



# WRDS BACKTESTER

---

## User Manual

Version 2.0

# Content

Executive Summary

Background

Methodology

Output Discussion

Appendix A: Output Description

Appendix B: List of Signals

# Executive Summary

The WRDS Backtester provides a platform for users to test the historical performance of common asset pricing signals for the U.S. equity market. It also gives the users the flexibility of uploading their own signal values and getting the output leveraging on the back test engine.

**Backtest Basic** – contains 103 signals that were created using common databases: CRSP Stocks, Compustat and CRSP CCM product.

**Backtest Plus** – contains the entire list of 134 signals that use additional databases including IBES, OptionMetrics, Thomson Reuters, and WRDS SEC Analytics.

**Upload Your Own Signal** – users can upload a txt file that contains the time series of customized signal values and test the signal performance.

# Background

Empirical asset pricing literature over the years documents many equity-based signals, such as accrual and book-to-market. These signals are used by not only asset pricing researchers as target variable of interest, they are also often used as control variables in a much wider area of academic study, including corporate finance, accounting and even management science.

While some of the signals can be constructed in a fairly straight forward manner (e.g. book-to-market and momentum), many require complicated data processing and deep knowledge of the relevant databases (e.g. distress and ownership breadth). Exploring research ideas involving these signals then becomes a relatively expensive exercise due to data processing and code generation. The WRDS Backtester aims to provide the research community a centralized platform aggregating the commonly used equity signals for the U.S. market. Researchers can have easy access to the list of signals as well as the historical returns and performance statistics through a web interface. Furthermore, we fully realize that the list of signals here is far from being complete, so we provide a functionality for researchers to upload their customized signal value and rely on the WRDS Backtester engine to produce the return performance and various summary statistics.

# Methodology

## Signal Construction

We follow methodology detailed in the published as well as working papers to construct the signals. For the signals listed under the **Backtest Basic** page, they are constructed by purely using data items from CRSP Stocks, Compustat North America as well as CRSP CCM databases. These signals tend to be the more traditional ones that rely on the pricing or the fundamentals of the underlying securities (e.g. book-to-market, beta, momentum). Currently there are around 100 of these signals.

**Backtest Plus** includes more complicated signals that utilizes data items from additional databases, including IBES (e.g. Analyst Forecast Dispersion and Earnings Surprise), OptionMetrics (e.g. Skewness), Institutional data from Thomson Reuters (e.g. Best Idea and Institution Flows), and even textual analysis data based on regulatory filings from WRDS SEC Analytics (e.g. Filing Length and Readability). Please see Appendix A for the complete list of signals.

## Portfolio Generation

The series of signal values (at Stock ID x Date x Signal Value combination) is fed into the WRDS Backtester to produce the various return series as well as portfolio statistics. Researchers have the following options to construct the portfolios.

- **Weight:**  
Assigned to stocks within portfolio when calculating returns
  - Equally-weighted
  - Value-weighted
- **Active Returns:**  
While traditional total returns of the portfolios are included by default, researchers sometimes would also like to examine the active component of the portfolio returns that is net of the risk factor adjustment:
  - Fama-French 4 Factors Adjusted Excess Returns<sup>1</sup>
  - DGTW Characteristics Based Excess Returns<sup>2</sup>
- **Number of Portfolios:**  
5 quintiles or 10 deciles
- **Stock Universe:**  
We use CRSP common stock space as default but plan to include in the future other equity universe (such as S&P500 or Russell Index family).

## Output Discussion

WRDS Backtester provides a series of metrics to evaluate the historical performance of the signals. These are metrics commonly examined by academics as well as industry practitioners.

### Returns

Depending on the portfolio specification, **Total Portfolio Return** by quintile (decile) is available at a monthly frequency. Portfolios are formed based on signal value at the end of month  $t-1$ , and equal- or value-weighted portfolio returns are calculated at the end of month  $t$ . The **Long-Short Portfolio Return** is then calculated as the difference between the highest signal value portfolio

<sup>1</sup> Eugene F. Fama and Kenneth R. French, the Cross-Section of Expected Stock Returns, *Journal of Finance*, Vol 47, No 2, 1992.

<sup>2</sup> Kent Daniel, Mark Grinblatt, Sheridan Titman and Russ Wermers, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *Journal of Finance*, Vol 52, No 3, 1997.

(quintile 5 or decile 10) and the lowest signal value portfolio (1). **Long-Only** portfolio is simply holding the portfolio with the highest signal value. And **130-30** portfolio is longing 130% of the portfolio value of the stocks in the highest signal value portfolio and shorting 30% of the total portfolio value among the lowest signal value stocks.

Instead of only focusing on the two tails of the portfolios based on the signal distribution and ignoring the middle 60%-80% of the universe, another angle to evaluate the signal performance is to include all stocks. This way the strength of the signal, if any, can be fully reflected in the portfolio composition. To do so, we create **Market Neutral** Portfolio, which is a signal-weighted portfolio that includes all the stocks in the universe. In order to make the weight based on signal value meaningful and well-behaved, we first feed the raw signal values through a standardization procedure<sup>3</sup> to arrive at the distribution with mean = 0 and standard deviation = 1. Stocks with positive normalized signal values are long in the portfolio, and stocks with negative normalized signal values are shorted. The higher the absolute value of the signal, the higher the weight a stock will take in the portfolio's long or short arm.

