

1. Some very points:

标普中国A股质量价值指数：方法论

目的：该指数旨在反映中国A股市场中质量高且估值较合理的100只股票的业绩。

合格备选集：

- 母指数：标普中国A股BMI指数+标普中国A股创业板指数。ST和*ST股除外。
- 自由流通市值：至少为10亿人民币(当前成份股至少为9亿人民币)。
- 三个月日均交易额：至少2千万人民币(当前成份股至少为1千8百万人民币)。

选股方法 挑选质量得分最高的200只股票组成质量200组合，再根据价值得分挑选价值得分最高的100只股票组成标普中国A股质量价值指数，剩下的100只股票组成质量高估值组合。调整缓冲为20%。

加权方式：基于自由流通市值*(1+质量得分)的乘积进行加权，个股及行业权重限制分别为5%和40%。

指数调整：半年调整

- 调整参考日：每年5月及11月月底。
- 调整生效日：每年6月及12月的第三个星期五收盘后。

S&P Dow Jones
Indices

- a) On the long-running question of market efficiency, we cannot confidently identify that and particular Return Predict Signals RPS (2006) reflects the rational pricing of economic risks over mispricing, or vice-versa.

2. Practically

- a) We seek to dampen the potential impacts of economically genuine and data-error-based outliers by using monthly cross-sectional scaled decile rankings of each continuous or non-indicator RPS in our return prediction regressions. We implement the scaled decile ranked approach at the end of every calendar month by ranking each non-indicator RPS into deciles where zero is the lowest decile and nine is the highest decile, and then dividing by nine. However, the scaled decile ranked RPS are only created after resetting missing RPS values to each RPS ' monthly mean.

3. Details

- a) Mean annualized long/short hedge returns (MALSRets) and associated t-statistics implied by slope coefficients estimated in Fama-MacBeth regressions of 1 month ahead, 2-12 months ahead and 13-36 months ahead firm-specific returns $RET = a + b_1 mve + b_2 .bm + b_3 .mom12m + e$ over 1980-2012. Each of the three RPS mve, bm, mom12m is recalculated monthly by first ranking the RPS into deciles (0-9) and then dividing by 9. Before

ranking, certain missing data items are set to zero (see Table 3), and missing RPS values are reset to that month's mean RPS value. Estimated MALSReTs are annualized by multiplying by 12 (12/11, 0.5) for 1 month (2-12, 13-36) ahead horizons. t-statistics use Newey-West adjustments of 12 (1) lags for 1 (2-12, 13-36) month ahead returns. Return windows are defined in non-overlapping calendar time, so that 2-12 months ahead returns cumulate from Aug. of year t thru June of year t+1, while 13-36 months ahead returns cumulate in non-overlapping two year periods starting Jan. 1981 - Dec. 1982. Large-Cap are the largest 1,000 companies by market cap; Mid-Cap are the next largest 2,000 companies; Small-Cap are all remaining firms.

4. Fama French⁷³

- a) The Fama-MacBeth regression is a method used to estimate parameters for asset pricing models such as the Capital asset pricing model (CAPM). The method estimates the betas and risk premia for any risk factors that are expected to determine asset prices. The method works with multiple assets across time (panel data). The parameters are estimated in two steps:
 - i. First regress each asset against the proposed risk factors to determine that asset's beta for that risk factor.
 - ii. Then regress all asset returns for a fixed time period against the estimated betas to determine the risk premium for each factor.
- b) Note that Fama MacBeth regressions provide standard errors corrected only for cross-sectional correlation. The standard errors from this method do not correct for time-series autocorrelation. This is usually not a problem for stock trading since stocks have weak time-series autocorrelation in daily and weekly holding periods, but autocorrelation is stronger over long horizons. This means Fama MacBeth regressions may be inappropriate to use in many corporate finance settings where project holding periods tend to be long. For alternative methods of correcting standard errors for time series and cross-sectional correlation in the error term look into double clustering by firm and year.
- c) Arithmetic return instead of logarithm return
<https://www.forbes.com/sites/frankarmstrong/2013/05/23/fama-french-three-factor-model/#106bcd6b26b3>

5. HAC covariance estimator

- a) One way empirical finance researchers can address two sources of correlation

is to parametrically estimate one of the dimensions (e.g., by including dummy variables). Since many panel data sets have more firms than years, a common approach is to include dummy variables for each time period (to absorb the time effect) and then cluster by firm (Lamont and Polk, 2001; Anderson and Reeb, 2004; Gross and Souleles, 2004; Sapienza, 2004; and Faulkender and Petersen, 2006).

