

# Open Source Cross-Sectional Asset Pricing Online Appendix

Andrew Y. Chen

Federal Reserve Board  
andrew.y.chen@frb.gov

Tom Zimmermann

University of Cologne  
tom.zimmermann@uni-koeln.de

May 18, 2020

**Table 1 Description of Anomaly Construction.** This table provides details of the construction of return predictors used in the paper. Data come from the CRSP stock return database, Compustat North America Annual and Quarterly databases, IBES earnings estimates database, OptionMetrics, Thomson SDC and a number of additional databases noted in the descriptions of specific anomalies. Our final database is set up at monthly frequency. We lag annual Compustat data by six months and quarterly Compustat data by 3 months to assure availability of relevant data at the time of trading.

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Abarbanell and Bushee	1998	Change in capital inv (ind adj)	ChInvIA	1974	1988	Growth in capital expenditure (capx) minus average growth in capital expenditure in the same industry (two-digit SIC). If capx is missing, capital expenditure is defined as the annual change in property, plant and equipment (ppent). Capital expenditure growth is defined as the percentage growth of capx today relative to the average capx over the previous two years $(.5*(capx_{t-1} + capx_{t-2}))$ , or as percentage growth relative to the previous year only if t-2 is missing.	clear	Original
Abarbanell and Bushee	1998	Gross Margin growth over sales growth	GrGMToGrSales	1974	1988	Define gross margin GM as revenue (sale) minus cost of goods sold (cogs). GrGMToGrSales is the percentage growth of GM relative to average GM in years t-1 and t-2, minus the percentage growth of revenue relative to average revenue in years t-1 and t-2. Replace growth rates with growth relative to the previous year only if data for t-2 are not available.	clear	Original
Abarbanell and Bushee	1998	Sales growth over inventory growth	GrSaleToGrInv	1974	1988	Percentage growth in sales (sale) relative to average sales of t-1 and t-2, minus percentage growth in inventory (invnt) relative to average inventory of t-1 and t-2. Both growth terms are calculated relative to t-1 only if t-2 is missing.	likely	Original
Abarbanell and Bushee	1998	Sales growth over overhead growth	GrSaleToGrOverhead	1974	1988	GrSaleToGrOverHead = Percentage growth in sales (sale) relative to average sales of t-1 and t-2, minus percentage growth in administrative expenses (xsga) relative to average administrative expenses of t-1 and t-2. Both growth terms are calculated relative to t-1 only if t-2 is missing. Remove if in the highest quintile of GrSaleToGrOverHead. Returns are nicely monotonic until the highest quintile, consistent with original paper's rank regressions.	likely	Original
Abarbanell and Bushee	1998	Effective Tax Rate	ETR	1974	1988	Let Tax be tax payments (txt) over pre-tax income (pi) and amortization of intangibles (am, 0 if missing). Let earnings per share be $epsx/ajex$ . ETR if Tax minus the average of Tax over the previous three years multiplied by one-year EPS growth scaled by by price (prcc_f).	maybe	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Abarbanell and Bushee	1998	Change in sales vs change in receivables	GrSaleToGrReceivables	1974	1988	Percentage growth in sales (sale) relative to average sales of t-1 and t-2, minus percentage growth in receivables (rect) relative to average receivables of t-1 and t-2. Both growth terms are calculated relative to t-1 only if t-2 is missing.	maybe	Original
Abarbanell and Bushee	1998	Laborforce efficiency	LaborforceEfficiency	1974	1988	One year growth of the ratio between sales (sale) and employment (emp).	maybe	Original
Abarbanell and Bushee	1998	Change in gross margin vs sales	pchgm_pchsale	1974	1988	Annual percentage change in revenue (sale) minus cost (cogs), minus annual percentage change in revenue.	maybe	Original
Acharya and Pedersen	2005	Illiquidity-illiquidity beta (beta2i)	betaCC	1964	1999	see monthly Code	maybe	Original
Acharya and Pedersen	2005	Illiquidity-market return beta (beta4i)	betaCR	1964	1999	see monthly Code	maybe	Original
Acharya and Pedersen	2005	Net liquidity beta (betanet,p)	betaNet	1964	1999	see monthly Code	maybe	Original
Acharya and Pedersen	2005	Return-market illiquidity beta (beta3i)	betaRC	1964	1999	see monthly Code	maybe	Original
Acharya and Pedersen	2005	Return-market return illiquidity beta (beta1i)	betaRR	1964	1999	see monthly Code	maybe	Original
Adrian, Etula and Muir	2014	Broker-Dealer Leverage Beta	BetaBDLeverage	1973	2009	Regress quarterly stock return minus 3 month treasury bill rate (tbillrate3m) on broker dealer leverage. Use a rolling window of 40 quarters (require at least 20 non-missing observations). BetaBDLeverage is the coefficient on broker-dealer leverage. Broker-dealer leverage is the seasonally adjusted ratio of assets (FRED series BOGZ1FL664090005Q) and liabilities (FRED series BOGZ1FL664190005Q).	clear	Original
Ali, Hwang, and Trombley	2003	Idiosyncratic risk (AHT)	IdioVolAHT	1976	1997	Standard deviation of residuals from CAPM regressions using the past year of daily data. Require at least 100 non-missing observations.	clear	Original
Alwathainani	2009	Earnings growth for consistent growers	EarningsConsistency	1971	2002	Average earnings growth over previous 48 months. Earnings growth is defined as EPS (epspx) minus EPS 12 months ago divided by average EPS 12 and 24 months ago. Exclude if price less than 5, absolute value of 12 month earnings growth greater 600%, or earnings growth and earnings growth 12 months ago have different signs.	clear	Original
Amihud	2002	Amihud's illiquidity	Illiquidity	1964	1997	Past twelve month average of: daily return (abs(ret)) divided by turnover((abs(prc)*vol)	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Amihud and Mendelsohn	1986	Bid-ask spread	BidAskSpread	1961	1980	Spread estimates from Shane Corwin's website ( <a href="https://www3.nd.edu/~scorwin/">https://www3.nd.edu/~scorwin/</a> ) divided by price (abs(prc)).	likely	Original
Anderson and Feijoo	Garcia-2006	Change in capex (two years)	grcapx	1976	1999	Growth rate of capital expenditures (capx) relative to two years ago. If capx is missing, replace with annual change in property, plant and equipment (ppent).	clear	Original
Anderson and Feijoo	Garcia-2006	Investment growth (1 year)	grcapx1y	1964	2003	Growth between one-year lagged capital expenditures (capx) and two-year lagged capital expenditures. Replace capx with the one year difference in property, plant and equipment (ppent) if capx is missing and the corresponding firm age is greater or equal than two years.	clear	Original
Anderson and Feijoo	Garcia-2006	Change in capex (three years)	grcapx3y	1976	1999	Growth rate of capital expenditures (capx) relative to three years ago. If capx is missing, replace with annual change in property, plant and equipment (ppent).	clear	Original
Anderson, Ghysels, and Juergens	2005	dispersion in long-term analyst forecasts	ForecastDispersionLT	1991	1997	Standard deviation of earnings estimates (stdev\_est) scaled by mean earnings estimate.	maybe	Original
Ang et al.	2006	Systematic volatility	betaVIX	1986	2000	Coefficient on daily change in the VIX of a 1-month rolling window regression of daily stock excess returns on market return and the daily change in the CBOE S&P 100 volatility index (downloaded from FRED). Require at least 15 non-missing observations.	clear	Original
Ang et al.	2006	Idiosyncratic risk	IdioRisk	1963	2000	Standard deviation of residuals from CAPM regressions using the past month of daily data. Value weighted	clear	Original
Ang et al.	2006	Idiosyncratic risk (3 factor)	IdioVol3F	1963	2000	Standard deviation of residuals from Fama-French three factor regressions using the past month of daily data. Value weighted	clear	Original
Ang et al.	2006	Idiosyncratic risk (CAPM)	IdioVolCAPM	1963	2000	Standard deviation of residuals from CAPM regressions using the past month of daily data. Value weighted	clear	Original
Ang et al.	2006	Idiosyncratic risk (q factor)	IdioVolQF	1967	2000	Standard deviation of residuals from q-factor regressions using the past month of daily data. Value weighted	clear	Risk Model
Ang, Chen and Xing	2006	Downside beta	DownsideBeta	1963	2001	Beta of Daily stock return (ret - rf) regressed on market return (mktrf) for days for which the market return was less than the average market return over the previous year. Rolling window of 252 trading days with at least 50 observations.	not	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Asquith Pathak and Ritter	2005	Inst own among high short interest	IO_ShortInterest	1980	2002	Exclude all stocks with short interest (ShortInterest) below .025. IO_ShortInterest is institutional ownership (instown\_perc). Keep NYSE Only.	clear	Original
Avramov et al	2007	Junk Stock Momentum	Mom6mJunk	1985	2003	Mom6m. Include only stocks with a credit rating (splticrm) of BBB or lower	clear	Original
Balakrishnan, and Faurel	Bartov 2010	Return on assets	Profitability	1976	2005	Quarterly earnings per share (epspxq) times quarterly shares outstanding used to calculate EPS (cshprq) divided by total assets (at). Exclude if price less than 1.	clear	Original
Balakrishnan, and Faurel	Bartov 2010	Return on assets incl extraordinary income	roaq	1976	2005	Quarterly net income (ibq) divided by lagged total assets (atq). Exclude if price less than 1.	clear	Original
Bali, Cakici, and Whitelaw	2010	Maximum return over month	MaxRet	1962	2005	Maximum of daily returns (ret) over the previous month	clear	Original
Bali, Engle and Murray	2015	Skewness of daily returns	ReturnSkew	1963	2012	Skewness of daily returns (ret) over previous month.	clear	Original
Bali, Engle and Murray	2015	Skewness of daily idiosyncratic returns (3F model)	ReturnSkew3F	1963	2012	Skewness of idiosyncratic returns computed as residuals from regression of daily excess returns (ret - rf) on Fama-French factors (mktrf, smb, hml) over the previous month. We require at least 15 non-missing observations.	clear	Original
Bali, Engle and Murray	2015	Skewness of daily idiosyncratic returns (CAPM)	ReturnSkewCAPM	1963	2012	Skewness of idiosyncratic returns computed as residuals from regression of daily excess returns (ret - rf) on market factor (mktrf) over the previous month. We require at least 15 non-missing observations.	clear	Risk Model
Bali, Engle and Murray	2015	Skewness of daily idiosyncratic returns (Q factor model)	ReturnSkewQF	1967	2012	Skewness of idiosyncratic returns computed as residuals from regression of daily excess returns (ret - rf) on q-factors (r\_mkt, r\_me, r\_ia, r\_roe) over the previous month. We require at least 15 non-missing observations. We download q-factor data from <a href="http://global-q.org/index.html">url{http://global-q.org/index.html}</a> .	clear	Risk Model

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Ball et al.	2016	Cash-based operating profitability	CBOperProfNoLag	1963	2014	Revenue (revt) minus cost (cogs) - (administrative expenses (xsga) - R&D expenses (xrd)) minus annual change in receivables (rect), annual change in investment (invt) and annual change in prepaid expenses, plus annual change in current deferred revenue (drc), long-term deferred revenue (drlt), accounts payable (ap) and accrued expenses (xacc), all divided by total assets (at) in year t. Replace all variables in the numerator with 0 if they are missing. Exclude if share code is greater 11, market value of equity, BM or total assets are missing, or if SIC code between 6000 and 6999.	clear	Original
Ball et al.	2016	Cash-based operating profitability	OperProfRDNoLag	1963	2014	Revenue (revt) minus cost (cogs) - (administrative expenses (xsga) - R&D expenses (xrd)) minus annual change in receivables (rect), annual change in investment (invt) and annual change in prepaid expenses, plus annual change in current deferred revenue (drc), long-term deferred revenue (drlt), accounts payable (ap) and accrued expenses (xacc), all divided by total assets (at) in year t. Replace all variables in the numerator with 0 if they are missing. Exclude if share code is greater 11, market value of equity, BM or total assets are missing, or if SIC code between 6000 and 6999.	likely	Original
Ball et al.	2016	Cash-based operating profitability	CBOperProf	1963	2014	Revenue (revt) minus cost (cogs) - (administrative expenses (xsga) - R&D expenses (xrd)) minus annual change in receivables (rect), annual change in investment (invt) and annual change in prepaid expenses, plus annual change in current deferred revenue (drc), long-term deferred revenue (drlt), accounts payable (ap) and accrued expenses (xacc), all divided by total assets (at) in year t-1. Replace all variables in the numerator with 0 if they are missing. Exclude if share code is greater 11, market value of equity, BM or total assets are missing, or if SIC code between 6000 and 6999.	clear	Lag structure

