhun (JFE, 1986)	1975-1981	Each month we add the total number of shares bought by insiders and divide by shares outstanding. Insider Buy is the natural log of one plus the sum of this measure over 5 months. This measured is delayed such that there is a least a full month between the trading data and the return forecast period.
Insider Sell	Seyhun (JFE, 1986)	1975-1981	Each month we add the total number of shares sold by insiders and divide by shares outstanding. Insider Sell is the natural log of one plus the sum of this measure over 5 months. This measured is delayed such that there is a least a full month between the trading data and the return forecast period.
Institutional ownership	Gompers and Metrick (QJE, 2001)	1980-1996	Total number of shares owned by institutions, dividend by shares outstanding. This measure is not used to predict returns until the beginning of the second month after the reporting days.
Institutional ownership, Change	Nosfinger and Sias (JF, 1999)	1980-1996	The three-month change in institutional ownership.
Inventory Growth Rate	Belo and Lin (RFS, 2012)	1965-2009	This year's inventory minus last year's inventory, dividend by one half the sum of these quantities
Low Recommendation	Barber, Lehavy, McNichols, anmd Trueman (JF, 2001)	1985-1996	An indicator variable equal to one if the average analyst recommendation is less than 3 (where 5 is a strong buy and 1 is a strong sell). If the recommendation greater than or equal to 3, the indicator is set to zero.
Misvalued Innovation	Cohen, Diether, Malloy (RFS, 2013)	1980-2009	R&D intensity is measured as the natural log of one plus the ratio of annual R&D to sales. For each firm, we estimate 5 regressions of annual sales growth on various lags R&D intensity, where the lags vary from one to 5 years. The slope coefficients on R&D intensity are averaged for regressions with 6 or more observations. The Misvalued Innovation is measure is an indicator variable that is equal to one if R&D intensity is greater than the cross-sectional 25th percentile and the ratio of annual R&D to sales is greater than the cross-sectional 25th percentile. Firm with both measures, but not in the 25th percentiles of both receive a Misvalued Innovation value of zero.

Number of Analysts, Change	Scherbina (ROF, 2008)	1983-2005	The number of analysts who in the current month issued a recommendation for a stock sometime in the previous 12 months, minus the number of analysts two months ago who issued a recommendation in the previous 12 month. This difference is dividend by one half of the sum of these quantities.
Opportunistic Buys	Cohen, Malloy, and Pomorski (2012)	1989-2007	We classify an insider trade as opportunistic if the insider has a history of over two years of stock trades, and the insider did trade in the same month for the last two years. Opportunistic Buys is the number of opportunistic buys in a given firm-month.
Opportunistic Sells	Cohen, Malloy, and Pomorski (2012)	1989-2007	We classify an insider trade as opportunistic if the insider has a history of over two years of stock trades, and the insider did trade in the same month for the last two years. Opportunistic Sells is the number of opportunistic sells in a given firm-month.
Percentage change in the number of Institutional Owners	Guo and Qiu (CAR, 2016)	1982-2010	The number of institutional owners minus the number of institutional owners 3 months earlier, dividend by one half of the sum of these quantities.
Projected Earnings Release	Barber, De George, Lehavy, and Truman (JFE, 2013)	1990-2009	We design an indicator variable to predict whether a firm will announce earnings next month. The value of one is assigned if Compustat records that earnings were released 11 months ago. Firms that do not have an indicator equal to one, for which the announced earnings in the current month, or one of the last 4 months, are assigned a value of 0.
Real Estate Holdings	Tuzel (RFS, 2010)	1971-2005	The sum of Property, Plant and Equipment-Buildings at cost and Property, Plant, and Equipment Leases at Cost, divided by gross Property, Plan, and Equipment. If gross PPE is unavailable, Net PPE is used instead. The value is demeaned at the Fama-French 49 industry level.
Reverse Stock Split	Desai and Jain (JB, 1997)	1976-1991	An indicator equal to one, if a CRSP firm has a share adjustment factor that is negative, it is designated as a reverse split for 12 months or until another split or reverse split occurs. All other observations have a value of zero.
Secured to Total Debt	Valta (JFQA, 2016)	1985-2012	Secured Debt, Mortgages, and Other, divided by Total Long-Term Debt
Sin Stock	Hong and Kacperczyk (JFE, 2009)	1965-2006	An indicator variable equal to one if the firm's SIC is between 2100 and 2199, or between 2080 and 2085, or if its NAICS is equal to either 7132, 71312, 713210, 713290, 72112, or 721120.