**Active Return** is another important perspective to analyze the signal performance above and beyond the embedded risk premiums intrinsic to the stock characteristics. Two styles of active returns are calculated in the WRDS Backtester:

- (i) Fama-French 4 Factor Adjusted Excess Return, which is the residual plus the alpha term from regressing stock returns on Fama-French 4 factors.
- (ii) DGTW Characteristics Based Excess Return, which is excess stock return by subtracting the characteristics based portfolio return from the individual stock return.

## Performance Statistics

With the return series, WRDS Backtester calculates various performance statistics to provide a more intuitive grasp of how the signal performs historically. These metrics are often used by industry practitioners to gauge the validity of the signal.

Volatility of the long-short portfolio: the return volatility of the long-short portfolio at both monthly and annualized level.

Best and worst month return: reporting the highest and lowest monthly signal return

Maximum Drawdown: measures the downside risk of the signal as the maximum loss from a peak to a trough of the portfolio cumulative return.

<sup>3</sup> SAS provides a PROC STAND KentARD procedure to normalize values.

Sharpe Ratio: measures the risk adjusted performance of the signal as the excess portfolio return over risk free rate normalized by portfolio volatility.

Information Ratio: measures the portfolio return in excess of the benchmark normalized by the volatility of the strategy returns.

Information Coefficient: measures the correlation between the predicted returns and the actual realized returns.

Breadth: measures the representation of stocks in the portfolio among the benchmark.

Hit Rate: measures the percentage of the months where signal outperforms the benchmark: for long-short and market neutral strategy, the outperformance is defined as absolute positive returns, for long-only and 130-30 strategy, the outperformance is defined as the chosen benchmark (CRSP common stock space).

Annualized Turnover: measures how much a portfolio's holding is replaced, and it is often used as a proxy for trading cost.

These statistics are reported for all styles of returns in the WRDS Backtester, including Total Return, Active Return for Long-Short, Long-Only, Market Neutral and 130-30 portfolios.

## Graphic Output

WRDS Backtester also provides a set of graphic output for the above mentioned performance metrics.

# Appendix A: Output Definition

## B.1 Total and Active Return & Cumulative Returns

Variable Name	Description
date	Portfolio Return Date
Signal_AR_1	Active Return - Portfolio 1 (Lowest Signal Value)
Signal_AR_2	Active Return - Portfolio 2
Signal_AR_3	Active Return - Portfolio 3
Signal_AR_4	Active Return - Portfolio 4
Signal_AR_5	Active Return - Portfolio 5 (Highest Signal Value)
Signal_TR_1	Total Return - Portfolio 1 (Lowest Signal Value)
Signal_TR_2	Total Return - Portfolio 2
Signal_TR_3	Total Return - Portfolio 3
Signal_TR_4	Total Return - Portfolio 4
Signal_TR_5	Total Return - Portfolio 5 (Highest Signal Value)
Signal_AR_Sig	Active Portfolio Return - Signal Weighted
Signal_TR_Sig	Total Portfolio Return - Signal Weighted
Signal_AR_130	Active Portfolio Return - Long 130 Short 30
Signal_TR_130	Total Portfolio Return - Long 130 Short 30
Signal_TR_BP	Total Return - Benchmark
Signal_TR_LO	Total Portfolio Return - Long Only
Signal_CAR_1	Cumulative Active Return - Portfolio 1 (Lowest Signal Value)
Signal_CAR_2	Cumulative Active Return - Portfolio 2
Signal_CAR_3	Cumulative Active Return - Portfolio 3
Signal_CAR_4	Cumulative Active Return - Portfolio 4
Signal_CAR_5	Cumulative Active Return - Portfolio 5 (Highest Signal Value)
Signal_CTR_1	Cumulative Total Return - Portfolio 1 (Lowest Signal Value)
Signal_CTR_2	Cumulative Total Return - Portfolio 2
Signal_CTR_3	Cumulative Total Return - Portfolio 3
Signal_CTR_4	Cumulative Total Return - Portfolio 4
Signal_CTR_5	Cumulative Total Return - Portfolio 5 (Highest Signal Value)
Signal_CAR_LS	Cumulative Active Return - Long Short
Signal_CTR_LS	Cumulative Total Return - Long Short
Signal_CAR_SW	Cumulative Active Return - Signal Weighted
Signal_CTR_SW	Cumulative Total Return - Signal Weighted
Signal_CTR_BP	Cumulative Total Return - Benchmark
Signal_CTR_LO	Cumulative Total Return - Long Only
Signal_CXR_LO	Cumulative Excess Return of Long Only - Benchmark
Signal_CTR_13	Cumulative Total Return - Long 130 Short 30



## A.2 Information Ratio

Variable Name	Description
LAG	Number of months lagged between portfolio formation date and return date
IR_Signal	Information Ratio
IC_Signal	Information Coefficient

## A.3 Information Coefficient and Information Ratio by Industry

Variable Name	Description
Ind	Fama-French 48 Industry Classification
IC_Signal	Information Coefficient
IRet_Signal	Industry Return

## A.4 Statistics of Various Portfolio Returns

Variable Name	Description
signal_name	signal_name
mindate	Start Date
maxdate	End Date
cumtret12	Cumulative Total Return 12 Month
cumtret36	Cumulative Total Return 36 Month
cumtret60	Cumulative Total Return 60 Month
avrtret_m	Total Return, Monthly
avrtret_a	Total Return, Annualized
stdt_a	Volatility of Long-Short Total Returns, Annualized
best_tmonth	Best-month Total Return
worst_tmonth	Worst-month Total Return
maxdrawdown_tret	Maximum Drawdown
Sharpe_tall	Sharpe Ratio
Sharpe_tmean	Sharpe Ratio, Moving Average
IR_tmean	Information Ratio
IC_raw	Information Coefficient
IC_tstat_raw	Information Coefficient, t-stat
breadth	Breadth
hitrate_tret	Hit Rate
turnover	Turnover, Annual