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Ball et al.	2016	Cash-based operating profitability	OperProfRD	1963	2014	Revenue (revt) minus cost (cogs) - (administrative expenses (xsga) - R&D expenses (xrd)) minus annual change in receivables (rect), annual change in investment (invt) and annual change in prepaid expenses, plus annual change in current deferred revenue (drc), long-term deferred revenue (drlt), accounts payable (ap) and accrued expenses (xacc), all divided by total assets (at) in year t. Replace all variables in the numerator with 0 if they are missing. Exclude if share code is greater 11, market value of equity, BM or total assets are missing, or if SIC code between 6000 and 6999.	likely	Lag structure
Ball et al.	2016	Cash-based operating profitability	CBOperProf_q	1963	2014	Revenue (revt) minus cost (cogs) - (administrative expenses (xsga) - R&D expenses (xrd)) minus annual change in receivables (rect), annual change in investment (invt) and annual change in prepaid expenses, plus annual change in current deferred revenue (drc), long-term deferred revenue (drlt), accounts payable (ap) and accrued expenses (xacc), all divided by total assets (at) in year t-1. Replace all variables in the numerator with 0 if they are missing. Exclude if share code is greater 11, market value of equity, BM or total assets are missing, or if SIC code between 6000 and 6999.	clear	Quarterly
Ball et al.	2016	Cash-based operating profitability	OperProfRD_q	1963	2014	Revenue (revt) minus cost (cogs) - (administrative expenses (xsga) - R&D expenses (xrd)) minus annual change in receivables (rect), annual change in investment (invt) and annual change in prepaid expenses, plus annual change in current deferred revenue (drc), long-term deferred revenue (drlt), accounts payable (ap) and accrued expenses (xacc), all divided by total assets (at) in year t. Replace all variables in the numerator with 0 if they are missing. Exclude if share code is greater 11, market value of equity, BM or total assets are missing, or if SIC code between 6000 and 6999.	likely	Quarterly
Banz	1981	Size	Size	1926	1975	Log of monthly market value of equity (abs(prc)*shrout)).	clear	Original
Barbee, Mukherji and Raines	1996	Sales-to-price	SP	1979	1991	Ratio of annual sales (sale) to market value of equity.	likely	Original
Barbee, Mukherji and Raines	1996	Sales-to-price quarterly	SP_q	1979	1991	Ratio of annual sales (sale) to market value of equity.	clear	Quarterly

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Barber et al.	2002	Consensus Recommendation	ConsRecomm	1994	1997	Binary variable if the monthly mean of recommendations (ireccd) over analysts is greater than 3, and 0 if it is less or equal than 1.5.	clear	Original
Barber et al.	2002	Down forecast EPS	DownForecast	1985	1997	Binary variable equal to 1 if mean earnings forecast (meanest) decreased over the past month.	clear	Original
Barber et al.	2002	Up Forecast	UpForecast	1985	1997	Binary variable equal to 1 if mean analyst earnings forecast for the next quarter (meanest) has improved over the previous month, and 0 otherwise.	clear	Original
Barry and Brown	1984	Firm age based on CRSP	FirmAge	1931	1980	Months since start of CRSP coverage.	maybe	Original
Barth and Hutton	2004	Change in Forecast and Accrual	ChForecastAccrual	1981	1996	Within upper half of Accruals distribution, equal to 1 if mean earnings estimate increased relative to the previous month. 0 if it decreased.	clear	Original
Bartov and Kim	2004	Book-to-market and accruals	AccrualsBM	1980	1998	Binary variable equal to 1 if stock is in the highest Accrual quintile and the lowest BM quintile, and equal to 0 if stock is in the lowest Accrual quintile and the highest BM quintile. Exclude if book equity (ceq) is negative.	clear	Original
Basu	1977	Earnings-to-Price Ratio	EP	1957	1971	ib / lag(market value of equity, 6 months). NYSE stocks only. Exclude if EP < 0. Lag simulates the Dec 31 market equity used in original paper	clear	Original
Basu	1977	Earnings-to-Price Ratio	EPq	1963	1971	ib / lag(market value of equity, 6 months). NYSE stocks only. Exclude if EP < 0. Lag simulates the Dec 31 market equity used in original paper	clear	Quarterly
Bazdresch, Belo and Lin	2014	Employment growth	hire	1965	2010	Change in number of employees (emp) between t-1 and t, scaled by average number of employees in t-1 and t. Replace hire with 0 if emp or lagged emp is missing.	clear	Original
Belo and Lin	2012	Inventory Growth	InvGrowth	1965	2009	Inventory (invt) growth rate from fiscal year t-2 to fiscal year t-1	clear	Original
Belo, Lin and Vitorino	2014	Brand capital investment	BrandInvest	1975	2010	Advertising expenses (xad) divided by Brand-Capital. Brand capital is computed by the perpetual inventory method. In the first year, brand capital is advertising expense divided by (.5 + .1). In subsequent years, we let brand capital depreciate with a rate of .5 and add current advertising expenses. Brand capital is scaled by total assets (at). Set to missing if advertising expense is missing.	likely	Original

Continued on next page



**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Belo, Lin and Vitorino	2014	Brand capital to assets	BrandCapital	1975	2010	Brand capital is computed by the perpetual inventory method. In the first year, brand capital is advertising expense divided by (.5 + .1). In subsequent years, we let brand capital depreciate with a rate of .5 and add current advertising expenses. Brand capital is scaled by total assets (at). Set to missing if advertising expense is missing.	maybe	Original
Bhandari	1988	Market leverage	Leverage	1952	1981	Total liabilities (lt) divided by market value of equity.	clear	Original
Bhandari	1988	Market leverage quarterly	Leverage_q	1952	1981	Total liabilities (lt) divided by market value of equity.	clear	Quarterly
Blitz, Huij and Martens	2011	11 month residual momentum	ResidualMomentum11m	1930	2009	Run a rolling regression over 36 months of excess return (retrf) on excess market return (mktrf), size and value factors (smb, hml) and compute idiosyncratic returns as the one-month lagged residual. ResidualMomentum is the rolling mean of the residual divided by the rolling standard deviation of the residual, both computed over the past 11 months.	clear	Original
6 Blitz, Huij and Martens	2011	6 month residual momentum	ResidualMomentum6m	1930	2009	Run a rolling regression over 36 months of excess return (retrf) on excess market return (mktrf), size and value factors (smb, hml) and compute idiosyncratic returns as the one-month lagged residual. ResidualMomentum is the rolling mean of the residual divided by the rolling standard deviation of the residual, both computed over the past 6 months.	clear	Original
Blume and Husic	1972	Price	Price	1932	1971	Log of absolute value of price (prc).	clear	Original
Boudoukh et al.	2007	Net Payout Yield	NetPayoutYield	1984	2003	Dividends (dvc) plus purchase of common and preferred stock (prstkc) minus sale of common and preferred stock (sstk), divided by market value of equity.	clear	Original
Boudoukh et al.	2007	Payout Yield	PayoutYield	1984	2003	Sum of dividends (dvc), purchase of common and preferred stock (prstkc) and max(preferred stock redemption value (pstkrv), 0), divided by lag(market value of equity, 6 months). Exclude if PayoutYield $\leq 0$ .	clear	Original
Boudoukh et al.	2007	Net Payout Yield quarterly	NetPayoutYield_q	1984	2003	Dividends (dvc) plus purchase of common and preferred stock (prstkc) minus sale of common and preferred stock (sstk), divided by market value of equity.	clear	Quarterly

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Boudoukh et al.	2007	Payout Yield quarterly	PayoutYield_q	1984	2003	Sum of dividends (dvc), purchase of common and preferred stock (prstk) and max(preferred stock redemption value (pstkrv), 0), divided by lag(market value of equity, 6 months). Exclude if PayoutYield $\leq$ 0.	clear	Quarterly
Bradshaw, Richardson and Sloan	2006	Net debt financing	NetDebtFinance	1971	2000	Long-term debt issuance (dltis) minus long-term debt reduction (dltr) minus current debt changes (dlcch), scaled by average total assets (at) in years t-1 and t. Replace missing values of dlcch with 0. Exclude if ratio is greater than 1.	clear	Original
Bradshaw, Richardson and Sloan	2006	Net equity financing	NetEquityFinance	1971	2000	Sale of common stock (sstk) minus purchase of common stock (prstk), scaled by average total assets (at) from years t and t-1. Exclude if absolute value of ratio is greater than 1.	clear	Original
Bradshaw, Richardson and Sloan	2006	Net external financing	XFIN	1971	2000	Sale of common stock (sstk) minus dividends (dv) minus purchase of common stock (prstk) plus long-term debt issuance (dltis) minus long-term debt reductions (dltr). Scaled by total assets (at).	clear	Original
Brennan, Chordia and Subrahmanyam	1998	Past trading volume	DoIVol	1966	1995	Log of two-month lagged trading volume (vol) times two-month lagged price (prc).	clear	Original
Brown and Rowe	2007	Return on invested capital	roic	1970	2005	EBIT (ebit) minus non-operating income (nopi) divided by the sum of equity (ceq), liabilities (lt) and cash (che).	not	Original
Callen, Khan and Lu	2013	Accounting component of price delay	DelayAcct	1981	2006	Monthly cross-sectional regression of PriceDelay on AccrualQuality, special items (si) scaled by average total assets (at) and earnings surprise (meanest - actual) scaled by its cross-sectional standard deviation. DelayAcct is the predicted value from that regression.	not	Original
Callen, Khan and Lu	2013	Non-accounting component of price delay	DelayNonAcct	1981	2006	Monthly cross-sectional regression of PriceDelay on AccrualQuality, special items (si) scaled by average total assets (at) and earnings surprise (meanest - actual) scaled by its cross-sectional standard deviation. DelayNonAcct is the residual from that regression.	not	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Campbell, Hilscher and Szilagyi	2008	Failure probability	FailureProbability	1981	2003	Failure probability is $-9.16 - 0.058*PRICE + 0.075*MB - 2.13*CASHMTA - 0.045*RSIZE + 1.41*IdioRisk - 7.13*EXRETAVG + 1.42*TLMTA - 20.26*NIMTAAVG$ . PRICE is $\log(\min(\text{abs}(\text{prc}), 15))$ ; MB is $\text{shrout}*\text{abs}(\text{prc})/\text{ceqq}$ ; CASHMTA is $\text{cheq}/(\text{shrout}*\text{abs}(\text{prc}) + \text{ltq})$ ; RSIZE is $\log(\text{shrout}*\text{abs}(\text{prc})/\text{sum of shrout}*\text{abs}(\text{prc}) \text{ for the largest 500 companies each month})$ ; IdioRisk is defined above, EXRETAVG is the weighted average excess return $(\log(1 + \text{ret}) - \log(1 + \text{mktrf}))$ over the previous 12 months, with weight on month $t-j$ being $\phi^j$ and the sum scaled by $\frac{1-\phi}{1-\phi^{12}}$ ; TLMTA is total liabilities $(\text{ltq}/(\text{shrout}*\text{abs}(\text{prc})))$ ; NIMTAAVG is a weighted average of net income over total assets $(\text{ibq}/(\text{shrout}*\text{abs}(\text{prc}) + \text{ltq}))$ over four quarters, with weight $\phi^q$ on quarter $t-q$ and the sum scaled by $\frac{1-\phi^3}{1-\phi^{12}}$ . $\phi = 2^{-\frac{1}{3}}$ . All input variables are winsorized at the 5th and 95th percentile. Exclude if price less than 1.	not	Original
11 Campbell, Hilscher and Szilagyi	2008	Failure probability	FailureProbabilityJune	1981	2003	Failure probability is $-9.16 - 0.058*PRICE + 0.075*MB - 2.13*CASHMTA - 0.045*RSIZE + 1.41*IdioRisk - 7.13*EXRETAVG + 1.42*TLMTA - 20.26*NIMTAAVG$ . PRICE is $\log(\min(\text{abs}(\text{prc}), 15))$ ; MB is $\text{shrout}*\text{abs}(\text{prc})/\text{ceqq}$ ; CASHMTA is $\text{cheq}/(\text{shrout}*\text{abs}(\text{prc}) + \text{ltq})$ ; RSIZE is $\log(\text{shrout}*\text{abs}(\text{prc})/\text{sum of shrout}*\text{abs}(\text{prc}) \text{ for the largest 500 companies each month})$ ; IdioRisk is defined above, EXRETAVG is the weighted average excess return $(\log(1 + \text{ret}) - \log(1 + \text{mktrf}))$ over the previous 12 months, with weight on month $t-j$ being $\phi^j$ and the sum scaled by $\frac{1-\phi}{1-\phi^{12}}$ ; TLMTA is total liabilities $(\text{ltq}/(\text{shrout}*\text{abs}(\text{prc})))$ ; NIMTAAVG is a weighted average of net income over total assets $(\text{ibq}/(\text{shrout}*\text{abs}(\text{prc}) + \text{ltq}))$ over four quarters, with weight $\phi^q$ on quarter $t-q$ and the sum scaled by $\frac{1-\phi^3}{1-\phi^{12}}$ . $\phi = 2^{-\frac{1}{3}}$ . All input variables are winsorized at the 5th and 95th percentile. Exclude if price less than 1.	not	Lag structure

