Introduction to Empirical Methods

Lars A. Lochstoer
UCLA Anderson School of Management

Winter 2025

Main actors in financial markets

- Institutional investors
- Retail investors
- Firms
- Banks (commercial and broker-dealer)
- Central banks (in U.S., the Federal Reserve)

Common problem

Make decisions today based on expectations about the future.

How do you form expectations about the future?

- Historical data and events?
- Theory?
- Gut feeling?
- All of the above?

Conditional expectations and predictions

Expectations today about a future outcome is a "Conditional Expectation"

- We typically write this as

$$E_t\left(x_{t+j}\right)$$

  where t is today (a variable that counts periods, e.g. daily) and x_{t+j} is the outcome j periods from now

This expectation can only be formed using information available up until time t

- If information arriving at time t is denoted I_t we could alternatively write the above as

$$E\left(x_{t+j}|I_t, I_{t-1}, I_{t-2}, ...\right)$$

We need to come up with the information we want to use $\left\{I_{t-j}\right\}_{j=0}^{\infty}$ and a model that maps the information into a prediction

- A typical example is linear regression: $E\left(y|X\right) = X\beta$, where y is the future outcome and X contains the lagged information you want to use

Investors: the benchmark case

Mean-variance problem at each time t. Find $(N_t \times 1)$ portfolio weights:

$$\min_{\omega_t} \omega_t' V_t \omega_t \quad \text{s.t. } \omega_t' \mu_t = \text{return target (e.g., 7\%)}$$

Notation:

$$\begin{aligned} \underset{(N_t \times N_t)}{V_t} &\equiv Var_t \left(\underset{(N_t \times 1)}{R_{t+1}^{excess}} \right) = E_t \left(\left(R_{t+1}^{excess} - E_t \left(R_{t+1}^{excess} \right) \right)^2 \right), \\ \underset{(N_t \times 1)}{\mu_t} &\equiv E_t \left(R_{t+1}^{excess} \right). \end{aligned}$$

- "Excess" means in excess of a reference return (e.g., 1-month T-bill rate)
- The t in Var_t and E_t refers to a *conditional* variance (covariance) and expectation, respectively
  - That is, as of time t, with the information available to you, what are your forward-looking expectations
- I.e., need to forecast expected excess returns on N_t assets and the $(N_t + 1) N_t / 2$ elements in the covariance matrix (variance and correlations)

Active mutual funds: A fundamentals-based manager

"Our investment professionals don't just sit behind their screens, they go out into the field to get the answers needed for your investments. They're constantly on the lookout, analyzing the markets and the companies within them.

Classic "stock-picking" talk. But, in the end they need to make decisions on portfolio weights

- Solution to last slide's problem:

$$\omega_t = \text{constant} \times V_t^{-1} \mu_t$$

- Thus, in the end, they make a quantitative decision on a forecasting problem

Active mutual funds: A classic quant-based manager

"*While we're known as a pioneer in quantitative investing, we also have considerable expertise in areas that involve fundamental analysis or portfolio manager discretion. We use a combination of quantitative and qualitative tools to uncover independent, hard-to-find sources of return across global public and private markets.*"

"*Within the D. E. Shaw group, we run on collaboration, not internal competition. Teams work together to share trade ideas, identify and address risks, build tools, and explore new opportunities. Our staff includes world-class mathematicians, physicists, computer scientists, economists, analysts, business-builders, and system architects relying on specialized trading, operational, and compliance expertise developed over 30 years.*"

Benchmarking and the portfolio choice problem

Most managers have a benchmark index against which their investments are measured

This leads a minor modification of the portfolio choice problem.

- Instead of $R_t^{excess} = R_t - R_{f,t}$, we let $R_t^{excess} = R_t - R_{idx,t}$

Thus, risk is meaured as "Tracking Error":

$$TrackingError = \sigma\left(R_{fund,t} - R_{idx,t}\right)$$

where $\sigma\left(\cdot\right)$ is short-hand for $\sqrt{Var\left(\cdot\right)}$

Risk-reward ratio often expressed as an "Information Ratio"

$$\text{Information Ratio } = \frac{E\left(R_{fund,t} - R_{idx,t}\right)}{\sigma\left(R_{fund,t} - R_{idx,t}\right)}$$

Project (or firm) valuations: capital budgeting

Classic NPV analysis:

$$Value_t = \sum_{j=1}^{\infty} \frac{E_t\left(CF_{t+j}\right)}{\left(1 + DiscountRate_{t+j}\right)^j}$$

