Empirical Finance: A Review

For Personal Reference

Sai Zhang

April 3, 2021

HERE WE GO!

Empirical finance is an absolutely fascinating field, with some of the most cutting-edge methodologies and the most exploratory techniques. Although it is not my speciality, I am always interested in this literature. During my pre-doc research fellowship at London Business School, I have had the previlege to study in the course *Financial Economics II: Empirical Finance*. The course instructor Dr. Svetlana Bryzgalova is absolutely one of the most brilliant scholars I have encountered. Thanks to her, I have got to understand this liturature more systematically. In this (personal) review, I summarize the most influential and inspirational works in this field and organize them by different topics. The structure of this review resembles the structure of Dr. Bryzgalova's course, while adjusted according to my personal research interest. I intend to review classic works and discuss some potential directions of future study regarding my personal interest in Behavioral Economics, Game Theory and Network.

Since this review is tailored according to my own research interest and experience, I will not only summarize the theoretical perspectives of the studies, present their findings and discuss how they fit into the literature, but document my replication attempts and pseudo codes as well. All the codes related to this review can be found on my Github page.

I thank Dr. Svetlana Bryzgalova for her valuable intuitions and impressive knowledge of the empirical finance literature. Building this review is truly a memorable journey for me. I would love to share this review and all the related materials to anyone that finds them useful. And unavoidably, I would make some typos and other minor mistakes (hopefully not big ones). So I'd really appreciate any correction. If you find any mistakes, please either set up a branch on Github or send the mistakes to this email address saizhang.econ@gmail.com, BIG thanks in advance!

1

Contents

CHAPTER 1

TIME-SERIES PREDICTABILITY

Contents

Every investor knows that trading in financial markets is to play games with time itself. Daily trades determine asset prices at every date and hence influence the random distribution of future prices as well as the initial level of prices. One would need "much more careful attention to the process by which both expected payoffs and required rates of return determine asset prices"[1].

In this chapter, I first, following Campbell (2017, Chapter 5), summarize models mapping cash flows and discount rates into prices using present value relations in Section 1.1. Then I discuss the early evidence for mean reversion in returns in Section

---

[1]See Campbell (2017, p. 121)

1.2. In Section 1.3, I examine the excess volatility puzzle in the predictability debate. To accomodate the stylized facts of time-series predictability, Section 1.4 presents two of the most influential approaches to decompose prices. In Section 1.5, I selectively summarize some researches from the so-called "Prediction Zoo", which satirically describes the floods of price predictors. Finally, I discuss the issues and extensions of time-series predictability in Section 1.6.

1.1   Concepts and models

In this section, I follow Campbell (2017, Chapter 5) and discuss some of the conceptual building blocks for the strand of time-series empirical finance literature.

1.1.1   Market Efficiency

An intuitive way of explaining *market efficiency* is that efficient markets are competitive and allow no easy ways to make economic profit. A more useful and testable definition was given by Malkiel (1989, p. 127):

> The market is said to be efficient with respect to some information set ϕ, if security prices would be unaffected by revealing that information to all participants.

Some event studies that measure market responses to news announcements can be interpreted as tests of market efficiency regarding the announced information, but in general, this definition is not easy to test. On the other hand, Malkiel (1989) gives a more testable alternative:

> Efficiency with respect to an information set ϕ implies that it is impossible to make economic profits by trading on the basis of ϕ.

This is the idea behind an enormous literature in empirical asset pricing: if an economic model defines the equilibrium return as $\Theta_{i,t}$, then the null hypothesis is

$$R_{i,t+1} = \Theta_{i,t} + U_{i,t+1} \tag{1.1}$$

where $U_{i,t+1}$ is a FAIR game regarding the information set at t, or $\mathbb{E}(U_{i,t+1}|\phi_t) = 0$. Notice that market efficiency is equivalent to rational expectations, one must text a model of expected returns as well when testing market efficiency. After defining a model of expected returns, the variables to be included in the information set must be specified. Malkiel and Fama (1970) define three forms of efficient market hypothesis and the corresponding information sets:

  - the *weak form*: past returns
  - the *semi-strong form*: publicly available information such as stock splits, dividends, or earnings

- the **strong form**: information available to some market participants, but NOT necessarily to all participants.

In the time-series literature, the simplest economic model is constant return: $\Theta_{i,t} = \Theta$. In Section 1.2, I summarize the early literature focusing on this model.