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Chan and Ko	2006	Momentum and LT Reversal	MomRev	1965	2001	Binary variable equal to 1 if firm is in the highest Mom6m quintile and the lowest Mom36m quintile, and equal to 0 if firm is in the lowest Mom6m quintile and the highest Mom36m quintile. Exclude if price less than 5.	clear	Original
Chan, Jegadeesh and Lakonishok	1996	Earnings forecast revisions	REV6	1977	1992	Define revisions as the change in the mean earnings estimate (meanest) for the next quarter from month t-1 to t, scaled by stock price in month t-1. REV6 is the sum of that variable from months t-6 to t.	clear	Original
Chan, Jegadeesh and Lakonishok	1996	Earnings announcement return	AnnouncementReturn	1977	1992	Get announcement date for quarterly earnings from IBES (fpi = 6). AnnouncementReturn is the sum of (ret - mktrf + rf) from one day before an earnings announcement to 2 days after the announcement.	clear	Original
Chan, Lakonishok and Sougiannis	2001	Advertising Expense	AdExp	1975	1996	Advertising expense (xad) over market value of equity (shrout*abs(prc))	clear	Original
Chan, Lakonishok and Sougiannis	2001	R&D over market cap	RD	1975	1995	R&D expense (xrd) over market value of equity.	clear	Original
Chan, Lakonishok and Sougiannis	2001	R&D to sales	rd_sale	1975	1995	One year lagged R&D (xrd) divided by one year lagged sales (sale).	not	Original
Chan, Lakonishok and Sougiannis	2001	R&D over market cap	RD_q	1975	1995	R&D expense (xrd) over market value of equity.	clear	Quarterly
Chan, Lakonishok and Sougiannis	2001	R&D to sales	rd_sale_q	1975	1995	One year lagged quarterly R&D (xrdq) divided by one year lagged quarterly sales (saleq).	not	Quarterly
Chandrashekar and Rao	2009	Cash Productivity	CashProd	1963	2003	Calculate market value of equity (mve_c) as absolute price (prc) times number of shares outstanding (shrout). Cash productivity is equal to the difference between mve_c and total assets (at) divided by cash and short-term investments (che).	clear	Original
Chen, Hong and Stein	2002	Breadth of ownership	DelBreadth	1979	1998	Quarterly change in the number of institutional owners (numinstowners) from 13F data. Exclude if in the lowest quintile of stocks by market value of equity (based on NYSE stocks only).	clear	Original
Chordia, Subrahmanyam and Anshuman	2001	Share turnover volatility	std_turn	1966	1995	Standard deviation of turnover (vol/shrout) over the past 36 months.	clear	Original
Chordia, Subrahmanyam and Anshuman	2001	Volume Variance	VolSD	1966	1995	Rolling standard deviation of monthly trading volume (vol) over the past 36 months (require at least 24 observations). Include only NYSE stocks.	clear	Original
Cohen and Frazzini	2008	Customer momentum	CustomerMomentum	1980	2004	Based on firms' principals customers from Compustat Segment data as in Cohen and Frazzini.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Cohen and Lou	2012	Conglomerate return	retConglomerate	1977	2009	Identify conglomerate firms as those with multiple OPSEG or BUSSEG entries in the Compustat segment data (and require that at least 80% of firm's total assets are covered by segment data). Compute monthly stock return at the 2-digit SIC level for stand-alone (non-conglomerate) firms only, and match those returns to conglomerates' segments. Compute weighted conglomerate return as the industry return of stand-alone companies, weighted with a conglomerate's total sales in each industry.	clear	Original
Cohen, Diether and Malloy	2013	R&D ability	RDAbility	1980	2009	Regress log of sales growth (sale over sale in previous year) on log of (1+xrd/sale) in 5 bivariate regressions with (1+xrd/sale) lagged by \$1, \dots, 5\$ years. Run regressions over previous 8 years and require at least 6 non-missing observations. Also require at least half of past research and development observations to be non-zero. RDAbility is the mean of the coefficients on the five lags of log(1+xrd/sale). Set to missing if firm is not in the highest tercile of xrd/sale in a year or if reseach and development expenses are non-positive.	clear	Original
Cooper, Schill, Gulen and	2008	Asset Growth	AssetGrowth	1968	2003	Annual growth rate of total assets (at)	clear	Original
Cooper, Schill, Gulen and	2008	Asset Growth quarterly	AssetGrowth_q	1968	2003	Annual growth rate of total assets (at)	clear	Quarterly
Cremers and Nair	2005	Shareholder activism 1	Activism1	1990	2001	24 minus Governance Index (G). Set to missing if G is missing, or if not in the highest quartile of institutional ownership (maxinstown\_perc), or if dual share class.	clear	Original
Cremers and Nair	2005	Shareholder activism 2	Activism2	1990	2001	Institutional ownership share (maxinstown\_perc) if share greater than 5 percent. Set to missing if G is missing, or if dual share class, or if 24 minus Governance index (G) is less than 19.	likely	Original
Cusatis, Miles and Woolridge	1993	Spinoffs	Spinoff	1965	1988	Spinoffs are identified as all observations in the CRSP acquisition file with valid acperm entry. Spinoff is a binary variable equal to 1 if a firm is identified in the CRSP Acquisition data and if it has at most one year of history in the CRSP stock return data. Spinoff is equal to 0 otherwise.	likely	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Da and Warachka	2011	Long vs short-term earnings expectations	EarningsForecastDisparity	1983	2006	Analyst forecasted 5-year earnings growth (fgr5yr) minus 100 times the difference between mean earnings forecast (meanest) and fiscal year earnings expectations (fy0a) scaled by the absolute value of fy0a.	clear	Original
Daniel and Titman	2006	Composite equity issuance	CompEquIss	1968	2003	5 year growth rate of market value of equity minus 5 year stock return.	clear	Original
Daniel and Titman	2006	Intangible return using BM	IntanBM	1968	2003	In each month, run a cross-sectional regression of a firm's five-year stock return on 5 year lagged BM (defined above) and a constructed regressor that is the change in BM from 5 years ago to today plus the five-year stock return. The residual from that regression is IntanBM.	clear	Original
Daniel and Titman	2006	Intangible return using CFtoP	IntanCFP	1968	2003	In each month, run a cross-sectional regression of a firm's five-year stock return on the 5 year lagged CFP = (net income (ni) plus depreciation (dp))/market value of equity and a constructed regressor that is the change in CFP from 5 years ago to today plus the five-year stock return. The residual from that regression is IntanCFP.	clear	Original
Daniel and Titman	2006	Intangible return using EP	IntanEP	1968	2003	In each month, run a cross-sectional regression of a firm's five-year stock return on the 5 year lagged EP = net income (ni)/market value of equity and a constructed regressor that is the change in EP from 5 years ago to today plus the five-year stock return. The residual from that regression is IntanEP.	clear	Original
Daniel and Titman	2006	Intangible return using Sale2P	IntanSP	1968	2003	In each month, run a cross-sectional regression of a firm's five-year stock return on 5 year lagged SP (defined above) and a constructed regressor that is the change in SP from 5 years ago to today plus the five-year stock return. The residual from that regression is IntanSP.	clear	Original
Daniel and Titman	2006	Share issuance (5 year)	ShareIss5Y	1968	2003	5-year growth in number of shares. Number of shares is calculated as shrou/cfacshr to adjust for splits.	clear	Original
Datar, Naik and Radcliffe	1998	Share Volume	ShareVol	1962	1991	Sum of monthly share trading volume (vol) over the previous three months, scaled by 3 times common shares outstanding (shrou). Drop if ShareVol is below its median	clear	Original
De Bondt and Thaler	1985	Momentum-Reversal	Mom18m13m	1933	1980	Stock return between months t-18 and t-13.	clear	Original
De Bondt and Thaler	1985	Long-run reversal	Mom36m	1929	1982	Stock return between months t-36 and t-13.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Dechow et al.	2001	Short Interest	ShortInterest	1976	1993	Short-interest from Compustat (shortint) scaled by shares outstanding (shrou). Short-interest data are available bi-weekly with a four day lag. We use the mid-month observation to make sure data would be available in real time.	clear	Original
Dechow, Sloan and Soliman	2004	Equity Duration	EquityDuration	1966	1999	see code	clear	Original
Desai, Rajgopal and Venkatachalam	2004	Operating Cash flows to price	cfp	1973	1997	Operating cash-flow (oancf) divided by market value of equity. If operating cash-flow is missing, replace by difference between net income (ib) and level of accruals, where the latter is the annual change in current assets (act) minus the annual change in cash and short-term investments (che), minus the annual change in current liabilities (lct) plus the annual change in debt in current liabilities (dlc) plus the annual change in payable income taxes (txp) plus depreciation (dp).	clear	Original
Desai, Rajgopal and Venkatachalam	2004	Operating Cash flows to price quarterly	cfpq	1973	1997	Operating cash-flow (oancf) divided by market value of equity. If operating cash-flow is missing, replace by difference between net income (ib) and level of accruals, where the latter is the annual change in current assets (act) minus the annual change in cash and short-term investments (che), minus the annual change in current liabilities (lct) plus the annual change in debt in current liabilities (dlc) plus the annual change in payable income taxes (txp) plus depreciation (dp).	clear	Quarterly
Dharan and Ikenberry	1995	Exchange Switch	ExchSwitch	1962	1990	Binary variable equal to 1 if a firm switched from AMEX or NASDAQ to NYSE within the past year, or from NASDAQ to AMEX within the past year.	clear	Original
Dichev	1998	O Score	OScore	1981	1995	OScore = $-1.32 - .407 \cdot \log(\text{at}/\text{GNP deflator}) + 6.03 \cdot (\text{lt}/\text{at}) - 1.43 \cdot (\text{act} - \text{lct})/\text{at} + .076 \cdot (\text{lct}/\text{act}) - 1.72 \cdot \text{I}(\text{lt} > \text{at}) - 2.37 \cdot (\text{ib}/\text{at}) - 1.83 \cdot (\text{fopt}/\text{lt}) + .285 \cdot (\text{ib} + \text{ib}_{\{t-12\}} + \text{ib}_{\{t-24\}} < 0) - .521 \cdot (\text{ib} - \text{ib}_{\{t-12\}})/(\text{abs}(\text{ib}) + \text{abs}(\text{ib}_{\{t-12\}}))$ . fopt = oancf if fopt is missing. Exclude Exclude if SIC code between 3999 and 4999, or greater than 5999. Exclude if price less than 5. Then exclude if OScore is in bottom quintile of OScore (original paper shows non-monotonic returns, as does our replication)	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Dichev	1998	Altman Z-Score	ZScore	1981	1995	1.2*(current assets (act) - current liabilities (lct))/total assets (at) + 1.4*(Retained earnings (re)/total assets (at)) + 3.3*(net income (ni) + interest expense (xint) + total taxes (txt))/total assets (at) + .6*(market value of equity/Total liabilities (lt)) + revenue (revt)/ total assets (at). Include only NYSE stocks. Exclude if SIC code between 4000 and 4999, or above 5999. Exclude if ZScore is in bottom quintile of ZScore (original paper shows non-monotonic returns, as does our replication)	clear	Original
Dichev	1998	O Score quarterly	OScore_q	1981	1995	OScore = -1.32 - .407*log(at/GNP deflator) + 6.03*(lt/at) - 1.43*(act - lct)/at + .076*(lct/act) - 1.72*I(lt > at) - 2.37*(ib/at) - 1.83*(fopt/lt) + .285*(ib + ib\$_{t-12}\$ + ib\$_{t-24}\$ < 0) - .521*(ib - ib\$_{t-12}\$)/(abs(ib) + abs(ib\$_{t-12}\$)). fopt = oancf if fopt is missing. Exclude Exclude if SIC code between 3999 and 4999, or greater than 5999. Exclude if price less than 5. Then exclude if OScore is in bottom quintile of OScore (original paper shows non-monotonic returns, as does our replication)	clear	Quarterly
Dichev	1998	Altman Z-Score quarterly	ZScore_q	1981	1995	1.2*(current assets (act) - current liabilities (lct))/total assets (at) + 1.4*(Retained earnings (re)/total assets (at)) + 3.3*(net income (ni) + interest expense (xint) + total taxes (txt))/total assets (at) + .6*(market value of equity/Total liabilities (lt)) + revenue (revt)/ total assets (at). Include only NYSE stocks. Exclude if SIC code between 4000 and 4999, or above 5999. Exclude if ZScore is in bottom quintile of ZScore (original paper shows non-monotonic returns, as does our replication)	clear	Quarterly
Dichev and Piotroski	2001	Credit Rating Downgrade	CredRatDG	1986	1998	A downgrade happens if credit rating (spltrcm) decreased by at least one notch relative to the previous month. CredRatDG = 1 if a downgrade happened over the past 3 months. OP studies Moody's ratings changes between 1970 and 1997, but our data doesn't begin in earnest until 1986, and our sample definitions apply to returns, and this predictor implicitly averages over the past 6 months.	clear	Original
Diether, Malloy and Scherbina	2002	EPS Forecast Dispersion	ForecastDispersion	1976	2000	Standard deviation of earnings estimates (stdev\_est) scaled by mean earnings estimate.	clear	Original

