A Data Appendix

The variables are arranged in the alphabetical order according to the abbreviated variable name used in the tables.

Amihud (Amihud illiquidity measure) - the average ratio of absolute return to dollar volume, both from CRSP. The ratio is computed daily and averaged within each firm-year (firms with less than 200 valid return observations in a year and the stock price of less than \$5 at the end of the previous year are excluded).

$\overline{\boldsymbol{\alpha}}_{CAPM}$ **(risk-adjusted return)** - in Table III, refers to the risk-adjusted return following Brennan et al. (1998): $\alpha_t^{CAPM} = Ret_t - \beta_{t-1;\ t-36} \cdot MKT_t$. $\beta_{t-1;\ t-36}$ is the firm-level beta estimated in each month t by regressing the firm's monthly excess returns on the market excess returns in months t-1 to t-36. $\overline{\alpha}_{FF}$ and $\overline{\alpha}_{ICAPM}$ are computed in a similar fashion.

CAR (announcement return) - size and book-to-market adjusted cumulative daily returns between the day prior to the earnings announcement and the day after the earnings announcement. Earnings announcement dates are from COMPUSTAT, daily returns are from CRSP daily files, size and book-to-market adjustment is performed following Daniel et al. (1997).

Cred (credit rating) - Standard and Poor's rating (spdr variable in the Compustat quarterly file). The credit rating is coded as 1=AAA, 2=AA+, 3=AA, ... , 21=C, 22=D.

CVEarn/CVCFO (earnings/cash flows volatility) - coefficient of variation (standard deviation over the average) of quarterly earnings/cash flows measured in the past twelve quarters. Earnings are EPS (epspiq over prccq lagged by one quarter). Cash flows are operating income before depreciation (oibdpq) less the change in current assets (actq) plus the change in current liabilities (lctq) less the change in short-term debt (dlcq) plus

the change in cash (cheq). The cash flows are scaled by average total assets (atq) in the past two years. All variables are from the Compustat quarterly file.

CVTurn (turnover variability) - coefficient of variation (standard deviation over the average) of monthly turnover measured between months t-2 and t-36. Turnover is dollar volume over market cap, both dollar volume and market cap are from CRSP.

CVLiq (variability of liquidity) - coefficient of variation (standard deviation over the average) of monthly liquidity measures. The coefficient of variation is measured between months t-2 and t-36. The liquidity measures include Amihud, EffTick, Gamma, Roll, Spread, Zero. The detailed definitions of the liquidity measures can be found in this Appendix.

Disp (analyst forecast dispersion) - the standard deviation of all outstanding earnings-per-share forecasts for the current fiscal year scaled by the absolute value of the outstanding earnings forecast (zero-mean forecasts and forecasts by only one analyst excluded). Earnings forecasts are from the IBES Summary file.

EffTick (effective tick size) - measure of effective spread from Holden (2009). On the simple $\$\frac{1}{8}$ grid, frequency of odd $\frac{1}{8}$s prices (prices that end with $\frac{1}{8}$, $\frac{3}{8}$, $\frac{5}{8}$, or $\frac{7}{8}$) measures the probability of the bid-ask spread being equal to $\$\frac{1}{8}$, the frequency of odd $\frac{1}{4}$s prices measures the probability of the bid-ask spread being equal to $\$\frac{1}{4}$, the frequency of the prices that end $\frac{1}{2}$ measures the probability of the bid-ask spread being $\$\frac{1}{2}$, and the frequency of whole-dollar prices measures the probability of the spread being \$1. For each firm-month, I estimate the probabilities of the spread as above and compute its expected value by multiplying the probabilities by the respective spread values. I use the $\$\frac{1}{16}$ grid before 2001 (decimalization) and the grid with clustering on dollars, half-dollars, quarters, dimes, nickels, and cents from 2001 on.