- b) One way empirical finance researchers can address two sources of correlation is to parametrically estimate one of the dimensions (e.g., by including dummy variables). Since many panel data sets have more firms than years, a common approach is to include dummy variables for each time period (to absorb the time effect) and then cluster by firm (Lamont and Polk, 2001; Anderson and Reeb, 2004; Gross and Souleles, 2004; Sapienza, 2004; and Faulkender and Petersen, 2006). If the time effect is fixed (e.g., Equation (15)), the time dummies completely remove the correlation between observations in the same time period. In this case, there is only a firm effect left in the data.
- c) if not full ranked:
 - i. Cameron, Gelbach, and Miller (2006), and Thompson (2006) proposed the following estimate of the variance-covariance matrix:
 - ii. $V_{\text{firm \& time}} = V_{\text{firm}} + V_{\text{time}} - V_{\text{white}}$
 - iii. Cluster methodologies: from Monte Carlo simulation this way has improved the accuracy of the estimated covariance sharply

6. Sort method?

- a) Why sort
- b) in FM93 the result sorted by size and beta

7. It often makes sense to add quadratic terms of any significant variables to a model.

- a) Often the relationship between y and x is nonlinear. There are a variety of solutions. One solution is to add polynomial terms and the first one to look at is usually x^2 . But you should first look at a scatterplot of x and y ; you should also look at the residuals from the linear model without the quadratic term. But it turns out that many relationships are pretty well fit by $y \sim b_0 + b_1x + b_2x^2$ (plus any other x variables, of course).
- b) It is also possible to add cubic, quartic and even higher order terms, but such models quickly become hard to interpret. Another possibility is to look at a spline regression

8. Long-only and long/short risk factors have not the same behavior. This is for example the case of BAB and WML factors.
 - a) WML is the equal-weight average of the returns for the two winner portfolios for a region minus the average of the returns for the two loser portfolios,
 - b) Betting against beta
9. Risk factors are local, not global. It means that risk factors are not homogeneous. For instance, the value factors in US and Japan cannot be compared (distressed stocks versus quality stocks).
10. Tangency portfolio, maybe one of the breakthrough
11. Risk factor/slope weighted methodologies
 - a) Value weight
 - b) Equally weight
 - c) Rank weight
12. Constant Mixed strategy?
 - a) The skewness premia of the factors, especially during the crisis
 - b) Scalability of risk factors.
13. How to find the basic factors?
 - a) ICAPM
 - b) APT theory
14. Combination of the factors?
 - a) How to lead the combo make sense? Find explainable risk premia?
 - b) Machine learning? GPU?
 - c) Reasonable v.s. statistically robustness
15. Special Literature review
 - a) FF93:
 - i. equally weighted the monthly return
 - ii. $Nb_underlyings/20$ for each portfolio for cross validation regression
 1. Comments: is it good for solving the regression model problem and full ranked problem??
 2. The estimation of the beta of the equally weighted portfolios is more accurate than the individual portfolio in the case.
 - iii. rearrange the regression period
 1. Comments: Why at that period?
 - a) 5 years covers each
 - b) pre WWI & WWII and post WWI & WWII

- iv. Design the hypothesis test for different distribution of the underlying returns?
- v. The study on the market performance/behavior -- market efficiency
- vi. R_f as the 1 month treasury bills
- b) Huatai securities
 - i. WLS for Heteroskedasticity?? where W as square root of the mkt value weighted beta
- c) Burton 2014
 - i. Statistical result
 - 1. Value Stocks Are to be Favored Over Growth Stocks
 - 2. The small firm effect
 - a) Internet Bubble
 - ii. The momentum and mean reversion
 - 1. ST - momentum
 - 2. LT - Mean reversion
 - 3. Comments: overreact? over confidence? underestimate the risk or change the risk appetites?
 - iii. Low volatility, however high return
 - a) Different weight methodologies
 - b) Quality and quantity
 - iv. RAFI Fundamental Index
 - v. According with RAFI, mkt cap weighted portfolio always over holding too much overpriced assets
 - vi. Design the framework
 - 1. Find the weight single factor
 - 2. Statistical test for the different weighted way for alpha (return of the factor weighted portfolio - return of the Mkt cap weighted portfolio)
 - 3. Matrix decomposition find most set of factors play as the orthogonal basis
 - 4. Any optimized way for the further study?
 - a) Minimized the volatility of the portfolio?
 - b) Minimized the tracking errors?
 - c) Maximized the expected sharp ratio?
 - d) Something else?
 - vii. Versus Equally weighted portfolio

- viii. Tax result? For invest in mutual fund? Some Special design the portfolio for the high cash and cash equivalent companies

16. Strongly related anomalies

- a) Low volatility anomaly
- b) Idiosyncratic volatility anomaly
- c) Low beta anomaly