Continued on next page



**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Dimson	1979	Dimson Beta	BetaDimson	1955	1974	Rolling regression of daily return (ret - rf) on the same-day, one-day ahead, and one-day lagged value of the market return (mktrf). Rolling regression with 20 observations (minimum 15). BetaDimson is the sum of the three coefficients.	maybe	Original
Doyle, Lundholm and Soliman	2003	Excluded Expenses	ExclExp	1988	1999	Difference between unadjusted earnings (EPS-ActualUnadj) from IBES and quarterly earnings per share (epspiq). Exclude the highest and lowest 1% of values.	clear	Original
Easley, Hvidkjaer and O'Hara	2002	Probability of Informed Trading	ProbInformedTrading	1984	1998	Downloaded from Soren Hvidkjaer's website: <a href="https://sites.google.com/site/hvidkjaer/data">\url{https://sites.google.com/site/hvidkjaer/data}</a>	clear	Original
Eberhart, Maxwell and Siddique	2004	Unexpected R&D increase	SurpriseRD	1974	2001	Binary variable equal to 1 if: R&D (xrd) scaled by revenue (revt) is positive, R&D scaled by total assets (at) is positive, annual R&D growth is greater than 5%, annual growth in R&D over total assets is greater than 5%. SurpriseRD is 0 otherwise.	clear	Original
Eisfeldt and Papanikolaou	2013	Organizational Capital	OrgCap	1970	2008	Defined recursively. Initialize with OrgCap = 4*general expenses (xsga) in the first year, and calculate as .85*OrgCap previous year + xsga current year thereafter. Scale by total assets (at).	clear	Original
Eisfeldt and Papanikolaou	2013	Organizational Capital industry adj	OrgCapAdj	1970	2008	OrgCap winsorized at the 1% level and standardized by subtracting the same month within-industry mean (Fama French 17 industries) and dividing by the same month within industry standard deviation of OrgCap.	clear	Original
Elgers, Lo and Pfeiffer	2001	Earnings Forecast to price	sfe	1982	1999	Mean earnings estimate (meanest) for next quarter's earnings divided by stock price (prc). Exclude if price less than 1.	likely	Original
Elgers, Lo and Pfeiffer	2001	Number of analysts	nanalyst	1982	1998	Number of estimates (numest) in IBES for one-quarter ahead earnings. Replace with 0 if missing after 1989.	maybe	Original
Fairfield, Whisenant and Yohn	2003	Growth in Long term net operating assets	GrLTNOA	1964	1993	Annual growth in net operating assets, minus accruals. Net operating assets are (rect + invt + ppent + aco + intan + ao- ap- lco- lo) / at. Accruals are ( rect-l12.rect + invt - l12.invt + aco - l12.aco - (ap - l12.ap + lco - l12.lco) - dp ) / ((at + l12.at)/2)	likely	Original
Fama and French	1992	Total assets to market	AM	1963	1990	Total assets (at) divided by market value of equity.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Fama and French	1992	Book leverage (annual)	BookLeverage	1963	1990	Total assets (at) divided by book value of equity plus deferred taxes (txditi) and preferred stock. Equity is shareholder equity (seq) if available, or book equity (ceq) plus preferred stock (pstk, if missing pstkrv, if missing pstkl), or total assets minus total liabilities (lt).	clear	Original
Fama and French	1992	Total assets to market (quarterly)	AMq	1975	1990	Total assets (at) divided by market value of equity.	clear	Quarterly
Fama and French	1992	Book leverage (quarterly)	BookLeverageQuarterly	1973	1990	Total assets (atq) divided by book value of equity plus deferred taxes (txditiq) and preferred stock. Equity is shareholder equity (seqq) if available, or book equity (ceqq) plus preferred stock (pstkq), or total assets minus total liabilities (ltq).	clear	Quarterly
Fama and French	2006	operating profits / book equity	OperProf	1977	2003	Revenue (revt) minus cost (cogs) - administrative expenses (xsga) - interest expenses (xint), scaled by book value of equity (ceq). Exclude smallest size tercile.	likely	Original
Fama and French	2006	operating profits / book equity	OperProfLag	1977	2003	Revenue (revt) minus cost (cogs) - administrative expenses (xsga) - interest expenses (xint), scaled by lagged book value of equity (ceq). Exclude smallest size tercile.	clear	Lag structure
Fama and French	2006	operating profits / book equity	OperProfLag_q	1977	2003	Revenue (revtq) minus cost (cogsq) - administrative expenses (xsgaq) - interest expenses (xintq), scaled by equity. Equity is shareholder equity (seqq) if not missing. Otherwise, the sum of book equity (ceqq) and preferred stock (pstkq) if not missing. Otherwise, assets (atq) minus liabilities (ltq).	clear	Quarterly
Fama and MacBeth	1973	CAPM beta	Beta	1929	1968	Coefficient of a 60-month rolling window regression of monthly stock returns minus the riskfree rate on market return minus the risk free rate (ewretd - rf). Exclude if estimate based on less than 20 months of returns.	likely	Original
Fama and MacBeth	1973	CAPM beta squared	BetaSquared	1929	1968	Square of Beta (defined above).	not	Original
Foster, Olsen and Shevlin	1984	Earnings Surprise	EarningsSurprise	1974	1981	EPS (epspxq) minus EPS twelve months ago - Drift, scaled by standard deviation of that expression. Drift is the average earnings growth (EPS - EPS twelve months ago) over the past two years. Exclude if price less than 5	clear	Original
Francis, Lafond, Olsson and Schipper	2004	Earnings persistence	EarningsPersistence	1975	2001	Earnings per share (epspx/ajex) regressed on its value from the previous year. Rolling regression in annual data with 10 observations. EarningsPersistence is the coefficient on the lagged value.	maybe	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Francis, Lafond, Olsson and Schipper	2004	Earnings Predictability	EarningsPredictability	1975	2001	Earnings per share (epspx/ajex) regressed on its value from the previous year. Rolling regression in annual data with 10 observations. EarningsPredictability is the R2 of this regression.	maybe	Original
Francis, Lafond, Olsson and Schipper	2004	Earnings Smoothness	EarningsSmoothness	1975	2001	10-year rolling standard deviation of (ib/at) divided by 10-year rolling standard deviation of cash-flow ((ib - (act - act_{t-1}) - (lct - lct_{t-1}) - (che - che_{t-1}) + (dlc - dlc_{t-1}) - dp))/at_{t-1})	maybe	Original
Francis, LaFond, Olsson and Schipper	2004	Earnings conservatism	EarningsConservatism	1975	2001	Earnings (ib) scaled by market capitalization (shrou*abs(prc)) regressed on the 15 month stock return from (t=-11 to t=+3), an indicator for whether that 15- month return is negative and the interaction of these two variables. Rolling regression in annual data with 10 observations. EarningsConservatism is the sum of the coefficients on the 15-month stock return and the interaction term divided by the coefficient on the 15-month stock return.	maybe	Original
Francis, LaFond, Olsson and Schipper	2004	Earnings timeliness	EarningsTimeliness	1975	2001	Earnings (ib) scaled by market capitalization (shrou*abs(prc)) regressed on the 15 month stock return from (t=-11 to t=+3), an indicator for whether that 15- month return is negative and the interaction of these two variables. Rolling regression in annual data with 10 observations. EarningsTimeliness is the R2 of this regression.	maybe	Original
Francis, LaFond, Olsson and Schipper	2004	Value relevance of earnings	EarningsValueRelevance	1975	2001	15 month stock return from (t=-11 to t=+3) regressed on Earnings (ib) and the change in earnings, both scaled by market capitalization (shrou*abs(prc)). Rolling regression in annual data with 10 observations. EarningsValueRelevance is the R2 of this regression.	maybe	Original
Francis, LaFond, Olsson and Schipper	2004	RoA volatility	roavol	1975	2001	Rolling standard deviation of quarterly return on assets (roaq) over 4 years (minimum 2 years).	maybe	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Francis, LaFond, Olsson and Schipper	2005	Accrual Quality	AccrualQuality	1971	2002	Define Accruals (tempAccruals) as (difference between current assets (act) and one-year lagged current assets) - (difference between cash and short-term investments (che) and one-year lagged cash and short-term investments) - (difference between current liabilities (lct) and one-year lagged current liabilities) + (difference between debt in current liabilities (dlc) and one-year lagged debt in current liabilities) - (depreciation and amortization (dp)) all divided by (mean of total assets (at) and one-year lagged total assets). Create tempCAcc as tempAccruals + dp / ((at + 1.at)/2), tempRev as sale / ((at + 1.at)/2), tempDelRev as tempRev - 1.tempRev, tempPPE as ppeg / ((at + 1.at)/2) where ppeg is total gross property, plant and equipment and tempCFO as ib / ((at + 1.at)/2) - tempAccruals where ib is the income before extraordinary items. Run a regression for each year and industry of tempCAcc on the current value and one year lead and lag of tempDelRev and tempPPE. Save the regression residuals and replace with missing if there are not at least 20 observations per year and industry. Calculate accrual quality (AQ) as the standard deviation of residuals over 4 years. If more than one observation is missing set AQ to missing. Replace AccrualQuality by the one-year lagged AQ to make sure the signal is available at time of investment.	maybe	Original
Francis, LaFond, Olsson and Schipper	2005	Accrual Quality in June	AccrualQualityJune	1971	2002	see AccrualQuality. Update only with June values for each variable.	maybe	Original
Frankel and Lee	1998	Predicted Analyst forecast error	PredictedFE	1979	1993	Define FROE as mean earnings estimate (meanest) times shares outstanding (shROUT), divided by book equity (ceq). Define the prediction error as net income (ib) over book equity (ceq), minus FROE. In each month t, regress the prediction error on 3 year lagged values of a firm's relative ranks in the cross-sectional revenue (sale), BM (defined above), AOP (defined above) and FROE distributions. PredictedFE is the fitted value from that regression. Update monthly.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Frankel and Lee	1998	Analyst Value	AnalystValue	1975	1993	Analyst value is $(1 + (FROE - .1)/1.1 + (FROE - .1)/(1.1*1.1)) * BM_{Ave}$ . FROE is the most recent mean analyst EPS forecast (mean-est) times shares outstanding (shROUT) divided by book value of common equity (ceq). BM <sub>Ave</sub> is average book to market equity ( $ceq / (shROUT * abs(prc))$ ) over the past two years. Exclude if FROE > 1 or book equity negative or $abs(prc) < 1$ or $div/earnings > 1$ or $ib/ceq > 1$ (Frankel and Lee, page 291).	likely	Original
Frankel and Lee	1998	Analyst Optimism	AOP	1975	1993	AnalystValue (defined above) minus IntrinsicValue (defined above), divided by $abs(IntrinsicValue)$ .	likely	Original
Frankel and Lee	1998	Intrinsic or historical value	IntrinsicValue	1975	1993	Define FROE as net income (ib) divided by book equity (ceq), and drop if $abs(FROE) > 1$ . Define AveBM as average book equity to market value of equity for years t-1 and t (in the first year of coverage, use book-to-market equity directly). IntrinsicValue is $\$(1 + \frac{FROE - .1}{1.1} + \frac{(FROE - .1)}{1.1} * AveBM)$ . Exclude if price less than 1 or book equity less than 0.	maybe	Original
Franzoni and Marin	2006	Pension Funding Status	FR	1980	2002	FR = (FVPA - PBO), scaled by market value of equity. FVPA is pbnaa from 1980 to 1986, pplao + pplao from 1987 to 1997, and pplao after 1997. PBO is pbnvv from 1980 to 1986, pbpro + pbpru from 1987 to 1997, and pbpro after 1997. Exclude if price less than 5 or shrcd > 11.	clear	Original
Franzoni and Marin	2006	Pension Funding Status	FRbook	1980	2002	FR = (FVPA - PBO), scaled by total assets (at). FVPA is pbnaa from 1980 to 1986, pplao + pplao from 1987 to 1997, and pplao after 1997. PBO is pbnvv from 1980 to 1986, pbpro + pbpru from 1987 to 1997, and pbpro after 1997. Exclude if price less than 5 or shrcd > 11.	likely	Original
Frazzini and Pedersen	2014	Frazzini-Pedersen Beta	BetaFP	1929	2012	see code	likely	Original
George and Hwang	2004	52 week high	High52	1963	2001	Price (prc/cfacshr) divided by the maximum price over the previous 12 months.	clear	Original
Gompers, Ishii and Metrick	2003	Governance Index	G_Binary	1990	1999	Index available from <a href="http://faculty.som.yale.edu/andrewmetrick/data.html">http://faculty.som.yale.edu/andrewmetrick/data.html</a> . The index is only available every 2-3 years for each firm, we replace intermediate missing values with the latest available one. Value-weighted.	clear	Original
Gou, Lev and Shi	2006	IPO and no R&D spending	RDIP0	1980	1995	Binary variable equal to 1 if R&D expense (xrd) = 0 and IndIPO = 1. 0 otherwise.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Grinblatt and Moskowitz	1999	Industry Momentum	IndMom	1963	1995	Weighted average of firm-level 6 month buy-and-hold return. Average is taken over two digit industries each month and weights are based on market value of equity.	clear	Original
Hafzalla, Lundholm and Van Winkle	2011	Percent Abnormal Accruals	AbnormalAccrualsPercent	1989	2008	Income before extraordinary items (ib) minus net cash flow (oancf) divided by absolute value of ib. If oancf is missing, PctAcc is defined as $(\text{act} - \text{act}_{t-12}) - (\text{che} - \text{che}_{t-12}) - ((\text{lct} - \text{lct}_{t-12}) - (\text{dlc} - \text{dlc}_{t-12}) - (\text{txp} - \text{txp}_{t-12}) - \text{dp}) / \text{abs}(\text{ib})$ . In either case, if ib is equal to 0, divide by .01 instead. Exclude if price less than 5.	clear	Original
Hafzalla, Lundholm and Van Winkle	2011	Percent Operating Accruals	PctAcc	1989	2008	Income before extraordinary items (ib) minus net cash flow (oancf) divided by absolute value of ib. If oancf is missing, PctAcc is defined as $(\text{act} - \text{act}_{t-12}) - (\text{che} - \text{che}_{t-12}) - ((\text{lct} - \text{lct}_{t-12}) - (\text{dlc} - \text{dlc}_{t-12}) - (\text{txp} - \text{txp}_{t-12}) - \text{dp}) / \text{abs}(\text{ib})$ . In either case, if ib is equal to 0, divide by .01 instead. Exclude if price less than 5.	clear	Original
Hafzalla, Lundholm and Van Winkle	2011	Percent Total Accruals	PctTotAcc	1989	2008	Net income (ni) minus (purchase of common and preferred stock (prstkcc) minus sale of common and preferred stock (sstk) plus dividends (dvt), cash flow from operations (oancf), from financing (fincf) and investment (ivncf)). Scaled by absolute value of net income.	clear	Original
Hahn and Lee	2009	Tangibility	tang	1973	2001	Cash and short-term investments (che) plus .715*receivables (rect) + .547*inventory (invt) + .535* property, plant and equipment (ppent), scaled by total assets (at). Only defined for manufacturing firms (SIC $\geq$ 2000 and SIC <4000). Exclude the lowest tercile of manufacturing firms by total assets.	clear	Original
Hahn and Lee	2009	Tangibility quarterly	tang_q	1973	2001	Cash and short-term investments (che) plus .715*receivables (rect) + .547*inventory (invt) + .535* property, plant and equipment (ppent), scaled by total assets (at). Only defined for manufacturing firms (SIC $\geq$ 2000 and SIC <4000). Exclude the lowest tercile of manufacturing firms by total assets.	clear	Quarterly
Hartzmark and lomon	Sa- 2013	Dividends	DivInd	1927	2011	Binary variable equal to 1 if return with dividends (ret) is greater than return without dividends (retx) 11 months ago or 2 months ago, and 0 otherwise or if price less than 5.	clear	Original
Harvey and Siddique	2000	Coskewness	Coskewness	1963	1993	see code	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Haugen and Baker	1996	Capital turnover	CapTurnover	1979	1993	Lagged sales (sale) divided by two-year lagged assets (at).	maybe	Original
Haugen and Baker	1996	net income / book equity	RoE	1979	1993	Net income (ni) over book value of equity (ceq). Exclude if price less than 5.	maybe	Original
Haugen and Baker	1996	Cash-flow to price variance	VarCF	1979	1993	Rolling variance of (ib+dp)/mve\_c over the past 60 months (minimum 24 months data required).	maybe	Original
Haugen and Baker	1996	Volume to market equity	VolMkt	1979	1993	Average monthly dollar trading volume (vol*abs(prc)) over the previous 12 months, scaled by market value of equity. Exclude if price less than 5.	maybe	Original
Haugen and Baker	1996	Volume Trend	VolumeTrend	1979	1993	Rolling coefficient from regressing monthly trading volume on a linear time trend over a window of 60 months (require that at least 30 exist). Scale coefficient by 60-month average of trading volume.	maybe	Original
Haugen and Baker	1996	Capital turnover (quarterly)	CapTurnover_q	1979	1993	Sales (saleq) divided by one quarter lagged assets (atq).	maybe	Quarterly
Heston and Sadka	2008	Return seasonality	MomSeas	1965	2002	Average return in the same month over the preceding 2-5 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Return seasonality years 11 to 15	MomSeasAlt11to15a	1965	2002	Average return in the same month over the preceding 11-15 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Return seasonality years 16 to 20	MomSeasAlt16to20a	1965	2002	Average return in the same month over the preceding 16-20 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Returns in not-same month years 16 to 20	MomSeasAlt16to20n	1965	2002	Average return in other months over the preceding 16-20 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Return seasonality last year	MomSeasAlt1a	1965	2002	Average return in the same month in the previous year. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Returns in not-same month last year	MomSeasAlt1n	1965	2002	Average return in other months over the previous year. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Returns in not-same years 2 to 5	MomSeasAlt2to5n	1965	2002	Average return in other months over the preceding 2-5 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Return seasonality years 6 to 10	MomSeasAlt6to10a	1965	2002	Average return in the same month over the preceding 6-10 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Returns in different months years 6 to 10	MomSeasAlt6to10n	1965	2002	Average return in other months over the preceding 6-10 years. Exclude NASDAQ stocks.	clear	Original
Heston and Sadka	2008	Returns in different months years 11 to 15	MomSeasAlt11to15n	1965	2002	Average return in other months over the preceding 11-15 years. Exclude NASDAQ stocks.	not	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Hirschleifer, Hsu and Li	2013	Citations to RD expenses	CitationsRD	1982	2008	Set expenses for research and development (xrd) and number of citations (ncit) to 0 if missing. Calculate Citations to RD as sum of citations over previous 5 years divided by sum of xrd over years $t-3, \dots, t-7$ . Set citations to RD to missing if the date is before 1982 or if firm is in the upper half of market cap (shares outstanding times price) distribution in a month.	clear	Original
Hirschleifer, Hsu and Li	2013	Patents to RD expenses	PatentsRD	1982	2008	Set expenses for research and development (xrd) and number of patents (npat) to 0 if missing. Calculate Patents to RD as patents in a year divided by $xrd_{t-2} + .8 xrd_{t-3} + .6 xrd_{t-4} + .4xrd_{t-5} + .2 xrd_{t-6}$ . Set patents to RD to missing if the date is before 1982 or if firm is in the upper half of market cap (shares outstanding times price) distribution in a month.	clear	Original
Hirshleifer et al.	2004	Net Operating Assets	NOA	1964	2002	Difference between operating assets and operating liabilities, scaled by lagged total assets. Operating assets are total assets (at) minus cash- and short-term investments (che), operating liabilities are total assets minus long-term debt (dltt), minority interest (mib), deferred charges (dc) and book equity (ceq).	clear	Original
Hirshleifer, Hou, Teoh, Zhang	2004	change in net operating assets	dNoa	1964	2002	12-month growth in Net Operating Assets scaled by lagged total assets (at). Net Operating assets are operating assets minus operating liabilities. Operating assets are total assets (at) minus cash- and short-term investments (che), operating liabilities are total assets minus long-term debt (dltt), minority interest (mib), deferred charges (dlc), book equity (ceq) and preferred stock (pstk), all items (except at and ceq) replaced with 0 if missing.	clear	Original
Holthausen and Larcker	1992	Depreciation to gross PPE	depr	1978	1988	Depreciation and amortization (dp) divided by property, plant and equipment net total (ppent).	maybe	Original
Holthausen and Larcker	1992	Change in depreciation to gross PPE	pchdepr	1978	1988	Annual percentage change in the ratio of depreciation (dp) to property, plant and equipment (ppent).	maybe	Original