## A.5 Graphic Output

<u>Page Number</u>	<u>Description</u>
1	Cumulative Active Portfolio Returns
2	Cumulative Active Returns for Long/Short Portfolio
3	Cumulative Active Returns for Market-Neutral Portfolio
4	Cumulative Total Returns for Size-Weighted Portfolios
5	Cumulative Total Returns for Long/Short Portfolio
6	Cumulative Total Returns for Market-Neutral Portfolio
7	IR and IC Horizon Plot
8	Performance by Industry

## Appendix B: List of Signals

<u>Variable Name</u>	<u>Variable Label</u>	<u>Description</u>	<u>Paper</u>
AbsAccrual	Absolute Accruals	The absolute value of accruals captures the effect of the accrual anomaly. Absolute accruals is consistent with Sloan's (1996) paper since the accrual anomaly is about the level. Also absolute accrual can be thought of as a one-period volatility measure.	Bandyopadhyay Huang and Wirjanto (2010) WP
Accrual	Accruals	Reflects the level of accrual of firms where firms with high accruals earn lower average returns than firms with low accruals.	Sloan (1996) TAR
Accrualpct	Accrual Percentage	Accruals divided by Earnings (rather than total assets) seems to improve the predictive power of the accrual signal.	Hafzalla, Lundholm and Van Winkle (2007)
AccrualVol	Accrual Volatility	Measuring the consistent deviation by accrual volatility they find a strong and long-lasting negative association between accrual volatility and future stock returns. In decile portfolios that rank accrual volatility a hedge portfolio that goes long in the lowest decile and short in the highest decile generates an annual risk-adjusted return in the order of 10 from one-month to five-year horizon.	Bandyopadhyay Huang and Wirjanto (2010) WP
active_flow	Active Flows		
ATG	Total Assets Growth	Reflects the speed of past asset growth where stocks with larger past asset growth tend to have lower returns in the future.	Cooper Gulen and Schill (2008) JF
AUE	Asset Use Efficiency	Benchmark companies in the same sector by the ability of their assets to generate revenues and contribute to future growth where firms with firms with high asset usage efficiency are expected to outperform in the future. WRDS.	
AUEG	Efficiency Gains	Gauges the improvements in revenues generated per unit of firm assets. Expects that such improvements will be persistent over time and firms with gain in asset use efficiency are expected to outperform in the future. WRDS.	
avgwordpara	Wording Complexity Index	Calculating average number of words per paragraph	
avgwordsent	Sentence Complexity	Calculating average number of words per sentence	
best_idea	Best Idea	Shows the best idea by portfolio managers where stocks with maximum tilt away from benchmark weight tend to systematically outperform in the future.	Cohen Polk and Silli (2010) WP
beta	12-Month Beta	Represents the level of beta of a stock where stocks with high beta tend to have low average returns.	Frazzini and Pedersen (2013) JFE

betasq	12-Month Beta Squared	Relation between expected return and risk (beta) may not be linear. Beta squared may add predictability to standard linear models.	Fama and MacBeth (1973) JPE
BtM	Fiscal Book-to-Market	Measures the book value firms' equity relative to the market cap of the firm at each fiscal year end, where value firms (low book-to-market ratio) are expected to outperform in the future.	
CAPEC	Cyclically-Adjusted Price to Cash Flows	Set of CAPE-like measures popularized by Robert Shiller constructs long-term proxies for firm relative valuation. Uses price scaled by 5-year quarterly average of (operating cash flows plus extra items and discontinued operations) as the basic valuation metric	Campbell and Shiller (1988)
CAPEF	Cyclically-Adjusted Price to Free Cash Flows	Set of CAPE-like measures popularized by Robert Shiller constructs long-term proxies for firm relative valuation. Uses price scaled by 5-year quarterly average free cash flow as the basic valuation metric. Free cash flow is the net operating cash flow plus extra items and discontinued operations minus capital expenditures and cash dividends	Campbell and Shiller (1988)
CAPEI	Cyclically-Adjusted Price-Earnings Ratio	Set of CAPE-like measures popularized by Robert Shiller constructs long-term proxies for firm relative valuation. Uses price scaled by 5-year quarterly net income before extraordinary items and discontinued operations as the basic valuation metric.	Campbell and Shiller (1988)
CAPES	Cyclically-Adjusted Price to Sales	Set of CAPE-like measures popularized by Robert Shiller constructs long-term proxies for firm relative valuation. Uses price scaled by 5-year quarterly sales as the basic valuation metric.	Campbell and Shiller (1988)
CAPXAG	Abnormal Capital Expenditures Growth	Abnormal capital investment may be negatively correlated with abnormal future returns.	Titman, Wei and Xie (2004) JFQA
CAPXG	Capital Expenditures Growth	Defined as industry adjusted percentage change in capital expenditures. Firms with CAPX growth outpacing the industry average are found to underperform, as it reflects a poor cost management when compared to industry peers.	Abarbanell & Bushee (1998), TAR
CAPXLTG	Long Term Growth in Capital Expenditures	Subsequent lower returns for firms that have recently accelerated investment spending. So bet against firms with strong investment growth. Likely to be very correlated with ITOA and CINVEST.	Anderson & Garcia-Feijoo (2006), JF
Cash_to_Debt	Cash Flow to Debt Ratio	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. Cash flow to Debt is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE

