Introduction to Financial Data Analysis

Week 14: Portfolio Analysis in Asset Pricing

Portfolio Analysis in Asset Pricing

- Portfolio analysis is one of the most commonly used statistical methodologies in empirical asset pricing.
- Aims: Examine the relationship between two or more variables.
  - Example: Examine the ability of one or more variables to predict future stock returns.
  - Example: Examine the relationship between different characteristics of stocks.
- Portfolio Analysis: Form portfolios of stocks, where the stocks in each portfolio have different levels of the variable/variables and to examine the returns of these portfolios.

Portfolio Analysis in Asset Pricing

- Portfolio analysis: Nonparametric approach.
- Nonparametric Approach: It does not make any assumptions about functional relationship between the variables under investigation.
- Parametric Approach: Methodologies that rely on some assumptions regarding the functional form of the relation between the variables being examined.
  - For example, linear regression analysis assumes that the relation between the dependent and independent variables is linear.
- Portfolio analysis can uncovering nonlinear relations between variables that are quite difficult to detect using parametric techniques.
- Drawback of Portfolio analysis: difficult to control for a large number of variables in examining the relationship compared to regression analysis.

Univariate Portfolio Analysis

- A univariate portfolio analysis: Only one sort variable X.
- Objective: Examine the relationship between X and the outcome variable Y.
- Procedure: 4 Steps
- 1. Calculate the breakpoints that will be used to divide the sample into portfolios.
- 2. Use these breakpoints to form the portfolios.
- 3. Calculate the average value of the outcome variable Y within each portfolio for each period t.
- 4. Examine the variation in these average values of Y across the different portfolios.

Univariate Portfolio Analysis

- Example: Consider all the stocks that were traded on A share market from 2000-Jan to 2015-Dec.
- Target: Examine the relationship between the month end market cap (size) and the next month returns.
- The sort variable X is the market cap of the all stocks.
- The dependent variable Y is the next month returns of all stocks.

Breakpoints

- Step 1: Calculate the periodic breakpoints that will be used to group the entities in the sample into portfolios based on values of the sort variable X.
  - Entities with values of X that are less than the first breakpoint will be placed into the first portfolio.
  - Entities with values of X that are between the first and second breakpoints will comprise the second portfolio, etc.
  - Finally, entities with X values higher than the highest breakpoint will be placed in the last portfolio.
- Denote the number of portfolios to be formed each time period as nP.
- The number of breakpoints that need to be calculated each period is therefore $nP - 1$.

Breakpoints

- In portfolio analysis, the portfolio are reformed every period.
- The number of portfolios to be formed is the same for all time periods.
- The value of the kth breakpoint varies from time period to time period.
- Denote the kth breakpoint for period t as $B_{k,t}$ for $k \in \{1, 2, \cdots, nP - 1\}$.

Breakpoints

- The breakpoints for period t are determined by percentiles of the time t distribution of the sort variable X among all assets.
- Let p_k be the percentile that determines the *k*th breakpoint, the *k*th breakpoint for period *t* is calculated as the p_k th percentile of the values of *X* across all entities in the sample in period *t*.
- Define the breakpoints as

$$B_{k,t} = Pctl_{p_k}(\{X_t\})$$

- The percentiles, and thus the breakpoints, increase as *k* increases for all periods t.

$$0 < p_1 < p_2 < \cdots < p_{n_P-1} \text{ and } B_{1,t} \leq B_{2,t} \leq \cdots \leq B_{n_P-1,t}$$

Breakpoints

- Discussion on breakpoints:
- 1. The actual breakpoints may not be strictly increasing
  - Example: There may be a large number of entities for which the values of X are the same, causing two or more of the breakpoints to be the same. Eg. Credit Rating of bonds.
- 2. Number of entities in each portfolio:
  - A small number of entities in each portfolio results in increased noise when using the sample mean value of Y as an estimate of the true mean.
  - The more entities we group into each portfolio, the smaller the number of portfolios. Few portfolio more difficult to detect the relations between X and Y.

Breakpoints

- Researchers usually form portfolios using breakpoints that represent evenly spaced percentiles of the cross-sectional distribution of the sort variable.
  - For example, split the sample into five portfolios, we may use the 20th, 40th, 60th, and 80th percentiles of the sort variable as the portfolio breakpoints.
  - For example, when splitting the sample into only three portfolios, it is common to use the 30th and 70th percentiles of the sort variable as the breakpoints. (As in FF)
- Almost all studies use between three and 20 portfolios, with most researchers choosing either five or 10.