Market efficiency has been widely tested and debated, now the most accepted view of market efficiency hypothesis is that it is a useful benchmark but does not hold perfectly. The debates between long-term versus short-term efficiency, micro versus macro efficiency are still and will continue to be heated. Some noticable alternative hypotheses are:

- **High-frequency noise**: market prices are contaminated by short-term noise, which can be caused by measurement errors or illiquidity (bid-ask bounce).
- **Inperfect information processing**: the market reacts sluggishly to information after its releasing
- **Persistent mispricing**: market prices deviate substantially from efficient levels in a LONG time
- **Disposition effect**: individual investors are more willing to sell winning stocks then losing stocks, see Shefrin and Statman (1985) for details.

← this could be tested by using measureable actions (trades or portfolio holdings) of the potentially better informed agents

1.1.2 Model: autocorrelation of returns

The most basic time-series test of market efficiency is to test "whether past deviations of returns from model-implied expected returns predict future return deviations" (See Campbell, 2017, p. 124). The leading approach to do so is to test the autocorrelations.

Starting points:

1. The null hypothesis H_0: the stock returns are i.i.d.
2. The standard error for any single sample autocorrelation equals asymptotically $1/\sqrt{T}$, see Box and Pierce (1970) for a detailed discussion.
3. The standard error would be large, (0.1 if $T = 100$), not so easy to detect small autocorrelation

Any autocorrelation test would have to solve these issues.

1.1.2.1 Q-statistics

Because the stock returns are i.i.d. (H_0), different autocorrelations are uncorrelated with one another. Box and Pierce (1970) calculates a sum of K squared sample auto-correlations:

$$Q_K = T \sum_{j=1}^{K} \hat{\rho}_j^2 \tag{1.2}$$

where $\hat{\rho}_j = \hat{Corr}(r_t, r_{t-j})$. Q is asymptotically distributed χ^2 with K degrees of freedom.

Pros: It solves the problem of the large standard errors.

Cons: It does NOT use the sign of the autocorrelations (squared). What could happen is that the expected reutrns are not constant, instead, they are each individually small but all have the same sign.

1.1.2.2 Variance ratio

One way to take the sign of autocorrelations into consideration is the variance ratio statistic. This statistic was introduced to the finance literature by Lo and MacKinlay (1988) and Poterba and Summers (1988).

The basic setting is: for a holding period K, the log return of this entire period $r_t(K)$ is the sum of all the one-period returns r_{t+i}:

$$r_t(K) \equiv r_t + r_{t+1} + \cdots + r_{t+K-1}$$

and the variance ratio over the period K would be defined as:

$$VR(K) = \frac{Var(r_t(K))}{K \cdot Var(r_t)}$$

If there are not autocorrelations, then the i.i.d. returns would have identical variance in each period from t to $t+K$, and $Var(r_t(K)) = Var(r_t + \cdots + r_{t+K-1}) = Var(r_t) + \cdots + Var(r_{t+K-1}) = K \cdot Var(r_t)$. Thus, $VR(K) = 1$. If we rewrite the definition of the variance ratio as:

$$VR(K) = \frac{Var(r_t(K))}{K \cdot Var(r_t)} = 1 + \underbrace{2 \sum_{j=1}^{K-1} \left(1 - \frac{j}{K}\right) \hat{\rho}_j}_{\text{weighted average of the first } K-1 \text{ sample autocorrelations}} \tag{1.3}$$

Then by comparing $VR(K)$ with 1, we can deduct the direction of the autocorrelations:

$VR(K) > 1$	predominantly **positive** autocorrelations
$VR(K) = 1$	no autocorrelations
$VR(K) < 1$	predominantly **negative** autocorrelations: mean reversion

Notice that the weight term $1 - \frac{j}{K}$ increases as j approaches K.

The asymptotic variance of the variance-ratio statistic, under H_0 (i.i.d. returns), is:

$$Var(\hat{V}R(K)) = \frac{4}{T} \sum_{j=1}^{K-1} \left(1 - \frac{j}{K}\right)^2 = \frac{2(2K-1)(K-1)}{3KT} \xrightarrow{K \to \infty} \frac{4K}{3T} \tag{1.4}$$

When $K \to \infty, T \to \infty$, and $K/T \to 0$ (Priestley, 1981, p. 463), the true return process can be serially correlated and heteroskedastic, but the variance of the variance-ratio is still given as:

$$Var(\hat{V}R(K)) = \frac{4K}{3T} \cdot VR(K)^2 \tag{1.5}$$

Notice that this can be quite large with a large $VR(K)$. This is due to the fact that $K/T \to 0$ is a dangerous assumption because in practice K is often large relative to the sample size. To tackle this, Lo and MacKinlay (1988) develop alternative asymptotics assuming $K/T \to \delta$ where $\delta > 0$.

