**Variables Definition**

96 Variables: 86 Firm Characteristics, 4 ownership, 4 Chinese

1. *acc*: We follow the definition of accruals in Sloan (1996) to construct *acc*, i.e.,

*acc* = [(∆CA − ∆CASH) − (∆CL − ∆STD − ∆TP) − Dep]*/*Total Assets*,*

where ∆ represents the difference between two consecutive periods, CA, CASH, CL, STD, TP, Dep, denote current assets, cash/cash equivalents, current liabilities, debt included in current liabilities, income tax payable, depreciation and amortization expense, respectively. These data are acquired from CSMAR.

1. *absacc*: Absolute value of *acc*.
2. *agr*: Annual percentage change in total assets. Data of total assets are acquired from CSMAR.
3. *beta*: We estimate stock-level beta using weekly returns and value-weighted market returns for three years ending month *t* − 1 with at least 52 weeks of returns. Stock returns are acquired from the WIND database.
4. *betasq*: Stock-level market beta squared.
5. *bm*: Book-to-market ratio, which equals the book value of equity divided by market capitalization. Data are acquired from CSMAR.
6. *bm ia*: This is the industry-adjusted book-to-market ratio introduced in Asness et al. (2000),

*bm iait* = *bmit* − *bmIit,*

where *bmIit* is the equally-weighted average book-to-market ratio of firms in firm *i*’s industry. As firms’ industries are reported annually in CSMAR, we let firm *i*’s current industry to be the one reported in the year prior to the current month.

1. *cash*: Cash and cash equivalents divided by average total assets. Related data are reported in quarterly reports and are acquired from CSMAR.
2. *cashdebt*: Earnings divided by total liabilities, which is defined similar to that in Ou and Penman (1989). Data are acquired from CSMAR.
3. *cashspr*: Cash productivity, which is defined as quarter-end market capitalization plus longterm debt minus total assets divided by cash and equivalents. Related data are contained in quarterly reports and acquired from CSMAR.
4. *cfp*: Operating cash flows divided by quarter-end market capitalization. Related data are contained in quarterly reports and acquired from CSMAR.
5. *cfp ia*: This is the industry-adjusted operating cash flows. The way of adjustment is similar to that for *bm ia*. Data are acquired from CSMAR.
6. *chato*: Change in sales divided by average total assets. Quarterly data on sales and total assets are acquired from CSMAR.
7. *chato ia*: Industry-adjusted change in sales divided by average total assets. Data are acquired from CSMAR.
8. *chcsho*: Monthly percentage change in shares outstanding. Monthly data on shares outstanding are acquired from CSMAR.
9. *chempia*: Industry-adjusted change in the number of employees. Related data are available annually on CSMAR and the way of industry adjustment is similar as that for *bm ia*.
10. *chinv*: Change in inventory scaled by total assets. Data are available quarterly on CSMAR.
11. *chmom*: Cumulative returns from months *t* − 6 to *t* − 1 minus months *t* − 12 to *t* − 7. Stock returns are acquired from WIND database.
12. *chpm*: Change in income before extraordinary items scaled by sales. Related data are acquired from CSMAR.
13. *chpm ia*: Industry-adjusted change in income before extraordinary items scaled by sales. Data are acquired from CSMAR.
14. *chtx*: Percentage change in taxes from quarter *t* − 1 to *t*. Data are acquired from CSMAR.
15. *cinvest*: Change over one quarter in fixed assets divided by sales - average of this variable for prior three quarters; if sales are zero, then scale by 0*.*01. Data are acquired from CSMAR.
16. *currat*: The ratio of current assets to current liabilities. Data are acquired from CSMAR.
17. *depr*: Depreciation divided by fixed assets. Data are acquired from CSMAR.
18. *divi*: A dummy variable that equals to 1 if company pays dividends this year but did not in prior year. Data are acquired from CSMAR.
19. *divo*: A dummy variable that equals to 1 if company does not pay dividends this year but did in prior year. Data are acquired from CSMAR.
20. *dolvol*: Natural logarithm of trading volume times price per share from month *t*−2. Data are acquired from CSMAR.
21. *dy*: Total dividends divided by market capitalization at year end. Data are acquired from

CSMAR.