Continued on next page



**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Hong and Kacperczyk	2009	Sin Stock (selection criteria)	sinAlgo	1926	2006	Using Compustat Segment data, sinAlgo is defined as a binary variable equal to 1 if at least one segment of a firm is listed as being in at least one of the following industries: sic $\geq$ 2100 & sic $\leq$ 2199, sic $\geq$ 2080 & sic $\leq$ 2085, NAICS in $\{7132, 71312, 713210, 71329, 713290, 72112, 721120\}$ . As in the original paper, we assume that the sin stock indicator applies to the entire history and future of the identified firm. sinAlgo is equal to 0 if the firm is not identified in the CS Segment data as a sin stock and if the firm is in one of the following industries: (sic $\geq$ 2000 & sic $\leq$ 2046) OR (sic $\geq$ 2050 & sic $\leq$ 2063) OR (sic $\geq$ 2070 & sic $\leq$ 2079) OR (sic $\geq$ 2090 & sic $\leq$ 2092) OR (sic $\geq$ 2095 & sic $\leq$ 2099) OR (sic $\geq$ 2064 & sic $\leq$ 2068) OR (sic $\geq$ 2086 & sic $\leq$ 2087) OR (sic $\geq$ 920 & sic $\leq$ 999) OR (sic $\geq$ 3650 & sic $\leq$ 3652) OR sic == 3732 OR (sic $\geq$ 3931 & sic $\leq$ 3932) OR (sic $\geq$ 3940 & sic $\leq$ 3949) OR (sic $\geq$ 7800 & sic $\leq$ 7833) OR (sic $\geq$ 7840 & sic $\leq$ 7841) OR (sic $\geq$ 7900 & sic $\leq$ 7911) OR (sic $\geq$ 7920 & sic $\leq$ 7933) OR (sic $\geq$ 7940 & sic $\leq$ 7949) OR sic == 7980 OR (sic $\geq$ 7990 & sic $\leq$ 7999)	likely	Original
Hou	2007	Earnings surprise of big firms	EarnSupBig	1972	2001	Average monthly value of EarningsSurprise (defined above) of the 30% largest companies by market value of equity in the same Fama-French 48 industry. Exclude the largest 30% of companies for EarnSupBig (not to compute the anomaly)	clear	Original
Hou	2007	Industry return of big firms	IndRetBig	1972	2001	Average monthly return (ret) of the 30% largest companies by market value of equity in the same Fama-French 48 industry. Exclude the largest 30% of companies for IndRetBig (not to compute the anomaly!)	clear	Original
Hou and Loh	2016	Bid-ask spread based on TAQ data	BidAskTAQ	1984	2012	TAQ-based trading costs estimates (SAS code).	maybe	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Hou and Moskowitz	2005	Price delay r square	PriceDelayRsq	1964	2001	Regress daily stock return (ret) on market return (mktrf) in $t, t-1, \dots, t-4$ with observations over the previous year. Trim the highest and lowest 1% of estimated coefficients. Define PriceDelay as the ratio of $1 \cdot \text{beta on mktrf}_{t-1} + 2 \cdot \text{beta on mktrf}_{t-2} + 3 \cdot \text{beta on mktrf}_{t-3} + 4 \cdot \text{beta on mktrf}_{t-4}$ , and beta on $\text{mktrf}_t + \text{beta on mktrf}_{t-1} + \text{beta on mktrf}_{t-2} + \text{beta on mktrf}_{t-3} + \text{beta on mktrf}_{t-4}$ . The final variable is the average of that ratio over the previous month.	clear	Original
Hou and Moskowitz	2005	Price delay coeff	PriceDelay	1964	2001	Regress daily stock return (ret) on market return (mktrf) in $t, t-1, \dots, t-4$ with observations over the previous year. Trim the highest and lowest 1% of estimated coefficients. Define PriceDelay as the ratio of $1 \cdot \text{beta on mktrf}_{t-1} + 2 \cdot \text{beta on mktrf}_{t-2} + 3 \cdot \text{beta on mktrf}_{t-3} + 4 \cdot \text{beta on mktrf}_{t-4}$ , and beta on $\text{mktrf}_t + \text{beta on mktrf}_{t-1} + \text{beta on mktrf}_{t-2} + \text{beta on mktrf}_{t-3} + \text{beta on mktrf}_{t-4}$ . The final variable is the average of that ratio over the previous month.	likely	Original
Hou and Moskowitz	2005	Price delay SE adjusted	PriceDelayAdj	1964	2001	Regress daily stock return (ret) on market return (mktrf) in $t, t-1, \dots, t-4$ with observations over the previous year. Trim the highest and lowest 1% of estimated coefficients. Define PriceDelay as the ratio of $1 \cdot \text{beta on mktrf}_{t-1} + 2 \cdot \text{beta on mktrf}_{t-2} + 3 \cdot \text{beta on mktrf}_{t-3} + 4 \cdot \text{beta on mktrf}_{t-4}$ , and beta on $\text{mktrf}_t + \text{beta on mktrf}_{t-1} + \text{beta on mktrf}_{t-2} + \text{beta on mktrf}_{t-3} + \text{beta on mktrf}_{t-4}$ . The final variable is the average of that ratio over the previous month.	likely	Original
Hou and Robinson	2006	Industry concentration (Herfindahl) sales	Herf	1963	2001	Three-year rolling average of the three digit industry Herfindahl index based on firm revenue (sale). Exclude regulated industries (4011, 4210, 4213 & year $\leq 1980$ ; 4512 & year $\leq 1978, 4812, 4813$ & year $\leq 1982, 4900-4999$ in any year)	clear	Original
Hou and Robinson	2006	Industry concentration (Herfindahl) assets	HerfAsset	1963	2001	Three-year rolling average of the three digit industry Herfindahl index based on firm assets (at). Exclude regulated industries (4011, 4210, 4213 & year $\leq 1980$ ; 4512 & year $\leq 1978, 4812, 4813$ & year $\leq 1982, 4900-4999$ in any year)	likely	Original

