December 12, 2023

1 Double regression analysis:

We start with the following data generating process:

$$y_t = z_t'\beta + \varepsilon_t$$

The variance is given by:

$$V(y) = \beta'\Sigma_Z\beta + \sigma_\varepsilon^2$$

Using the fitted values, we now run a second regression:

$$\hat{y}_t = x_t'\gamma + \eta_t$$

The variance is therefore given by:

$$\begin{aligned} V(\hat{y}_t) &= V(z_t'\beta) \\ &= \beta'\Sigma_Z\beta \\ &= \gamma'\Sigma_X\gamma + \sigma_\eta^2 \end{aligned}$$

This implies:

$$\begin{aligned} V(y_t) &= V(\hat{y}_t + \varepsilon_t) \\ &= \gamma'\Sigma_X\gamma + \sigma_\eta^2 + \sigma_\varepsilon^2 \end{aligned}$$

Note that this applies only to the backcasted part of the track record. The plug-in estimator for the total volatility of an asset will need to account for both backcasted and non-backcasted portions.

Start by making the strong assumption that the idiosyncratic variance ε_t is independent of the factors in the second regression. This means that

$$E[\hat{y}_t|x_t] = E[y_t|x_t] = x_t'\gamma$$

Then systematic risk is estimated as

$$\sigma_{SYS}^2 = \hat{\gamma}'\hat{\Sigma}_X\hat{\gamma}$$

The idiosyncratic risk then depends on whether a data point uses fitted values or actual data:

$$\hat{\sigma}_{IVOL}^2 = \frac{1}{T}\sum_{t\in 1:T}\begin{cases}(y_t - x_t'\hat{\gamma})^2 & t \in data \\ (\hat{y}_t - x_t'\hat{\gamma})^2 + \hat{\sigma}_\varepsilon^2 & t \in backcast\end{cases}$$

These are equivalent because:

$$
\begin{aligned}
E\left[\left(y_t - x_t'\gamma\right)^2\right] =& E\left[\left(x_t'\gamma + \eta_t + \varepsilon_t - x_t'\gamma\right)\left(x_t'\gamma + \eta_t + \varepsilon_t - x_t'\gamma\right)\right] \\
=& \sigma_\eta^2 + \sigma_\varepsilon^2
\end{aligned}
$$

Note the above also implies that the Morningstar idiosyncratic variance can be adjusted as follows:

$$
\begin{aligned}
\hat{\sigma}^2_{IVOL*} =& \frac{1}{T}\sum_{t\in 1:T}\begin{cases}\left(y_t - x_t'\hat{\gamma}\right)^2 & t \in data \\ \left(\hat{y}_t - x_t'\hat{\gamma}\right)^2 & t \in backcast\end{cases} \\
=& \hat{\sigma}^2_{IVOL*} + \frac{T_{back}}{T}\sigma_\varepsilon^2
\end{aligned}
$$