1. *ear*: Sum of daily returns in three days around earnings announcement. Data are acquired from CSMAR.
2. *egr*: Quarterly percentage change in book value of equity. Data are acquired from CSMAR.
3. *gma*: Revenue minus cost of goods sold divided by lagged total assets. Quarterly data are acquired from CSMAR.
4. *grCAPX*: Percentage change in capital expenditures from year *t*−2 to year *t*. Data are acquired from CSMAR.
5. *herf* : Sum of squared percentage sales in industry for each company. Sales data and industry code are acquired from CSMAR.
6. *hire*: Percentage change in number of employees. Related data are acquired from CSMAR.
7. *idiovol*: Standard deviation of residuals of weekly returns on weekly equally-weighted market returns for three years prior to month end. Data are acquired from the WIND database.
8. *ill*: Average of daily (absolute return/RMB volume) in month *t*. Daily data are acquired from

WIND database.

1. *invest*: The sum of annual change in fixed assets and annual change in inventories divided by lagged total assets. Data are acquired from CSMAR.
2. *lev*: Total liabilities divided by quarter-end market capitalization. Quarterly data are acquired

from CSMAR.

1. *lgr*: Quarterly percentage change in total liabilities. Data are acquired from CSMAR.
2. *maxret*: Maximum daily return from returns during month *t* − 1. Daily returns are acqired from WIND database.
3. *mom12m*: 11-month cumulative returns ending one month before month end. Stock returns are acqired from WIND database.
4. *mom1m*: 1-month cumulative return. Stock returns are acqired from WIND database.
5. *mom6m*: 5-month cumulative returns ending one month before month end. Stock returns are acqired from WIND database.
6. *mom36m*: Cumulative returns from months *t* − 36 to *t* − 13. Stock returns are acqired from

WIND database.

1. *ms*: Sum of eight indicator variables for fundamental performance following the corresponding definitions in Mohanram (2005). Data are acquired from CSMAR.
2. *mve*: Natural log of market capitalization at end of month *t* − 1. Related data are acquired from CSMAR.
3. *mve ia*: Industry adjusted natural log of market capitalization at end of month *t* − 1. Related data are acquired from CSMAR.
4. *nincr*: Number of consecutive quarters (up to eight quarters) with an increase in earnings. Earnings data are acquired from CSMAR.
5. *operprof* : Quarterly operating profit divided by lagged common shareholders’ equity. Related data are acquired from CSMAR.
6. *orgcap*: Capitalized management expenses. This characteristic uses expense data acquired from

CSMAR and is constructed according to the definition in Eisfeldt and Papanikolaou (2013). Data are acquired from CSMAR.

1. *pchcapx ia*: Industry adjusted percentage change in capital expenditure. Data are acquired from CSMAR.
2. *pchcurrat*: Percentage change in current ratio (current liabilities divided by current assets).

Data are acquired from CSMAR.

1. *pchdepr*: Percentage change in depreciation. Data are acquired from CSMAR.
2. *pchgm pchsale*: Percentage change in gross margin minus Percentage change in sales. Data are acquired from CSMAR.
3. *pchquick*: Percentage change in quick ratio. Data are acquired from CSMAR.
4. *pchsale pchinvt*: Quarterly percentage change in sales minus quarterly percentage change in inventory. Data are acquired from CSMAR.
5. *pchsale pchrect*: Quarterly percentage change in sales minus quarterly percentage change in receivables. Data are acquired from CSMAR.
6. *pchsale pchxsga*: Quarterly percentage change in sales minus quarterly percentage change in management expenses. Data are acquired from CSMAR.
7. *pchsaleinv*: Quarterly percentage change in sales-to-inventory. Data are acquired from CS-

MAR.

1. *pctacc*: Same as *acc* except that the numerator is divided by the absolute value of net income; if net income = 0 then net income set to 0*.*01 for denominator. Data are acquired from CSMAR.
2. *pricedelay*: The proportion of variation in weekly returns for 36 months ending in month *t* explained by four lags of weekly market returns incremental to contemporaneous market return. Stock returns are acquired from WIND database.
3. *ps*: Sum of nine indicator variables that are defined similarly as in Piotroski (2000). Related data are acquired from CSMAR.
4. *quick*: Quick ratio = (current assets - inventory) / current liabilities. Data are acquired from

CSMAR.