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Hou and Robinson	2006	Industry concentration (Herfindahl) book	HerfBE	1963	2001	Three-year rolling average of the three digit industry Herfindahl index based on firm book equity. Exclude regulated industries (4011, 4210, 4213 & year $\leq$ 1980; 4512 & year $\leq$ 1978, 4812, 4813 & year $\leq$ 1982, 4900-4999 in any year)	likely	Original
Hou, Xue and Zhang	2018	Change in Return on assets	ChangeRoA	1973	2016	Quarterly return on assets (rdq/atq) minus its value four quarters ago.	clear	Original
Hou, Xue and Zhang	2018	Change in Return on equity	ChangeRoE	1973	2016	Quarterly return on equity (ceqq/atq) minus its value four quarters ago.	clear	Original
Ikenberry, Lakonishok and Vermaelen	1995	Share repurchases	ShareRepurchase	1980	1990	Binary variable equal to 1 if stock repurchase indicated in cash flow statement (prstk > 0), and 0 if prstk = 0.	clear	Original
Jegadeesh	1989	Short term reversal	Mom1m	1934	1987	Stock return (ret) over the previous month.	clear	Original
Jegadeesh and Livnat	2006	Revenue Surprise	RevenueSurprise	1987	2003	Define revenue per share as quarterly revenue (revtq) divided by quarterly common shares outstanding (cshprq). RevenueSurprise is the 4-quarter change in revenue per share minus the average 4-quarter change in revenue per share over the previous 2 years. RevenueSurprise is scaled by its standard deviation over the previous 2 years. Exclude if price less than 5.	clear	Original
Jegadeesh and Titman	1993	Momentum (12 month)	Mom12m	1964	1989	Stock return between months t-12 and t-1.	clear	Original
Jegadeesh and Titman	1993	Momentum (6 month)	Mom6m	1964	1989	Stock return between months t-6 and t-1. Exclude if price less than 5.	clear	Original
Jegadeesh et al.	2004	Change in recommendation	ChangeInRecommendation	1994	1998	(As in MP). If an analyst issues a new strong buy recommendation (ireccd == 1), we assign a value of 1 to that event, if an analyst issues any other change in recommendation, we assign a value of -1; we assign 0 if the recommendation is unchanged. The final variable is the average over the constructed variable over all analysts each month.	clear	Original
Johnson and So	2012	Option Volume to Stock Volume	OptionVolume1	1996	2010	Total monthly option volume (volume) over all puts and calls, divided by monthly stock trading volume (vol). Exclude if price less than 1 or share code greater 11 or option volume or stock volume data are missing for the previous month.	clear	Original
Johnson and So	2012	Option Volume relative to recent average	OptionVolume2	1996	2010	Based off of OptionVolume1. OptionVolume2 = OptionVolume1 / average of OptionVolume1 from months t-6 to t-1.	likely	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Kelly and Jiang	2014	Tail risk beta	BetaTailRisk	1963	2010	Each month, compute the 5th percentile over daily returns over all firms. For all daily return observations with return below that 5th percentile, compute the average of (log(ret/5th percentile of cross-sectional return distribution). Call that average tailEX. BetaTailRisk is the coefficient of a 120-month rolling regression of a firm's stock return on tailEX. Exclude if price less than 5 or share code greater than 11.	clear	Original
La Porta	1996	Long-term EPS forecast	fgr5yrLag	1983	1990	Long-term earnings forecast (fgr5yr) lagged by twelve months. Exclude if book equity (ceq), net income (ib), deferred taxes (txdi), dividends (dvp), revenue (sale) or depreciation (dp) is missing.	clear	Original
La Porta	1996	Long-term EPS forecast from June	fgr5yrLagJune	1983	1990	Long-term earnings forecast (fgr5yr) lagged by twelve months. Exclude if book equity (ceq), net income (ib), deferred taxes (txdi), dividends (dvp), revenue (sale) or depreciation (dp) is missing.	clear	Lag structure
Lakonishok, and Vishny	Shleifer 1994	Cash flow to market	CF	1968	1990	Net income (ib) plus depreciation (dp) divided by market equity. Exclude NASDAQ stocks.	clear	Original
Lakonishok, and Vishny	Shleifer 1994	Revenue Growth Rank	MeanRankRevGrowth	1968	1990	Rank firms by their annual revenue growth each year over the past 5 years. MeanRankRevGrowth is the weighted average of ranks over the past 5 years, that is, $MeanRankRevGrowth = (5*Rank_{t-1} + 4*Rank_{t-2} + 3*Rank_{t-3} + 2*Rank_{t-4} + 1*Rank_{t-5})/15$ . Exclude NASDAQ stocks.	clear	Original
Lakonishok, and Vishny	Shleifer 1994	Cash flow to market quarterly	CFq	1968	1990	Net income (ib) plus depreciation (dp) divided by market equity. Exclude NASDAQ stocks.	clear	Quarterly
Lakonishok, and Vishny	Shleifer 1994	Annual sales growth	sgr	1968	1990	Sales (sale) relative to t-1.	clear	Quarterly
Lakonishok, and Vishny	Shleifer 1994	Annual sales growth quarterly	sgr_q	1968	1990	Sales (sale) relative to t-1.	clear	Quarterly

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Lamont, Polk and Saa- Requejo	2001	Kaplan Zingales index	KZ	1968	1997	-1.002* (net income (ni) + depreciation (dp))/total assets (at) + .283*(total assets (at) + market value of equity - book value of equity (ceq) - deferred taxes (txdi))/total assets (at) + 3.319*(debt in current liabilities (dlc) + long-term debt (dltt))/(debt in current liabilities + long-term debt + book value of equity) - 39.368*(Dividends (divamt)/total assets) - 1.315*(cash and short-term investments (che)/total assets). Replace txdi and divamt with 0 if missing.	clear	Original
Lamont, Polk and Saa- Requejo	2001	Kaplan Zingales index quarterly	KZ_q	1968	1997	-1.002* (net income (ni) + depreciation (dp))/total assets (at) + .283*(total assets (at) + market value of equity - book value of equity (ceq) - deferred taxes (txdi))/total assets (at) + 3.319*(debt in current liabilities (dlc) + long-term debt (dltt))/(debt in current liabilities + long-term debt + book value of equity) - 39.368*(Dividends (divamt)/total assets) - 1.315*(cash and short-term investments (che)/total assets). Replace txdi and divamt with 0 if missing.	clear	Quarterly
Landsman et al.	2011	Real dirty surplus	RDS	1976	2003	Define Dirty Surplus as annual change in marketable securities adjustment msa plus annual change in retained earnings adjustment (recta) + .65 times the annual change in min(Unrecognized prior service cost (pcupsu) - Pension additional minimum liability (pad-dml),0). Real dirty surplus is the annual change in book equity (ceq) minus dirty surplus minus (net income (ni) minus dividends preferred (dvp)) + dividends (divamt) - end-of-fiscal-year-stock-price (prcc\_f)*annual change in common shares outstanding (csho).	clear	Original
Lee and Swaminathan	2000	Momentum and Volume	MomVol	1965	1995	Mom6m. Include only stocks in the highest quintile of average trading volume (vol) over the previous 6 months. Exclude NASDAQ stocks, if price less than 1 or if stock has been trading for less than 24 months.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Lev and Nissim	2004	Taxable income to income	Tax	1973	2000	Ratio of Taxes paid and tax share of net income. Numerator is defined as the sum of foreign (txfo) and federal (txfed) income taxes. If either one is missing, numerator is defined as total taxes (txt) minus deferred taxes (txdi). Denominator is the product of the prevailing tax rate and net income (ib). Tax rate is .48 before 1979, .46 from 1979 to 1986, .4 in 1987, .34 between 1988 and 1992 and .35 from 1993 onwards. If net income is negative, and the numerator is positive, tax is defined as 1. Exclude if price less than 5.	clear	Original
Lev and Nissim	2004	Taxable income to income quarterly	Tax_q	1973	2000	Ratio of Taxes paid and tax share of net income. Numerator is defined as the sum of foreign (txfo) and federal (txfed) income taxes. If either one is missing, numerator is defined as total taxes (txt) minus deferred taxes (txdi). Denominator is the product of the prevailing tax rate and net income (ib). Tax rate is .48 before 1979, .46 from 1979 to 1986, .4 in 1987, .34 between 1988 and 1992 and .35 from 1993 onwards. If net income is negative, and the numerator is positive, tax is defined as 1. Exclude if price less than 5.	clear	Quarterly
Li	2011	R&D capital-to-assets	RDcap	1980	2007	R&D capital to assets is the weighted sum of lagged R&D expenditures scaled by the assets (at). We replace xrd with 0 if xrd is missing and compute the numerator as $\$xrd_t + .8 xrd_{t-1} + .6xrd_{t-2} + .4 xrd_{t-3} + .2 xrd_{t-4}$ . Replace RDcap with missing before 1980 or if firm is in upper two thirds of market cap ( $shrout * abs(prc)$ ) distribution in a month.	clear	Original
Liu	2006	Days with zero trades	zerotrade	1960	2003	In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover ( $vol/shrout$ ) divided by $48 * 10^5$ ), multiplied by 21/number of trading days per month. Zerotrade is the 6-month average of that variable.	clear	Original
Liu	2006	Days with zero trades	zerotradeAlt1	1960	2003	In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover ( $vol/shrout$ ) divided by $48 * 10^5$ ), multiplied by 21/number of trading days per month. Zerotrade is the 6-month average of that variable.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Liu	2006	Days with zero trades	zerotradeAlt12	1960	2003	In each month, count the number of days with no trades. Define zerotrade as the number of days without trades plus (the sum of monthly turnover (vol/shrout) divided by $48 \times 10^5$ ), multiplied by 21/number of trading days per month. Zerotrade is the 6-month average of that variable.	clear	Original
Lockwood and Prombutr	2010	Sustainable Growth	ChEQ	1964	2007	Ratio of book equity (ceq) to book equity in the previous year. Include only if book equity is positive this year and last year.	clear	Original
Loh and Warachka	2012	Earnings streak indicator	EarnIncrease	1987	2009	Binary variable equal to 1 if the change in quarterly net income (ibq) from t-1 to 1 was positive in quarters t, t-1, t-2, t-3 and t-4, and 0 otherwise.	clear	Original
Loh and Warachka	2012	Number of consecutive earnings increases	NumEarnIncrease	1987	2009	Number of 4-quarter net income (ibq) increases over the previous 2 years.	clear	Original
Lou	2014	Growth in advertising expenses	GrAdExp	1974	2010	Log of advertising expense (xad) minus log of advertising expense last year. Exclude if price less than 5, xad less than .1 or stock in the lowest decile of market value of equity.	clear	Original
Loughran and Wellman	2011	Enterprise Multiple	EntMult	1963	2009	Market value of equity + long-term debt (dltt) + debt in current liabilities (dlc) + deferred charges (dc) - cash and short-term investments (che) , divided by operating income (oibdp). Exclude if missing book equity or negative operating income.	clear	Original
Loughran and Wellman	2011	Enterprise Multiple	EntMult_q	1963	2009	Market value of equity + long-term debt (dltt) + debt in current liabilities (dlc) + deferred charges (dc) - cash and short-term investments (che) , divided by operating income (oibdp). Exclude if missing book equity or negative operating income.	clear	Quarterly
Lyandres, Sun and Zhang	2008	Composite debt issuance	CompositeDebtIssuance	1970	2005	Log of long-term debt (dltt) plus debt in current liabilities (dlc) minus log of the same variable 5 years ago.	clear	Original
Lyandres, Sun and Zhang	2008	change in ppe and inv/assets	InvestPPEInv	1970	2005	One-year change in property, plants and equipment (ppeg) plus one year change in inventory (inv), scaled by one-year lagged assets (at).	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Menzly and Ozbas	2010	Customers momentum	iomom_cust	1986	2005	We download BEA data as follows: Pre-1996: <a href="https://www.bea.gov/industry/input-output-accounts-data">https://www.bea.gov/industry/input-output-accounts-data</a> , Click: "Historical Make-Use Tables". 1997-present: download zip file: <a href="https://apps.bea.gov//industry/iTables%20Static%20Files/AllTablesSUP.zip">https://apps.bea.gov//industry/iTables%20Static%20Files/AllTablesSUP.zip</a> . But keep only Supply_1997-2018_SUM.xlsx (a.k.a. Make) and Use_SUT_Framework_1997-2018_SUM.xlsx (a.k.a. Use). Assume these data are available 5 years after the survey year in the spreadsheet. Match Compustat firm-years to the BEA data by NAICS code. Compute returns within each BEA industry (we're using the roughly 70 industry version). Then match each industry-year to "matched-industry weights" corresponding to IO tables as follows: for supplier momentum, industry comes from cols of "Use" table, matched industries from the rows. For customer momentum: industry comes from rows of make (a.k.a. supply) table, matched industries come from cols. Weights exclude own-industry entries. For each industry-month, compute the weighted average of returns of matched industries using previously mentioned weights. Then we sort industries into deciles based on matched returns. Then we assign each firm-month to an industry decile. Finally, we long stocks in industry deciles 9 or 10 and short stocks in industry deciles 1 or 2. Drop pre-1986 due to NAICS availability.	clear	Original