CASHPROD	Cash Productivity	They studied the relationship between the productivity of cash (the rents that the firm generates per dollar of cash holdings) and expected stock returns and find that the productivity of cash is a highly significant and negative predictor of stock returns. It implies an inverse relation between the productivity of cash measure and the firm's financial risk. Productivity of cash is defined as the ratio of "Rents" and Cash. Rents are the present value of all firm projects. So it can be proxy by the difference between the market value of the firm and the market value of its investments. A long short portfolio that goes long in the decile with low productivity and short in the decile with high productivity produces an alpha of 1.0 per month (t-statistics 6.36).	Chandrashekar and Rao (2009) WP
CCH	Change in Cash	Author found that stocks with a positive change in company cash holdings have significantly higher risk-adjusted returns than stocks with a negative change in cash holdings. This effect is distinct from the effect of cash holdings, absent among cash-rich companies, stronger among small-cap stocks and limited to non-January months.	Sodjahn (2013), FAJ
CEI	Composite Equity Issuance	Measures the amount of equity the firm issues (or retires) in exchange for cash services where stocks with high level of composite equity issuance tend to underperform.	Daniel and Titman (2006) JF
CEQG	Growth in Common Stockholder Equity	Likely to be very related to NSI and CEI signals. Increases in Common Equity are negatively related to the next years accounting rate of return (ROA). Bet against increases in common equity	Richardson, Sloan, Soliman & Tuna (2005), JAE
CF	Cash Flow Efficiency	Provides systematic exposure to stocks with high real cash-based profitability relative to firm assets. Measures the cash flows generated by unit of firm's assets. Firms with higher cash-flow based profitability are expected to outperform. WRDS.	
CF_P	Cash Flows to Price	One of the metrics to capture value premium. Firms with high (low) cash flow-to-price are known to earn positive (negative) future abnormal returns	Desai, Rajgopal & Venkatachalam (2004), TAR
CFM	Cash Flow Margins	Represents the level of real cash-based profitability (net operating cash flows plus extra items and discontinued operations) scaled by overall reported revenues as a more accurate and untainted measure of profit margin. Firms with high CFM are more profitable and expected to outperform in the future. WRDS.	

CFVol	Cash Flow Volatility	Huang shows that historical cashflow volatility is negatively related to future returns. According to his research, the negative association is large and last up to five years. The least volatile decile portfolio outperforms the most volatile decile portfolio by 13% a year relative to the Fama-French four factors. The cashflow volatility effect is closely related to the idiosyncratic return volatility effect documented in Ang et al. [Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X. "The cross-section of volatility and expected returns." Journal of Finance, 51 (2006), 259-299.]. However, in portfolios simultaneously sorted on both cashflow and return volatilities, and in cross sectional regressions of returns at the firm level, these two effects neither drive out nor dominate each other. While the pricing of idiosyncratic cashflow volatility represents an anomaly against the traditional asset pricing theories, the pricing of historical cashflow uncertainty sheds light on potential fundamental risks embodied in the Fama-French HML and SMB factors. Nor dominate each other. While the pricing of idiosyncratic cashflow volatility represents an anomaly against the traditional asset pricing theories, the pricing of historical cashflow uncertainty sheds light on potential fundamental risks embodied in the Fama-French HML and SMB factors.	Huang (2009), JEF
CFVoIS	Cash Flow Margin Volatility	Represents the quarterly volatility (over 4 year horizons) of cash-based operating profitability scaled by overall reported revenues. Historical cash flow volatility is negatively related to future returns with relationship lasting up to 5 years. The result is robust to size, value, price and earnings momentum and illiquidity and extends both to systematic and idiosyncratic cash flow volatilities.	Huang (2009), JEF
CHETURN	Cash Turnover	Measures firms' ability to generate sales per unit of cash. Firms with better ability of generating sales are expected to have higher future returns	Ou and Penman (1989), JAE
CHETURNG	Cash Turnover Growth	Companies that are able to grow the ability to generate sales per unit of cash are expected to outperform	Ou and Penman (1989), JAE
chs	CHS Distress	Captures the likelihood of a company experiencing financial distress in the near future where firms with high level of financial distress are expected to underperform in the future hence referred to as an asset pricing anomaly.	Campbell Hilscher and Szilagyi (2008)
coleman	Coleman Readability Index	Coleman-Liau Readability Index was first developed in 1967 to gauge the readability of a document based on inputs such as character, word, and sentence counts	
CONVDEBT	Convertible Debt	Valta's model predicts that expected stock returns are higher for firms that face high debt renegotiation difficulties and that have a large fraction of secured or convertible debt.	Valta (2014), JFQA
CORPINVEST	Corporate Investments	Quality Signal. Measures abnormal investment. PPENTQ CAPXQ and SALEQ are quarterly alternatives. SALE and	Titman Wei and Xie (2004) JFQA

