

# Understanding of Factor Strength

Zhiyuan Jiang

I.D:28710967

April 17, 2020

## How the factor strength is defined and calculated

Factor strength has been elaborated by Bailey, Kapetanios, and Pesaran (2020), The following section will re-iterate their definition of factor strength as well as the method to estimate it.

To start with, we define a single factor CAPM model:

$$y_{it} = \beta_i + \theta_i x_t + \varepsilon_{it} \quad (1)$$

Assume we have  $n$  different assets (for instance,  $n = 500$  if using data from S&P 500 index). Collecting and calculating those assets returns from  $T$  different observations.  $y_{it}$  on the left hand side of equation (1) is the excess return of asset  $i$  at time  $t$ , The excess return equals to the asset return minus the risk free return.  $x_t$  in the right hand side is the factor with interest at time  $t$ . Therefore,  $\theta_i$  is the loading of factor  $x_t$ .  $\beta_i$  is the constant term, represent the asset's ability to generate abnormal return from the market.  $\varepsilon_{it}$  as the idiosyncratic error term has been assumed to follow independent, identical distribution, with zero mean and time invariant variance  $\sigma_i^2$ ,

After settle down, we run OLS for this model and obtain the results:

$$y_{it} = \hat{\beta}_i + \hat{\theta}_i x_t + \hat{\varepsilon}_{it}, \quad t = 1, 2, 3, \dots, T$$

Both  $\hat{\beta}_i$  and  $\hat{\theta}_i$  are the OLS estimation results of equation (1). Because we want to investigate the differences between estimated factor loading  $\hat{\theta}_i$  and zero, we can construct a t-test with  $t_i = \frac{\hat{\theta}_i - 0}{\hat{\varsigma}_i}$  where  $\hat{\varsigma}_i$  is the standard error of  $\hat{\theta}_i$ . Then we defined  $\pi_{nT}$  as the proportion of significant factor's amount to the total observations amount:

$$\hat{\pi}_{nT} = \frac{\sum_{i=1}^n \hat{\ell}_{i,nT}}{n} \quad (2)$$

$\ell_{i,nT}$  is an indicator function as:  $\ell_{i,nT} := \mathbf{1}[|t_i| > c(n)]$ . If the t-statistic  $t_i$  is greater than the critical value  $c_p(n)$ ,  $\hat{\ell}_{i,nT} = 1$ . In other word, we will count one if the factor loading  $\hat{\theta}_i$  is significant.  $c_p(n)$  represent the critical value of a test with test size  $p$ . The critical value is calculated by:

$$c_p(n) = \Phi^{-1}\left(1 - \frac{p}{2n^\delta}\right) \quad (3)$$

Here,  $\Phi^{-1}(\cdot)$  is the inverse cumulative distribution function of a standard normal distribution, and  $\delta$  is a non-negative value represent the critical value exponent. The traditional method to calculate critical value has not fixed the multiple testing problem. One of the most commonly used adjustment for multiple testing problem is Bonferroni correction. When  $n$  as sample size goes to infinity, however, the Bonferroni correction can not yield satisfying results since the  $\frac{p}{2n^\delta} \rightarrow 0$  when  $n \rightarrow \infty$ . Therefore, Bailey, Kapetanios, and Pesaran (2016) provides another adjustment with additional exponent  $\delta$  to constrain  $n$ .

After obtain the  $\hat{\pi}_{nT}$ , we can use the following formula to estimate our strength indicator  $\alpha$ :

$$\hat{\alpha} = \begin{cases} 1 + \frac{\ln(\hat{\pi}_{nT})}{\ln n} & \text{if } \hat{\pi}_{nT} > 0, \\ 0, & \text{if } \hat{\pi}_{nT} = 0. \end{cases}$$

From the estimation, we can find out that  $\hat{\alpha} \in [0, 1]$

$\hat{\alpha}$  represent the pervasiveness of a factor. Here we denote  $[n^\alpha]$ ,  $[\cdot]$  will take the integer part of number inside. For factor  $\theta_i$ :

$$\begin{aligned} |\theta_i| &> c_p(n) \quad i = 1, 2, \dots, [n^\alpha] \\ |\theta_i| &= 0 \quad i = [n^\alpha] + 1, [n^\alpha] + 2, \dots, n \end{aligned}$$

For a factor has strength  $\alpha = 1$ , factor will be significant for every assets at every time. The more observation the factor can significantly influence, the stronger the factor is, and vice versa.

## References

- Bailey, N., Kapetanios, G., & Pesaran, M. H. (2016, sep). Exponent of Cross-Sectional Dependence: Estimation and Inference. *Journal of Applied Econometrics*, 31(6), 929–960. Retrieved from <http://doi.wiley.com/10.1002/jae.2476> doi: 10.1002/jae.2476
- Bailey, N., Kapetanios, G., & Pesaran, M. H. (2020). *Measurement of Factor Strength: Theory and Practice*.