Estimates of Expected Skewness

This document explains the measures of expected idiosyncratic stock return skewness posted online, which are developed based on the methodology of Boyer, Mitton and Vorkink (2010).

Suppose we are interested in forecasting skewness for the distribution of daily returns over the next T months. Let t denote the current month, let $S(t)$ denote the set of trading days from the first day of month $t-T+1$ through the end of month t, and let $N(t)$ denote the number of days in this set. Let $\epsilon_{i,d}$ be the regression residual using the Fama-French three-factor model on day d for firm i, where the regression coefficients that define this residual are estimated using daily data for days in $S(t)$. In addition, let $iv_{i,t}$ and $is_{i,t}$ denote historical estimates of idiosyncratic volatility and skewness (respectively) for firm i using daily data for all days in $S(t)$. We can then define $iv_{i,t}$ and $is_{i,t}$ as:

$$iv_{i,t} = \left(\frac{1}{N(t)-1} \sum_{d \in S(t)} \epsilon_{i,d}^2 \right)^{1/2}, \tag{1}$$

$$is_{i,t} = \frac{1}{N(t)-2} \frac{\sum_{d \in S(t)} \epsilon_{i,d}^3}{iv_{i,t}^{3/2}}. \tag{2}$$

By replacing $\epsilon_{i,d}$ with the total return minus an estimate of the mean return on the right of equations (1) and (2) we can, of course, also obtain estimates of historical total volatility and historical total skewness.

We need measures of *expected* skewness over a horizon of T months for firm i at the end of month t, $E_t[is_{i,t+T}]$, rather than measures of historical skewness as defined in equation (2). These estimates of expected skewness should be feasible in that they use information available to investors at the end of month t. To model investor perceptions of expected skewness in a feasible manner, we first estimate cross-sectional regressions separately at the end of each month t in our sample,

$$is_{i,t} = \beta_0^t + \beta_1^t is_{i,t-T} + \beta_2^t iv_{i,t-T} + \boldsymbol{\gamma}^{\mathbf{t}} \mathbf{X}_{i,t-T} + \varepsilon_{i,t}, \tag{3}$$

where $\mathbf{X}_{i,t-T}$ is a vector of additional firm-specific variables observable at the end of month $t-T$. Superscripts on regression parameters are included to emphasize that we estimate these parameters using information observable at the end of month t. Equation (3) is similar to the panel estimations conducted in Chen, Hong, and Stein (2001) with the exception that we estimate the model separately each month. We then use the regression parameters from equation (3), along with information observable at the end of each month t, to estimate expected skewness for each firm,

$$E_t[is_{i,t+T}] = \beta_0^t + \beta_1^t is_{i,t} + \beta_2^t iv_{i,t} + \boldsymbol{\gamma}^{\mathbf{t}}\mathbf{X}_{i,t}. \tag{4}$$

This approach not only allows the relation between firm-specific variables and skewness to vary across time, but also provides feasible estimates of expected skewness each month. We can also obtain estimates of expected *total* skewness by replacing measures of idiosyncratic skewness and volatility in equations (3) and (4) with measures of total skewness and volatility.

The skewness estimates posted online are for a horizon of T=60 months and the firm specific variables, $\mathbf{X}_{i,t-T}$, are the same as in Boyer, Mitton, and Vorkink (2010) with the exception that industries are defined in a slightly different manner. In particular, the firm specific variables for the online estimates are

Momentum ($mom_{i,t-T}$) : the cumulative return for firm i from the end of month $t-T-12$ through the end of month $t-T-1$.

Turnover ($turn_{i,t-T}$) : the sum of daily turnover for firm i over month $t-T$. Daily turnover for day d is defined as volume for day d divided by shares outstanding reported on day d.

Nasdaq ($Nasd_{i,t-T}$) : dummy variable indicating firms listed on Nasdaq (CRSP exchange code = 3 for month $t-T$).

Small ($Small_{i,t-T}$) : dummy variable indicating firms in the bottom tercile ranked by size at the end of month $t-T$.

Medium ($Med_{i,t-T}$:) dummy variable indicating firms in the middle tercile ranked by size at the end of month $t-T$.

Industry Dummies: dummies for 16 of the 17 industries defined by Ken French to create the "17 Industry Portfolios" on his website. Boyer, Mitton, and Vorkink (2010) originally developed their own method to define industries parsimoniously using information on Ken French's website in order to better investigate the skewness of specific industries. Using Ken French's 17-industry classification makes the expected skewness estimates posted online more transparent and easier to replicate.

CRSP reports turnover for NASDAQ on a widespread basis beginning January 1983. Hence, Boyer, Mitton, and Vorkink (2010) recommend estimation of equation (3) beginning with $T - t$=January 1983, implying that the initial estimate of predicted skewness given by equation (4) is observable at the end of January 1987. However, estimates of predicted skewness posted online begin much earlier. Other important data procedures include the following:

Ordinary Common Shares: We provide predicted skewness estimates only for ordinary common shares (CRSP share code 10 or 11). We also limit the universe of securities to ordinary common shares before sorting into size terciles to define the size dummies.

Positive Shares Outstanding: We eliminate securities in month t if CRSP reports zero shares outstanding in month t.

Sufficient Data to Calculate Momentum : The variable $mom_{i,t-T}$ is set to missing at the end of month $t-T$ if a security does not have a complete set of 11 monthly returns in CRSP from $t-T-11$ through month $t-T-1$.

Sufficient Data to Calculate Skewness : The variable $is_{i,t}$ is sent to missing at the end of month t if the number of observable returns in the set $S(t)$ is less than 250.

References

[1] Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169-202.

[2] Chen, Joseph, Harrison Hong, and Jeremy Stein, 2001, Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices, *Journal of Financial Economics* 61, 345-381.

[3] Hong, Harrison and Jeremy Stein, 2003, Differences of opinion, short-sales constraints and market crashes, *Review of Financial Studies* 16, 487-525.

[4] Zhang, Yijie, 2005, Individual skewness and the cross-section of average stock returns, Working paper, Yale University.