Stock Split	Desai and Jain (JB, 1997)	1976-1991	An indicator equal to one, if a CRSP firm has a share adjustment factor that is greater than 1.25, it is designated as split for 12 months or until another split or reverse split occurs. All other observations have a value of zero.
Tangible Assets	Hahn and Lee (JOF, 2009)	1974-2003	The sum of cash and short-term investments + 0.715 times total receivables + 0.547 times total inventory + 0.535 times total property, plant and equipment, divided by total assets. Tangible Assets is only used to predict returns on the subset of constrained firms. We define a constrained firm as a firm whose book value is the bottom 30 percentile, or a firm whose ratio of the sum of dividends and stock repurchases, divided by net income is in the bottom 30 percentile, orfirms with debt and neither a S & P bond rating nor a S & P commercial paper rating.
Target Price Forecast, Change	Brav and Lehavy (JF, 2013)	1997-1999	Using all target forecast prices from all analysts who issued a target within the last 12 months, we divide the current month's change in average target by the current stock price.
Target Price Forecasted Return	Da and Schaumburg (JFM, 2011)	1997-2004	Using all target forecast prices from all analysts who issued a target within the last 12 months, we divide the difference of median target forecast price and the current stock price, by the current stock price.
Tax to Assets, Change	Thomas and Zhang (JAR, 2010)	1977-2006	The ratio of annual income tax to assets, minus the ratio of last year's annual income tax to assets.
Zero Trading Days, Augmented	Liu (JFE, 2006)	1963-2003	The augmented zero trading days measure is calculated using data from the last 6 months. It equates to, (NZVD + (1/Turnover)/11000) x (126/TNTD). NZVD is number of zero volume trading days. Turnover is the total turnover. TNTD is the total number of trading days.
Bid-Ask Spread Volatility	Blau and Whitby (FM, 2012)	1993-2012	Using CRSP closing bid and ask data, we calculate the daily spread as the difference the bid and ask, divided by half of the sum of the bid and ask. We discard observations with missing bids or asks, or observations with ask values that are lower than the bid value. Bid-Ask Spread Volatility is the standard deviation of daily bid-ask spreads for a particular firm in a particular month. This measure is only included for a given month if data is available for more than 20% of CRSP listings. Data availability prevents construction of this variable between November 1960 and December 1984.
Prospect Value	Barberis, Mukherjee, and Wang (RFS, 2016)	1931-2010	This measure is computed for CRSP stocks with a full 5 years for return data, and is based on returns in excess of the CRSP equal-weighted return. For details, please see equations (5), (6), and (8) of Barberis, Mukherjee, and Wang (RFS, 2016).

Table IA.VI
Cross-sectional Predictors used to Make Times-Series Predictors

The table lists the 26 cross-sectional predictors that have been used to form time-series predictors in the existing literature, along with the 23 time-series papers that each was featured in.

Predictor	First Cross-sectional Paper	Associated Time-series Papers
E/P	Basu (1977)	Lewellen (2004)
M/B	FF (1992)	Pontiff-Schall (1998); Lewellen (2004)
Momentum	Jegadeesh and Titman (1993)	Lansing and Tubbs (2018); Hurst, Ooi, and Pedersen (2017); Moskowitz, Ooi, & Pedersen (2012)
LT reversal	Debondt and Thaler (1985)	Balvers and Wu (2006); Bhojraj and Swaminathan (2006); Malin and Bornholt (2013)
ST reversal	Jegadeesh (1989)	Li and Yu (2012)
Dividends	Naranjo, Nimalendran, & Ryngaert (1998)	Li, Ng, & Swaminathan (2018); Lewellen (2004)
Leverage	Bhandari (1988)	Baker Wurgler (2000)
Volume	Datar, Naik, & Radcliffe (1998)	Baker Wurgler (2007)
Investment	Titman, Wei, & Xie (2004)	Bulter, Cornaggia, Grullon, & Weston (2011)
Accruals	Sloan (1996)	Hirshleifer et al (2009)
Post Earnings Drift	Foster, Olsen, & Shevlin (1984)	Kothari et al (2006)
Asset Growth	Cooper, Guylen, & Schill (2008)	Wen (2018)
ROE	Haugen and Baker (1996)	Campbell and Thomson (2006)
Amihud's Measure	Amihud (2002)	Chen, Eaton, & Paye (2018)
IPOs	Ritter (1991)	Baker and Wurgler (2007)
Short Interest	Dechow et al. (2001)	Lynch et al (2014), Rapach, Ringgenberg, & Zhou (2016)
ΔRecommendation	Jegadeesh et al. (2004)	Howe et al (2009)
Spreads	Amihud and Mendelsohn (1986)	Chen, Eaton, & Paye (2018)
Hybrid Tail Risk	Bali, Cakici, & Whitelaw (2014)	Chevapatrakul et al (2019)
Insider Buy	Seyhun (E1986)	Seyhun (1992)
Insider Sell	Seyhun (1986)	Seyhun (1992)
Zero Trading Days	Liu (2006)	Chen, Eaton, & Paye (2018)
Debt issues	Spiess and Affleck-Graves (1999)	Baker Wurgler (2000)
Repurchases	Ikenberry, Lakonishok, & Vermaelen (1995)	Baker Wurgler (2000)
SEO	Loughran and Ritter (1995)	Baker Wurgler (2000)
Total XFIN	Bradshaw, Richardson, & Sloan (2006)	Bulter, Cornaggia, Grullon, & Weston (2011)