		CAPX need to be quarterized using SALEY and CAPXY. Bet against high abnormal investment	
CR	Short Term Leverage	Expected common stock returns are positively related to the ratio of debt (non-common equity liabilities) to equity, controlling for the beta and firm size and including as well as excluding January, though the relation is much large in January. The relationship is not sensitive to variations in the market proxy, estimation technique, etc. the evidence suggests that the premium associated with the debt/equity ratio is not likely to be just some kind of risk premium	Bhandari (1988), JF
CURRENT	Current Ratio	Ou and Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. As far as I can see here there are no definitive and significant relationships between ratios and the hedged portfolio returns in general. The established relationship is in fact the between the instrumental or projected EPS and portfolio returns which is in the line with most academic findings.	Ou and Penman (1989) JAE
CURRENTG	Current Ratio Growth	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. As far as I can see here, there are no definitive and significant relationships between ratios and the h+J74edged portfolio returns in general. The established relationship is in fact the between the instrumental or projected EPS and portfolio returns, which is in the line with most academic findings.	Ou and Penman (1989), JAE
dBlock_N	Blockholders	Measures the change in the number of block holders, where block holders are defined as institutional investors holding more than 5% of the shares outstanding.	
DBREADTH	Institutional Change in Breadth	Measures changes in number of institutional owners, reduction in breadth of institutional ownership of a stock tend to lead to lower future returns.	Chen, Hong and Stein (2002) Lehavy and Sloan (2008)
DefTax	Deferred Taxes	Measures the firms' tax liability defined as deferred tax as a fraction of firms' total asset and is correlated with other accrual-like measures that predict negative future returns. Authors have shown that a negative association exists between deferred taxes and security returns after controlling for pseudo-net income, at least for some subsets of firms	Chaney and Jeter (1994), Journal of Accounting, Auditing and Finance
DEPRATE	Depreciation Rate	Depreciation of Gross PPE is used as the only one of accounting ratio variables used in a logit model. (The trading strategy takes a long position in firms predicted to have positive excess returns and a short position in firms predicted to have negative excess returns). As a result, this	Holthausen and Larcker (1992), JAE

		variable enters the models positively six times and negatively twice.	
DEPRATEG	Depreciation Rate Growth	Depreciation of Gross PPE is used as the only one of accounting ratio variables used in a logit model. (The trading strategy takes a long position in firms predicted to have positive excess returns and a short position in firms predicted to have negative excess returns). As a result, this variable enters the models positively six times and negatively twice.	Holthausen and Larcker (1992), JAE
dHHI	Ownership Breadth	Measures the change in the Herfindahl Index of company's institutional ownership.	
dIOR	Institutional Flows	Measures the changes in the institutional ownership of a company as a fraction of total shares outstanding.	
dispersion	Analyst Forecast Dispersion	Measures the level of dispersion in analysts' opinion regarding future stock performance where stocks with higher level of dispersion in analysts' opinion tend to underperform and this effect is more pronounced among smaller cap stocks.	Diether Malloy and Scherbina (2002) JF
DIV_P	Cash Dividends	Measures the level of dividend payment of a firm defined as level of cash dividend payment on common stocks as a fraction of firms' market cap where stocks with high dividend yield tend to outperform in the future.	Choudhury (2003)
divinc_prob	Probability of dividend increase	Following Bessembinder et al. (2013) special dividend calculation, extend the hazard probabilistic models to predict company's probability of increasing dividend.	
DIVYield	Dividend Yield	Measures the firms' level of dividend yield, defined as level of dividend payment as a fraction of average past stock prices, where stocks with high dividend yield tend to outperform in the future.	
DNOA	Change in Net Operating Assets	Predicts that increases in net operating assets (scaled by total assets) are bad omens for company's productivity and will reflect in negative future returns where the level itself serves as a strong negative predictor of long-run stock returns	Hirshleifer Hou Teoh and Zhang (2004) JAE
DR	Long Term Leverage	Expected common stock returns are positively related to the ratio of debt (non-common equity liabilities) to equity, controlling for the beta and firm size and including as well as excluding January, though the relation is much large in January. The relationship is not sensitive to variations in the market proxy, estimation technique, etc. the evidence suggests that the premium associated with the debt/equity ratio is not likely to be just some kind of risk premium	Bhandari (1988), JF



DRG	Leverage Expansion	Expected common stock returns are positively related to the ratio of debt (non-common equity liabilities) to equity, controlling for the beta and firm size and including as well as excluding January, though the relation is much large in January. The relationship is not sensitive to variations in the market proxy, estimation technique, etc. the evidence suggests that the premium associated with the debt/equity ratio is not likely to be just some kind of risk premium	Bhandari (1988), JF
EV_EBITDA	Enterprise Value to EBITDA	One of the metrics to capture value premium. Authors argue that the Enterprise Multiple, calculated as (equity value+debt value+preferred stock-cash)/EBITDA is better than book-to-market in cross-sectional monthly regressions over 1963-2008 period. Factor-mimicking portfolio long stocks with low enterprise value and short stocks with high enterprise multiple generates a return premium of 5.76% per year	Loughran and Wellman (2009)
FCF	Free Cash Flow Efficiency	Similar to CF measure, but uses free cash flows net of capital expenditures and cash dividends. Firms with higher FCF are more profitable and expected to outperform in the future. WRDS.	
FCF_P	Free Cash Flows	One of the metrics to capture value premium. Firms with high (low) free cash flow-to-price are known to earn positive (negative) future abnormal returns. Free Cash flows are net operating cash flows net of capital expenditures and cash dividends	Desai, Rajgopal & Venkatachalam (2004), TAR
FilingSize	Filing Length and Complexity	Measures the readability of the annual filing by its length.	Li (2008) JAE
FIRMTANG	Debt Capacity to Firm Tangibility	Debt capacity is a significant determinant of stock returns only in the cross section of financially constrained firms, after controlling for beta, size, book-to-market, leverage, and momentum. Debt capacity is proxied by the firm-level tangibility measure of Almeida and Campello (2007) which is the expected asset liquidation value of a firm. Accordingly, a dollar book value is equivalent to 72 cents for receivables, 55 cents for inventory and 54 cents for fixed assets. Ordering firms by low and high tangibility measure,	Hahn & Lee (2009), JF
FOG_Complex	Change in Gunning-Fog Readability Index	Measures the change in the FOG readability index, which estimates the years of formal education a person needs to understand the text on the first reading.	
ft_mstrong	Loughran-McDonald modal strong word proportion	Loughran-McDonald Modal Strong Word proportion reflects the fraction of financial-modal-strong words in the document based on the Loughran-McDonald financial dictionary.	Loughran and McDonald (2011) JF
geo_seg	Geographical Segment	Measures the companies' exposure to each macro market and hence infers the future sales and profitability, where stocks with high sales exposure to fast growing economies tend to have higher future returns	Li, Richardson and Tuna (2012)