1.1.2.3   Regression approach

Fama and French (1988) established a regression approach to test AR(K). The basic idea is to regress the K-period return on the lagged K-period return:

$$r_t(K) = \alpha_K + \beta_K r_{t-K}(K) + \epsilon_t^K$$

The coefficient β_K would then be:

$$\beta_K = \frac{Cov[r_t(K), r_{t-K}(K)]}{Var[r_{t-K}(K)]} = 2 \left[\frac{VR(2K)}{VR(K)} - 1 \right] = \frac{2 \sum_{j=1}^{K-1} \left(\frac{\min(j, 2K-j)}{K} \right) \rho_j}{VR(K)} \qquad (1.6)$$

It is clear to see that:

$\beta_K > 0$	predominantly **positive** autocorrelations
$\beta_K = 0$	no autocorrelations
$\beta_K < 0$	predominantly **negative** autocorrelations: mean reversion

1.1.2.4   Other variance-ratio tests

See Charles and Darné (2009) for a review.

1.2   Mean reversion in returns: early evidence

With the models introduced in Section 1.1.2, the mean reversion in stock returns has been examined empirically

1.3   Excess volatility puzzle

[insert text]

1.4   Decomposing prices

[insert text]

1.4.1   Campbell-Schiller decomposition

[insert text]

1.4.2   Lettau-Ludvigson decomposition

[insert text]

1.5   Prediction zoo

[insert text]

1.6   Issues and extensions

[insert text]

1.6.1   Persistency of most regressors

[insert text]

1.6.2   Aggregate predictors without ex-ante choice

[insert text]

1.6.3   Instability in the prediction relation

[insert text]

1.6.4   Measurement

[insert text]

CHAPTER 2

CROSS-SECTION PREDICTABILITY

Contents

Intro:

2.1 Section 1

Section 1:

CHAPTER 3

GMM AND CROSS-SECTION TEST

Contents

Intro:

3.1 Section 1

Section 1:

CHAPTER 4

ADVANCES IN CROSS-SECTION ASSET PRICING

Contents

Intro:

4.1 Section 1

Section 1:

CHAPTER 5

CONSUMPTION-BASED ASSET PRICING

Contents

Intro:

5.1   Section 1

Section 1:

CHAPTER 6

TERM STRUCTURE OF RETURNS

Contents

In this chapter, I summarize the stylized facts and models of intrest rates, and, combining with the time-series and cross-sectional properties of equities, discuss how the term structure of equity can be incorporated into the asset pricing dynamic. Instead of assuming the risk-free rate to be one period, as classic asset pricing models implying in the Euler equations and SDFs, one would expect that an ideal asset pricing model could not only explain the dynamic of equity, but reconcile the property of the term structure of interest rates as well.

The first part of this chapter summarizes studies of risk free bonds and the term structure of this asset class.

6.1 Section 1

Section 1:

CHAPTER 7

LEARNING

Contents

In this chapter, I summarize the learning in empirical finance. This is one of the most cutting-edge research area now.

7.1 Section 1

Section 1:

CHAPTER 8

CURRENCIES: TIME-SERIES AND CROSS-SECTION

Contents

Intro:

8.1   Section 1

Section 1:

CHAPTER 9

INTERMEDIARY-BASED ASSET PRICING

Contents

In this chapter, I summarize the

9.1 Section 1

Section 1:

CHAPTER 10

FINANCE AND BIG DATA

Contents

Intro:

10.1 Section 1

Section 1:

BIBLIOGRAPHY

George EP Box and David A Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332):1509–1526, 1970.

John Y Campbell. *Financial decisions and markets: a course in asset pricing*. Princeton University Press, 2017.

Amélie Charles and Olivier Darné. Variance-ratio tests of random walk: an overview. *Journal of Economic Surveys*, 23(3):503–527, 2009.

Eugene F Fama and Kenneth R French. Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2):246–273, 1988.

Andrew W Lo and A Craig MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1): 41–66, 1988.

Burton G Malkiel. Efficient market hypothesis. In *Finance*, pages 127–134. Springer, 1989.

Burton G Malkiel and Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

James M Poterba and Lawrence H Summers. Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1):27–59, 1988.

Maurice Bertram Priestley. *Spectral analysis and time series: probability and mathematical statistics*. Number 04; QA280, P7. 1981.

Hersh Shefrin and Meir Statman. The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance*, 40(3):777–790, 1985.