Error (analyst forecast error) - the absolute value of the difference between the

one-year-ahead consensus forecast and actual earnings divided by actual earnings. All variables are from the IBES Summary file.

FVIX (aggregate volatility risk factor - factor-mimicking portfolio that tracks the daily changes in the VIX index. Following Ang, Hodrick, Xing, and Zhang (2006), I regress the daily changes in VIX on the daily excess returns to the five portfolios sorted on past sensitivity to VIX changes:

$$\Delta VIX_t = \gamma_0 + \gamma_1 \cdot (VIX1_t - RF_t) + \gamma_2 \cdot (VIX2_t - RF_t) +$$
$$+\gamma_3 \cdot (VIX3_t - RF_t) + \gamma_4 \cdot (VIX4_t - RF_t) + \gamma_5 \cdot (VIX5_t - RF_t), \qquad \text{(A-1)}$$

where $VIX1_t, \ldots, VIX5_t$ are the VIX sensitivity quintiles described below, with $VIX1_t$ being the quintile with the most negative sensitivity. The VIX sensitivity quintiles in month t are formed using information from month t-1 and are rebalanced monthly.

The fitted part of the regression above less the constant is my aggregate volatility risk factor (FVIX factor):

$$\begin{aligned} FVIX_t = & -0.043 \cdot (VIX1_t - RF_t) - 0.667 \cdot (VIX2_t - RF_t) \\ & - 0.273 \cdot (VIX3_t - RF_t) - 0.742 \cdot (VIX4_t - RF_t) + 0.197 \cdot (VIX5_t - RF_t) \end{aligned} \qquad \text{(A-2)}$$

The daily returns to FVIX are then cumulated within each month to get the monthly return to FVIX.

The return sensitivity to VIX changes ($\beta_{\Delta VIX}$) I use to form the base assets is measured separately for each firm-month by regressing daily stock excess returns in the past month on daily market excess returns and the VIX index change using daily data (at least 15 non-missing returns are required):

$$Ret_t - RF_t = \alpha + \beta_{MKT} \cdot (MKT_t - RF_t) + \beta_{\Delta VIX} \cdot \Delta VIX_t. \qquad \text{(A-3)}$$

Gamma (Pastor-Stambaugh gamma) - the firm return sensitivity to the firm previous day dollar volume times the sign of the previous date return, from

$$R_{t+1} = \theta + \phi R_t + \gamma_{PS} \cdot \text{sign}(\text{R}_\text{t}) \cdot \text{Vol}_\text{t}. \tag{A-4}$$

Both the returns and the volume are from CRSP. The dollar volume is scaled by the ratio of the current total market value of NYSE and AMEX shares to the total market value of NYSE and AMEX shares in January 1963. γ_{PS} is computed only for NYSE (exchcd=1) and AMEX (exchcd=2) shares.

Gibbs (Gibbs measure) - the slope from the regression $\Delta P_t = a + c\Delta Q_t$, where P_t is the stock price and Q_t is the trade direction indicator. The values of the Gibbs measure come from the website of Joel Hasbrouck.

IG (investment growth) - the change in capital expenditures (capx item from Compustat) in percentage of last year capital expenditures: $IG_t = \dfrac{capx_t - capx_{t-1}}{capx_{t-1}}$

IVol (idiosyncratic volatility) - the standard deviation of residuals from the Fama-French model, fitted to the daily data for each month (at least 15 valid observations are required).

MB (market-to-book) - equity value (share price, prcc, times number of shares outstanding, csho) divided by book equity (ceq) plus deferred taxes (txdb), all items from Compustat annual files.

Mom (cumulative past return) - cumulative return to the stock between month t-2 and t-12.

Lev (leverage) - long-term debt (dltt) plus short-term debt (dlc) divided by equity value, all items from Compustat annual.

#An (number of analysts; analyst coverage) - the number of analysts covering the firm (from IBES).

