A Data Appendix

Age – the number of months the stock has appeared on CRSP.

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 a stock price of less than \$5 at the end of the previous year are excluded).

Beta (market beta) – market beta from fitting the CAPM to monthly stock returns between months t and t-36.

Cumulative issuance – the difference between the log market capitalization growth and the log cumulative returns in the past five years.

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 in $\frac{1}{2}$ 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.

E(Rank) (expected probability of having a high-rank underwriter) – for IPOs, fitted value from

$$E(Rank) = \Phi\big[c_0 + c_1 \cdot \log(ME) + c_2 \cdot \log(IVol) + \\ + c_3 \cdot \log(Proceeds) + c_4 \cdot VC + c_5 \cdot Loss + c_6 \cdot Liq\big] \tag{A-1}$$

where Φ is the normal c.d.f., ME is post-issue market cap (from CRSP), IVol is post-issue idiosyncratic volatility, Proceeds is the difference between pre-issue and post-issue market

cap (from SDC), VC is a dummy variable that equals 1 if the IPO is backed by a venture capital firm and 0 otherwise (from SDC), Loss is a dummy variable equal to 1 if the IPO incurred an operating loss in the pre-issue quarter, 0 otherwise, Liq is one of the trading costs measures or turnover. For SEOs and convertible debt issuers, E(Rank) is fitted value from

$$E(Rank) = \Phi\big[c_0 + c_1 \cdot \log(ME) + c_2 \cdot \log(IVol) + c_3 \cdot \log(Proceeds) + \\ + c_4 \cdot DIV + c_5 \cdot Loss + c_6 \cdot \#An + c_7 \cdot Liq\big] \qquad \text{(A-2)}$$

where ME is pre-issue market cap (from CRSP), IVol is pre-issue idiosyncratic volatility, DIV is a dummy variable that equals 1 if the company pays dividends (reports *dvc¿*0 on Compustat), 0 otherwise, # An is the number of analysts covering the firm (from IBES).

E(Under) (expected probability of underpricing – fitted value from

$$E(Under) = \Phi\big[c_0 + c_1 \cdot \log(ME) + c_2 \cdot \log(IVol) + c_3 \cdot \log(MB) + \\ + c_4 \cdot \log(Lev) + c_5 \cdot R\&D + c_6 \cdot Liq\big] \qquad \text{(A-3)}$$

where ME, IVol, Liq are as defined above (see E(Rank)), MB is market-to-book (see definition below), Lev is market leverage equal to long-term debt (dltt) plus short-term debt (dlc) divided by equity value (prcc times csho), all items from Compustat annual file, R&D is R&D-to-assets ratio (item xrd over item at, both from Compustat annual).

E(VC) (expected probability of venture capital backing – fitted value from

$$E(VC) = \Phi\big[c_0 + c_1 \cdot \log(ME) + c_2 \cdot \log(IVol) + c_3 \cdot \log(Proceeds) + \\ + c_4 \cdot \log(MB) + c_5 \cdot Distress + c_6 \cdot Liq\big] \qquad \text{(A-4)}$$

where ME, IVol, Proceeds, MB, Liq are as defined above (see E(Rank) and E(Under)), and Distress is a dummy variable that equals 1 if oibdp item from Compustat annual is negative, 0 otherwise.

FVIX (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:

$$\Delta VIX_t = \gamma_0 + \gamma_1 \cdot (VIX1_t - RF_t) + \gamma_2 \cdot (VIX2_t - RF_t) +$$
$$+\gamma_3 \cdot (VIX3_t - RF_t) + \gamma_4 \cdot (VIX4_t - RF_t) + \gamma_5 \cdot (VIX5_t - RF_t), \qquad \text{(A-5)}$$

where $VIX1_t, \ldots, VIX5_t$ are the VIX sensitivity quintiles described above, with $VIX1_t$ being the quintile with the most negative sensitivity.