GProf	Gross Profitability	Indicates the clean representation of firm's level of profitability and hence future stock returns	Novy-Marx (2012) JFE
GProfG	Gross Profitability Growth	Indicates the clean representation of firm's level of profitability and hence future stock returns	Novy-Marx (2012) JFE
HIRE	Growth in Employee Base	Firm level hiring rate is negatively related with future stock returns in the cross-section of US publicly traded firms even after controlling for investment, size, book-to-market and momentum as well as other known predictors of stock returns. The predictability shows up in both Fama-MacBeth cross sectional regressions and in portfolio sorts and it is robust to the exclusion of micro-cap firms from the sample. Firm level hiring rate is informative about the firms' expectations about future cash-flows and risk-adjusted discount rates	Bazdresch, Belo and Lin (2009), WP
INVG	Change in Inventory Turnover	Captures the growth rate in firms' inventory relative to total assets, and firms with higher growth in the inventory imply lower ability to generate revenue and hence are expected to have lower future returns.	Ou and Penman (1989, JAE)
INVTURN	Inventory Turnover	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. Inventory turnover is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE
INVTURNG	Inventory Turnover Growth	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. Inventory turnover is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE
IOR	Institutional Ownership Ratio	Measures the percentage of a company's common shares that are held by institutional investors.	
ITOA	Investment to Assets	Measures the level of firms' investment relative to firm size where stocks with high investment to asset ratio tend to underperform in the future.	Lyandres Sun and Zhang (2008) RFS; Novy-Marx and Zhang (20
ivol	60-Day IVol	Measures the level of idiosyncratic volatility of stocks where stocks with high level of idiosyncratic volatility tend to underperform the peers in the future.	Ang Hodrick Xing and Zhang (2006) JF
LiqVol	Volatility of Liquidity	Measures the volatility of a company's liquidity.	
LTDG	Total Liabilities Growth	Measured as year-over-year change in total liabilities. Firms with higher growth in total debt are expected to have lower future returns. When increase in debt is used to finance increases in operating assets, earnings persistence is lower. The negative relationship between accruals and future stock returns will be stronger for less reliable accruals that result in lower earnings persistence.	Richardson, Sloan, Soliman & Tuna (2005), JAE, Nissim and Penman (2003)

ltr36	36-Month LTR	Measures stocks' return performance in past 36 months and aims to capture the long term reversal trend.	Poterba and Summers (1988) JFE
ltr60	60-Month LTR	Measures stocks' return performance in past 60 months and aims to capture the long term reversal trend.	De Bondt and Thaler (1985) JF
maxret	Maximum Return	Reflects recent extreme returns at stock level. Stocks that experience positive extreme returns are anticipated to underperform going forward.	Bali Cakici and Whitelaw (2011) JFE
momaccel	Momentum Acceleration	Stocks with extremely high price acceleration outperform stocks with extremely low acceleration. Focusing on stocks with both high six-month momentum and rapidly increasing six-month momentum offers significant excess returns.	Gettleman & Marks (2006), WP
momentum	6-Month Momentum	Measures the traditional price momentum of stocks based on the return performance in the past 6 months where stocks with high past performance tend to continue to outperform in the future.	Jegadeesh and Titman (1993) JF
momentum12	12-Month Momentum	Measures the traditional price momentum of stocks based on the return performance in the past 12 months where stocks with high past performance tend to continue to outperform in the future.	Jegadeesh and Titman (1993) JF
momentum36	36-Month Momentum	Measures the traditional price momentum of stocks based on the return performance in the past 36 months where stocks with high past performance tend to continue to outperform in the future.	Jegadeesh and Titman (1993) JF
NOA	Net Operating Assets	Measures the level of net operating assets (scaled by total assets) which serves as a strong negative predictor of long-run stock returns	Hirshleifer Hou Teoh and Zhang (2004) JAE
nsi	Net Stock Issuance	Indicates the level of stock issuance by a firm where companies with extreme positive issues see underperformance in the stocks.	Fama and French (2008) JF
numest	Number of valid forecasts at time t	Captures the breadth of coverage of a given stock by analysts, to overweight stocks with more analyst attention in a particular sector.	
O_score	Ohlson Score	Measures the probability of financial distress of a company where stocks with high Ohlson score tend to produce lower future returns.	Ohlson (1980) Griffin and Lemmon (2002) JF
OIVol	Excess Cash Flow Volatility	Uses instead of real operating cash flow (as in cfvol) OIBDP in the numerator, i.e., measures volatility of Operating Income Before Depreciation as a % of Sales. The underlying intuition is the same as in CFVOL	Huang (2009), JEF
OPLEV	Operating Liabilities Leverage	Suppliers who advance the payables reduce the net investment required to run the operations and so lever up the operating profitability	Nissim and Penman (2001) RAS
OProf	Operating Profitability	Operating profitability is similar in spirit to gross profitability of Novy-Marx (2013), but it is constructed to better match current expenses with current revenues. This measure exhibits a far stronger link with expected returns than either net income or gross profit. It predicts returns	Ball, Gerakos, Linnainmaa and Nikolaev (2015), JFE