- Need to forecast current expecations of cash flows and the right discount rate at each horizon j

This is also relevant for Private Equity and Venture Capital analysis, obviously

Banks

Is lending or trading worth the cost of capital (often regulatory constraints)?

$$\text{Value of position} = \text{expected benefit from trade - cost of financing}$$

For instance, deleveraging due to cost of leverage (market and/or regulatory) can force sales of assets which can drive leverage cycles

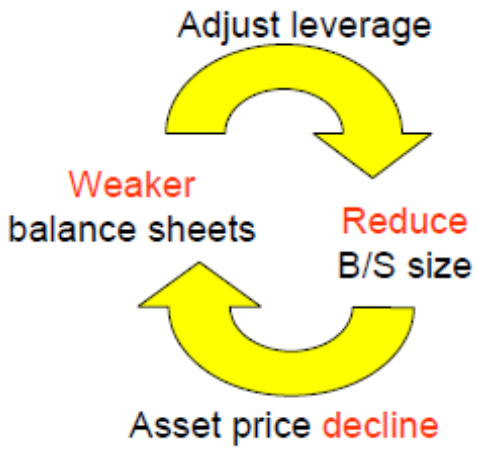


Figure 2.6: Leverage Adjustment in Downturn

Central Banks

The Fed has a joint employment, inflation, and financial stability goal

- Forecast expected future aggregate economic activity
- Input to FOMC decision on whether to adjust interest rates or not
- Forecast effect of interest changes on economic activity (need model here)

Like standard NPV analysis, need to forecast distribution of outcomes not just one period ahead, but for multiple periods!

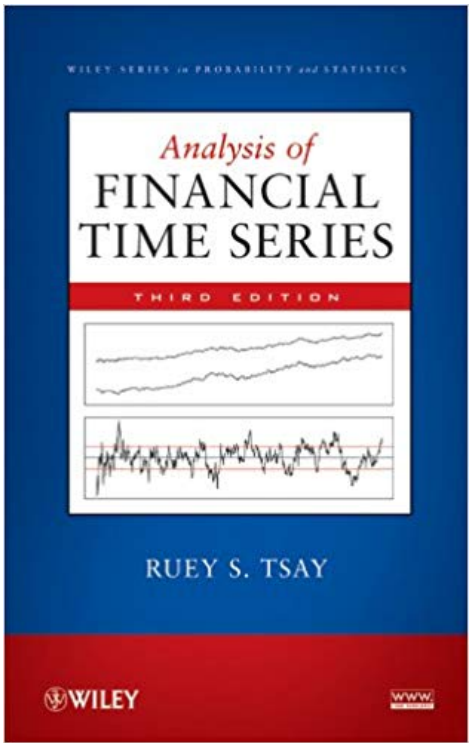
This class

Forecasting

- Multiple variables
- Multiple horizons

Base material: standard quant time series analysis techniques

- Book for class



Topic 1: Fundamental facts and concepts

Time series properties of stock market returns and prices

- Class intro: Forecasting and Finance
- Properties of asset returns
- Risk and leverage
- Estimating means and variances: sample length and frequency
- The random walk hypothesis
- Stationarity
- Time-varying volatility and General Least Squares
- Robust standard errors and OLS

Topic 2: Time-dependence and predictability

- ARMA models
- The likelihood function, exact and conditional likelihood estimation
- Predictive regressions, autocorrelation robust standard errors
- The Campbell-Shiller decomposition: a useful present value model
- Multivariate analysis
  - Vector Autoregression (VAR) models, the Kalman Filter
  - Cointegration

Topic 3: Volatility (Heteroskedasticity)

- Time-varying volatility in the data
- Realized Variance
- ARCH and GARCH models
- State-of-the-art multi-frequency models

Topic 4: The cross-section of stock returns

- Single- and multifactor models
- Economic factors: Models and data exploration
- Statistical factors: Principal Components Analysis
- Fama-MacBeth regressions and characteristics-based factors

Class admin

Grade

- 60% final, 32% homeworks, 8% class participation

Class materials

- Textbook
- Class notes posted on BruinLearn each week
- Reference to textbook posted on BruinLearn each week

Homeworks

- Use your pre-assigned groups
- *Everyone* must have done homework even though you submit it as a group
  - I discuss the homework each week *in class*, so having done it is an integral part of learning in the lectures