Financial Econometrics

Bo Hu*

*Center for New Structural Economics, Peking University. Email: bohu@nsd.pku.edu.cn

1 Introduction

1.1 Financial Econometrics

Financial Econometrics is the study of finance using statistical tools. The study of finance has very strong empirical nature, and most of the financial data are generated in non-experimental environments. This calls for model-based statistical inference in the empirical study of financial markets, which is the main task of financial econometrics. Financial econometrics can be used to discover facts in financial markets, to test finance theories, to predict asset returns, to study short-run and long-run relationships between different markets and between financial markets and the macro economy, and to make financial decisions. Brooks (2014) in his introductory textbook listed ten topics as examples of the capability of financial econometrics.

1. Testing whether financial markets are weak-form informationally efficient.
2. Testing whether the Capital Asset Pricing Model (CAPM) or Arbitrage Pricing Theory (APT) represent superior models for the determination of returns on risky assets.
3. Measuring and forecasting the volatility of bond returns.
4. Explaining the determinants of bond credit ratings used by the ratings agencies.
5. Modeling the long-term relationships between prices and exchange rates.
6. Determining the optimal hedge ratio for a spot position in oil.
7. Testing technical trading rules to determine which makes the most money.
8. Testing the hypothesis that earnings or dividend announcements have no effect on stock prices.
9. Testing whether spot or futures markets react more rapidly to news.
10. Forecasting the correlation between the stock indices of two countries.

The above is of course not an exhaustive list and we may easily add more.

1.2 Financial Data

The tools used in financial econometrics is hardly different from those in the "general econometrics". However, financial data have some special characteristics. In finance, we mostly deal with time series data or panel data, and usually both the cross-sectional dimension and the time dimension of the data can be reasonably large. Sometimes we have high frequency data, i.e., data sampled with very small time intervals. Unlike many economic data which suffer from the problems of small sample, missing values, measurement error and data revision, financial data are usually large in scale, readily available, and well reflect the true transaction prices. However, financial data are usually considered "noisy", meaning that it is more difficult to separate the underlying patterns from random features, and contains "patterns" as a result of the way the market/price recording system works (market micro structure).

1.3 Financial Returns

In finance, most of the time we are dealing with prices and returns on assets. Suppose that P_t is the price of an asset at date t with no dividend payments. The simple net return, R_t, on the asset between dates $t-1$ and t is

$$R_t = \frac{P_t}{P_{t-1}} - 1.$$

The simple gross return, given by the ratio of the price at time t to the price at time $t-1$, is $1 + R_t$. The compound return, $R_t(k)$, on the asset between dates $t-k$ and t, is

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1.$$

The compound return connects to the simple net returns through

$$1 + R_t(k) = (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}).$$

To be able to compare asset returns corresponding to different time intervals, the returns are usually annualized. The annualized compound return is given by

$$\text{Annualized } (R_t(k)) = (1 + R_t(k))^{1/k} - 1 = \left(\prod_{i=0}^{k-1} (1 + R_{t-i}) \right)^{1/k} - 1.$$

Calculating the k-th root could be challenging to human beings (of course not to computers), so we may approximate the annualized returns when the single-period returns are small. We define

$$f(R_t, R_{t-1}, \ldots, R_{t-k+1}) = \left(\prod_{i=0}^{k-1} (1 + R_{t-i}) \right)^{1/k} - 1.$$

Taylor expansion at around $(0, 0, \ldots, 0)$ yields

$$\begin{aligned} f(R_t, R_{t-1}, \ldots, R_{t-k+1}) &\approx f(0, 0, \ldots, 0) + \sum_{i=0}^{k-1} \frac{\partial f(0, 0, \ldots, 0)}{\partial R_{t-i}} R_{t-i} \\ &= \frac{1}{k} \sum_{i=0}^{k-1} R_{t-i}. \end{aligned}$$

That is, we may approximate the annualized compound return as the arithmetic average of the simple net returns.

There is another frequently used compounding method, called continuous compounding. The return r_t corresponding to continuous compounding is the continuously compounded return, or the log return, which is defined through

$$e^{r_t} = 1 + R_t.$$

Consequently, we have

$$r_t = \ln(1+R_t) = \ln \frac{P_t}{P_{t-1}} = \ln P_t - \ln P_{t-1} := p_t - p_{t-1}.$$

The advantage of using log returns is that it has the additivity property:

$$r_t(k) = \ln(1+R_t(k)) = r_t + r_{t-1} + \cdots + r_{t-k+1}.$$

However, a disadvantage of using log returns is that the log return on a portfolio of assets is not exactly the weighted average of the log returns on individual assets. Suppose there are N assets in the portfolio, each with a weight w_{it} and a simple net return R_{it} at time t. The simple return on this portfolio of assets at time t is simply

$$R_t = \sum_{i=1}^{N} w_{it} R_{it}.$$

However, in general,

$$r_t = \ln\left(\sum_{i=1}^{N} w_{it}(1+R_{it})\right) \neq \sum_{i=1}^{N} w_{it} \ln(1+R_{it}) = \sum_{i=1}^{N} w_{it} r_{it}.$$

Fortunately, this problem is mitigated when the log returns are small. Note that we have

$$r_t = \ln\left(\sum_{i=1}^{N} w_{it}(1+R_{it})\right) = \ln\left(\sum_{i=1}^{N} w_{it} e^{r_{it}}\right).$$

Let

$$g(r_{1t}, r_{2t}, \ldots, r_{Nt}) = \ln\left(\sum_{i=1}^{N} w_{it} e^{r_{it}}\right).$$

Taylor expansion at around $(0, 0, \ldots, 0)$ yields

$$g(r_{1t}, r_{2t}, \ldots, r_{Nt}) \approx g(0, 0, \ldots, 0) + \sum_{i=1}^{N} \frac{\partial g(0, 0, \ldots, 0)}{\partial r_{it}} r_{it}$$

$$= \sum_{i=1}^{N} w_{it} r_{it}.$$

Above we assume that there are no dividend payments. When there are, we may adjust the formula for the returns accordingly. For example, if D_t is the dividend payment at time t, the simple net return R_t is

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1.$$

When we talk about returns, we always need to adjust for risks. Therefore, we need a "benchmark" return. The risk-free asset return, R_{0t}, at time t, is the return on an asset that guarantees future repayments. Although there is no asset that is completely risk free in a strict sense, in reality we usually take the return on a short term Treasury bill as the risk-free return. With the risk-free return, we may define the excess return and the log excess return respectively as

$$Z_{it} = R_{it} - R_{0t}$$

and

$$z_{it} = r_{it} - r_{0t}.$$

In a rational market, the excess return on an asset should reflect the risk of that asset.