		as far as ten years ahead, seemingly inconsistent with irrational pricing explanations	
Parity_VSpread	Put-Call Parity Volatility Spread	Measures the spread between implied volatility between paired Call and Put options of stocks, where stocks with relatively expensive Put options tend to underperform.	Cremers and Weinbaum (2010)
PC_Divergence	Option-Based Divergence of Opinion	Measures the level of divergence in opinion in the option market, to be against signals with high level of opposing views reflect in high directional open interest in various exercise prices using total open interest of Call options and Put options.	
PC_Ratio	Put-Call Ratio	Measures the open interest of Call options relative the total open interest of Calls and Puts combined, where stocks with higher level of open interest in put options tend to underperform the peers in the future.	Pan and Potesman (2006)
PE	Price Earnings Ratio	One of the metrics to capture value premium. Basu finds that the common stock of high E/P firms earn, on average, higher risk-adjusted returns than the common stock of low E/P firms and that this effect is clearly significant even if experimental control is exercised over differences in firm size.	Basu (1977), JFE
PFS_Score	Piotroski Financial Statement Score	Defined as a composite score of nine individual dummy variables, each reflecting different aspects of firms' financial strength. Firms with high P-score is expected to have better future performance and hence higher returns.	Piotroski 2000, JAR
PM_P	Profit Margin to Price	One of the metrics to capture value premium and profitability premia. Measures profitability of a firm defined as firm's gross profitability as a fraction of firm's market cap where more profitable firms ones with higher level of gross profit relative to market cap tend to outperform in the future. Measure is supposed to provide a simultaneous bet on profitability and value. WRDS.	
PMG	Abnormal Profit Margin Growth	Captures growth in gross profitability in excess of asset growth. Firms that are able to grow their gross profit faster than their asset base are expected to outperform. WRDS.	
PMSG	Gross Margin Growth	Defined as percentage change in gross margin relative to percentage change in sales. Firms with faster growth in gross margin relative to sales are expected to outperform. A faster growth in gross margin relative to sales reflects an improvement in the firm's operating performance, and hence should lead to higher stock returns.	Abarbanell & Bushee (1998), TAR
prcdelay1	Price Delay 1	The degree of market frictions affecting a stock is approximated by the delay with which its price responds to information. The most delayed firms seems to have a large return premium. Delay of one month.	Hou & Moskowitz (2005), RFS

prcdelay2	Price Delay 2	The degree of market frictions affecting a stock is approximated by the delay with which its price responds to information. The most delayed firms seems to have a large return premium. Delay of two months.	Hou & Moskowitz (2005), RFS
prcdelay3	Price Delay 3	The degree of market frictions affecting a stock is approximated by the delay with which its price responds to information. The most delayed firms seems to have a large return premium. Delay of three months.	Hou & Moskowitz (2005), RFS
PtB	Price-to-Book	Price-to-Book measures the market capitalization of the firm relative to the book value of firms' equity. This strategy aims at profiting on buying value stocks (low price-to-book ratio) and selling growth stocks (high price-to-book ratio). Fama and French (1992) highlighted the relevance of value premium in explaining the cross section of stock returns.	Fama and French (1992), JF
QUICK	Quick Ratio	Ou and Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. As far as I can see here there are no definitive and significant relationships between ratios and the h+J74edged portfolio returns in general. The established relationship is in fact the between the instrumental or projected EPS and portfolio returns which is in the line with most academic findings.	Ou and Penman (1989) JAE
QUICKG	Quick Ratio Growth	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. Quick Ratio is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE
RD_P	Research and Development scaled by Market Cap	R&D intensity is strongly and positively related to the future returns	Guo, Lev & Shi (2006), JBFA
RD_SALE	Research and Development Expense Scaled by Sales	R&D intensity is strongly and positively related to the future returns	Guo, Lev & Shi (2006), JBFA
RDG	Research and Development Growth	Market is slow to recognize the extent of R&D increases (consistent with investor under reaction). Bet on R&D increases	Eberhart, Maxwell & Siddique (2004), JF
RECTURN	Receivable Turnover	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. Sales to receivables is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE

RECTURNG	Receivable Turnover Growth	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. Receivable turnover is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE
REHOLDING	Real Estate Holdings	Based on a general equilibrium production economy where firms use two factors: real estate and other capital, investment is subject to asymmetric adjustment costs. Because real estate depreciates slowly, firms with high real estate holdings are more vulnerable to bad productivity shocks and hence are riskier and have higher expected returns. Returns of firms with a high share of real estate capital exceed that of low real estate firms by 3 to 6% annually, adjusted for exposures to the market return, size, value, and momentum factors.	Tuzel (2010), RFS
REVSUP	Revenue Surprise	Revenue surprises (Sales) may have different dynamic than earnings surprises. Kama found for example that revenue surprises is higher in industry with oligopolistic competition. Kama also reported that market reaction to earnings surprises is not higher than to revenues surprises. The paper reported the persistence of the post-announcement drift six months after the earnings announcement date.	Kama (2009) JBFA
ROA	Returns on Assets	Measures the profitability of a firm defined as income before extraordinary items as a fraction of total assets where stocks with higher level of profitability are expected to continue to outperform. Five-factor model directed at capturing size, value, profitability and investment patterns in average stock returns performs better than the 3-factor model of Fama and French (1993)	Fama and French (2006, 2008, 2014)
ROAG	Growth in Returns on Assets	Ou & Penman (1989) paper looks into the relationship between accounting ratios and portfolio returns through an instrumental EPS approach within a logit regression setting. ROA is one of these variables. As far as the paper results, there are no definitive and significant relationships between ratios and the hedged portfolio returns in general.	Ou & Penman (1989), JAE
ROAVol	ROA Volatility	Measures earnings volatility, where earnings are defined as ROA. Stocks with higher level of earnings volatility are expected to underperform.	Gow and Taylor (2009)
ROE	Returns on Equity	Authors form portfolios by sorting directly on past ROE and find that firms with high ROE earn substantially higher returns during subsequent periods. Also show that ROE-based factor can help explain a large set of asset-pricing anomalies.	Hou, Xue and Zhang (2012), WP
RV_IV	Excess Implied Volatility	Stocks with low realized-implied volatility spread are found to outperform ones with high relative volatility spread.	Bali and Hovakimian (2009)