The fitted part of the regression above less the constant is my volatility risk factor (FVIX factor). The returns are then cumulated to the monthly level to get the monthly return to FVIX.

Leverage – the ratio of the sum of short-term debt (dlc item from Compustat) and long-term debt (dltt) to market value of equity (share price, prcc, times number of shares outstanding, csho).

LMH (turnover factor) – the arbitrage portfolio that buys firms in the bottom turnover quintile and shorts firms in the top turnover quintile (NYSE quintile breakpoints are used).

Negative return – in Tables 2 and 3, a control that equals the value of the monthly stock return if the value is negative and zero otherwise.

MB (market-to-book) – For IPOs, I use after-issue market cap and divide it by total common equity (both from SDC). For other firms, MB equals market value of equity (share price, prcc, times number of shares outstanding, csho) divided by book equity (ceq) plus deferred taxes (txdb), all items from Compustat annual files.

Mills (inverse Mills ratio) – inverse Mills ratio from the respective regression (see E(Rank), E(Under), E(VC)), estimated as

$$Mills = \frac{\phi(X\beta)}{\Phi(X\beta)} \qquad \text{(A-6)}$$

where ϕ is normal p.d.f., Φ is normal c.d.f., and $X\beta$ is the sum of products of all regressors and respective slopes.

Positive return – in Tables 2 and 3, a control that equals the value of the monthly stock return if the value is positive and zero otherwise.

Press (price pressure dummy) – Mutual fund flow is calculated for each fund as

$$\text{Flow}_\text{t} = \frac{\text{TNA}_\text{t} - (1 + \text{R}_\text{t}) \cdot \text{TNA}_{\text{t}-1}}{\text{TNA}_{\text{t}-1}} \tag{A-7}$$

where TNA stands for total net assets. Funds with flows in the top/bottom decile are assumed to exhibit non-information-driven positive/negative price pressure. For each firm, price pressure equals the number of its shares bought (from change in holdings) by funds creating positive price pressure minus the number of its shares sold by funds creating negative price pressure, normalized by shares outstanding. If price pressure is in the top decile among all firms with valid mutual fund flow data on the CRSP mutual fund database, Press dummy equals one, otherwise Press dummy equals zero.

Price (stock price) – stock price from CRSP monthly file.

Roll (Roll measure) – $Roll_t = 200 \cdot \sqrt{abs(Cov(R_t, R_{t-1}))}$, where R_t are daily stock returns. Roll measure is computed within each firm-year.

SG1/SG2 (the smallest/second smallest growth portfolio) – the intersection of the bottom (second bottom) size quintile and the top market-to-book quintile.

Size (market capitalization) – shares outstanding times price. For new issues, I use after-issue market cap from SDC. For other firms, both shares outstanding and the stock price come 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 (2012):

$$\text{Spread} = \frac{2 \cdot (\exp^{\alpha} - 1)}{1 + \exp^{\alpha}}, \quad \text{where} \tag{A-8}$$

$$\alpha = \frac{\sqrt{\beta} \cdot (\sqrt{2} - 1)}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}, \quad \text{where} \tag{A-9}$$

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

where HI_t (LO_t) is the highest (lowest) price of the stock on day t. Spread is computed within each firm-year.

Turn (turnover) – trading volume divided by shares outstanding (both from CRSP monthly data). The monthly turnover is then averaged in each calendar year with at least 5 valid observations. To make comparisons across exchanges more meaningful, I adjust NASDAQ volume for the double counting following Gao and Ritter (2010): NASDAQ volume is divided by 2 for the period from 1983 to January 2001, by 1.8 for the rest of 2001, by 1.6 for 2002–2003, and is unchanged after that. A firm is classified as a NASDAQ firm if its CRSP events file listing indicator (exchcd) is equal to 3.

VIX – the VIX index, defined as the implied volatility of at-the-money options on the S&P 100 (current ticker VXO). VIX is computed by CBOE and obtained from WRDS.

Zero (zero frequency) – the fraction of zero-return days within each firm-year.