Continued on next page



**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Menzly and Ozbas	2010	Suppliers momentum	iomom_supp	1986	2005	We download BEA data as follows: Pre-1996: <a href="https://www.bea.gov/industry/input-output-accounts-data">https://www.bea.gov/industry/input-output-accounts-data</a> , Click: "Historical Make-Use Tables". 1997-present: download zip file: <a href="https://apps.bea.gov//industry/iTables%20Static%20Files/AllTablesSUP.zip">https://apps.bea.gov//industry/iTables%20Static%20Files/AllTablesSUP.zip</a> . But keep only Supply_1997-2018_SUM.xlsx (a.k.a. Make) and Use_SUT_Framework_1997-2018_SUM.xlsx (a.k.a. Use). Assume these data are available 5 years after the survey year in the spreadsheet. Match Compustat firm-years to the BEA data by NAICS code. Compute returns within each BEA industry (we're using the roughly 70 industry version). Then match each industry-year to "matched-industry weights" corresponding to IO tables as follows: for supplier momentum, industry comes from cols of "Use" table, matched industries from the rows. For customer momentum: industry comes from rows of make (a.k.a. supply) table, matched industries come from cols. Weights exclude own-industry entries. For each industry-month, compute the weighted average of returns of matched industries using previously mentioned weights. Then we sort industries into deciles based on matched returns. Then we assign each firm-month to an industry decile. Finally, we long stocks in industry deciles 9 or 10 and short stocks in industry deciles 1 or 2. Drop pre-1986 due to NAICS availability.	clear	Original
Michaely, Thaler and Womack	1995	Dividend Omission	DivOmit	1964	1988	Define dividend omission as not having paid a dividend in the current month or the two preceding months, but having paid dividends in the 3, 6, 9, 12, 15, 18 months before. DivOmit is equal to 1 if a dividend was omitted in the previous 12 months and 0 otherwise.	clear	Original
Michaely, Thaler and Womack	1995	Dividend Initiation	DivInit	1964	1988	Define dividend initiation as having paid a dividend in month $t$ ( $divamt > 0$ ) and not having paid a dividend in the 24 preceding months. DivInit is equal to 1 if a dividend was initiated in the past 12 months and 0 otherwise. Exclude if share code greater 11 and use NYSE stocks only.	likely	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Mohanram	2005	Mohanram G-score	MS	1978	2001	Examine only stocks in lowest BM quintile. Binary variable based on sum of eight indicator variables which are: 1 if return on assets (ni/average assets) above the two digit industry median; 1 if net cash flow to assets (oancf/average assets) above the two digit industry median; 1 if net cash flow greater than net income; 1 if R&D expense to assets (xrd/average assets) greater than two digit industry median; 1 if capital expenditure (capx/average assets) greater than two digit industry median; 1 if advertising expenses (xad/average assets) greater than two digit industry median; 1 if the volatility of net income over the past 3 years is below the two digit industry median, 1 if the volatility of revenue (revt) over the past 3 years is below the two digit industry median. The final variable is equal to 1 if the sum of the above 8 indicators is greater than 5 and 0 if the sum is less than 2.	clear	Original
Nagel	2005	Inst Own and BM	RIO_BM	1980	2003	Residual institutional ownership (RIO) is defined as $\log(\text{institutional ownership} / (1 - \text{institutional ownership})) + 23.6 - 2.89 * \log(\text{market value of equity}) + .08 * \log(\text{market value of equity})^2$ . Replace instown\_perc with 0 if it is missing, with .9999 if it's above .9999, and with .0001 if it's below .0001. RIO\_BM is a binary variable equal to 1 if a firm is in the highest quintile of the monthly RIO distribution and has BM below the cross-sectional median, and 0 if a firm is in the lowest quintile of RIO and has BM below the median.	clear	Original
Nagel	2005	Inst Own and Forecast Dispersion	RIO_Dis	1980	2003	Binary variable equal to 1 if RIO (defined above) is in the highest quintile and ForecastDispersion (defined above) is above the median, 0 if RIO is in the lowest quintile and ForecastDispersion is above the median.	clear	Original
Nagel	2005	Inst Own and Idio Vol	RIO_IdioRisk	1980	2003	Binary variable equal to 1 if RIO (defined above) is in the highest quintile and monthly IdioRisk (defined above) is above the median, 0 if RIO is in the lowest quintile and IdioRisk is above the median.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Nagel	2005	Inst Own and Turnover	RIO_Turnover	1980	2003	Binary variable equal to 1 if RIO (defined above) is in the highest quintile and monthly turnover (vol/shrout) is above the median, 0 if RIO is in the lowest quintile and turnover is above the median.	clear	Original
Naranjo, Nimalendran and Ryngaert	1998	Dividend Yield	DivYield	1963	1994	Define tempDY as 4 times latest dividend (divamt) divided by price (prc). Define positive yield stocks as those which paid a dividend in all of the past 4 quarters. Set DivYield = 1 if tempDY is above the median among positive yield stocks, and set DivYield = 0 if a stock in the top quartile of among positive tempDY stocks for the month, and DivYield = 0 if tempDY = 0. Set DivYield to missing if stock is above the median firm size. This procedure is based on Table 1B.	likely	Original
Naranjo, Nimalendran and Ryngaert	1998	Dividend Yield	DivYield_q	1963	1994	Define tempDY as 4 times latest dividend (divamt) divided by price (prc). Define positive yield stocks as those which paid a dividend in all of the past 4 quarters. Set DivYield = 1 if tempDY is above the median among positive yield stocks, and set DivYield = 0 if a stock in the top quartile of among positive tempDY stocks for the month, and DivYield = 0 if tempDY = 0. Set DivYield to missing if stock is above the median firm size.	clear	Quarterly
Nguyen and Swanson	2009	Efficient frontier index	Frontier	1980	2003	Frontier is the residual of a regression of log(BM) on log(book equity (ceq)), long-term debt (dltt) to assets (at), capital expenditures (capx) to revenue (sale), R&D expense (xrd) to revenue, advertising expense (xad) to revenue, property plant and equipment (ppent) to assets, EBIT (ebitda) to assets, and dummies for Fama-French's 48 industry definitions. Regression is updated each month with a rolling window of 60 months.	clear	Original
Novy-Marx	2010	Operating Leverage	OPLeverage	1963	2008	Sum of administrative expenses (xsga) and cost of goods sold (cogs), scaled by total assets (at). Use xsga = 0 if xsga is missing.	clear	Original
Novy-Marx	2010	Operating Leverage	OPLeverage_q	1963	2008	Sum of administrative expenses (xsga) and cost of goods sold (cogs), scaled by total assets (at). Use xsga = 0 if xsga is missing.	clear	Quarterly
Novy-Marx	2012	Intermediate Momentum	IntMom	1927	2010	Stock return between months t-12 and t-6	clear	Original
Novy-Marx	2013	gross profits / total assets	GP	1963	2010	Revenue (sale) - cost of goods solds (cogs), divided by total assets (at).	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Novy-Marx	2013	gross profits / total assets	GPlag	1963	2010	Revenue (sale) - cost of goods solds (cogs), divided by 12 months lagged total assets.	clear	Lag structure
Novy-Marx	2013	gross profits / total assets	GPlag_q	1963	2010	Revenue (revtq) - cost of goods solds (cogsq), divided by one quarter lagged total assets (atq).	clear	Quarterly
Ortiz-Molina Phillips	and 2014	Asset liquidity scaled by book assets	AssetLiquidityBook	1984	2006	(assets (at) + 0.75*(at - short term investments(che)) + 0.5*(at- curren total assets (act) - goodwill(gdwl) - intangibles (itan)) scaled by the one-month lagged book assets (at).	maybe	Original
Ortiz-Molina Phillips	and 2014	Asset liquidity scaled by market value of assets	AssetLiquidityMarket	1984	2006	(assets (at) + 0.75*(at - short term investments(che)) + 0.5*(at- curren total assets (act) - goodwill(gdwl) - intangibles (itan)) scaled by the one-month lagged market assets (at + end-of-fiscal-year-stock-price (prcc_f) * common shares outstanding (csho) - book equity (ceq)).	maybe	Original
Ortiz-Molina Phillips	and 2014	Quarterly asset liquidity (scaled by book value of assets)	AssetLiquidityBookQuart	1984	2006	(short term investments (cheq) + 0.75*(actq - short term investments(cheq)) + 0.5*(atq- curren total assets (actq) - goodwill(gdwlq) - intangibles (itanq)) scaled by the one-month lagged book assets (atq). Replace goodwill and intangibles with 0 if they are missing.	maybe	Quarterly
Ortiz-Molina Phillips	and 2014	Quarterly asset liquidity (scaled by market value of assets)	AssetLiquidityMarketQuart	1984	2006	(short term investments (cheq) + 0.75*(actq - short term investments(cheq)) + 0.5*(atq- curren total assets (actq) - goodwill(gdwlq) - intangibles (itanq)) scaled by the one-month lagged book assets (atq) + lagged closing price (prccq) times lagged common shares outstanding (cshoq) - lagged book equity (ceqq). Replace goodwill and intangibles with 0 if they are missing.	maybe	Quarterly
Ou and Penman	1989	CF to debt	cashdebt	1973	1983	Sum of income before extraordinary items (ib) and depreciation and amortization (dp) divided by total liabilities (lt) averaged over the current and previous fiscal years.	maybe	Original
Ou and Penman	1989	Current Ratio	currat	1973	1983	Current total assets (act) divided by current total liabilities (lct). If missing, replace current total assets by cash and short-term investments (che) plus total receivables (rect) plus total inventory (invt), and replace current total liabilities (lct) with accounts payable (ap).	maybe	Original
Ou and Penman	1989	Change in Current Ratio	pchcurrat	1973	1983	One year growth of the current ratio (currat) defined as current total assets (act) divided by current total liabilities (lct).	maybe	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Ou and Penman	1989	Change in quick ratio	pchquick	1973	1983	One year growth in quick ratio, defined as change between total current assets (act) and inventory (invt) scaled by the total current liabilities (lct).	maybe	Original
Ou and Penman	1989	Change in sales to inventory	pchsaleinv	1973	1983	One year growth in sales (sale) to inventory (invt) ratio.	maybe	Original
Ou and Penman	1989	Quick ratio	quick	1973	1983	Difference between current total assets (act) and total inventory (invt) all divided by the current total liabilities (lct).	maybe	Original
Ou and Penman	1989	Sales to cash ratio	salecash	1973	1983	Sales (sale) divided by cash and short-term investments (che).	maybe	Original
Ou and Penman	1989	Sales to inventory	saleinv	1973	1983	Sales (sale) divided by total inventory (invt).	maybe	Original
Ou and Penman	1989	Sales to receivables	salerec	1973	1983	Sales (sale) divided by total receivables (rect).	maybe	Original
Palazzo	2012	Cash to assets	Cash	1972	2009	Ratio of quarterly cash and short-term investments (cheq) and total assets (atq).	clear	Original
Pastor and Stambaugh	2003	Pastor-Stambaugh liquidity beta	BetaLiquidityPS	1968	1999	Monthly excess return (ret -rf) regressed on innovations in liquidity from Pastor's website ( <a href="https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt">\url{https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt}</a> ). Use 60 month rolling window regression, and require at least 36 non-missing observations.	clear	Original
Penman, and Tuna	Richardson 2007	Leverage component of BM	BPEBM	1963	2001	BP - EBM, where BP = $(ceq + tstkp - clear\ dvpa)/(shrout*abs(prc))$ , and EBM is defined above. Exclude if price less than 5.	clear	Original
Penman, and Tuna	Richardson 2007	Enterprise component of BM	EBM	1963	2001	$(ceq + che - dltt - dlc - dc - dvpa + tstkp) / (mve\_c + che - dltt - dlc - dc - dvpa + tstkp)$ . Exclude if price less than 5.	clear	Original
Penman, and Tuna	Richardson 2007	Net debt to price	NetDebtPrice	1963	2001	Long-term debt (dltt) plus debt in current liabilities (dlc) plus preferred stock (pstk) plus preferred dividends in arrears (dvpa) minus treasury stock (tstkp) minus cash and short-term investments (che), scaled by market value of equity. Exclude if SIC between 6000 and 6999, or if missing value for total assets (at), net income (ib), common shares outstanding (csho), book value of equity (ceq) or price close fiscal year (prcc\_f). Keep only 3rd B/M Quintile, following Table 4 (and in contrast to Table 1).	clear	Original
Penman, and Tuna	Richardson 2007	Enterprise component of BM	EBM_q	1963	2001	$(ceq + che - dltt - dlc - dc - dvpa + tstkp) / (mve\_c + che - dltt - dlc - dc - dvpa + tstkp)$ . Exclude if price less than 5.	clear	Quarterly