Sale_P	Sales to Price	One of the metrics to capture value premium. Measures the profitability of a firm defined as seasonally smoothed quarterly sales to market cap where firms with higher level of sales are expected to outperform. Sales-Price ratio captures the role of the debt-equity ratio in explaining stock returns. Neither the book-to-market value of equity nor the market value of equity has consistent explanatory power for stock returns, and sales-price ratio is a more reliable explanatory factor	Barbee, Mukherji and Raines (1996), FAJ
SALEG	Sales Growth	Firms with higher sales growth underperform the ones with lower sales growth, after adjusting for various risk factors, such as size. This is explained by the contrarian nature of the value strategies, which longs stocks that are overlooked by naïve investors and shorts ones overpriced by the market.	Lakonishok, Shleifer and Vishny (1994), JF
SCR	Short Squeeze Probability		
seasonality	Seasonality	Captures the seasonality in stock return performance where stocks tend to exhibit cross-sectional autocorrelation pattern at lags of 12 24 and 36 months where stocks with high past seasonally adjusted return tend to display similar return patterns.	Heston and Sadka (2008) JFE
SECUDEBT	Secured Long-Term Debt		
SG	Abnormal Sales Growth	Captures growth in sales in excess of growth in net operating assets. First that are able to grow their revenues faster than their net operating assets are expected to outperform.	Hirshleifer Hou Teoh and Zhang (2004) JAE
SGAG	Growth in Selling and General Expenses to Sales	Authors find that increases (decreases) in SG&A expenses relative to the % change in sales, predicts increases (decreases) in the indirect cost structure of the firm, and, thus, decreases (increases) in future earnings. Therefore, firms that are able to grow their sales faster than their SG&A expenses are expected to outperform in the future	Abarbanell & Bushee (1998), TAR
SIO	Short Interest Scaled by Supply	Captures the security shorting activity. Signal level is expressed as fraction of shares held by institutional investors that are sold short, where stocks with high level of shorting are expected to underperform.	Negal (2005) Asquith, Pathak and Ritter (2005)
SIR	Short Interest Ratio	Captures the security shorting activity. Signal level is expressed as fraction of total shares outstanding that are sold short, where stocks with high level of shorting are expected to underperform.	
Skew_ATM	ATM Put Volatility Skewness	Measures the spread between implied volatilities of at-the-money Call options and Put options, where stocks with relative expensive Call options tend to outperform.	Cremers and Weinbaum (2010)
Skew_OTM	OTM Put Volatility Skewness	Measures the spread between implied volatilities of at-the-money Call options and out-of-the-money Put options, where stocks with relative expensive out-of-the-money Put options tend to underperform.	Xing, Zhang and Zhao (2010)



Smirk	Abnormal Put Volatility Smirk	Measures the second order volatility of stocks, as excess in skew in out-of-the-money Put options relative to historical realized volatility, where stocks with relative expensive put options tend to underperform going forward.	
SMOG_Complex	Change in Smog Readability Index	Measures the change in the SMOG readability Index, which is a measure of readability that estimates the years of education needed to understand a piece of writing.	
specdiv_prob	Probability of special dividend	Hazard probabilistic Models to predict Companies' distributions	Bessembinder et al (2013)
split_prob	Probability of split	Following Bessembinder et al. (2013) special dividend calculation, extend the hazard probabilistic models to predict company's probability of stock split.	
stockdiv_prob	Probability of stock dividend	Following Bessembinder et al. (2013) special dividend calculation, extend the hazard probabilistic models to predict company's stock dividend.	
str_mod	Modified Short-term Reversal	Stock returns unexplained by "fundamentals," such as cash flow news, are more likely to reverse in the short run than those linked to fundamental news. Making novel use of analyst forecast revisions to measure cash flow news, a simple enhanced reversal strategy generates a risk-adjusted return four times the size of the standard reversal strategy	Da, Liu, and Schaumburg (2014), Management Science
sue	Earnings Surprise	Measures the level of earnings surprises where stocks with high level of earnings surprise tend to continue to drift up following the announcement day (phenomenon also known as post-earnings announcement drift)	Bernard and Thomas (1990), JAE
TaxExp	Changes in Tax Expense	Measures the change in firms' tax expenses where stocks with positive tax expense surprise tend to outperform stocks with negative tax expense shocks.	Thomas and Zhang (2011) JAR
TAXINC	Taxable Income	Quality signal. Taxable income information about future earnings is incremental to that in accruals and cash flows. Bet on firms with relatively large amounts of industry-adjusted taxable income relative to earnings	Lev and Nissim (2004), TAR
trend_factor	Trend Factor	A factor that is created from the information in the three anomalous price patterns: the short-term reversal effects, the momentum effects and the long-term reversal, is correlated with future abnormal returns.	Han, Zhou and Zhu (2015)