1. *rd*: An indicator variable equal to 1 if R&D expense as a percentage of total assets has an increase greater than 5%. Data are acquired from CSMAR.
2. *rd mve*: R&D expense divided by end-of-quarter market capitalization. Data are acquired from

CSMAR.

1. *rd sale*: R&D expense divided by quarterly sales. Data are acquired from CSMAR.
2. *realestate*: Investment in real estates divided by fixed assets. Data are acquired from CSMAR.
3. *volatility*: Standard deviation of daily returns from month *t* − 1. Stock returns are acquired from WIND.
4. *roaq*: Income before extraordinary items divided by one quarter lagged total assets. Related data are acquired from CSMAR.
5. *roavol*: Standard deviation of 16 quarters of income before extraordinary items divided by average total assets. Data are acquired from CSMAR.
6. *roeq*: Income before extraordinary items divided by lagged common shareholders’ equity. Related data are acquired from CSMAR.
7. *roic*: Quarterly earnings before interest and taxes minus nonoperating income divided by noncash enterprise value. Related data are acquired from CSMAR.
8. *rsup*: Sales from quarter *t* minus sales from quarter *t* − 1 divided by quarter-end market capitalization. Related data are acquired from CSMAR.
9. *salecash*: Quarterly sales divided by cash and cash equivalents. Data are acquired from CS-

MAR.

1. *saleinv*: Quarterly sales divided by total inventory. Data are acquired from CSMAR.
2. *salerev*: Quarterly sales divided by accounts receivable. Data are acquired from CSMAR.
3. *sgr*: Quarterly percentage change in sales. Data are acquired from CSMAR.
4. *sp*: Quarterly sales divided by quarter-end market capitalization. Data are acquired from

CSMAR.

1. *std dolvol*: Monthly standard deviation of daily RMB trading volume. Data are acquired from

CSMAR.

1. *std turn*: Monthly standard deviation of daily share turnover. Data are acquired from CSMAR.
2. *stdacc*: Standard deviation of 16 quarters of accruals from month *t* − 16 to *t* − 1. Data are acquired from CSMAR.
3. *stdcf* : Standard deviation for 16 quarters of net cash flows divided by sales. Data are acquired from CSMAR.
4. *tang*: Cash holdings + 0.715 × receivables + 0.547× inventory + 0.535 × fixed assets*/*total assets. Data are acquired from CSMAR.
5. *tb*: Tax income, defined as current tax expense divided by enterprise income tax rate in China

(25%), divided by total income. Data are acquired from CSMAR.

1. *turn*: Average monthly trading volume for month *t* − 3 to *t* − 1 scaled by number of shares outstanding in month *t*. Related data are acquired from CSMAR.
2. *zerotrade*: Turnover weighted number of zero trading days in month *t* − 1. Related data are acquired from CSMAR.
3. *atr*: The abnormal turnover ratio (*atr*) is constructed following the definition in Pan et al.

(2016). Specifically, for stock *i* in month *t*, we run the following regression using daily data

from month *t* − 7 to *t* − 1,

*K*

DTR*i,t* = *β*1 + *β*2 × DMRT*t* + X*cj* × Dummy Event(*,*

*j*=1

where DTR*i,t* is stock *i*’s daily turnover ratio, DMRT*t* is the market turnover ratio, Dummy Event(*j*), *j* = 1*,...,K* is a sequence of event dummy variables. The aggregated ˆ*i,t* for the entire month *t* is defined to be the abnormal turnover ratio (*atr*). Data are acquired from both WIND database and CSMAR.

1. *er trend*: This trend factor is constructed following the definition in Liu et al. (2020). For the sake of simplicity, we refer readers to the original paper for more details. Data are acquired from both WIND database and CSMAR.
2. *largestholderrate*: Percentage of common shares owned by the largest shareholder. Data are acquired from CSMAR.
3. *top10holderrate*: Percentage of common shares owned by top 10 shareholders. Data are acquired from CSMAR.
4. *soe*: A dummy variables that equals 1 if the firm is state-owned. Data are acquired from

CSMAR.

1. *private*: A dummy variable that equals 1 if the firm is privately-owned. Data are acquired from

CSMAR.

1. *foreign*: A dummy variable that equals 1 if the firm is controlled by foreign investors. Data are acquired from CSMAR.
2. *others*: A dummy variable that equals 1 if the firm is not state-owned, private, or foreign. Data are acquired from CSMAR.