O-score - the probability of bankruptcy measure from Ohlson (1980), computed as

$$O = -1.32 - 0.407 \cdot \ln TA + 6.03 \cdot \frac{TL}{TA} - 1.43 \cdot \frac{WC}{TA} + 0.076 \cdot \frac{CL}{CA} - \\ -1.72 \cdot I(TL > TA) - 2.37 \cdot \frac{NI}{TA} - 1.83 \cdot \frac{FFO}{TA} + \\ +0.285 \cdot I(NI < 0 \ \& \ NI_{-1} < 0) - 0.521 \cdot \frac{NI - NI_{-1}}{|NI| + |NI_{-1}|}. \qquad \text{(A-5)}$$

where TA is the book value of total assets (Compustat item at), TL is the book value of total liabilities (lt), WC is working capital (wcap), CL are current liabilities (lct), CA are current assets (act), NI is net income (ni), NI_{-1} is the previous year net income, FFO are funds from operation (pi plus dp), $I(TL > TA)$ is the dummy variable equal to 1 if the book value of total liabilities exceeds the book value of total assets, and equal to 0 otherwise, $I(NI < 0 \ \& \ NI_{-1} < 0)$ is the dummy variable equal to 1 if the net income was negative in the two most recent years, and equal to 0 otherwise.

RI (residual IO) - the residual (ϵ) from the logistic regression of IO on log Size and its square

$$log(\frac{Inst}{1 - Inst}) = \gamma_0 + \gamma_1 \cdot log(Size) + \gamma_2 \cdot log^2(Size) + \epsilon. \qquad \text{(A-6)}$$

IO is the sum of institutional holdings from Thompson Financial 13F database, divided by the shares outstanding from CRSP. All stocks below the 20th NYSE/AMEX size percentile are dropped. If the stock is not dropped, appears on CRSP, but not on Thompson Financial 13Fs, it is assumed to have zero IO.

Roll (Roll measure) - $\text{Roll}_\text{t} = 200 \cdot \sqrt{-\text{Cov}(\text{R}_\text{t}, \text{R}_{\text{t}-1})}$ if the covariance is positive and 0 otherwise.

RSI (relative short interest) – outstanding shorts reported by NYSE and NASDAQ divided by the number of shares outstanding. The data are monthly and reported on the 15th calendar day of each month.

SG (sales growth) - the change in sales (sale item from Compustat) in percentage of last year sales: $SG_t = \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$

Size (market cap) - shares outstanding times price, both from the CRSP monthly returns file.

Spread - the spread implied by the daily high and low prices. Spread is calculated by the formula from Corwin and Schultz (2011):

$$\text{Spread} = \frac{2 \cdot (\exp^{\alpha} - 1)}{1 + \exp^{\alpha}}, \tag{A-7}$$

where

$$\alpha = \frac{\sqrt{\beta} \cdot (\sqrt{2} - 1)}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}, \tag{A-8}$$

where

$$\beta = \log^2\left(\frac{HI_t}{LO_t}\right) + \log^2\left(\frac{HI_{t+1}}{LO_{t+1}}\right) \tag{A-9}$$

and

$$\gamma = \log^2\left(\frac{\max(HI_t,\ HI_{t+1})}{\min(LO_t,\ LO_{t+1})}\right), \tag{A-10}$$

where HI_t (LO_t) is the highest (lowest) price of the stock in day t.

Turn (turnover) - monthly dollar trading volume over market capitalization at the end of the month (both from CRSP).

Volume (dollar trading volume) - price per share at the end of a month times shares traded during the month (both from CRSP), averaged in each firm-year. Following Gao and Ritter (2010), the NASDAQ volume is divided by 2 before February 2001, by 1.8 for the rest of 2001, by 1.6 for 2002-2003, and are unchanged after that. A firm is classified as a NASDAQ firm if its CRSP events file listing indicator - exchcd - is equal to 3.

Zero (zero frequency) - the fraction of zero-return days within each month.