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Penman, Richardson and Tuna	2007	Net debt to price	NetDebtPrice_q	1963	2001	Long-term debt (dltt) plus debt in current liabilities (dlc) plus preferred stock (pstk) plus preferred dividends in arrears (dvpa) minus treasury stock (tstkp) minus cash and short-term investments (che), scaled by market value of equity. Exclude if SIC between 6000 and 6999, or if missing value for total assets (at), net income (ib), common shares outstanding (csho), book value of equity (ceq) or price close fiscal year (prcc\_f). Keep only 3rd B/M Quintile, following Table 4 (and in contrast to Table 1).	clear	Quarterly
Piotroski	2000	Piotroski F-score	PS	1976	1996	Sum of nine indicator variables which are: 1 if net income (ib) greater 0; 1 if net cash flow (oanctf) greater 0; 1 if return on assets (ib/at) increased relative to previous year; 1 if net cash flow greater net income; 1 if long-term debt to assets (dltt/at) declined over the previous year; if current assets to current liabilities (act/lct) increased over the previous year; 1 if ebit/sale (ebit = ib + txt + xint) increased over the previous year; 1 if revenue to assets increased over the previous year; 1 if shroust $\leq$ shroust last year. Include highest quintile of book-to-market only. Exclude if missing any of the input variables.	clear	Original
Piotroski	2000	Piotroski F-score	PS_q	1976	1996	Sum of nine indicator variables which are: 1 if net income (ib) greater 0; 1 if net cash flow (oanctf) greater 0; 1 if return on assets (ib/at) increased relative to previous year; 1 if net cash flow greater net income; 1 if long-term debt to assets (dltt/at) declined over the previous year; if current assets to current liabilities (act/lct) increased over the previous year; 1 if ebit/sale (ebit = ib + txt + xint) increased over the previous year; 1 if revenue to assets increased over the previous year; 1 if shroust $\leq$ shroust last year. Include highest quintile of book-to-market only. Exclude if missing any of the input variables.	clear	Quarterly
Pontiff and Woodgate	2008	Share issuance (1 year)	ShareIss1Y	1970	2003	Growth in number of shares between t-18 and t-6. Number of shares is calculated as shroust/cfacshr to adjust for splits.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Prakash and Sinha	2012	Deferred Revenue	DelDRC	2002	2007	Annual change in deferred revenue (drc) scaled by average total assets (at) in t-1 and t. Exclude if negative book equity (ceq), deferred revenue equal to 0 in both years, revenue less than 5m, or SIC code between 6000 and 6999.	likely	Original
Rajgopal, Shevlin and Venkatachalam	2003	Order backlog	OrderBacklog	1981	1999	Order backlog (ob) divided by average total assets (at) in years t-1 and t. Exclude if order backlog is 0.	clear	Original
Richardson et al.	2005	Change in current operating assets	DelCOA	1962	2001	Difference in current operating assets (total current assets (act) minus cash and short-term investments (che)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t.	clear	Original
Richardson et al.	2005	Change in current operating liabilities	DelCOL	1962	2001	Difference in current operating liabilities (total current liabilities (lct) minus debt in current liabilities (dlc)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t.	clear	Original
Richardson et al.	2005	Change in equity to assets	DelEqu	1963	2001	Difference in book equity (ceq) between years t-1 and t, scaled by average total assets (at) in years t-1 and t.	clear	Original
Richardson et al.	2005	Change in financial liabilities	DelFINL	1962	2001	Difference in financial liabilities (sum of long-term debt (dltt), current liabilities (dlc) and preferred stock (pstk)) between years t-1 and t, scaled by average total assets (at) in years t-1 and t.	clear	Original
Richardson et al.	2005	Change in long-term investment	DelLTI	1962	2001	Difference in investment and advances (ivao) between years t-1 and t, scaled by average total assets (at) in years t-1 and t.	clear	Original
Richardson et al.	2005	Change in net financial assets	DelNetFin	1962	2001	Compute the sum of short-term investments (ivst) and investments and advances (ivao) minus the sum of long-term debt (dltt), debt in current liabilities (dlc) and preferred stock capital. Divide the difference between the current and one-year lagged sum by total assets (at) averaged over the current and previous fiscal years.	clear	Original
Richardson et al.	2005	Total accruals	TotalAccruals	1962	2001	Before 1988: Change in net working capital ((act - che) - (lct - dlc)) plus change in net noncurrent assets ( (at - act - ivao) - (lt - dlc - dltt)) plus change in net financial assets ( (ivst + ivao - (dltt + dlc + pstk)). Starting in 1988: net income (ni) minus total, operating and investment cashflows (oanct, ivnct, finct) plus stock sales minus repurchases and dividends (sstk, prstkc, dv)). Scaled by lagged total assets (at). Replace missings in ivao, ivst, dltt, dlc, pstk sstk, prstkc, dv with 0.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Richardson et al.	2005	Change in short-term investment	DelSTI	1962	2001	Difference of the current and one-year lagged short-term investments (ivst) scaled by the mean of current and one-year lagged total assets (at).	not clear	Original
Ritter	1991	IPO and age	AgeIPO	1981	1984	Age is (current year - founding year from Jay Ritter's dataset). Exclude if IndIPO == 0. AgeIPO = 1 if Age is	clear	Original
Ritter	1991	Initial Public Offerings	IndIPO	1975	1987	1 if IPO in the past 6-36 months. otherwise. IPO dates are taken from Jay Ritter's IPO data available at: <a href="http://bear.warrington.ufl.edu/ritter/ipodata.htm">http://bear.warrington.ufl.edu/ritter/ipodata.htm</a> . Missing IPO dates imply IndIPO = 0	clear	Original
Rosenberg, Lanstein, Reid, and	1985	Book to market using most recent ME	BM	1973	1984	Log of annual book equity (ceq) over market equity (see above).	clear	Original
Rosenberg, Lanstein, Reid, and	1985	Book to market using December ME	BMdec	1973	1984	BM using December value of market equity.	clear	Original
Rosenberg, Lanstein, Reid, and	1985	Book to market (quarterly)	BMq	1973	1984	Log of annual book equity (ceq) over market equity (see above).	clear	Quarterly
Scherbina	2008	Decline in Analyst Coverage	ChNAnalyst	1982	2005	Binary variable equal to 1 if the number of analysts (numest) for next quarter's EPS estimate decreased relative to three months ago, and 0 if it increased.	clear	Original
Sloan	1996	Accruals	Accruals	1962	1991	Annual change in current total assets (act) minus annual change in cash and short-term investments (che) minus annual change in current liabilities (lct) minus annual change in debt in current liabilities (dlc) minus change in income taxes (txp). All divided by average total assets (at) over this year and last year. Exclude if abs(prc) < 5.	clear	Original
Soliman	2008	Change in Net Noncurrent Operating Assets	ChNNCOA	1984	2002	Twelve-month change in noncurrent operating assets. Noncurrent operating assets is ( (at - act - iva0) - (lt - dlc - dltt) )/at.	clear	Original
Soliman	2008	Change in Net Working Capital	ChNWC	1984	2002	Twelve-month change in net working capital. Net working capital is ( (act - che) - (lct - dlc) )/at	clear	Original
Soliman	2008	Change in Asset Turnover	ChAssetTurnover	1984	2002	Annual change in AssetTurnover (defined above). Exclude if price less than 5.	likely	Original

Continued on next page



**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Soliman	2008	Asset Turnover	AssetTurnover	1984	2002	Sales (sale) divided by two year average of net operating assets. Net operating assets is the sum of receivables (rect), inventories (invt), current assets other (aco), net property, plants and equipment (ppent) and intangibles (intan), minus accounts payable (ap), other current liabilities (lco) and other liabilities (lo). Exclude if $abs(prc) < 5$ or $AssetTurnover < 0$ .	maybe	Original
Soliman	2008	Change in Noncurrent Operating Assets	ChNCOA	1984	2002	Twelve-month change in noncurrent operating assets. Noncurrent operating assets is $(at - act - ivao) - (lt - dlc - dltt) / at$ .	maybe	Original
Soliman	2008	Change in Noncurrent Operating Liabilities	ChNCOL	1984	2002	Twelve-month change in noncurrent operating liabilities. Noncurrent operating liabilities is $(lt - dlc - dltt)$ .	maybe	Original
Soliman	2008	Change in Profit Margin	ChPM	1984	2002	Annual change in profit margin PM (profit margin defined below). Exclude if price less than 5.	maybe	Original
Soliman	2008	Profit Margin	PM	1984	2002	Net income (ni) over revenue (revt). Exclude if price less than 5.	maybe	Original
Soliman	2008	Return on Net Operating Assets	RetNOA	1984	2002	Net income (ni) over revenue (revt). Exclude if price less than 5.	maybe	Original
Soliman	2008	Profit Margin	PM_q	1984	2002	Net income (ni) over revenue (revt). Exclude if price less than 5.	clear	Quarterly
Soliman	2008	Asset Turnover	AssetTurnover_q	1984	2002	Sales (sale) divided by two year average of net operating assets. Net operating assets is the sum of receivables (rect), inventories (invt), current assets other (aco), net property, plants and equipment (ppent) and intangibles (intan), minus accounts payable (ap), other current liabilities (lco) and other liabilities (lo). Exclude if $abs(prc) < 5$ or $AssetTurnover < 0$ .	maybe	Quarterly
Soliman	2008	Return on Net Operating Assets	RetNOA_q	1984	2002	Net income (ni) over revenue (revt). Exclude if price less than 5.	not	Quarterly
Spieß and Affleck-Graves	1999	Debt Issuance	DebtIssuance	1975	1989	Equal to 1 if debt issuance (dltis) greater 0 and 0 otherwise. Exclude if share code > 11 or missing book-to-market.	likely	Original
Thomas and Zhang	2002	Inventory Growth	ChInv	1970	1997	12 month change in inventory (invt) divided by average total assets.	clear	Original
Thomas and Zhang	2011	Change in Taxes	ChTax	1977	2006	4-quarter change in quarterly total taxes (txtq), scaled by lagged total assets (at).	clear	Original
Titman, Wei and Xie	2004	Investment to revenue	Investment	1973	1996	Ratio of capital investment (capx) to revenue (revt) divided by the firm-specific 36-month rolling mean of that ratio. Exclude if revenue less than \$10m.	clear	Original

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Tuzel	2010	Real estate holdings	realestate	1971	2005	Industry-adjusted value of real estate holdings. Real estate holdings are calculated as: PPE Buildings at cost (fatb) plus PPE Leases at cost (fatl), divided by PPE (ppeg). Use ppeg if ppeg is missing. Subtract monthly industry-mean at the 2 digit SIC level.	clear	Original
Valta	2016	Convertible debt indicator	ConvDebt	1985	2012	Binary variable equal to 1 if deferred charges (dc) greater than 0 or common shares reserved for convertible debt (cshrc) greater than 0.	clear	Original
Valta	2016	Secured debt	secured	1985	2012	Debt/mortgages and other secured (dm) divided by long-term liabilities (dltt) plus current liabilities (dlc). Replace with 0 if missing.	likely	Original
Valta	2016	Secured debt indicator	securedind	1985	2012	Binary version of secured. 1 if secured greater than 0, 0 otherwise.	likely	Original
Whited and Wu	2006	Whited-Wu index	WW	1975	2001	Group data by 3 digit SIC code and month to compute total sales (sale) by industry. Calculate the Whited-Wu index as $-0.091 * (\text{income before extraordinary items (ib)} + \text{depreciation and amortization (dp)}) / (4 * \text{assets(at)}) - 0.062$ (if the dividends per share (dvpsx_c) is not missing and greater than 0) $+ 0.021 * \text{long-term debt (dltt)} / \text{at} - 0.044 * \log(\text{at}) + 0.102 * (\text{twelve-month growth of Industry Sales}) / 4 - 0.035 * (\text{one-month growth of sales}) / 4$ .	not	Original
Whited and Wu	2006	Whited-Wu index	WW_Q	1975	2001	Group data by 3 digit SIC code and month to compute total sales (saleq) by industry. Calculate the Whited-Wu index as $-0.091 * (\text{quarterly income before extraordinary items (ibq)} + \text{quarterly depreciation and amortization (dpq)}) / (\text{quarterly assets(atq)}) - 0.062$ (if the quarterly dividends per share (dvpsxq) is not missing and greater than 0) $+ 0.021 * \text{quarterly long-term debt (dlttq)} / \text{atq} - 0.044 * \log(\text{atq}) + 0.102 * (\text{four-quarter growth of quarterly Industry Sales}) - 0.035 * (\text{one-quarter growth of quarterly sales (saleq)})$ .	not	Quarterly

Continued on next page

**Table 1** (continued)

Author(s)	Pub year	Description	Acronym	Sample Start	Sample End	Construction	Category	Variant
Xie	2001	Abnormal Accruals	AbnormalAccruals	1971	1992	Define Accruals as net income (ib) minus operating cash flow (oancf), divided by average total assets (at) for years t-1 and t. If oancf is missing, replace operating cash flow with funds from operations (fopt) minus the annual change in total current assets (act) plus the annual change in cash and short-term investments (che) plus the annual change in current liabilities (lct) minus the annual change in debt in current liabilities (dlc). For each year t, regress Accruals on: the inverse of average total assets for years t-1 and t, the change in revenue (sale) from year t-1 to t divided by average total assets, property plant and equipment (ppeg) divided by average total assets, industry dummies for Fama-French's 48 industry classification. AbnormalAccrual is the residual from this cross-sectional regression.	clear	Original
Xing, Zhang and Zhao	2010	Volatility smirk near the money	skew1	1996	2005	Using OptionMetrics data, among options with duration between 10 and 60 days, implied volatility of put option with moneyness closest to but above 1 minus implied volatility of call option with moneyness closest to but below 1.	likely	Original
Yan	2011	Put volatility minus call volatility	SmileSlope	1996	2005	Using OptionMetrics data, average implied volatility of put options with duration between 15 and 30 days and rounded delta of -.5 minus average implied volatility of call options with duration between 15 and 30 days and rounded delta of .5.	clear	Original
Zhang	2004	Firm Age - Momentum	FirmAgeMom	1983	2001	6 month return, restricted to the bottom quintile of the cross-sectional firm age distribution. Exclude if price less than 5 or firm younger than 12 months.	clear	Original