Breakpoints

- See “Size_Ret.html”
- Example: Form the portfolios of stocks traded in China A shares market based on the market cap(size) of stocks using data from 2010-01 to 2016-12.
- The breakpoints of market cap are the 10th, 20th, 40th, 60th, 80th, and 90th percentiles of size.
- Use uneven breakpoints simply to exemplify the flexibility of the portfolio procedure.
  - Researchers choose to make the distance between the percentiles that determine the breakpoints smaller for the lowest and highest portfolios because doing so can help us understand whether the relation under investigation is stronger for entities with extreme (low or high) values of the sort variables X.

Portfolio Formation

- Step 2: Group the entities in the sample into portfolios.
  - Each time period t, all entities in the sample with values of the sort variable X that are less than or equal to the first breakpoint, are put in portfolio one.
  - Portfolio two holds entities with values of X that are greater than the first breakpoint and less than or equal to the second breakpoint.
  - Portfolio three holds entities with values of X greater than the second breakpoint and less than or equal to the third breakpoint, and so on.
- The number of entities in each of the portfolios should be approximately dictated by the percentiles used to calculate the breakpoints and the number of stocks in the sample during the given period t.
- See “Size_Ret.html”

Average Portfolio Values

- Step 3: First, calculate the average value of the outcome variable *Y* for each of the portfolios in each time period *t*.
  - Equal-weighted average
  - Value-weighted average: use market cap as weights in the calculation of averages.

$$\overline{Y}_{k,t} = \frac{\sum_{i \in P_{k,t}} W_{i,t} Y_{i,t}}{\sum_{i \in P_{k,t}} W_{i,t}}$$

  - Where k represents the kth portfolio, i is the entity (stock).

Average Portfolio Values

- Discussion:
- 1. The value of Y might not be available though the stock exist in the portfolio. In this case, the set over which the average is taken should be the set of entities for which a value of Y is available.
- 2. Similar, if W is missing, the summation is taken over all entities i for which values of both W and Y are available.

Average Portfolio Values

- Remember: Target of Portfolio Analysis: To detect the relationship between the sort variable X and the outcome variable Y. This is done through the analysis of the difference portfolio.
- Second, calculate the difference in average values between portfolio last portfolio and the first portfolio.
  - Define the difference in the average outcome variable between the highest and lowest portfolios to be

$$\overline{Y}_{Diff,t} = \overline{Y}_{n_P,t} - \overline{Y}_{1,t}.$$

  - The difference in the average value of the outcome variable *Y* for entities with high values of the sort variable compared to those with low values of the sort variable.
  - Use this value to detect the relationship between the sort variable X and the outcome variable Y.

Average Portfolio Values

- Example: See “Size_Ret.html”
- Use the 1 month ahead raw returns r(t+1) as Y.
- The average stock returns in the next month represent the returns that would be realized by an investor who, at the end of month t, created the portfolios as described previously and held the portfolios without further trading for the next entire month t+1.
- The investor change his portfolio at each month end.
- The difference portfolio return is the return the investor could earn by long the highest category portfolio while short the lowest category portfolio with the same amount of capital.

Summarizing and Interpreting the Results

- To examine whether there is a relation between the sort variable *X* and the outcome variable *Y*.
- First, calculate the time-series means of each portfolio's Y and the difference portfolio's time-series average

$$\overline{Y}_k = \frac{\sum_{t=1}^{T} \overline{Y}_{k,t}}{T}$$

$$\overline{Y}_{Diff} = \frac{\sum_{t=1}^{T} \overline{Y}_{Diff,t}}{T}$$

Summarizing and Interpreting the Results

- Second, test whether the time-series mean for each of the portfolios statistically distinguishable differs from some null hypothesis mean value. (Usually 0)
- Statistically nonzero in the difference portfolio mean indicates a relationship between X and Y.
- Calculate the standard errors, t-statistics, and p-values for each portfolio.
- Due to the existence of autocorrelation in time series, the standard errors are not accurate and are frequently adjusted following Newey and West (1987).
- Finally, also examine whether the average values of Y across the portfolios from lowest to highest is monotonic.

Summarizing and Interpreting the Results

- Example: See “Size_Ret.html”